

# John Benjamins Publishing Company



This is a contribution from *Studies in Language 35:1*  
© 2011. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

# Bare classifiers and definiteness

## A cross-linguistic investigation

Andrew Simpson, Hooi Ling Soh and Hiroki Nomoto

University of Southern California / University of Minnesota / Tokyo  
University of Foreign Studies\*

In some (numeral) classifier languages, a classifier may occur “bare” (i.e. with a noun but without a numeral) and the nominal expression receives a definite interpretation. On the basis of evidence from Cantonese and Mandarin Chinese, Cheng and Sybesma (1999) hypothesize that classifier languages exhibit either the bare classifier or the bare noun pattern for definite reference, but not both. To evaluate this hypothesis against more typologically diverse languages, a parallel elicitation study of three non-Sinitic languages was conducted — Vietnamese, Hmong and Bangla — as well as two geographical varieties of Cantonese, focusing on the definite interpretation of bare classifier and bare noun patterns. The results show that although the use of bare classifier patterns for definite reference is a cross-linguistically connected phenomenon, there is more variation than previously described in the alternation between definite bare classifier and bare noun patterns, and that the preference for one pattern over another may receive functional/ pragmatic explanations.

**Keywords:** classifiers, definiteness, bare nouns, grammaticalization

### 1. Introduction

While the prototypical cross-linguistic function of numeral classifiers (henceforth “classifiers”) is to support the numerical quantification of nouns (e.g. Gil 2005), in certain languages of East, Southeast and South Asia, such as Cantonese, Vietnamese, Hmong and Bangla, such classifiers have grammaticalized an additional function which has sometimes been likened to the emergence of new definite determiners.<sup>1</sup> In these languages, numeral classifiers may occur paired with nouns in the absence of any numeral and give rise to interpretations of definiteness. Henceforth we will use the term “bare classifier pattern/form” to refer to forms in which a classifier occurs without a numeral but with a following noun. Examples of bare classifier patterns in Cantonese are given in (1).<sup>2</sup>

## (1) Cantonese

- a. *Zek3 gau2 gam1jat6 dak6bit6 teng1waa6.*  
 CLF dog today especially obedient  
 ‘The dog is especially obedient today.’ (Cheng and Sybesma 1999)
- b. *Keoi5 maai6-zo2 gaa3 ce1.*  
 he sell-ASP CLF car  
 ‘He sold the car.’ (Matthews and Pacioni 1996)
- c. *Ngo5 soeng2 maai5 bun2 syu1.*  
 I want buy CLF book  
 ‘I want to buy a book.’

In Cantonese, the bare classifier form may receive a definite interpretation in both the subject and object position as shown in (1a) and (1b), and it may also receive an indefinite interpretation in an object position as shown in (1c).

The existence of the bare classifier pattern in Cantonese has been commented on in a number of recent works on the form and interpretation of nominal phrases in Chinese (Cheng and Sybesma 1999, 2005, Simpson 2005, Wu and Bodomo 2009, among others).<sup>3</sup> These works highlight a systematic difference between Cantonese and Mandarin in the distribution and interpretation of bare classifier forms. Unlike Cantonese, a bare classifier form in Mandarin may not receive a definite interpretation. It may not occur in subject position, and may only receive an indefinite interpretation when in object position, as shown in (2).

## (2) Mandarin

- a. \**Zhī gǒu jīntiān tèbié tīnghuà.*  
 CLF dog today especially obedient  
 Intended: ‘The dog is especially obedient today.’
- b. *Wǒ xiǎng mǎi běn shū.*  
 I want buy CLF book  
 Only: ‘I want to buy a book.’  
 Not: ‘I want to buy the book.’

Definite reference equivalents to the Cantonese bare classifier pattern commonly take the form of bare nouns in Mandarin, as illustrated in (3):

## (3) Mandarin

- Gǒu jīntiān tèbié tīnghuà.*  
 dog today especially obedient  
 ‘The dog is especially obedient today.’

This systematic difference between Mandarin and Cantonese has led to the hypothesis that numeral classifiers in Cantonese have developed a special reference-denoting function which is comparable to that of definite determiners in various Indo-European languages. Cheng and Sybesma (2005:279) suggest that:

‘...all the functions the determiner takes care of in French are performed by the classifier in Cantonese: definiteness, number, individuation, noun classification, subordination, and deictism.’

In a particularly prominent and engaging work on Cantonese and Mandarin, Cheng and Sybesma (1999) report that the occurrence of bare nouns is unavailable in Cantonese, but fully common in Mandarin in contexts of definite reference. This reported difference leads to the conclusion that languages should consistently exhibit either the bare noun pattern (Mandarin) or bare classifier pattern (Cantonese) for the expression of definite discourse referents, but not both. Likening the occurrence of bare classifiers to the use of definite determiners, Cheng and Sybesma suggest that if a language has available an overt way to express definiteness with nouns, it must necessarily make use of this. Hence, definite noun phrases in Cantonese should make use of the bare classifier strategy and not permit the use of bare nouns, as is common in Mandarin.

In comparison with Cantonese, much less attention has been paid to the existence of bare classifier patterns in Vietnamese, Hmong and Bangla, which like Cantonese also result in interpretations of definiteness, as illustrated below.

(4) Vietnamese

*Cái máy ảnh rất đắt tiền.*

CLF camera very expensive

‘The camera was very expensive.’

(5) Hmong

*Tus kws hais plaub tub tub nkeeg.*

CLF lawyer RED lazy

‘The lawyer is very lazy.’

(6) Bangla

*kEmera-Ta khub dami.*

camera-CLF very expensive

‘The camera was/is very expensive.’

The present paper presents the results of a comparative study of such definite “bare classifier” phenomena in Vietnamese, Hmong, Bangla and Cantonese, four genetically-unrelated languages.<sup>4</sup> A primary goal of the investigation was to significantly increase the level of empirical coverage and description of the definite bare classifier construction available in the literature, both expanding on current descriptions of Cantonese and presenting an initial cross-linguistic comparison of such forms in Vietnamese, Hmong and Bangla. The study set out to provide at least preliminary answers to the following questions:

- (7) a. Are bare classifier patterns in Vietnamese, Hmong, Bangla and Cantonese instantiations of the same linguistic phenomenon? How are the languages similar or different in terms of the range of “definite” interpretations that may be expressed by bare classifier patterns?
- b. To what extent does the availability of the bare classifier pattern in a language preclude the potential use of a bare noun to represent a discourse referent that is “definite”? When alternative ways of representing “definite” discourse referents are available, what factors affect the use of the bare classifier pattern as opposed to alternative forms that may also refer to “definite” referents?

On the basis of the analysis of extended parallel data gathered through elicitation in the four languages, the paper concludes that bare classifier patterns in these languages do constitute a unified phenomenon, finding considerable similarity in the use of bare classifier patterns in contexts of definiteness. The study also reveals, however, that the availability of the bare classifier pattern in a language does not necessarily preclude the use of a bare noun to represent a discourse referent that is “definite”. In Vietnamese, Hmong and Bangla, it is found that bare classifier and bare noun patterns both occur as possible expressions of definite noun phrases, indicating that an alternating pattern is indeed possible in certain languages. In considering such alternations, it is noted that the apparent optionality in reference form is not random and unconditioned free variation, and various functional and pragmatic factors play a role in the preferred use of bare classifier forms versus bare nouns for definite reference. Our investigation also shows that patterns in Cantonese are more complex than previously described, with speakers often permitting the use of a bare noun strategy to represent “definite” discourse referents as well as bare classifier forms. This leads to the conclusion that there is a range of variation in the grammaticalization of the definite bare classifier pattern, with the enforced use of such forms in all contexts of definiteness representing just one extreme point on a continuum of speakers’ repertoires in languages which have developed this novel use of numeral classifiers.

The structure of the paper is as follows. Section 2 describes the investigation carried out with Vietnamese, Hmong, Bangla and Cantonese which considers the range of use of bare classifier forms for definite reference in the four languages, as well as the potential use of bare nouns for similar occurrences of definite reference. The degree to which such strategies may compete with each other is also reported on. In Section 3, we revisit Cantonese, presenting a follow-up study on the use of the bare noun pattern for encoding definite referents, observing that for many speakers there is actually a broader patterning of definite nominal reference than is often described in the literature. Section 4 then discusses the set of functional and pragmatic factors which seem to lead to the preferred use of bare classifiers or

bare nouns for definite reference in different contexts. Section 5 closes the paper with a general summary of its findings.

## 2. “Definiteness,” bare classifiers and alternative expressions of “definiteness”

In attempting to explore how cross-linguistically similar the use of bare classifier forms in contexts of **definiteness** might be, an initial investigative goal of the planned comparison of Cantonese, Vietnamese, Hmong, and Bangla was to establish what range of “definite” interpretations could potentially be instantiated by bare classifier patterns in the four languages. It is generally recognized that the ways in which “definiteness” surfaces across languages may be similar, but not completely identical (cf. Lyons 1999, Chen 2004). Given the frequent likening of bare classifier patterns to the use of definite articles in contexts of definite reference, the present study set out to determine the extent to which bare classifiers may occur where definite articles are used in languages such as English and French for reference to entities that are unique and familiar/identifiable. Five primary sub-types of situations licensing the use of definite determiners in English, French etc. were used to probe the occurrence of the bare classifier pattern in definite contexts, as detailed below:

- (8) A. **Discourse-anaphoric** definite readings. Contexts in which a definite determiner is used in reference to an antecedent which has previously been introduced by means of some overt linguistic item.  
Example: *John has a dog and a cat. **The dog** is very old.*
- B. Use of a definite determiner for referents which are **visible** and **uniquely identifiable** for both speaker and hearer, but which have not been mentioned explicitly in the ongoing discourse.  
Example: *Pass me **the hammer**.*
- C. Definiteness by **association/inference**. Instances where the definite identity of a referent results from its association with some other entity mentioned in the discourse.  
Example: *We went to a wedding yesterday. **The bride** was very beautiful.*
- D. Use of the definite determiner with **contextually unique** items that are **not (necessarily) visible**.  
Example: *Where’s **the car/the key**?*
- C. Use of the definite determiner with **culturally unique, familiar** entities.  
Examples: ***The president** has died.  
**The sun** rises at 6:00 every morning.*

In order to investigate these patterns of definite reference in Cantonese, Vietnamese, Hmong and Bangla, parallel elicitations for these languages were constructed based on the same questionnaire translated by native speakers from English into Cantonese, Vietnamese, Hmong and Bangla. The questionnaire consisted of 56 sentences/pairs of sentences containing 77 noun phrases marked with the definite determiner. The composition of the data set in terms of the five subcategories of definiteness noted in (8A–E) was as follows:

- |  |                                |
|--|--------------------------------|
| (9) (A) Discourse-anaphoric:               | 30 noun phrases (16 sentences) |
| (B) Visible, uniquely identifiable:        | 13 noun phrases (12 sentences) |
| (C) Definiteness by association/inference: | 18 noun phrases (15 sentences) |
| (D) Not visible, contextually unique:      | 9 noun phrases (7 sentences)   |
| (E) Culturally unique, familiar:           | 7 noun phrases (7 sentences)   |

To ensure that the context of use was clear to the speakers, an explicit background context was supplied for most sentences. The sentences were also constructed in a way that made it natural for a bare classifier pattern or a bare noun pattern to be used for reference purposes in the translation, as opposed to the use of pronouns or demonstrative-modified nouns, which in certain (other) contexts might be more naturally used for anaphoric reference. This was achieved in certain instances by constructing sentences where two referents were introduced in one sentence and subsequently referred back to anaphorically in a follow-on sentence, as in (10a) below. If the first sentence in such pairs contains only a single new referent, as in (10b), speakers tend to favor the use of pronouns rather than bare classifier/bare noun patterns. The study therefore aimed at collecting data that in all instances maximized the possibility for bare classifier/bare noun patterns to occur in natural speech in the translations of the English questionnaire.

- (10) a. *Mary has a dog and a cat. **The dog** is very old, but **the cat** is still quite young.*  
 b. *Mary has a dog. It/#**The dog** is very old.*

For all four languages, three native speakers were interviewed individually by one of the authors. The speakers were university students in the United States, originally from Vietnam (speakers of Vietnamese), Laos/the United States (Minnesota) (Hmong speakers), India (West Bengal) (Bangla speakers), and Hong Kong (Cantonese speakers). The speakers were first asked to translate the English sentences in the questionnaire into their native languages. They were then asked for acceptability judgments for relevant alternative forms of reference. For example, if a bare classifier pattern was provided in the translation, speakers would be asked if a bare noun pattern could be substituted for the bare classifier pattern with equal or less naturalness. Conversely, if a bare noun pattern was the first transla-

tion choice, speakers were asked whether a bare classifier pattern might or might not be used, and how natural this would be. The responses of the speakers were coded according to five degrees of acceptability, as follows. 5 = fully natural/acceptable/provided by the speaker in the translation, 4 = acceptable, but somewhat less natural than the pattern rated '5' for this sentence, 3 = unnatural for this sentence/native speakers would not use this pattern in this way, 2 = very unnatural, 1 = completely unacceptable.

The rationale for adopting such a semi-experimental approach was to attempt to satisfy three specific operational goals. First, it was considered important that the data examined in the four languages be as parallel as possible, so as to strengthen any conclusions drawn from the comparison of patterns in the four languages and minimize interference potentially introduced by contexts which were not closely equivalent. Second, it was decided that the establishment of a substantial body of data bearing on the patterns being investigated would help minimize any accidental skewing of the conclusions which might result from the consideration of a more restricted quantity of token sentences.

## 2.2 Results and discussions

The results of the present study will be discussed in two parts: the first concerns the extent to which the range of definite interpretations that can be instantiated by bare classifier forms in these four languages is the same; the second concerns the extent to which an alternative way of encoding definite referents exists in Vietnamese, Hmong and Bangla, and the degree to which the alternative strategy occurs as a competitor to the bare classifier pattern.

### 2.2.1 *The bare classifier pattern and definiteness in Vietnamese, Hmong, Bangla and Cantonese*

Initial investigation into the kinds of definite reference contexts that characteristically lead to the use of a definite determiner in west European languages indicates that the use of the bare classifier pattern occurs and is common in Vietnamese, Hmong, Bangla and Cantonese in all five contexts described in (8) above. Table 1 summarizes the average of the speakers' responses relating to the use of the bare classifier pattern in each type of "definiteness" context, for each of the four languages being investigated. Note that the means given here are across items as well as speakers. The superscript "+" marks instances where the standard deviation from the mean was larger than 1.00, this being affected by both the factors of items and speakers.

Selected representative illustrations are given for the different contexts in (11)–(15).

**Table 1.** The definite interpretation of the bare classifier (CL + NP) pattern

Subcategories of definiteness	Vietnamese	Hmong	Bangla	Cantonese
(A) Discourse-anaphoric	4.75	5.00	4.81	5.00
(B) Visible, uniquely identifiable	4.82	5.00	4.82	4.60
(C) Association/inference	4.76	4.92	4.35 <sup>+</sup>	4.67
(D) Invisible, contextually unique	4.76	4.85	3.54 <sup>+</sup>	5.00
(E) Culturally unique, familiar	4.57	4.81	2.07 <sup>+</sup>	5.00
(A)–(E)	4.74	4.95	4.19 <sup>+</sup>	4.83

## (11) (A) Discourse-anaphoric definiteness

## a. Vietnamese

*Nga có một con chó và một con mèo.*

Nga have 1 CLF dog and 1 CLF cat

*Con chó rất già nhưng con mèo thì vẫn còn trẻ.*

CLF dog very old but CLF cat PRT still young

‘Nga has a dog and a cat. The dog is very old, but the cat is still quite young.’

## b. Bangla

*ami Sobe Ek-Ta notun radio ar ghoRi kinechi.*

I just 1-CLF new radio and clock bought

*Radio-Ta ranna ghore ar ghoRi-Ta khabar ghore ache.*

radio-CLF kitchen-in and clock-CLF dining-room-in be

‘I just bought a radio and a new clock. The radio is in the kitchen and the clock is in the dining-room.’

## (12) (B) Visible, uniquely identifiable elements

## a. Hmong

*Lub qhov rai puas tau kaw?*

CLF window Q ASP close

‘Did you shut the window?’

## b. Cantonese

*m4-goi1 dai6 go3 din6si6-jiu4-hong3-hei3 bei2 ngo5.*

please hand CLF TV-remote-control to I

‘Please give me the TV remote control.’

## (13) (C) Definiteness by association/inference

## a. Cantonese

*cam4-jat6 ngo5-dei6 heoi3 jat1 go3 fan1-lai5.*

yesterday we go 1 CLF wedding

*go3 san1-loeng4 hou2 leng3.*

CLF bride very pretty

‘Yesterday we went to a wedding. The bride was beautiful.’

- b. Vietnamese  
*Nam vừa mua một chiếc TV mới. Cái màn ảnh rất rộng.*  
 Nam just buy 1 CLF TV new CLF screen very large  
 ‘Nam just bought a new TV. The screen is very large.’
- (14) (D) Contextually unique items which are not (necessarily) visible
- a. Vietnamese  
*Nam và Nga đang ở trong một bãi đậu xe rất lớn.*  
 Nam and Nga ASP be in 1 car park very large  
*Nam hỏi: ‘Vậy là chiếc xe ở đâu?’*  
 Nam ask so be CLF car be where  
 ‘Nam and Nga are in a large car park. Nam says: ‘So, where’s the car?’
- b. Cantonese  
*Mary tung4 keoi5 go3 zai2 gong2: ‘nei5 ho2-m4-ho2ji5*  
 Mary with 3 CLF son say you can-not-can  
*bong1 ngo5 lo2 go3 tai1 gwo3 lei4?’*  
 help I take CLF ladder come come  
 ‘Mary says to her son: “Can you please bring me the ladder?”’
- (15) (E) Culturally unique, familiar entities
- a. Bangla  
*gOto ratey caMd-Ta khub ujjOl chilo.*  
 last night-in moon-CLF very bright was  
 ‘The moon was very bright last night.’
- b. Hmong  
*Lub lim tiam tom ntej tus president yuav mus saib China no.*  
 CLF week to front CLF president will go look China PRF  
 ‘This week the president will visit China.’

Broadly speaking, all four languages exhibit high acceptability rates for “definite” interpretations of bare classifier forms. That is, in all four languages, the average acceptability rates are all between 5 (“fully natural/acceptable/provided by the speaker in the translation”) and 4 (“acceptable, but somewhat less natural than the pattern rated ‘5’ for this sentence”) when different contexts of definiteness are abstracted away (see the last line of Table 1). This overall similarity among the four languages consequently justifies treating the definite interpretation of the bare classifier pattern in these four languages as a shared, connected linguistic phenomenon.

Variation is found when different definiteness contexts are taken into consideration. While Vietnamese, Hmong and Cantonese consistently show high acceptability rates in all five contexts of definiteness, Bangla shows low acceptability rating for (D) contextually unique items that are not visible (e.g., ‘the key’ referred to in *Where’s the key?*) and even lower rating for (E) culturally unique, familiar

entities (e.g., ‘the president’, ‘the sun’). The acceptability rates of these two subcategories of definiteness in Bangla are 3.54 and 2.07 respectively (cf. 3 = “unnatural for this sentence/native speakers would not use this pattern in this way,” 2 = “very unnatural”). It is arguably not surprising to find cross-linguistic variation with these particular categories, as the intended unique/familiar referent is less accessible to the speaker and hearer in these contexts, requiring inference based on elements not directly visible (D) or involving culture-specific conventions (E). In this sense, these two categories may be viewed as peripheral members of the various subcategories of definiteness and therefore might be coded differently from definite reference in contexts (A)–(C), where the intended unique/familiar referent is more directly accessible.

### 2.2.2 Bare nouns and “definiteness” in Vietnamese, Hmong and Bangla

Given the clear similarity among the four languages with regard to the use of bare classifier forms in contexts of definiteness, it is natural to ask whether Vietnamese, Hmong and Bangla also pattern like what has been regularly reported for Cantonese, in not allowing the use of bare noun patterns to represent definite referents. Our results here indicate that the expression of definiteness in Vietnamese, Hmong and Bangla does **not** appear to be fully parallel to that reported for Cantonese in previous literature. In Vietnamese, Hmong and Bangla, there exists a second mode of expression for definite referents which has been described as unavailable in Cantonese but fully common in Mandarin, namely the use of bare nouns in contexts of definite reference. This is illustrated in (16a)–(16c), with Vietnamese, Hmong and Bangla respectively:

- (16) a. Vietnamese  
*Cửa sổ có bị đóng không?*  
 window be PASS close Q  
 ‘Is the window closed?’
- b. Hmong  
*Koj pab taws teeb yod?*  
 you help turn.on light PRT  
 ‘Can you turn on the light, please?’
- c. Bangla  
*Mina dorja bOndho, cabi kothay?*  
 Mina door closed key where  
 ‘Mina, the door is locked. Where’s the key?’

Furthermore, and rather surprisingly, the initial parallel investigation of Cantonese indicated that the use of a bare noun pattern to represent a definite referent actually is possible in certain instances, contrary to what has regularly been reported

for this variety of Chinese. This unexpected result prompted a second follow-up study aimed at better understanding the use of the bare noun pattern in Cantonese, which is reported on in Section 3. The remainder of Section 2 thus concentrates on results on Vietnamese, Hmong and Bangla, deferring integration of the supplemental study of Cantonese until later.

The clear availability of bare noun forms for definite reference in Vietnamese, Hmong and Bangla found in examples such as (16a–c) raises the question of whether alternation between the bare classifier and bare noun pattern is simply free variation or whether there are any differences in the availability and use of the two patterns in contexts of definite reference. The results of the second stage of investigation which set out to probe this issue are summarized in Table 2 below.

**Table 2.** Comparative use of the bare classifier (CL+NP) and bare noun (NP) patterns in Vietnamese, Hmong and Bangla in contexts of definiteness

Subcategories of definiteness	Vietnamese		Hmong		Bangla	
	CL+NP	NP	CL+NP	NP	CL+NP	NP
(A) Discourse-anaphoric	4.75	3.60 <sup>+</sup>	5.00	1.42 <sup>+</sup>	4.81	3.84 <sup>+</sup>
(B) Visible, uniquely identifiable	4.82	4.82	5.00	2.62 <sup>+</sup>	4.82	3.48 <sup>+</sup>
(C) Association/inference	4.76	4.98	4.92	3.12 <sup>+</sup>	4.35 <sup>+</sup>	3.73 <sup>+</sup>
(D) Invisible, contextually unique	4.76	4.81	4.85	2.63 <sup>+</sup>	3.54 <sup>+</sup>	4.18 <sup>+</sup>
(E) Culturally unique, familiar	4.57	5.00	4.81	2.62 <sup>+</sup>	2.07 <sup>+</sup>	4.86
(A)–(E)	4.74	4.54	4.95	2.26 <sup>+</sup>	4.19 <sup>+</sup>	3.91 <sup>+</sup>

As seen in Table 2, the three languages exhibit a common patterning only with regard to category (A) — discourse-anaphoric definite reference — where the uniqueness/familiarity of a referent is established by previous mention in the discourse. In such a context, speakers of all three languages show a preference for the bare classifier pattern over the bare noun pattern. This preference is most clearly observed in Hmong. While bare classifier forms were judged perfectly acceptable by all speakers for all test sentences (hence the maximum rate of 5.00), bare nouns receive the low average rate of 1.42, meaning that the speakers' responses were mostly "1 (= completely unacceptable)" or "2 (= very unnatural)."

In the other four contexts of definiteness investigated, Vietnamese appears to contrast with Hmong and Bangla in that no major difference in acceptability of the two patterns was observed. Both bare classifier and bare noun forms were rated as readily available in these contexts. This may relate to the fact that, in comparison with Bangla and Hmong, the use of classifiers is less automatic in other contexts in Vietnamese, for example in numeric quantification and with nouns of certain

types, so that “classifier-less” nouns may more generally be a frequent occurrence (see Löbel 2000 among others).

Hmong and Bangla behave similarly for context (B) referents that are visible and uniquely identifiable (e.g., a ‘hammer’ referred to in a situation in which both the speaker and hearer can see such an item) and context (C) referents whose uniqueness/familiarity is based on other previously-mentioned, associated entities (e.g., a ‘bride’ when the word ‘wedding’ has been mentioned in the immediate discourse). Both languages prefer bare classifier forms to bare nouns in these contexts. However, they show an opposite patterning with regard to context (D), contextually unique entities that are not visible, and context (E), culturally unique, familiar entities. While Hmong speakers showed a preference for bare classifier forms over bare nouns in these two contexts, Bangla speakers demonstrated a preference for the use of the bare noun pattern.

These results from Vietnamese, Hmong and Bangla appear to be at odds with Cheng and Sybesma’s (1999, 2005) generalization about the expected distribution of definite bare classifier and bare noun forms, based on patterns in Mandarin and Cantonese. As noted in Section 1, Cheng and Sybesma (1999) suggest that the two forms are expected to occur in complementary distribution across languages: if a bare classifier form is available to refer to a definite referent in a particular language, it is anticipated (for various theoretical reasons) that a corresponding bare noun form should not be available for such use. However, as is evident from the Table 2 figures for Vietnamese, the distribution of the two forms is not necessarily complementary. Furthermore, if we interpret a rating of “3 (= unnatural)” as indicating that speakers are uncertain whether a particular form is grammatical or ungrammatical, it can be said that both patterns are available in at least one context of definiteness in all three languages: (A)–(E) in Vietnamese, (C) in Hmong, and (A)–(C), (D) in Bangla. While this is so, it should also not be overlooked that some clear difference in occurrence between the bare classifier and bare noun patterns still holds to varying degrees — very clearly in Hmong, less so in Bangla and only for context (A) in Vietnamese. Thus, the distribution and availability of definite bare classifier and bare noun forms instantiates **patterned variation**, and is not simply random fluctuation between two equivalent forms. Additionally, how the two patterns complement each other in the different sub-categories of definite reference may be partly language-specific. For contexts (D) and (E), Hmong prefers the bare classifier pattern just as reported for Cantonese by Cheng and Sybesma, but Bangla regularly uses the bare noun pattern. The detailed results from Vietnamese, Hmong and Bangla therefore demonstrate the descriptive advantage of distinguishing different contexts of “definiteness” in attempting to approach the phenomenon of definite bare classifier patterns.

Quite generally, then, the results of the study of Vietnamese, Hmong and Bangla lead to a more nuanced understanding of the complementarity of the bare classifier and bare noun patterns for definite reference. In the next section, we turn to Cantonese and show that a broader investigation of this variety of Chinese leads to conclusions about Cantonese which are more in line with what has been described for Vietnamese, Hmong and Bangla than in previous characterizations of the language.

### 3. Bare nouns and “definiteness” in Cantonese

As noted briefly above in Section 2.2.2, the initial parallel study of Cantonese, Vietnamese, Hmong and Bangla resulted in the unexpected finding that use of the bare noun pattern to represent a definite referent is sometimes possible for speakers of Cantonese, contrary to what has been reported in the literature. In order to better understand the patterning of bare classifier and bare noun forms in definite contexts, we conducted a follow-up study, expanding on the number of speakers and the geographic areas they are from to include Cantonese speakers from both Hong Kong and Malaysia. The expanded follow-up study made use of the same set of data as the original study and brought the total number of native Cantonese speakers participating in the investigation to six, with three from Hong Kong and three from Malaysia.

#### 3.1 Results

The results of the investigation of Cantonese indicate much greater variation with regard to the use of bare noun forms in instances of definite reference than previously reported for Cantonese. This variation cannot be attributed neatly to the existence of different regional varieties of Cantonese. While all speakers confirmed that bare classifier sequences could be used for the same range/sub-types of definite reference as in Vietnamese, Hmong and Bangla (which was not a surprise), only one of the six Cantonese speakers interviewed (Speaker 3 in Table 3 below) disallowed the use of bare nouns as a means to encode definite reference.<sup>5</sup> The other five speakers all regularly allowed bare nouns as a potentially available alternative to bare classifier patterns in instances of definite reference.

Illustrative examples are given below. In each case, either the bare noun or the bare classifier pattern was indicated to be an acceptable option by all six speakers:

- (17) (Go3) *coeng1 hai6 mai6 saan1 zo2?*  
 CLF window is-not-is closed ASP  
 ‘Is the window closed?’

- (18) *Mary tung4 John gong2: Bong1 ngo5 jing2 zoeng1 soeng2.*  
 Mary with John say help me take CLF photo  
*John gong2: m4 dak1. (Go3) soeng2-gei1 waa16 zo2.*  
 John say not can CLF camera broken ASP  
 ‘Mary says to John: Take a photograph of me.’  
 ‘John says: I can’t. The camera’s broken.’
- (19) *Cam4 jat6 ngo5-dei6 heoi3 jat1 go3 fan1-lai5. (Go3) san1-loeng4 hou2 leng3.*  
 yesterday we go 1 CLF wedding CLF bride very pretty  
 ‘Yesterday we went to a wedding. The bride was beautiful.’
- (20) *Soeng6 go3 lai5 baai3, ngo5 tai2 zo2 jat1 bun2 hou2 hou2 tai2*  
 last CLF week I read ASP 1 CLF very readable  
*ge3 syu1. (Go3) zok3-ze2 hai6 jan3-dou6 jan4.*  
 MOD book CLF author be India person  
 ‘Last week I read a really good book. The author was from India.’
- (21) *Ngo5 teng1 gong2 (go3) zung2-tung2 haa6 go3 sing1-kei4 wui5 heoi3*  
 I hear-say CLF president next CLF week will go  
*zung1-gwok3.*  
 China  
 ‘I heard that the president is going to visit China next week.’
- (22) *John tung4 Mary hai6 jat1 go3 hou2 daai6 ge3 ting4-ce1-coeng4.*  
 John and Mary be 1 CLF very big MOD car-park  
*John waa6: (gaa3) ce1 hai6 bin1 aa3?*  
 John say CLF car be where Q  
 ‘John and Mary are in a large car park. John says: “So, where’s the car?”’

The results for each speaker are summarized below. Table 3 presents results for Hong Kong Cantonese speakers, while Table 4 presents results for Malaysian Cantonese speakers.

Only Hong Kong Speaker 3 responded in the manner reported in previous studies on Cantonese. For this speaker, only bare classifier forms are acceptable for definite reference, with the exception of category (E) — culturally unique, familiar items — for which bare noun forms are also available, though less natural than bare classifier forms. The responses given by the other two Hong Kong Cantonese speakers actually resemble those of Vietnamese (cf. Table 2), i.e., bare classifier forms are clearly preferred only for (A) discourse-anaphoric definite reference whilst both bare classifier and bare noun forms are available in the other definiteness contexts with no strong preference for either of them. As for Malaysian Cantonese, Speaker 1 and 2’s responses pattern more similarly to Bangla (cf. Table 2) than Hong Kong Cantonese. A dividing line could be drawn between categories

**Table 3.** The definite interpretation of the bare classifier (CL+NP) and bare noun (NP) patterns in Hong Kong Cantonese

Subcategories of definiteness	Speaker 1		Speaker 2		Speaker 3		All speakers	
	CL+NP	NP	CL+NP	NP	CL+NP	NP	CL+NP	NP
(A) Discourse-anaphoric	5.00	3.80 <sup>+</sup>	5.00	3.13 <sup>+</sup>	5.00	1.00	5.00	2.70 <sup>+</sup>
(B) Visible, uniquely identifiable	5.00	4.27 <sup>+</sup>	4.81	5.00	4.00 <sup>+</sup>	1.67	4.61 <sup>+</sup>	3.73 <sup>+</sup>
(C) Association/inference	5.00	4.20 <sup>+</sup>	5.00	4.50 <sup>+</sup>	4.00 <sup>+</sup>	1.14	4.67	3.30 <sup>+</sup>
(D) Invisible, contextually unique	5.00	5.00	5.00	5.00	5.00	1.25	5.00	3.93 <sup>+</sup>
(E) Culturally unique, familiar	5.00	5.00	5.00	5.00	5.00	3.75 <sup>+</sup>	5.00	4.62
(A)–(E)	5.00	4.23 <sup>+</sup>	4.96	4.23 <sup>+</sup>	4.50 <sup>+</sup>	1.42 <sup>+</sup>	4.85	3.35 <sup>+</sup>

**Table 4.** The definite interpretation of the bare classifier (CL+NP) and bare noun (NP) patterns in Malaysian Cantonese

Subcategories of definiteness	Speaker 1		Speaker 2		Speaker 3		All speakers	
	CL+NP	NP	CL+NP	NP	CL+NP	NP	CL+NP	NP
(A) Discourse-anaphoric	5.00	3.97 <sup>+</sup>	5.00	3.77 <sup>+</sup>	4.74	4.61 <sup>+</sup>	4.91	4.12 <sup>+</sup>
(B) Visible, uniquely identifiable	5.00	3.85 <sup>+</sup>	4.92	3.92 <sup>+</sup>	4.69 <sup>+</sup>	4.77	4.87	4.18 <sup>+</sup>
(C) Association/inference	4.06 <sup>+</sup>	2.78 <sup>+</sup>	5.00	3.82 <sup>+</sup>	4.56 <sup>+</sup>	4.56 <sup>+</sup>	4.53 <sup>+</sup>	3.72 <sup>+</sup>
(D) Invisible, contextually unique	3.00 <sup>+</sup>	5.00	4.63 <sup>+</sup>	4.63 <sup>+</sup>	4.63 <sup>+</sup>	5.00	4.08 <sup>+</sup>	4.88
(E) Culturally unique, familiar	3.29 <sup>+</sup>	4.43 <sup>+</sup>	4.57 <sup>+</sup>	4.57 <sup>+</sup>	4.71	5.00	4.19 <sup>+</sup>	4.67 <sup>+</sup>
(A)–(E)	4.41 <sup>+</sup>	3.80 <sup>+</sup>	4.91	3.96 <sup>+</sup>	4.67 <sup>+</sup>	4.70 <sup>+</sup>	4.66 <sup>+</sup>	4.15 <sup>+</sup>

(A)–(C) and (D)–(E); while the bare classifier pattern is clearly preferred to the bare noun pattern in the former contexts, in the latter contexts either the bare classifier pattern is preferred over the bare noun pattern (Speaker 1) or the two patterns are equally acceptable (Speaker 2). The response pattern of Malaysian Cantonese Speaker 3 resembles those of Hong Kong Cantonese Speakers 1 and 2, except that the former speaker does not show any strong preference for any one form, not even for category (A) discourse-anaphoric definites.

### 3.2 Discussion

In Cheng and Sybesma (1999), Cantonese and Mandarin Chinese are described as varieties of Chinese that make use of significantly different strategies in the linguistic encoding of definite noun phrases. In situations where English, French and other similar languages combine a definite determiner with a noun, Mandarin makes exclusive use of bare nouns, while Cantonese is described as adopting the bare classifier strategy. A principled general account is then developed which likens the apparent enforced use of bare classifiers in Cantonese to the obligatory use of definite determiners with nouns in English, French etc., predicting that there should not be any optional alternation in the use of the bare classifier strategy in languages like Cantonese which exhibit it. In subsequent work, Cheng and Sybesma (2005) actually became aware of a variety of Chinese that does permit an alternation between a bare noun and a bare classifier pattern with definite referents — Wenzhou Chinese, one of the Wu dialects. In order to accommodate Wenzhou Chinese with the general theory developed in their earlier work, Cheng and Sybesma (2005) draw attention to the fact that classifiers in contexts of definite reference in Wenzhou carry a special low tone not present when a noun phrase is interpreted as being indefinite. Cheng and Sybesma suggest that it is this low tone rather than the classifier which expresses definiteness in Wenzhou, and that a non-segmental, tonal marking of definiteness may not count as sufficiently overt to preclude the use of an alternate mode of representation with definite noun phrases — the use of a bare noun.<sup>6</sup>

For principled theoretical reasons, then, the “rigid” use of the bare classifier pattern in Cantonese is viewed in Cheng and Sybesma (2005) as a prototypical instantiation of the phenomenon of definite bare classifiers, and other languages with genuine bare classifier patterns are expected to show the same rigidity as reported for Cantonese and exclusive use of the bare classifier pattern in contexts of definiteness. Indeed, with Cantonese as one’s starting point, and the model of languages such as English, French, etc. enforcing the obligatory use of definite determiners with definite referents, it may be natural to assume that bare classifier forms would generally share the mandatory-use property of definite determiners. The present work, however, adds new perspective, changing this expectation. What emerges from a rather more extended comparison of different languages is that a Wenzhou-type pattern of optionally permitting both bare classifier and bare noun forms for definite referents appears to be quite widespread, both geographically and across language families, and shows initial signs of being more common than the rigid use of the bare classifier pattern in Cantonese described in Cheng and Sybesma (1999, 2005). Furthermore, whether it is the special, modified tone on the classifier or the classifier itself which signals the definiteness of a noun phrase in Wenzhou

(and certain other languages, see footnote 6), in Vietnamese, Hmong and Bangla, it appears to be the presence of a classifier alone (with no tonal/non-segmental modification) which regularly expresses/corresponds with interpretations of definiteness in noun phrases in alternation with a definite bare noun strategy.

Finally, it seems that Cantonese is itself much less monolithic than previously described in the literature; for many speakers “Mandarin-style” bare noun patterns are actually possible for use in definite reference. This observation further buttresses the assumption based on Vietnamese, Hmong and Bangla that the potential availability of both the bare classifier and bare noun patterns may be more of a norm for languages (with bare classifier patterns) than the rigid use of only the bare classifier pattern. Cantonese, as it occurs in the speaker variety described by Cheng and Sybesma (1999, 2005), may represent an endpoint of grammatical development rather than a typical canonical patterning of bare classifiers cross-linguistically. At this point we would like to speculate that the following stages of development may occur grammaticalizing definite bare classifier patterns: (i) only bare noun patterns are well-formed and are used for both definite and indefinite reference (as currently is the case in many classifier languages, e.g. Thai, Burmese, Japanese, etc.); (ii) bare classifier patterns emerge and are used interchangeably with bare nouns for nominals with singular reference (as is the case, for example, with indefinite object noun phrases in Mandarin Chinese); (iii) the bare classifier pattern starts to be specialized for definite reference as certain groups of speakers increasingly exclude contexts licensing the use of bare nouns for definite nominals, this being the broad stage of development assumed to be present in Vietnamese, Hmong, Bangla and Cantonese for certain speakers, and (iv) at its endpoint of development the bare classifier pattern becomes the only means for definite reference, with the result that bare nouns may not be used to encode definite reference — the “rigid” bare-classifier-only patterning reported in Cheng and Sybesma (1999). What may be interesting to probe in future investigations of Cantonese is the geographical spread of the latter bare-classifier-only variety, and its historical development as a process of grammaticalization akin to the systematization of definite determiner use in European languages. Leaving such a major diachronic project aside for future research, in Section 4 we now turn to consider certain factors which currently appear to condition choices between bare classifier and bare noun forms where both such patterns are available for definite reference.

#### 4. The choice of bare classifiers versus bare nouns

In examining alternations and preferences in the use of bare classifier and bare noun forms in contexts of definiteness, a number of general functional and

pragmatic pressures were noted to regularly affect the selection of the two patterns. While these factors, described below, by no means exhaustively account for alternations between the two forms, they do provide some insight into how choices in the pattern used to represent a definite referent may be made in certain instances where both patterns might be expected to be available.

#### 4.1 The role of contrast

The clearest set of data investigated with regard to a near universal strong preference of one pattern over another across the four languages involved sentences in which two definite referents occurred in a situation of contrast, as for example in equivalents to English sentences such as (23):

(23) *Mary has a dog and a cat. The dog is very old, but the cat is still quite young.*

In all four languages, there is a very strong pressure to encode such definite noun phrases with a bare classifier pattern, and a bare noun is often rejected or classed as inappropriate and unnatural in such contexts. A further, relevant observation underlining the important role of contrast in the selection of paired bare classifier forms is that in Bangla, bare nouns may potentially be used in place of bare classifier forms if heavy contrastive stress is added to the noun itself. Such stress is not necessary when bare classifier forms occur, which indicates that the presence of the classifier in the bare classifier pattern is important for the communication of contrast and regularly substitutes for the use of contrastive stress.<sup>7</sup>

In Vietnamese, a further special pattern was observed in sentences with pairs of definite noun phrases. Quite generally in Vietnamese, with certain nouns a classifier may often not occur in contexts of counting or use of 'one' as an indefinite article. This is frequently the case with human referents, where the classifier *người* is often omitted with numerically-quantified human nouns. Interestingly, when pairs of human referents occur in sentences which contrast certain properties of the referents, as in (24) below, the classifier *người* was spontaneously introduced by speakers in the bare classifier pattern, even though this classifier did not occur in the preceding sentence which introduced the noun. When definite noun phrases occur in contrasting anaphoric pairs, there is consequently a strong natural pressure to use the bare classifier pattern, and this may even override a tendency to omit classifiers with nouns in other contexts.

(24) *Thư viện vừa có thêm một kế toán và một luật sư.*  
 library just have add one accountant and one lawyer  
*Người kế toán rất chăm chỉ, nhưng người luật sư rất lười.*  
 CLF accountant very diligent but CLF lawyer very lazy

‘The library hired/has a new accountant and a new lawyer. The accountant is hard-working, but the lawyer is quite lazy.’

#### 4.2 The role of relative sentential prominence

In various instances in the investigation, it was noted that the use of (and preference for) the bare classifier pattern coincides with a certain clear prominence being associated with a definite referent — as, for example, in the contrastive pairs of referents noted above. Furthermore, when there is more than one contrastive pair of referents, it was found that sometimes only the most focally prominent of such pairs occurs in the bare classifier pattern whilst others may be substituted by the bare noun pattern. For example, in the Vietnamese passage in (25) below, ‘the table’ and ‘the chair’ in the final sentence were naturally represented with the bare noun pattern, as this sentence contains a second, new contrastive pair of referents with a higher level of prominence — the noun phrases ‘a book’ and ‘a glass’. The same effect occurs in (26).

- (25) *Trong một buổi biểu diễn ảo thuật ở Las Vegas,*  
 in one CLF performance magic in Las Vegas  
*nhà ảo thuật hướng dẫn người khán giả tình nguyện như sau:*  
 CLF magic instruct CLF audience volunteer like follow  
*‘Đây là một cái bàn và một cái ghế. Xin hãy đặt một quyển sách*  
 this be one CLF table and one CLF chair please place one CLF book  
*lên bàn, và đặt một ly nước lên ghế.’*  
 on table and place one CLF water on chair  
 ‘During a magic show in Las Vegas, a magician gives instructions to an audience volunteer/member of the audience: “Here is a table and a chair. Please put a book on the table, and put a glass on the chair.”’

- (26) *Hôm qua cảnh sát tìm thấy một chiếc xe hơi và một chiếc xe tải gần*  
 yesterday police find one CLF car and one CLF truck near  
*nhà ga Boston. Có một quả bom trong xe hơi và rất nhiều súng trong*  
 station Boston be one CLF bomb in car and very many gun in  
*xe tải.*  
 truck  
 ‘Yesterday the police found a car and a truck near Boston train station. There was a bomb in the car and many guns in the truck.’

To the extent that high relative sentential prominence leads to the use of the bare classifier pattern rather than the bare noun pattern, such patterns support observations in Daley (1998) and Bisang (1993) for Vietnamese and Hmong respectively (see also Aikhenvald 2000: 321–329 for some discussion). Daley (1998) in

particular attributes an important role to prominence in the favoring of bare classifier over bare noun forms in Vietnamese, whereas Bisang suggests that initial anaphoric mentions of a definite referent in Hmong may favor use of bare classifier forms, and subsequent references just bare nouns (or null pronominals). The present study was not able to strongly confirm such suggested generalizations, although prominence and activation-type effects were detected in various instances, as described above. In other instances, however, the importance of prominence did not appear to be so clear or decisive in the selection of the bare classifier over the bare noun pattern, and other factors may come into play. More work needs to be done in this regard to firm up conclusions.

If it is assumed that both contrast and high levels of prominence often lead to the use of the bare classifier rather than the bare noun pattern, one question that naturally arises is whether definite bare classifier forms functionally correspond to sentence topics or foci. We use the term sentence topic to refer to the referent which the proposition expressed by the sentence is about (Lambrecht 1994: 118). More specifically, “a referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee’s knowledge of this referent” (Lambrecht 1994: 131). By way of contrast, the focus of the proposition expressed by a sentence in a given utterance context is “the element of information whereby the presupposition and the assertion DIFFER from each other... It is the UNPREDICTABLE or pragmatically NON-RECOVERABLE element in an utterance” (Lambrecht 1994: 207). Here, three relevant observations can be made. First, bare classifier forms are often anaphoric repetitions of elements previously introduced in the discourse. In this sense they are topic-like (given/old information). Second, it is sometimes found that bare classifier forms also introduce new referents (as in instances of inferential definiteness in equivalents to English *I went to a wedding yesterday. The bride was very pretty.*), and so the bare classifier pattern might seem to instantiate an information focus. However, in this case, one can argue that the existence of a bride is at least partially presupposed given the previous sentence, which refers to a wedding. Thus, the bride may not necessarily constitute a new information focus in the second sentence. Third, in cases of ostensive first reference (equivalent to English: *Oh, (look) the door is locked*), where the bare classifier form is used for a contextually identifiable referent in its first explicit mention, it can be argued that direct visual identification (possibly accompanied by manual gesturing and eye movements) licenses the introduction of a new sentential topic. On balance, then, it can be suggested that the bare classifier pattern may regularly be used to represent topics rather than foci (even if prominent and potentially discourse-new), and these topics may either contrast with other topical referents, or be individual new/switch topics.

### 4.3 The role of disambiguation

A third functional factor at play in the favored selection of one pattern over the other is the disambiguating role that can be effected by the use of classifiers in the bare classifier pattern. The bare classifier pattern sometimes appears to be forced/heavily preferred by speakers in order to eliminate an unintended but salient interpretation that would naturally arise if the bare noun pattern were to be employed instead. Two particular instances of this phenomenon were noted to occur regularly in the investigation.

1. Singular vs. plural distinctions. In various cases, the use of bare nouns may result in a definite referent being given a plural interpretation. If unambiguous reference to a single definite individual is intended by a speaker, the bare classifier pattern needs to be used, as this eliminates any plural interpretation and can only be used to refer to single entities. For example, in the Bangla sentence in (27) below, if no classifier occurs, speakers indicate that the noun *alo* 'light' is naturally interpreted as a plural referent 'the lights' (or 'lights'). Use of the bare classifier form disambiguates (27) and signals that a single, referentially familiar light is being referred to:

- (27) *Tumi ki alo-Ta jele dite parbe please?*  
 you Q light-CLF turn.on give can please  
 'Can you turn on the light, please?'

2. Elimination of salient generic interpretations with stative predicates. The bare classifier pattern may also be selected to disambiguate sentences in which use of the bare noun pattern would result in a strong generic rather than individual interpretation of a referent/pair of referents. In the Vietnamese examples in (28) and (29) below, if a bare noun is used in the second sentences, speakers attest that it will naturally result in generic interpretations of the relevant noun phrases and the meaning: 'Cameras are expensive, but computers are cheap' and 'I like dogs but I don't like cats too much.'

- (28) *Hôm qua, Nam mua một cái máy ảnh và một cái máy vi tính.*  
 yesterday Nam buy one CLF camera and one CLF computer  
*Cái máy ảnh rất đắt tiền, nhưng cái máy tính thì khá rẻ.*  
 CLF camera very expensive but CLF computer PRT quite cheap  
 'Yesterday, Nam bought a camera and a computer. The camera was very expensive, but the computer was quite cheap.'
- (29) *Nga có một con chó và một con mèo. Tôi thích con chó, nhưng tôi không thích con mèo lắm.*  
 Nga has one CLF dog and one CLF cat I like CLF dog but I  
 không thích con mèo lắm.

not like CLF cat much  
 'Nga has a dog and a cat. I like the dog, but I don't like the cat too much.'

#### 4.4 The role of (contextual) uniqueness of the referent(s)

Certain apparent optionality between the bare noun and the bare classifier patterns in a “single” context — instances where speakers suggest that either the bare noun or the bare classifier is acceptable — appears to be due to differences in the background assumptions speakers hold about the contextual uniqueness of reference. In a number of sentences tested in the data set, somewhat different assumptions can be made about whether a referent is the unique entity of the type identified by the noun present in the context, or one of a set of similar entities. For example, speakers consulted in the investigation noted that sentences equivalent to English *Is the window closed?* might be used in a context in which a room is assumed to have only one window or alternatively a set of windows, with one particular window being identified as under discussion by speaker and hearer. This difference in background assumption may lead to different choices in the use of the bare noun and the bare classifier patterns. In Bangla, for example, speakers report that a bare noun is naturally used if the speaker assumes that there is only one window in the room, and a bare classifier pattern if it is assumed that there are several windows, hence the occurrence of the classifier in (30) appears to be optional but corresponds to rather different contexts in the mind of the speaker:

- (30) *janla(-Ta) ki bOndho ache?*  
 window(-CLF) Q closed be  
 'Is the window closed?'

A further example with a similar alternation in Bangla is (31), where the use of a bare noun occurs when the speaker assumes that there is a single guard present in the bank, and a classifier is inserted if it is assumed that there is more than one guard (and one particular guard from that set is under discussion):

- (31) *Bank-e Ramu Mina-ke bole: 'ami thik ki gOndogol bujhchi na.*  
 bank-LOC Ramu Mina-to say I rightly what problem understand not  
*kintu garD-(Ta) khub ghabre ache mone-hoche.*  
 but guard-CLF very scared is seems  
 'In a bank. Ramu says to Mina: "I wonder what's wrong. The guard seems to be very nervous."'

There is an interesting link that can be established between such alternations and the role of contrast previously mentioned. Intuitions offered by informants here suggest that the bare noun pattern seems to be licensed by **the absence of**

**contrast** — it occurs when a referent is the unique entity of the type referred to by the noun, and the bare classifier pattern occurs when an element is plausibly in contrast with other similar elements. It is also possible that the absence of any obvious contrast may explain the preference for the bare noun pattern in instances of definiteness licensed by association and inference, as in the Bangla example (32). Here it will be natural for speakers to assume that there is a unique bride present in the wedding situation, and this results in the natural use of a bare noun rather than a bare classifier form. The notion of contrast (and an absence of highlighted contrast) may thus be a common link in licensing the choice of the bare noun or the bare classifier pattern in various situations.

- (32) *GOtokal amra Ek-Ta biyebaRi giyechilam. Kone khub Sundori.*  
 yesterday we 1-CLF wedding went bride very pretty  
 ‘Yesterday we went to a wedding. The bride was beautiful.’

Quite generally, it can be claimed that two broad, important points are established by a closer consideration of the “optional” use of the bare noun or the bare classifier pattern in various instances. First, it emerges that the two patterns are frequently not simple functional equivalents but regularly associated with different aspects of meaning and contextual presupposition in the minds of speakers. The alternations between definite bare classifier and bare noun forms may therefore be predictable to a considerable extent. Second, in globally assessing the availability of the two patterns across a broad range of contexts, it is the bare noun pattern which turns out to be more regularly restricted in its occurrence and more frequently judged to be inappropriate when compared with the bare classifier pattern. Consequently, it may be that it is ultimately the availability of the bare noun pattern in contexts of definiteness that needs to be given some special explanation rather than the occurrence of the bare classifier pattern, which now functions more as a default form for definite referents, naturally licensed in a wider range of commonly-occurring contexts.

## 5. Conclusion

One of the goals of the current investigation of bare classifier patterns was to determine whether there is a sufficient degree of similarity in definite bare classifier patterns in Vietnamese, Hmong, Bangla and Cantonese that they should be viewed as constituting a single, cross-linguistic phenomenon. Having considered a wide range of patterns in the four languages, the conclusion is that there is significant similarity in the occurrence and use of definite bare classifier patterns in these four unrelated languages, and that they should therefore be grouped as instantiations of

a single, developing phenomenon. The investigation also examined alternations between bare classifier and bare noun forms in contexts of definiteness and concluded that these two patterns are (frequently) not direct equivalents and that a variety of factors may cause the use of either bare classifier or bare noun forms in different discourse contexts. There may also appear to be a spectrum of variation in potential alternations between bare classifier and bare noun patterns. Within a “single” language variety such as Cantonese, some speakers appear to be more permissive of the bare noun pattern than others and allow the bare noun pattern in a wider range of contexts. Quite possibly, the “rigid bare-classifier-only” form of Cantonese, which does not permit the use of bare nouns for definite referents, is an endpoint of development with bare classifiers — the bare classifier pattern has become fully grammaticalized for certain speakers as the marking for definite noun phrases, and these speakers no longer allow for contextual factors to potentially license an alternative bare noun pattern. In future comparative work charting the development and use of bare classifiers as potential markers of definiteness it will be useful to extend the range of study initiated here to other language families to see how and to what extent there are parallels or, alternatively, divergence from the patterns reported in the four language study initiated here.<sup>8</sup> The results gathered from experimental grammaticality judgments would also certainly benefit from testing against corpus studies for all languages being investigated, where such materials are available.

## Notes

\* The research on Hmong and Cantonese for this paper was supported by the College of Liberal Arts, University of Minnesota, through an Imagine Fund Award (2009-10) to Hooi Ling Soh. The authors would like to thank all the native speakers of Bangla, Vietnamese, Hmong and Cantonese who participated in the study, and to audiences in Ho Chi Minh City, Harvard, and Hong Kong, where this earlier versions of this work was presented in workshops and conferences.

1. Numeral classifiers represent only one type of classifier. Other types of classifier include noun classifiers, genitive classifiers and verbal classifiers. See Grinvald (2000) for an overview of the different classifier types and how they are different from gender and noun classes.
2. The following abbreviations are used: ACC: accusative case; ASP: aspectual marker; CLF: classifier; LOC: locative; MOD: modification marker; POSS: possessive marker; PRT: particle; Q: question particle; RED: reduplicant. To represent certain sounds in Bangla, the following symbols are used: T, D, R represent retroflex /t/, /d/, /r/. S is palato-alveolar /s/. N is a velar nasal. M is nasalization. E is an open mid front vowel.
3. Research on other varieties of Chinese is now revealing that the definite bare classifier pattern also occurs in several regional forms of Chinese (Cheng and Sybesma 2005, Zhang and Tang 2010).

4. Vietnamese and Hmong are SVO head-initial, isolating languages of Southeast Asia, genetically belonging to the Austro-Asiatic and the Hmong-Mien family respectively. The dialect of Hmong discussed in this paper is White Hmong. Bangla/Bengali is an SOV head-final language from the Indo-Aryan branch of Indo-European, in South Asia, with considerably more inflectional morphology than Cantonese, Vietnamese and Hmong.
5. Also, this speaker actually did allow for bare noun forms to encode definite reference with certain cultural or contextual unique entities: 'the president', 'the general manager' and 'the sun'.
6. Such tonal modification patterns with classifiers are interesting and have since been reported to occur in definite reference both with other varieties of Chinese (Zhang and Tang 2010), and with non-Sinitic languages such as the Tibeto-Burman language Yi (Hu and Jiang 2010).
7. The relevance of contrast is also found in the use of the pattern N-Adj-CLF in Thai (Hundius and Kolver 1983: 174, cited in Bisang 1999: 151–152).
8. For example, it is known that definite bare classifier forms are attested in the Tai-Kadai language family (Nung — see Saul and Wilson 1980) and Tibeto-Burman (Hu and Jiang 2010), but it is not yet clear how deeply grammaticalized the patterns currently are, and whether they alternate with the use of bare nouns as in Vietnamese, Hmong, etc.

## References

- Aikhenvald, Alexandra. 2000. *Classifiers*. Oxford: Oxford University Press.
- Bisang, Walter. 1993. Classifiers, quantifiers and class nouns in Hmong. *Studies in Language* 17. 1–51.
- Bisang, Walter. 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In Jadranka Gvozdanovic (ed.), *Numeral types and changes worldwide*, 113–185. Berlin: Mouton de Gruyter.
- Chen, Ping. 2004. Identifiability and definiteness in Chinese. *Linguistics* 42. 1129–1184.
- Cheng, Lisa L-S. & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30. 509–542.
- Cheng, Lisa L-S. & Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In Richard Kayne & Guglielmo Cinque (eds.), *The Oxford handbook of comparative syntax*, 259–292. Oxford: Oxford University Press.
- Daley, Karen. 1998. *Vietnamese classifiers in narrative texts*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Gil, David. 2005. Numeral classifiers. In Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds.), *The World Atlas of Language Structures*, Oxford: Oxford University Press, 226–29.
- Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In Gunter Senft (ed.), *Systems of nominal classification*, 50–92. Cambridge: Cambridge University Press.
- Hu, Suhua & Julie Li Jiang. 2010. The morpheme SU — determiner and complementizer in Nuosu Yi language. Paper presented at the 18th conference of the International Association of Chinese Linguistics, Harvard University.

- Hundius, Harald & Ulrike Kölver. 1983. Syntax and semantics of numeral classifiers in Thai. *Studies in Language* 7. 165–214.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Löbel, Elisabeth. 2000. Classifiers versus genders and noun classes: A case study in Vietnamese. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen & Mirja Saari (eds.), *Gender in grammar and cognition*, 259–320. Berlin: Mouton de Gruyter.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Matthews, Stephen & Patrizia Pacioni. 1996. Specificity and genericity of NPs in Cantonese and Mandarin. Paper presented at the Symposium on Referential Properties of Chinese Noun Phrases, City University of Hong Kong.
- Saul, Janice E. & Nancy F. Wilson. 1980. *Nung grammar*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asian languages. In Richard Kayne & Guglielmo Cinque (eds.), *The Oxford handbook of comparative syntax*, 806–838. Oxford: Oxford University Press.
- Wu, Yicheng & Adams Bodomo. 2009. Classifiers ≠ determiners. *Linguistic Inquiry* 40. 487–503.
- Zhang, Qingwen & Sze-Wing Tang. 2010. Cl-NP phrases in five Chinese dialects. Paper presented at the 18th conference of the International Association of Chinese Linguistics, Harvard University.

### *Authors' addresses*

Andrew Simpson  
Department of Linguistics  
GFS 301D  
University of Southern  
California  
Los Angeles,  
CA 90089-1693  
USA

andrew.simpson@usc.edu

Hooi Ling Soh  
Institute of Linguistics  
University of Minnesota  
214 Nolte Center  
315 Pillsbury Drive S.E.  
Minneapolis,  
MN 55455  
USA

sohxx001@umn.edu

Hiroki Nomoto  
Department of Southeast Asian  
Studies  
Tokyo University of Foreign Studies  
3-11-1 Asahi-cho,  
Fuchu,  
Tokyo 183-8534  
Japan

nomoto@tufs.ac.jp