

Varieties of indefinites in Cantonese¹

Christopher DAVIS — *University of the Ryukyus*

Zoe Pei-Sui LUK — *The Education University of Hong Kong*

Grégoire WINTERSTEIN — *Université du Québec à Montréal*

Abstract. We focus on the semantics of three types of Cantonese nominal constructions that can refer to indefinite referents. We argue that the indefinite interpretation is derived by a different semantic mechanism in each construction. The evidence for this claim comes from the different behavior of these constructions in terms of their scope-taking characteristics and their (in)compatibility with specific indefinite interpretations. Specifically, we make the following claims: (i) [BARE N] phrases denote type $\langle e, t \rangle$ properties, and get an indefinite interpretation via type-shifting, (ii) [CL N] and [*jat1* CL N] phrases are choice-functional indefinites, and (iii) the choice-function variable in [CL N] phrases can be left unbound, allowing for definite as well as (specific) indefinite uses, depending on context.

Keywords: Cantonese, Indefinites, Choice Functions, Type-Shifting, Bare Nouns, Bare Classifiers.

1. Introduction

This paper focuses on three types of nominal expression in Cantonese, which we refer to as *bare noun phrases* [BARE N], *bare classifier phrases* [CL N], and *jat1 phrases* [*jat1* CL N]. All three constructions are compatible with indefinite interpretations. We argue that each type of nominal gets an indefinite interpretation via a distinct semantic route, the details and motivations of which are spelled out in sections 3 and 4. In section 2, we begin by laying the empirical domain of our study and give more details about each construction, including other interpretations available to each. Section 5 situates the current work within the literature on Cantonese nominal expressions, especially in the light of the indefinite or definite interpretation allowed by those constructions. Section 6 concludes, notably by laying the ground for future work on non-indefinite interpretations of these constructions.

2. Empirical domain

The sentences in (1) form a minimal triplet exemplifying the three constructions whose interpretations are the focus of this paper. As indicated by their translations, each of the target structures, highlighted in bold in the examples, is compatible with an indefinite interpretation, though each construction allows for additional interpretations which we also indicate in the translations.

¹We would like to thank the audiences at the Meaning & Grammar Research Group of University of Edinburgh and Sinn und Bedeutung 28 for their invaluable comments on this work, as well as Teddy Robin for inspiration.

- (1) a. Bare noun phrase ([BARE N])
 我 有 扇 呀
 ngo5 jau5 **sin3** aa3
 1SG have fan SFP
 ‘I have a/some fan(s).’
- b. Bare classifier phrase ([CL N])
 我 有 把 扇 呀
 ngo5 jau5 **baa2 sin3** aa3
 1SG have CL fan SFP
 ‘I have a/the fan.’
- c. *jat1* phrase ([*jat1* CL N])
 我 有 一 把 扇 呀
 ngo5 jau5 **jat1 baa2 sin3** aa3
 1SG have JAT1 CL fan SFP
 ‘I have a fan.’

As will become clear in the course of the paper, Cantonese does not have any element that can be described as a marker of definiteness akin to English *the*, French *le • la*, or any comparable element studied in the vast literature on (in)definiteness. Therefore, giving a definition of (in)definiteness in Cantonese is not straightforward, since those values cannot be defined on the basis of specific forms (though we eventually argue that one variety of indefiniteness is formally marked in Cantonese). There is ample debate in the literature about how to approach these notions semantically and pragmatically. Here, we will define (in)definiteness in terms of the status of the referent of a nominal expression. Essentially, we will consider a referent to be indefinite if it is not shared by the hearer. In other terms, and much in line with notions in dynamic semantics (e.g. Heim (1983); Groenendijk and Stokhof (1989); Kamp and Reyle (1993)), we take nominal expressions to be indefinite if their referent is newly introduced in the discourse (and thus to the hearer), irrespective of whether the speaker has a specific referent in mind.

On a related terminological note, when discussing the meaning of the constructions under study, we will discuss their *interpretations*, by which we mean the way a putative hearer is able to understand the informational status of the referent of the targeted nominal. We also discuss *indefinite uses* of these phrases, by which we mean their ability to be interpreted indefinitely. Crucially, what we are interested is whether particular constructions are compatible with particular kinds of contexts, especially in terms of the properties of their discourse referents. Thus when we mention that a nominal phrase has an indefinite interpretation or an indefinite use, this is to be understood as shorthand for the fact that the referent of that phrase can be interpreted indefinitely in a particular context. We use *reading* as a technical term that applies to the distinct meanings of a semantically (and often syntactically) ambiguous sentence. One goal of this paper is to argue that distinct interpretations of the constructions under consideration are not in general due to semantic or syntactic ambiguity, but rather to underspecification.

A final note: the judgments regarding (un)available interpretations reported in this paper are based on the intuitions of the second author, who is a native speaker of Hong Kong Cantonese.

In the following subsections we discuss the possible interpretations of each of the three constructions exemplified above in more detail, and provide attested examples of indefinite uses of each construction.

2.1. Bare noun phrases

[BARE N] phrases are often used for indefinite reference of an underspecified number. For example, the sentence in (1a) could be used in a context in which the speaker has bought one or more than one fan, and in which these fans are not familiar to the hearer. A corpus example is shown in (2), for which, in the context of the utterance, the speaker might have kept one rabbit or more as (a) pet(s).²

- (2) 我 養.過 兔仔 啊
 ngo5 jeong5.gwo3 **tou3zai2** aa3
 1SG raise.EXP rabbit SFP
 ‘I once kept a rabbit.’ HKCanCorp (Luke and Wong, 2015)

[BARE N] phrases can also be interpreted as indefinite and non-specific. In (3), *pang4jau5* ‘friend’ does not refer to any specific individual, but to a class of individuals.

- (3) 唔.鍾意 話 好似 賣 高 些少 嘅 賣 俾 人, 嗰啲
 m4.zung1ji3 waa6 hou2ci5 maai6 gou1 se1siu2 maai6 gam2 bei2 jan4, go2di1
 NEG.like say like sell high a.bit sell like to people those
 好似, 好似 喺 朋友 身上 搵錢 嘅 嘅啫
 hou2ci5, hou2ci5 hai2 **pang4jau5** san1seong6 wan2cin2 gam2 ge3 ze1
 like like on friend body make.money SFP SFP SFP
 ‘(I) don’ t want to be, like, selling it at a higher price, like ripping off a friend.’ HK-CanCorp (Luke and Wong, 2015)

When the referent is unique (e.g. astral objects), bare nouns often receive a definite interpretation. *Jyut6kau4* ‘moon’ in (4) is a bare noun and it receives a definite interpretation.

- (4) 阿姆斯壯 喺 1969 年 登陸 月球
 aa3mou5si1zong1 hai2 1969 nin dang1luk6 **jyut6kau4**
 Armstrong in 1969 year land moon
 ‘Armstrong landed on the moon in 1969.’

Finally, as shown in (5), bare nouns can also receive a kind or generic interpretation.

- (5) 貓 食 老鼠
 maau1 sik6 lou4syu2
 cat eat mouse
 ‘Cats eat mice.’

²Most of the examples in this section are taken from the HKCanCorp corpus of Luke and Wong (2015). This is indicated with each example.

2.2. Bare classifier phrases

[CL N] phrases can receive both indefinite and definite interpretations. With most classifiers (except for the classifier *di1*, which will be discussed below), [CL N] phrases are interpreted as singular. In (6) and (7), the noun phrases *zek3syun4* ‘CL ship’ and *zoeng1pei5* ‘CL quilt’ are both interpreted as singular indefinites.

- (6) 即係 佢 會 有-有 隻 船 出海
 zik1hai6 keoi5 wui5 jau5-jau5 zek3 syun4 ceot1hoi2
 that.is 3SG will have-have CL ship cruise
 ‘... that is, they will have a ship cruising.’ HKCanCorp (Luke and Wong, 2015)
- (7) 唉, 我 攞 張 被 俾 你 啊!
 ai1 ngo5 lo2 zoeng1 pei5 bei2 nei5 aa1
 Oh.dear 1SG bring CL quilt to you SFP
 ‘Oh dear. I’ ll bring you a quilt.’ HKCanCorp (Luke and Wong, 2015)

On the other hand, the noun phrase *gaa3ce1* ‘CL car’ in (8) would typically be given a singular definite interpretation. The listener is likely to interpret the car as being the only (unique) car that the speaker owns. However, if the speaker is known to own many cars, the indefinite interpretation becomes stronger than the definite interpretation. In (9) it is also the definite interpretation that is the most obvious.

- (8) 我 賣咗 架 車
 ngo5 maai6zo2 gaa3 ce1
 I sell.PFV CL car
 ‘I sold (the/my) car.’
- (9) 你 喺 個 海 嗰度 你 梗係 着 泳褲 喇, 唔係
 nei5 hai2 go3 hoi2 go2dou6 nei5 gang2hai6 zok6 wing6fu3 laa1 m4hai6
 2SG COP CL sea there 2SG of.course wear swimming.trunks SFP NEG.COP
 着 乜嘢 啊 .
 zok6 mat1je5 aa3
 wear what SFP
 ‘If you’re at the sea, of course you’ll be wearing swimming trunks, if not, what else (would you wear)?’ HKCanCorp (Luke and Wong, 2015)

For the purposes of our paper, we do not distinguish between sortal and measure classifiers (Ahrens and Huang, 2016), given that they do not differ in terms of how they affect the informational status of the referents of the targeted constructions.³ Nevertheless, one classifier deserves special mention: the so-called ‘plural classifier’ *di1* (啲). Indeed, if we use that classifier instead of the sortal *gaa3* in (10), the phrase is interpreted as plural:

³This does not mean that we believe that the two types of classifiers necessarily have the same kind of semantics, e.g. in the mechanics of their atomizing and selection functions.

- (10) 我 賣咗 啲 車
 ngo5 maai6zo2 gaa3 ce1
 I sell.PFV DI1 car
 ‘I sold (the/my) cars.’

Contrary to the other classifiers, *di1* does not have selectional properties: it can combine with any noun, including mass nouns. Like many other markers of plurality cross-linguistically, *di1* is also semantically an inclusive plural. This can be seen in (11), where the sentence is compatible with situations in which the speaker’s mother occasionally brings only a single fan back from her Kyoto trips.

- (11) 我 阿媽 每次 去 京都 都會 買 啲 扇 返嚟
 ngo5 aa3maa1 mui5.ci3 heoi3 ging1dou1 dou1 wui5 maai5 di1 sin3 faan1lai4.
 my mother every.time go Kyoto PRT will buy DI1 fan back.come
 ‘Every time my mother goes to Kyoto she brings back fans.’

2.3. *jat1* phrases

[*jat1* CL N] phrases are built by combining a classifier and noun with the term *jat1*, which at first glance corresponds to the numeral ‘one’. Unlike bare nouns and bare classifiers, *jat1* phrases can only receive an indefinite interpretation, which can be either specific or non-specific. In (12), the speaker draws the hearer’s attention to a new mosquito bite on the hearer and introduces it into the conversation, so that *jat1 dat3 man1laan3* ‘one CL mosquito bite’ is indefinite and specific.

- (12) 嘩 你 隻手 已經 搵到 - 死喇, 有 一 隻 蚊 嚙
 waa3 nei5 zek3 sau2 ji5ging1 we2dou3 sei2 laa1 jau5 jat1 dat3 man1laan3 laa3
 wow you CL hand already scratch.to die SFP have JAT1 CL mosquito.bite SFP
 ‘Wow, your hand is already scratched to... Geez, you have a mosquito bite.’ HKCan-Corp (Luke and Wong, 2015)

In (13), *jat1 coet1 hei3* ‘one CL movie’ is indefinite and non-specific, as the speaker is not referring to a particular movie, but describing a hypothetical situation.

- (13) 即係 你 唔係 成日 睇 呢, 睇 一 齣 戲 呢, 就
 zik1hai6 nei5 m4hai6 sing4jat6 tai2 ne1 tai2 jat1 coet1 hei3 ne1, zau6
 that.is 2SG NEG.be always watch SFP watch JAT1 CL movie SFP then
 覺得 佢 好 正
 gok3dak1 keoi5 hou2 zeng3
 think 3SG very amazing
 ‘... that is, you do not always watch a movie and instantly think it is amazing.’ HK-CanCorp (Luke and Wong, 2015)

Note that with most classifiers, a *jat1* phrase gives rise to a singular interpretation. When the classifier and noun are combined with other numerals, the cardinality changes accordingly, cf. (14).

- (14) 我 有 三 把 扇 呀
 ngo5 jau5 **saam1** **baa2** **sin3** aa3
 1SG have three CL fan SFP
 ‘I have three fans.’

In these cases, *jat1* can be considered to have its ‘standard’ numeral semantics, meaning ‘one’ (and getting the same range of indefinite interpretations available to other numerals). However, unlike other numerals, *jat1* can also be combined with the plural classifier *dil*, in which case the phrase is no longer singular. An example is given in (15) which involves reference to a plurality of doctors.

- (15) 阿明 睇.過 一 啲 醫生
 aa3ming4 tai2.gwo3 **jat1** **dil** **jilsang1**
 A-Ming see.EXP JAT1 DI1 doctor
 ‘A-Ming saw a few doctors.’

As noted above, [CL N] phrases formed with classifiers other than *dil* are themselves necessarily singular. We thus consider the singular interpretation to stem from the use of the classifier itself (as in [CL N] constructions) rather than directly from the semantics of *jat1*, in spite of its relationship with the numeral ‘one’. In short, we think *jat1* is ambiguous between a ‘normal’ numeral meaning and a formal element that, as we discuss in more detail below, gives rise to a particular variety of indefinite interpretation.

2.4. Summary

The observations above are summarized in table 1.

	Number	Indefinite	Definite
[BARE N]	Undersp.	✓	~ (for unique ref.)
[CL N]	SG with sortal classifier	✓	✓
	PL with <i>dil</i> classifier	✓	✓
[<i>jat1</i> CL N]	SG with sortal classifier	✓	✗
	PL with <i>dil</i> classifier	✓	✗

Table 1: Summary of the possible interpretations for the three target nominal constructions

The table makes it clear that even though the three constructions differ in terms of their compatibility with singular, plural, and definite interpretations, they all share the possibility of being interpreted indefinitely. These indefinite interpretations and the details of their semantics are the focus of the rest of the paper.

3. Bare nouns get low-scope indefinite interpretations via type-shifting

This section contrasts the interpretative possibilities of [BARE N] phrases with those of [CL N] and [*jat1* CL N] phrases, arguing that the former are necessarily low-scope, while the latter are semantically compatible with a full range of scopal possibilities. These facts are laid out in section 3.1. The fact that [BARE N] gets only a low-scope indefinite interpretation is in line with cross-linguistic observations in the literature,⁴ in which indefinitely interpreted bare nouns receive only low-scope interpretations, contra other types of indefinites. This fact in turn is derived by treating the indefinite interpretation of bare nouns as arising from a type-shifting rule, which applies in such a way that non-surface scope interpretations are never generated. The details of such a proposal are given in section 3.2. The semantic details of [CL N] and [*jat1* CL N] are taken up in section 4.

3.1. Bare nouns as low-scope indefinites

Empirical support for the low-scope restriction on [BARE N] indefinites begins with the behavior of the three constructions in negated sentences like the one in (16). The felicity of each construction was checked relative to a context in which the speaker failed to bring any fans at all (the low-scope context), and one in which the speaker brought some fans but failed to bring some other(s) (the wide-scope context). As seen in table 2, the bare noun is compatible only with a low-scope interpretation relative to negation, while the other two constructions receive only a wide-scope interpretation relative to negation.

(16) Negation and scope

我 冇 帶 ((一) 把) 扇 嚟
 ngo5 mou5 daai3 ((jat1) baa2) sin3 lai4
 1SG NEG bring ((JAT1) CL) fan come

	[BARE N]	[CL N]	[JAT1 CL N]
<i>'I did not bring any fans.'</i> (low scope)	✓	#	#
<i>'There is a fan I did not bring.'</i> (high scope)	#	✓	✓

Table 2: Scope possibilities for the three nominal constructions with descriptive negation (free translations meant as a way to indicate the target interpretation)

As already noted, we analyze the indefinite interpretation of bare nouns as deriving from a type-shifting rule that is required to apply locally, giving only a low-scope reading. Turning to the other two constructions, the *unavailability* of the low-scope interpretation we might analyze as deriving from competition with the (unambiguously low-scope) bare noun alternative. Assuming for the moment that the [CL N] and [*jat1* CL N] constructions are *semantically* compatible with both a low-scope and a high-scope interpretation, we might derive the *pragmatic*

⁴See e.g. Dayal and Saĝ (2020) for discussion and references

unavailability of the low-scope interpretation through competition. Since the bare noun is unambiguously low-scope, and also less formally marked, the preference for its being used in the low-scope context seems reasonable from a pragmatic perspective.

However, examples like the one in (17) make this explanation untenable. Here, the nominal appears with an intensional predicate, and once again the bare noun is compatible with only a low-scope (i.e. *de se*) interpretation. The other two constructions are compatible with *both* low and wide-scope (i.e. *de re*) interpretations. Concretely, the low-scope interpretation is tested against a context in which A-Ming does not have particular marriage prospect in mind, but loves astronauts and wants whomever he marries to be one. The wide-scope interpretation is checked relative to a context in which A-Ming wants to marry a particular person, who happens to be an astronaut, despite A-Ming being unhappy with that career choice (it being more dangerous than A-Ming would prefer). The judgments relative to these contexts are summarized in table 3.

(17) Scope with intensional predicates

阿明 想 娶 ((一) 個) 太空人
 aa3ming4 soeng2 ceoi2 ((jat1) go3) taai3hung1jan4
 A-Ming want marry ((JAT1) CL) astronaut

‘A-Ming wants to marry an astronaut.’

	[BARE N]	[CL N]	[JAT1 CL N]
‘A-Ming loves astronauts.’ (low scope)	✓	✓	✓
‘A particular person happens to be an astronaut.’ (high scope)	#	✓	✓

Table 3: Scope possibilities with an intensional predicate

Thus, rather than a more general competition with the bare noun construction (which would lead, incorrectly, to a prediction that they would be infelicitous with low-scope interpretations in (18)), we think that, like English indefinites headed by ‘some’, [CL N] and [jat1 CL N] are subject to a PPI-like restriction on their distribution, leading to incompatibility with low-scope interpretations in sentences with negation like (17). We leave to future research further explication of this restriction.

Turning to sentences with universal quantification and modality, the contrast seen in (16) and (17) seems to disappear:

(18) a. Universal quantifier

個個人 都 睇咗 ((一) 本) 書
 go3go3jan4 dou1 tai2zo2 jat1 bun2 syu1
 every.person all read.PRF ((jat1) CL) book

‘Everyone read a book.’

b. Modal

我 要 喺 星期日 之前 睇 ((一) 本) 書
 ngo5 yiu3 hai2 sing1kei4jat6 zi1cin4 tai2 jat1 bun2 syu1
 I need by sunday before read ((jat1) CL) book
 ‘I have to read a book by Sunday.’

In these two examples, we can imagine a context in which the particular book (to be) read is not specified; in (18a) each person can have read a (potentially) different book, and in (18b) the speaker is able to choose what book they will read. These are the low-scope contexts, and all three constructions are compatible with these contexts. Alternatively, we might be in a wide-scope context, where there is a particular book that everyone read, or in which there is a particular book the speaker must read. Again, all three constructions are compatible with these contexts. In other words, there seems to be no distinction in these examples between the three constructions.

To account for the contrast in judgment patterns between the examples in (18) and those in (16-17), we argue that the apparently ‘wide-scope’ interpretation of the bare noun in (18) is an illusion. Semantically, as suggested above, we propose that [BARE N] phrases are *always* low-scope, resulting from the inherently low-scope semantics of indefinite type-shifting. In (16), the semantics of [BARE N] results in a truth-conditional incompatibility with the wide-scope context. The semantics of [BARE N] here requires that there not exist any fan that the speaker brought. The wide-scope context is one in which the speaker has brought at least some fans, but not all. The low-scope semantics of [BARE N] thus gives rise to falsity in this context, and hence the sentence is incompatible with this interpretation. Similarly in (17), we take it that a low-scope indefinite *reading* is necessarily interpreted *de se* with respect to the intensional predicate. To get a *de re* interpretation like that implied by the wide-scope context, we require (by hypothesis) the indefinite to scope over the intensional predicate. Bare nouns are thus predicted, correctly, to be incompatible with wide-scope/*de re* interpretations.

But this truth-conditional incompatibility does not hold for the examples in (18). The low-scope semantics of the bare noun construction will not lead to falsity in wide-scope contexts here (i.e., ones where there is a particular book that everyone read, or in which there is a particular book the speaker must read). Another way to say this is that the wide-scope *reading* of such a sentence would *entail* the low-scope reading; as such, any situation verifying the wide-scope reading will also verify the low-scope reading, and thus the low-scope semantics posited for the bare noun construction will be true in any situation where a wide-scope semantics would be true.

If we alter the examples in (18) in such a way that the the wide-scope reading no longer entails the low-scope reading, then our original contrast reemerges. This is exemplified by example (19), about which the judgments given in table 4 mirror those in table 3.

- (19) 阿明 要 娶 ((一) 個) 美國人
 aa3ming4 jiu3 ceoi2 ((jat1) go3) mei5gwok3jan4
 A-Ming need marry ((JAT1) CL) american
 ‘A-Ming has to marry an American.’

	[BARE N]	[CL N]	[JAT1 CL N]
<i>‘Marrying an American is A-Ming’s only solution to stay in the USA.’</i> (low scope)	✓	✓	✓
<i>‘A-Ming has been engaged at birth with an American.’</i> (high scope)	#	✓	✓

Table 4: Scope possibilities for the three nominal constructions with an intensional predicate as in (19)

3.2. Formal analysis: bare nouns and type-shifting

We begin our analysis by positing that nouns in Cantonese denote number-neutral properties. That is, a noun *N* will denote a type $\langle e, t \rangle$ property P^* that is true of any atomic individual for which $P(x)$ holds, as well as sums of such individuals. Cantonese nouns are thus in effect inclusive plurals (cf. the discussion and references in Little et al. (2022)). This kind of denotation applies straightforwardly to bare nominals in predicate position, where they take the type e subject referent as argument. Since the property is number-neutral, the subject referent can be either singular (20a) or plural (20b):

- (20) a. 我 係 道士
ngo5 hai6 dou6si2
1SG COP daoist.priest
‘I am a Daoist priest.’
- b. 我哋 係 道士
ngo5dei6 hai6 dou6si2
1PL COP daoist.priest
‘We are Daoist priests.’

Kind readings (which are beyond the scope of this paper) can be derived through type-shifting, using the down operator \cap of Chierchia (1998):

- (21) For any property P and world/situation s ,
- $$\cap P = \begin{cases} \lambda s \lambda x [P_s(x)], & \text{if } \lambda s \lambda x [P_s(x)] \text{ is in the set of worlds,} \\ \text{undefined otherwise.} \end{cases}$$
- where P_s is the extension of P in s .

In our analysis, low-scope existential readings of bare nouns are also derived by type-shifting. One concrete option from the literature is to use the Derived Kind Predication (DKP) rule of Chierchia (1998) (see Deal and Nee 2018 for an accessible overview and summary). Since we treat Cantonese nouns as type $\langle e, t \rangle$ properties, this would be a two-step process, whereby the $\langle e, t \rangle$ property would first be type-shifted to a kind-denoting type e individual, using Chierchia’s down operator \cap . When fed to a non-kind-selecting predicate, we would then apply the DKP rule in (22a), which involves mapping the kind back to a property using the complementary operator \cup , defined in (22b). This results in a lowest-scope indefinite interpretation, under the

assumption that type-shifting rules like DKP apply as locally as possible (see Chierchia 1998 and Krifka 2003 for discussion).

- (22) a. Derived Kind Predication (DKP)
 If P applies to objects and k denotes a kind, then $P(k) = \exists x[\cup k(x) \wedge P(x)]$
- b. Let d be a kind. Then for any world s ,
- $$\cup d = \begin{cases} \lambda x[x \leq d_s], & \text{if } d_s \text{ is defined} \\ \lambda x[\text{FALSE}] & \text{otherwise.} \end{cases}$$
- where d_s is the plural individual that comprises all of the atomic members of the kind.

The route mapped above is a rather circuitous one, and has been criticized by Krifka (2003), who argues instead that indefinite bare nouns (which for him denote properties, as here) are derived directly by an existential type-shifting rule (cf. Partee 1987), to resolve type clashes when a type $\langle e, t \rangle$ nominal appears in a type e argument position. To ensure a low-scope semantics, the existential type-shift operation is, like DKP above, required to apply maximally locally (or, procedurally speaking, as late possible in the derivation).

As far as we can tell, either of the two paths above will get us where we need for Cantonese, although we think Krifka's approach is more straightforward. We could even (as noted by Krifka 2003) follow van Geenhoven (1998) and have type-shifting apply to the predicate instead, again resulting in a low-scope existential interpretation. For us, the important points are (i) nouns denote type $\langle e, t \rangle$ properties, and (ii) [BARE N] in argument position is interpreted via type-shifting, which in the case at hand leads to a necessarily low-scope indefinite interpretation. There are of course other type-shifting options available. As already noted, the \sqcap operator can be used to give a kind interpretation. [BARE N] can also receive a unique definite interpretation, as noted in the previous section. We take this interpretation to be derived by type-shifting as well, but leave details for future research.

4. Classifiers and choice functions

The interpretative possibilities seen so far empirically distinguish [BARE N] phrases on the one hand from [CL N] and [*jat1* CL N] phrases on the other. We have seen that [CL N] and [*jat1* CL N] phrases show two systematic differences with [BARE N] phrases. First, [CL N] and [*jat1* CL N] exhibit a number restriction. With a 'standard' classifier (that is, a classifier that would be used in a numeral construction), the interpretation is necessarily singular, in contrast to [BARE N], which is number neutral. A weak plural interpretation can be generated for [CL N] and [*jat1* CL N] phrases by instead using the 'plural' classifier *dil*, which is not possible in a numeral construction. Second, [CL N] and [*jat1* CL N] have been seen to exhibit a full range of scopal possibilities, with the exception that they cannot appear under the scope of negation, which we attributed to a positive polarity effect. We take these two differences (number and scope) in turn, and sketch an analysis of [CL N] and [*jat1* CL N] phrases that accounts for these differences. Space limitations prevent us from rendering the sketch fully, an enterprise we plan to undertake in future work.

4.1. Number properties

We take it that classifiers in Cantonese combine, syntactically and semantically, with nouns, with Cantonese thus being a ‘classifiers-for-nouns’ language (see Little et al. (2022) for a cross-linguistic discussion). Space limitations prevent us from justifying this stance in detail, but the very existence of [CL N] phrases is good preliminary evidence for this view (since we would otherwise need to posit a null numeral in these structures). The fact that [CL N] and [*jat1* CL N] phrases with standard classifiers are necessarily singular we take to be a function of the semantics of the classifiers themselves. Following Bale et al. (2019) and Little et al. (2022), we take it that classifiers in Cantonese have an ‘atomizing’ function, taking the number-neutral property denoted by its nominal complement and returning the set of atoms, in effect making the denotation singular. In a numeral classifier construction, this atomizing is a necessary ingredient to allow for counting (following Bale et al. (2019), who adopt the theory of numerals proposed by Ionin and Matushansky (2006). Here, however, where no numerals are used, the effect is simply to make the resulting set atomic, and thus singular. [CL N] and [*jat1* CL N] phrases in Cantonese thus provide evidence for this approach to the semantics of classifiers in their canonical use in numeral classifier constructions.

As discussed earlier, [CL N] and [*jat1* CL N] phrases with the ‘plural’ classifier *di1* are interpreted as (weak) plurals. The classifier *di1* is unusual in that (i) it cannot be used with numerals, and (ii) it does not select for noun type (i.e. it does not really ‘classify’ the nouns at all). These properties, we posit, are linked to the semantic plurality that *di1* marks in bare classifier constructions. Given that atomization is necessary for counting in Cantonese (again following the proposals in Bale et al. (2019) and Little et al. (2022)), a ‘plural’ classifier (which does not restrict its complement to a set of atoms) will not be useable for counting. Moreover, given that the ‘classification’ function of classifiers (i.e. their sensitivity to the semantic class of the noun with which they can be combined) is itself related to this atomizing function, a non-atomizing classifier such as *di1* will be useable with all types nouns.

Semantically, then, we think that in both [CL N] and [*jat1* CL N] phrases, the classifier combines with a type $\langle e, t \rangle$ noun (which denotes both atoms and sums), and returns either (i) the subset of atomic entities, in the case of a standard sortal classifier, or (ii) the original set, in the case of *di1* (we put aside the semantics of measure classifiers here). Note that this proposal makes *di1* semantically null, with its function presumably a syntactic one, given that these constructions syntactically require a classifier.

4.2. Scope properties

We have already seen that [CL N] and [*jat1* CL N] phrases have a range of scopal interpretations available. Turning to sentences involving scope islands, we find that both [CL N] and [*jat1* CL N] can receive island-violating wide-scope interpretations, as illustrated in (23).

- (23) 如果 你 上完 ((一)個) 課程, 你 就 會 畢到業
 jyu4gwo2 nei5 soeng5jyun4 jat1 go3 fo3cing4, nei2 zau6 wui5 bat1dou2jip6
 if you take.finish JAT1 CL course you then will able.to graduate

- ‘If you finish a course, you will be able to graduate.’
 ‘Any one will do.’ ([BARE N],[CL N], [*jat1* CL N])
 ‘But I don’ t know which course it is.’ ([CL N] or [*jat1* CL N])

The relevant contexts here are one where the addressee can graduate if they finish any course (the low-scope context), and one where there is a particular course they have to finish in order to graduate (the wide-scope context). All three constructions are compatible with the low scope interpretation, although this interpretation seems to be marked/dispreferred for [CL N]. For the wide-scope interpretation, the impossibility of [BARE N] is as expected. Both [CL N] and [*jat1* CL N] are compatible with the wide scope interpretation, showing that both of these constructions can be given island-violating scopal interpretations.

The example in (24) shows that the antecedent of the conditional in (23) is indeed an island for quantifiers in Cantonese. Here, a universally quantified noun phrase appears in the antecedent of the conditional, and only the low-scope interpretation is possible.

- (24) 如果 你 上完 每 個 課程, 你 就 會 畢到業
 jyu4gwo2 nei5 soeng5jyun4 mui5 go3 fo3cing4, nei2 zau6 wui5 bat1dou2jip6
 if you take. every CL course you then will able.to graduate
 ‘If you finish every course, you will be able to graduate.’

To account for the exceptional scope-taking properties of [CL N] and [*jat1* CL N], we argue that their indefinite interpretations are derived via choice functions, inspired in particular by the foundational of Reinhart (1997), Winter (1997), Kratzer (1998), and Matthewson (1999), as well as that of Fodor and Sag (1982). Space considerations prevent us from going deeply into the technical details of the analysis. Intuitively, the idea is this: a choice-functional indefinite is derived by having a choice function variable, f , apply in-situ to a set-denoting NP. This variable is of type $\langle et, e \rangle$, and returns some entity from the set characterized by the NP.⁵

Up to this point, the literature cited above is in broad agreement. Analyses diverge in terms of what subsequently happens to this variable. For Reinhart and Winter, f is existentially bound, with binding occurring freely at different scope positions. This gives rise not only to wide-scope indefinite readings, including the apparently island-violating ones, but to low and intermediate scope readings as well. For choice-functional indefinites in St’ át’ imcets, Matthewson argues that only wide-scope interpretations are available, and thus that only wide-scope existential binding of the choice function variable is possible. Finally, Kratzer argues that the variable is not bound at all, but is left free, giving rise to a specific/referential indefinite interpretation (cf. Fodor and Sag (1982)). Another point of difference involves the question of ambiguity. Both Reinhart and Kratzer argue that English indefinites are ambiguous between a choice-functional and quantifier semantics, while Winter argues for a uniform choice-functional account of English indefinites.

Turning back to Cantonese, we propose that [CL N] and [*jat1* CL N] are both uniformly interpreted via choice function variables, and that these variables can be existentially bound at

⁵Slightly more technically: A function f is a choice function just in case for any non-empty set S , $f(S) = x$ for some x , where $x \in S$. There are technical issues that arise in case S is empty, which we set aside; see Winter (1997) for details and one solution.

any point, giving rise to various scopal possibilities, including the island-violating wide-scope interpretations seen above. To make things concrete, we posit a two-step process, whereby the classifier first combines with the noun phrase, deriving number distinctions, with the resulting phrase then being fed to a choice function variable. In the [*jat1* CL N] construction, we might treat *jat1* itself as the formal reflex of this variable, while in the [CL N] construction the variable is phonologically null.⁶ Similar to Reinhart (1997) and Winter (1997), we posit that existential binding of the choice function variable can take place freely in Cantonese, giving rise to a range of scopal possibilities, including the island-violating ones.

4.3. Deriving the definite interpretation of bare classifier phrases

As noted in section 2, [CL N] phrases can be interpreted definitely as well as indefinitely. They thus contrast with [*jat1* CL N] phrases, which can only be interpreted indefinitely. Previous literature (a.o. Cheng and Sybesma (1999); Jenks (2018) which we review in the next section) has posited various analyses to derive the definite interpretation(s) of [CL N] phrases. Despite their differences, one thing these proposals share is the assumption that the definite interpretation(s) associated with [CL N] phrases is derived from a definite *semantics*; that is, [CL N] phrases are assumed to have one or more definite *readings*, which would make definitely interpreted [CL N] phrases at least two-ways ambiguous, with distinct indefinite and definite semantic readings derived from distinct semantic and/or syntactic structures. We close this section by suggesting an alternative, according to which the choice-functional semantics sketched above can uniformly derive both indefinite and definite interpretations of [CL N] phrases, with no need for positing any semantic or syntactic ambiguity.

With *jat1*, the choice function variable *must* be existentially bound, deriving the fact that [*jat1* CL N] phrases can receive *only* an indefinite interpretation.⁷ We can thus think of *jat1* as a way of marking the existential closure of the choice function variable. For [CL N] phrases, we propose that existential binding is *optional*; it *can* be bound, and at all the same locations as can [*jat1* CL N]. But it need not be. When left unbound, we get (following Kratzer 1998) a specific/referential interpretation. In some contexts (namely, those in which the addressee is understood not to be familiar with the individual picked out by the choice function, or with the choice function itself), this specific interpretation will be indefinite. As Kratzer puts it, the value is “often intended by the speaker, but not revealed to the audience.” But nothing in the semantics forces an indefinite interpretation.

We propose that the definite interpretation is just a pragmatic variant of the specific indefinite interpretation, reflected in contexts where the referent returned by a free choice function variable is familiar to both the speaker and the addressee. We leave the formal details to future work, but we are inspired here by von Stechow (1997), who treats both definites and indefinites as term-denoting (type *e*) expressions, doing so through the use of choice functions. Roughly, in both cases a type $\langle e, t \rangle$ NP is converted to a type *e* expression by a choice function *f*. The definite/indefinite distinction, signaled in English by the choice of ‘the’ versus ‘a’, is

⁶We might instead take the classifier itself to contribute this variable (thus converting the NP from a property to an individual).

⁷More generally, true numerals (including *jat1* when functioning as a numeral) seem to come with this meaning component, driving the fact that numeral phrases in Cantonese are interpreted indefinitely.

handled (in a dynamic semantics) by resolving the choice function to an existing globally specified choice function (in the case of definites), or locally introducing a new choice function (in the case of indefinites) which then updates the global choice function.

In a language like English, we follow the proposal in Heim (1990, 2011) that the definite interpretation of an indefinite is blocked by the existence of the definite article. Since (contra literature we discuss in the next section) we propose that Cantonese lacks a definite article (overt or covert), there is nothing to prevent a choice-functional (type *e*) ‘indefinite’ from getting a definite interpretation (i.e. from being associated with a hearer-familiar specific referent). In short, the choice-function variable in [CL N] phrases can be left unbound, and in this case we get a specific referential interpretation that is underspecified for definiteness.

5. Previous analyses of Cantonese nominal expressions

Previous analyses of Cantonese nominal expressions have largely focused on the [CL N] construction, and on their definite, rather than indefinite, interpretations. We think this is because bare classifier constructions are one of the (many) areas where Cantonese differs strikingly from Mandarin. Though the construction is very frequent in Cantonese, it is rare and much more constrained in Mandarin.

Cheng and Sybesma (1999, 2008) are among the first authors to discuss the case of Cantonese nominal phrases from a formal perspective.⁸ The core of their argument is that [CL N] phrases are syntactically ambiguous. Those that are interpreted as indefinites are Numeral Phrases with an empty numeral head which provides an existential quantifier (this also applies to indefinite bare nouns, which, in addition, have an empty classifier head). Definite [CL N] phrases are instead Classifier Phrases, in which the classifier essentially functions like an *t* operator.

Wu and Bodomo (2009) treat the problem differently. In their account, the classifier always moves to the head of a DP in [CL N] phrases. Since a [CL N] phrase always has the same structure, definiteness does not come from any syntactic differences, as Cheng and Sybesma (1999) argue, but is “contextually constrained” (p. 495). Specifically, they claim the following: “(the definite interpretation) arises (a) when the referent has already been mentioned in the discourse context. . . (b) if not, the referent must be close by, so that the hearer can easily identify the referent.” (p. 495-496). Though that characterization is probably too restrictive (e.g. definite [CL N] can be accommodated), it is close in spirit to our approach.

Li and Bisang (2012) offer a view that comprises a pragmatic, syntactic, and semantic perspective. They first argue that the information structure and word order of Chinese languages force the preverbal NP to receive a definite reading, a hypothesis that we fully subscribe to. Syntactically, they argue that an indefinite [CL N] is maximally a CLP, which deviates from the proposal of Cheng and Sybesma (1999). A definite [CL N], on the other hand, is a DP, in which the classifier has moved to the head of the DP, as in the analysis of Wu and Bodomo (2009). This is facilitated by the fact that there is no numeral in between to interfere with the movement. Thus, in their analysis, classifiers in Cantonese are not definite articles. They only

⁸These authors argue that Cantonese bare NPs cannot be interpreted as definites. The data, and most of the literature, suggest otherwise.

serve the function of individualization. Interestingly, however, in their proposal, the mechanism through which the interpretation of definiteness of a [CL N] phrase is obtained is not a product of their syntactic structure. Instead, Li and Bisang argue that the definite interpretation is “familiarity-based” (p. 350). The difference between a definite and an indefinite interpretation is whether there is a relevant context that is familiar to both the speaker and the hearer such that the hearer can pick out a referent. They discuss three contexts in which a definite interpretation may obtain: when the referent is visible in the context of utterance, when the referent is known by the interlocuters, or when the referent is identifiable via a bridging inference.

Jenks (2018) starts with an analysis of definiteness in Mandarin, which formally distinguishes unique definites (in the form of bare nouns) and anaphoric definites (marked by a demonstrative). Jenks takes this as evidence for the existence of two distinct forms of definiteness, for which he provides a formal description. Turning to Cantonese, he claims that the compatibility of [CL N] phrases with both unique and anaphoric definite referents is evidence that those phrases are ambiguous in their definite readings and that Cantonese has a semantically ambiguous null definite determiner. This element is comparable to English *the*, which is claimed to be ambiguous between a unique definite and an anaphoric definite determiner. It differs from English *the* in that it is silent and not directly observable. Under this analysis, [CL N] phrases are (at least) triply ambiguous between an indefinite reading and two definite ones.

In summary, previous studies mostly focus on how a definite interpretation is derived, and indefiniteness is analyzed in negative terms (i.e. what is not definite), but rarely considered on its own. Furthermore, these studies typically analyze the different interpretations of the structures at hand as the result of ambiguities. Our work focuses on the indefinite interpretations instead, especially by looking at how indefinites behave in different environments and how to account for the low and wide-scope interpretations of these elements in various contexts. Our analysis takes a semantic approach without making any claims about the syntax of the nominal expressions discussed. In particular, the choice functional semantics we propose for [CL N] phrases makes them compatible with both indefinite and definite interpretations (cf. *infra*), without having to postulate any form of syntactic or semantic ambiguity. This does not mean that we argue against such an ambiguity: it might be warranted by other syntactic facts, but we claim that the semantic evidence at hand does not itself require such an analysis.

6. Conclusion and openings

Our discussion has largely focussed on indefinite interpretations, but many issues remain open. We mention these here, as a roadmap for future work on these topics.

First, we left the non-indefinite readings of [BARE N] phrases to the side. As pointed out in section 2, [BARE N] phrases have definite and kind/generic readings along with indefinite ones. The details of these other readings should be explored in more detail in future work. The definite interpretation, and its restriction to unique definite reference, requires further exploration. Our suggestion that the definite interpretation of [CL N] phrases can be unified with their indefinite interpretation also requires further elaboration.

Another open issue has to do with the precise semantics of the classifier *dil*. Unlike other

classifiers, *dil* does not have selectional restrictions (it combines with most, if not all, nouns), nor does it seem to atomize the denotation of the noun it combines with (given its compatibility with plural denotations). As noted earlier, our current proposal makes it semantically vacuous, but further investigation might call this conclusion into question.

As is well known, the (in)definiteness of noun phrases in Chinese languages is affected by the position of the noun phrase relative to the verb. This effect (which we have set aside in this paper by focusing on post-verbal noun phrases) calls for an explanation, in particular for [CL N] and [*jat1* CL N] phrases. Pre-verbal phrases cannot be interpreted indefinitely, which bars [*jat1* CL N] phrases in those environments (unless they're introduced with an existential construction), and forces [CL N] ones to be given a definite interpretation. One option to account for this is to follow Cheng and Sybesma (1999) and consider that existential closure happens at the VP level. In our analysis this would mean that choice functions can only be existentially bound if they appear in the VP domain. Given that [*jat1* CL N] phrases require such binding, this would explain why they are not licensed pre-verbally, and why [CL N] have to be interpreted definitely, i.e. left unbound, when appearing pre-verbally.

Finally, we might consider the idea that *jat1* functions as an indefinite determiner in Cantonese. In many ways, it resembles indefinite determiners in languages like French or German, for which indefinite determiners are identical to the numeral one. There is thus a case to be made that Cantonese does have an indefinite determiner, though it lacks a definite counterpart. This would have repercussions for principles like *Maximize Presupposition!* (Heim, 1990), which deserve closer investigations.

References

- Ahrens, K. and C.-R. Huang (2016). Classifiers. In C.-R. Huang and S.-t. Xu (Eds.), *A reference grammar of Chinese language*, pp. 169–198. Cambridge: Cambridge University Press.
- Bale, A., J. Coon, and N. Arcos (2019). Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa: a journal of general linguistics* 4(1).
- Cheng, L. L.-s. and R. Sybesma (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4), 509–542.
- Cheng, L. L.-s. and R. Sybesma (2008). Classifiers in four varieties of Chinese. In G. Cinque and R. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax*, pp. 259–292. Oxford University Press.
- Chierchia, G. (1998). Reference to kinds across language. *Natural language semantics* 6(4), 339–405.
- Dayal, V. and Y. Sağ (2020). Determiners and bare nouns. *Annual Review of Linguistics* 6(1), 173–194.
- Deal, A. R. and J. Nee (2018). Bare nouns, number, and definiteness in Teotitlán del Valle Zapotec. In R. Truswell, C. Cummins, C. Heycock, B. Rabern, and H. Rohde (Eds.), *Proceedings of Sinn Und Bedeutung 21*, Volume 21, pp. 317–334.
- Fodor, J. D. and I. Sag (1982). Referential and quantificational indefinites. *Linguistics and Philosophy* 5, 355–398.
- Groenendijk, J. and M. Stokhof (1989). Context and information in dynamic semantics. In

- H. Bouma and B. Elsendoorn (Eds.), *Working Models of Human Perception*, pp. 457–486. London: Academic Press.
- Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*, pp. 164–189. Berlin: De Gruyter.
- Heim, I. (1990). Artikel und Definitheit. In A. v. Stechow and D. Wunderlich (Eds.), *Handbuch der Semantik*, pp. 487–535. Berlin: de Gruyter.
- Heim, I. (2011). Definiteness and indefiniteness. In C. Maienborn, K. von Heusinger, and P. Portner (Eds.), *Semantics An International Handbook of Natural Language Meaning*, Volume 1, pp. 996–1025. Berlin: De Gruyter Mouton.
- Ionin, T. and O. Matushansky (2006, July). The Composition of Complex Cardinals. *Journal of Semantics* 23(4), 315–360.
- Jenks, P. (2018). Articulated definiteness without articles. *Linguistic Inquiry* 49(3), 501–536.
- Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer.
- Kratzer, A. (1998). Scope or pseudoscope? Are there wide scope indefinites? In G. Chierchia, P. Jacobson, F. J. Pelletier, and S. Rothstein (Eds.), *Events and Grammar*, Volume 70, pp. 163–196. Dordrecht: Springer Netherlands.
- Krifka, M. (2003). Bare NPs: Kind-referring, indefinites, both, or neither? In *Semantics and Linguistic Theory*, Volume 13, pp. 180–203.
- Li, X.-P. and W. Bisang (2012). Classifiers in Sinitic languages: From individuation to definiteness-marking. *Lingua* 122(4), 335–355.
- Little, C.-R., M. Moroney, and J. Royer (2022, November). Classifiers can be for numerals or nouns: Two strategies for numeral modification. *Glossa: a journal of general linguistics* 7(1).
- Luke, K. K. and M. L. Wong (2015). The Hong Kong Cantonese Corpus: Design and Uses. *Journal of Chinese Linguistics* 25, 312–333.
- Matthewson, L. (1999). On the interpretation of wide-scope indefinites. *Natural language semantics* 7, 79–134.
- Partee, B. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk (Ed.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, pp. 115–143. Dordrecht: Foris.
- Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20(4), 335–397.
- van Geenhoven, V. (1998). *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Dissertations in Linguistics. Stanford, Calif: CSLI Publications.
- von Heusinger, K. (1997). Definite descriptions and choice functions. In S. Akama (Ed.), *Logic Language and Computation*, pp. 61–91. Dordrecht: Kluwer.
- Winter (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20, 399–467.
- Wu, Y. and A. Bodomo (2009). Classifiers \neq Determiners. *Linguistic Inquiry* 40(3), 487–503.