

IDENTIFYING (IN)DEFINITENESS IN THE VIETNAMESE NOUN PHRASE

Trang Phan
VNU University of Languages & International
Studies, Vietnam National University, Hanoi
trangphan@vnu.edu.vn

Gennaro Chierchia
Harvard University
chierch@fas.harvard.edu

Abstract

This paper aims to settle the issue of whether *những*, *các*, *một* are articles in Vietnamese as argued by Nguyen T. C. (1975), Nguyen H. T. (2004), a.o. First, we adopt Dayal (in prep.)'s cross-linguistic questionnaire of (in)definiteness since this questionnaire offers us a set of useful tests to diagnose definiteness and indefiniteness from a crosslinguistic perspective. Second, we broaden up the empirical landscape by contrasting the interpretation of nominal constructions which have the so-called overt (in)definite markers (*các*-CLF-N, *những*-CLF-N, and *một*-CLF-N) with that of nominal constructions without them (including bare N and CLF-N, numeral(>1)-CLF-N), in order to see if the (in)definiteness effect truly comes from the presence or absence of these three markers, or from something else. We then conclude that (i) *những* and *các* are plural markers, (ii) only *một* seems to be a likely candidate for an indefinite article, and (iii) bare nouns and numerals are not genuine indefinites: the former denotes kinds, while the latter can be interpreted as definite, which sets Vietnamese apart cross-linguistically.

Keywords: definite, indefinite, bare noun, numeral, plural, classifier, article, Vietnamese
ISO 639-3 codes: vie

1 Introduction

This paper returns to the still-hotly-debated issue of the existence of lexical articles in Vietnamese. Traditionally, most researchers (Emeneau 1951, Thompson 1965, Nguyen K. T. 1981, Nguyen D. H. 1997, among others) hold the view that there are no lexical articles in Vietnamese. Nguyen T. C. (1975) is one exception, claiming that *những*, *các*, and *một* are articles.

“Trước từ loại danh từ của tiếng Việt thường có thể gặp ba từ công cụ có tác dụng ngữ pháp và ý nghĩa ngữ pháp khá gần với các article ở các tiếng Pháp, Ý, Anh, Đức v.v.: đó là ba từ “những”, “các”, “một” mà chúng tôi tạm gọi là quán từ” (Nguyen T. C. 1975:251)

[Before a noun in Vietnamese, there are three functional words whose grammatical function and grammatical meaning are pretty close to that of articles in languages like French, Italian, English, German, *etc.* They are “*những*”, “*các*”, “*một*”, i.e., by which we tentatively call articles] (Translation ours).

The question of whether Vietnamese has articles was revived for modern discussion by Nguyen H. T. (2004). Building on Nguyen T. C. (1975)'s insights, Nguyen H. T. (2004:42) posits that Vietnamese has three articles, which differ from one another in terms of definiteness and number.

- (1) *một* 'one': [-Plural, -Definite]
những: [+Plural, -Definite]
các: [+Plural, +Definite] (Nguyen H. T. 2004:42)

One crucial piece of evidence for the definiteness split, between *một* and *những* on the one hand and *các* on the other, comes from the observation that only the former is able to appear in existential constructions, which are typical contexts for indefinites:

- (2) a. Ngày xưa có **một** bà hoàng hậu độc ác
 day old have one CLF¹ queen wicked
 'Once upon a time there was a wicked queen.' (Nguyen H. T. 2004:32)
- b. Có **những** bản nhạc nghe mãi không biết chán
 have PL CLF music listen forever not know bored
 'There are those musical pieces you can listen to forever without feeling bored.'
 (Nguyen H. T. 2004:37)
- c. *Có **các** bản nhạc nghe mãi không biết chán
 have PL CLF music listen forever not know bored
 'There are the musical pieces you can listen to forever without feeling bored.'
 (Nguyen H. T. 2004:41)

To this, we add a new observation. Looking more closely at the examples provided by Nguyen H. T. in (2), if we remove the postnominal modifier (i.e., a reduced relative clause, in this case), we can make an even more fine-grained distinction, between *một* on the one hand and *những* and *các* on the other: only the former survives in bare existential constructions.

- (3) a. Có *một* bà hoàng hậu
 have one CLF queen
 'There is a queen.'
- b. *Có **những/các** bản nhạc
 have PL / PL CLF music
 *'There are those/the musical pieces.'

The full picture that emerges is as follows: *một*-CLF-N can appear in existential constructions, with or without a postnominal modifier; *các* can never appear in existential constructions, with or without a postnominal modifier; and *những*-CLF-N can appear in existential construction only in the presence of a postnominal modifier (which rescues the grammaticality of *những* in 2b).

Looking beyond Vietnamese, we also observe that the obligatory presence of a modifier phrase in existential constructions is established in French as well, as shown in (4).

- (4) Il y a le livre *(que j'ai lu) sur la table
 there Y has the book that I.have read on the table
 'There's a book I've read on the table.' (Ihsane and Puskas 2001:52)

Thus, the fact that *những* can appear in existential constructions only in the presence of modification such as a relative clause does not *per se* mean that *những* is indefinite. After all, the comparable example in (4) from French has the definite article *le*. In English, we also have a minimal contrast between the grammatical 'There are those musical pieces you can listen to forever without feeling bored' and the ungrammatical *'There are the musical pieces you can listen to forever without feeling bored'. That is to say, what we learn

¹ Abbreviations used in the paper: ANT: anterior, CLF: classifier, DUR: durative, FUT: future, NEG: negation, PL: plural, PROG: progressive, 2SG: second singular pronoun.

here is that *những* actually behaves similar to the French definite article and the English demonstrative in existential constructions and *những* seems to require a restriction on the noun it modifies, which clearly deserves closer examination.² It suffices for now to conclude that the existential construction is not a reliable test for (in)definiteness and that the distinctions within *những*, *các*, and *một* need to be fine-tuned.

Recently, Phan & Lam (2021) provide several objections to Nguyen H. T. (2004)'s proposal that *những*, *các*, *một* form a paradigm of genuine lexical articles in Vietnamese. The authors argue that *những* and *các* are different from the English definite article *the* in terms of their optionality and their incompatibility with numerals. For instance, the nominal phrase is definite either without or with *những* and *các*, as seen in (5a) and (5b) respectively.

- (5) a. *Cuốn sách rất cũ*³
 CLF book very old
 'The book is very old.'
- b. *Những/Các cuốn sách rất cũ*
 PL / PL CLF book very old
 'The books are very old.'

Crucially, unlike English *the*, Vietnamese *những* and *các* cannot precede the numerals as in (6b).

- (6) a. **The three** books
 b. **Những/Các ba cuốn sách*
 PL / PL three CLF book
 Intended meaning: 'The three books'

Phan & Lam (2021:6) also observe that *một* is sharply distinguished from *những* and *các* in being excluded from those strong definite contexts as in (7-8).

- (7) a. *Những/Các cuốn sách ấy*
 PL / PL CLF book that
 'Those books'
- b. **Một cuốn sách ấy*
 One CLF book that

² In addition to existential constructions, Nguyen H. T. (2004) also lists *wh*-pronouns as another indefinite context which is compatible with *những*:

(i) *Anh gặp những ai? Anh làm những gì?*
 2SG meet PL who 2SG do PL what
 'Who are those/the people that you meet? What are those/the things that you do?'
 (Nguyen H.T.'s example 2004: 38, translation ours)

As can be seen from the translation of (i) even in this *wh*-pronoun context, *những* still behaves like a definite article or a demonstrative and it also requires a restriction on the noun it modifies. Why it is so must be left for future research. See Phan & Lam (2021), however, for a possible interpretation, namely *những*-CLF-N-Modifier encodes the so-called 'activated givenness' in the sense of Gundel et al. (1993), i.e. its givenness retrieves activation from immediate linguistic contexts. See also Le & Schmitt (2018) for further discussion.

³ As can be seen from example (5a) and subsequent discussion, CLF-N sequence is definite-oriented in Vietnamese (see also Bisang and Quang 2020 for a corpus-based piece of evidence). However, we do not take classifiers as definite articles in Vietnamese since the definite reading of the CLF-N sequence can be overridden by the context. According to Sudo & Trinh (2009), Vietnamese CLF-N can be interpreted as indefinite in object positions, i.e., under the scope of Existential closure at the VP level. Interested readers are referred to Nguyen H. T. (2004), Simpson & Ngo (2018), Đoàn et al. (2019), Phan & Lam (2021) for further discussion and to Phan & Chierchia (submitted) for an explanation. We thank a reviewer for raising this point.

- (8) a. *Những / Các* *cuốn* *sách* *cũ* *nhất*
 PL / PL CLF book old most
 ‘The oldest books’
- b. **Một* *cuốn* *sách* *cũ* *nhất*
 One CLF book old most

Phan & Lam carefully state that ‘Vietnamese has no genuine lexical definite articles’ (2021:5) and ‘only *một* can be characterized as an indefinite marker’ (2021:6). These insightful observations are worth pursuing further.

Since Phan & Lam (2021) only focus on definiteness, more tests for indefiniteness are needed. Furthermore, more nominal constructions need to be investigated in order to see whether the presence or absence of the so-called overt (in)definite markers truly makes any difference for the definite as well as indefinite interpretation of the nominal phrase.

In order to fix the first problem, we adopt Dayal (in prep.)’s cross-linguistic questionnaire of (in)definiteness since this questionnaire offers us a set of useful tests to diagnose both definiteness and indefiniteness from a crosslinguistic perspective.

In order to address the second concern, we broaden up the empirical landscape by contrasting the interpretation of nominal constructions which have the three elements (*các*-CLF-N, *những*-CLF-N, and *một*-CLF-N) with that of nominal constructions without them (including bare N and CLF-N, numeral(>1)-CLF-N), in order to see if the (in)definiteness effect truly comes from the presence or absence of these three markers, or from something else.

- (9) Types of nominal constructions under investigation:
- a. bare nouns
 - b. *các* – classifier – noun phrases
 - c. *những* – classifier – noun phrases
 - d. classifier – noun phrases
 - e. *một* – classifier – noun phrases
 - f. numeral (>1) – classifier – noun phrases

We will carefully examine both sides of the coin: whether the listed nominal constructions in (9) survive over different definiteness AND indefiniteness diagnostics.

Our study, therefore, aims to address the following research questions:

- (10) Research questions:
- (i) Do *những*, *các*, and *một* (in comparison to bare nouns, bare classifiers, and numerals larger than one) survive the definiteness diagnostics?
 - (ii) Do *những*, *các*, and *một* (in comparison to bare nouns, bare classifiers, and numerals larger than one) survive the indefiniteness diagnostics?
 - (iii) What can we conclude about (in)definiteness marking in Vietnamese noun phrase?⁴

The rest of the paper is organized as follows. Section 2 tests the definiteness of the six nominal constructions under investigation. Section 3 diagnoses the indefiniteness of these constructions, as well as expanding the empirical coverage by looking at the kind-denoting reading and the genericity reading. Section 4 draws conclusions regarding (in)definite marking in Vietnamese noun phrase and crosslinguistic implications.

2 Testing for Definiteness

Dayal (in prep.) offers four diagnostics to detect whether a nominal phrase is definite or not. Specifically, a nominal phrase is interpreted as definite if (i) it refers to unique referent(s), (ii) not just to a member or some partitive members of a set, and (iii) it can refer to previously mentioned referents while (iv) it cannot hold of

⁴ (In)definiteness can be determined by a variety of non-nominal factors including structural position, verbal aspect, etc. (Diesing 1992, Abusch 1994, among others). In this paper, we are only concerned with the (in)definiteness of the noun phrase.

two predicates at the same time. Let us examine the behaviour of six different nominal constructions in Vietnamese in each context.

2.1 Uniqueness/maximality and partitivity

The first two diagnostics involve uniqueness (and its plural counterpart: maximality) and partitivity in different types of contexts including deictic contexts, wider contexts, and bridging contexts.

In the first place, we would like to see if in deictic contexts, the nominal construction under discussion refers to the unique individual or the maximal set of individuals (Context a below), OR refers to a subset of individuals (Context b below). Here we repeat Polinsky, Soloveva & Dayal (2020)'s contexts given for Russian, but we differ from their study in testing uniqueness/maximality and partitivity at the same time, since it is reasonable to expect that there is a complementary distribution between the two readings. Specifically, those nominals which refer to unique/maximal referents should not be interpreted as partitive, and vice versa, those which have a partitive reading should not be interpreted as unique/maximal.

This complementarity prediction is indeed borne out in Vietnamese.

- (11) *Tôi lấy hoa hồng*
I take flower rose

'I take the rose.'

a. AVAILABLE: spoken by a customer who entered a flower shop, pointing at the only rose in the set of flowers on sale or the only bunch of roses among various bunches of flowers on sale.

✓unique/✓maximal

b. UNAVAILABLE: spoken by a customer who entered a flower shop, pointing at one rose out of several roses on sale/ one bunch of roses out of several bunches of roses on sale.

*partitive

- (12) *Tôi lấy những bông hoa hồng*
I take PL CLF flower rose

'I take the roses.'

a. AVAILABLE: spoken by a customer who entered a flower shop, pointing at the only bunch of roses among various bunches of flowers on sale.

✓maximal

b. UNAVAILABLE: spoken by a customer who entered a flower shop, pointing at one bunch of roses out of several bunches of roses on sale.

*partitive

- (13) *Tôi lấy các bông hoa hồng*
I take PL CLF flower rose

'I take the roses.'

a. AVAILABLE: spoken by a customer who entered a flower shop, pointing at the only bunch of roses among various bunches of flowers on sale.

✓maximal

b. UNAVAILABLE: spoken by a customer who entered a flower shop, pointing at one bunch of roses out of several bunches of roses on sale.

*partitive

- (14) *Tôi lấy bông hoa hồng*
I take CLF flower rose

'I take the rose.'

a. AVAILABLE: spoken by a customer who entered a flower shop, pointing at the only rose in the set of flowers on sale.

✓unique

b. UNAVAILABLE: spoken by a customer who entered a flower shop, pointing at one rose out of several roses on sale.

*partitive

- (15) *Tôi lấy một bông hoa hồng*
 I take one CLF flower rose
 ‘I take one rose.’
 a. UNAVAILABLE: spoken by a customer who entered a flower shop, pointing at the only rose in the set of flowers on sale. *unique
 b. AVAILABLE: spoken by a customer who entered a flower shop, pointing at one rose out of several roses on sale. ✓partitive
- (16) *Tôi lấy hai bông hoa hồng*
 I take two CLF flower rose
 ‘I take (the) two roses.’
 a. AVAILABLE: spoken by a customer who entered a flower shop, pointing at the only two roses in the set of flowers on sale. ✓maximal
 b. AVAILABLE: spoken by a customer who entered a flower shop, pointing at two roses out of several roses on sale. ✓partitive

The uniqueness/maximality reading obtains with bare nouns, *những*-CLF-N, *các*-CLF-N, and CLF-N, as in (11a), (12a), (13a), and (14a) respectively; but is unavailable for *một*-CLF-N, as in (15a). The opposite pattern holds for the partitive reading: *một*-CLF-N must be interpreted partitively, as in (15b); whereas bare nouns, *những*-CLF-N, *các*-CLF-N, and CLF-N cannot receive the partitive reading, as seen in (11b), (12b), (13b), and (14b).

The same complementarity between the unique/maximal versus the partitive reading is also observed in the so-called ‘wider contexts’, in which world knowledge tells us that presidents of any country are unique, as in (17):

- (17)a. *Thủ tướng Nhật Bản đã soạn thảo bộ luật di trú*
 President Japan ANT compose CLF law immigration
 ✓ ‘The president(s) of Japan composed the Immigration Law.’ ✓unique/✓maximal
 * ‘One/Some of the president(s) of Japan composed the Immigration Law.’ *partitive
- b. *Những thủ tướng Nhật Bản đã soạn thảo bộ luật di trú*
 PL president Japan ANT compose CLF law immigration
 ✓ ‘The presidents of Japan composed the Immigration Law.’ ✓maximal
 * ‘Some of the presidents of Japan composed the Immigration Law.’ *partitive
- c. *Các thủ tướng Nhật Bản đã soạn thảo bộ luật di trú*
 PL president Japan ANT compose CLF law immigration
 ✓ ‘The presidents of Japan composed the Immigration Law.’ ✓maximal
 * ‘Some of the presidents of Japan composed the Immigration Law.’ *partitive
- d. *Ngài thủ tướng Nhật Bản đã soạn thảo bộ luật di trú*
 CLF president Japan ANT compose CLF law immigration
 ✓ ‘The president of Japan composed the Immigration Law.’ ✓unique
 * ‘One of the presidents of Japan composed the Immigration Law.’ *partitive
- e. *Một thủ tướng Nhật Bản đã soạn thảo bộ luật di trú*
 One president Japan ANT compose CLF law immigration
 * ‘The president of Japan composed the Immigration Law.’ *unique
 ✓ ‘A president of Japan composed the Immigration Law.’ ✓partitive

- f. **Hai** **thủ tướng** **Nhật Bản** **đã** **soạn thảo** **bộ** **luật** **di trú**
 two president Japan ANT compose CLF law immigration
 ✓ ‘Two presidents of Japan composed the Immigration Law.’ ✓ maximal
 ✓ ‘The two presidents of Japan composed the Immigration Law.’ ✓ partitive

as well as in “bridging” contexts (Clark 1975, Lyons 1999) in which the linking between the bedroom and the discourse-familiar apartment is a part-whole relation, as in (18):

- (18) **Nam** **mới** **mua** **một** **căn** **chung cư**
 Nam just buy one CLF apartment
 ‘Nam just bought an apartment.’
- a. **Phòng ngủ** **rất** **rộng**
 Bedroom very spacious
 ✓ ‘The bedroom(s) is/are very large.’ ✓unique/✓maximal
 * ‘One/Some of the bedroom(s) is/are very large.’ *partitive
- b. **Những** **cái** **phòng ngủ** **rất** **rộng**
 PL CLF bedroom very spacious
 ✓ ‘(All) The bedrooms are very large.’ ✓maximal
 * ‘Some of the bedrooms are very large.’ *partitive
- c. **Các** **cái** **phòng ngủ** **rất** **rộng**
 PL CLF bedroom very spacious
 ✓ ‘(All) The bedrooms are very large.’ ✓maximal
 * ‘Some of the bedrooms are very large.’ *partitive
- d. **Cái** **phòng ngủ** **rất** **rộng**
 CLF bedroom very spacious
 ✓ ‘The (only) bedroom is very large.’ ✓unique
 * ‘One of the bedrooms is very large.’ *partitive
- e. **Một** **cái** **phòng ngủ** **rất** **rộng**
 One CLF bedroom very spacious
 * ‘The (only) bedroom is very large.’ *unique
 ✓ ‘One of the bedrooms is very large.’ ✓partitive
- f. **Hai** **cái** **phòng ngủ** **rất** **rộng**
 two CLF bedroom very spacious
 ✓ ‘The two bedrooms are very large.’ ✓ maximal
 ✓ ‘Two bedrooms are very large.’ ✓partitive

A few comments are in order.

First, the uniqueness/maximality test allows us to make the first cut among the six nominal constructions in terms of definiteness: between *một*-CLF-N and the other constructions, with only the former failing in this context. Furthermore, *một*-CLF-N is indeed the mirror image of CLF-N. Both refer to a singular referent, *một*-CLF-N must not be unique, whereas CLF-N must not be partitive (compare (14) vs (15), (17d) vs (17e), (18d) vs (18e) all pointing to the same contrast).

Second, there is a number-sensitive distinction throughout the three contexts: the referent of bare nouns can be either unique or maximal, as in (11), (14a), (15a) whereas that of *những/các*-CLF-N must be maximal, as in (12-13), (17b-c), (18b-c), and that of CLF-N must be unique, as in (14), (17d), (18d).

Third, unlike what is expected from the paradigm in (1) proposed by Nguyen H. T. (2004: 118), *những*-CLF-N does not behave like an indefinite nominal construction: it allows only the maximal reading, not the partitive reading, as seen in (12), (17b), (18b).

2.2 Anaphoricity

Let us now turn to a strong definite context, namely anaphoric contexts, to see which nominal constructions do or do not survive the anaphoricity diagnostic.

(19) *Đầu năm mẹ mua cho Nam một quyển vở*
 Early year mom buy for Nam one CLF notebook
 ‘Mum bought for Nam one notebook earlier this year.’

a. *Vở giờ vẫn còn mới*
 notebook now still still new
 ‘The notebook is now still new.’ ✓anaphorically definite

b. *Quyển vở giờ vẫn còn mới*
 CLF notebook now still still new
 ‘The notebook is now still new.’ ✓anaphorically definite

c. **Một quyển vở giờ vẫn còn mới*
 one CLF notebook now still still new
 *‘The notebook is now still new.’ *anaphorically definite

(20) *Đầu năm mẹ mua cho Nam mười quyển vở*
 Early year mom buy for Nam ten CLF notebook
 ‘Mum bought for Nam ten notebooks earlier this year.’

a. *Vở giờ vẫn còn mới*
 notebook now still still new
 ‘The notebooks are now still new.’ ✓anaphorically definite

b. *Những quyển vở giờ vẫn còn mới*
 PL CLF notebook now still still new
 ‘The notebooks are now still new.’ ✓anaphorically definite

c. *Các quyển vở giờ vẫn còn mới*
 PL CLF notebook now still still new
 ‘The notebooks are now still new.’ ✓anaphorically definite

d. *Mười quyển vở giờ vẫn còn mới*
 Ten CLF notebook now still still new
 ‘The ten notebooks are now still new.’ ✓anaphorically definite

The anaphoricity test gives us the same cut as the uniqueness/maximality test, also singling out *một*-CLF-N from other constructions: only *một*-CLF-N fails the anaphoric definite diagnostic, as seen in the unavailability of the definite reading of *một*-CLF-N in (19c).

Furthermore, similar to uniqueness/maximality, anaphoricity is also sensitive to number. In particular, for the anaphorically definite constructions, the bare N is open to both singular and plural readings as seen in (19a-20a), whereas the CLF-N must be interpreted as singular as seen in (19b), and the *những*-CLF-N, *các*-CLF-N and numeral (>1)-CLF-N must be interpreted plural, as in (20b-c-d).

One thing to emphasize here is that the Vietnamese numeral (>1)-CLF-N construction passes the anaphoricity test, as in (20d). This is a typologically distinctive feature of the Vietnamese numeral construction in comparison to its counterparts in better-studied languages like English and Chinese. It is well-documented in the literature that English and Chinese numeral constructions fail the anaphoricity test (Jiang 2012):

(21) John bought three dogs and five cats. #(The/Those) three dogs and five cats are very expensive.
(Example of Jiang 2012: 75)

(22) *jiaoshi li zuo zhe san ge nansheng he wu*
classroom inside sit PROG three CLF boy and five
ge nusheng, #(na) wu ge nusheng hen chao-si le
CLF girl that/those five CLF girl very noisy-dead SFP
‘There are three boys and five girls in the classroom. Those five girls are so noisy.’
(Example from Jiang 2012: 121)

English and Chinese numeral constructions cannot be interpreted as anaphorically definite without the accompany of a definite article or a demonstrative, as in (21-22), whereas Vietnamese numeral (>1)-CLF-N can perfectly be interpreted as anaphorically definite on its own, as in (20d). We will come back to this later, but it suffices to note for now that numeral constructions with numbers larger than one can be interpreted as maximally definite and anaphorically definite in Vietnamese.

2.3 The compatible predicate test

The last definiteness diagnostic involves what we tentatively call the ‘compatible predicate test’,⁵ which asks whether the candidate for definiteness can take two different predicates which cannot hold of a single individual (or of a single set of individuals) at the same time.

- (23) a. **Chó đang ngủ còn chó đang chạy*
dog DUR sleep and dog DUR run
*‘The dogs are sleeping and the dogs are running.’ *compatible predicate test
- b. **Những con chó đang ngủ còn những con chó đang chạy*
PL CLF dog DUR sleep and PL CLF dog DUR run
*‘The dogs are sleeping and the dogs are running.’ *compatible predicate test
- c. **Các con chó đang ngủ còn các con chó đang chạy*
PL CLF dog DUR sleep and PL CLF dog DUR run
*‘The dogs are sleeping and the dogs are running.’ *compatible predicate test
- d. **Con chó đang ngủ còn con chó đang chạy*
CLF dog DUR sleep and CLF dog DUR run
*‘The dog is sleeping and the dog is running.’ *compatible predicate test
- e. *Một con chó đang ngủ còn một con chó đang chạy*
one CLF dog DUR sleep and one CLF dog DUR run
✓ ‘One dog is sleeping and one dog is running.’ ✓ compatible predicate test

⁵ Originally Dayal (in prep.) dubbed this as ‘the homogeneity test’. However, in order to avoid unwarranted terminological confusion and to stay theory neutral, we decided to rename the test as ‘the compatible predicate test’.

- f. *Hai con chó đang ngủ còn hai con chó đang chạy*
 two CLF dog DUR sleep and two CLF dog DUR run
 ✓ ‘Two dogs are sleeping and two dogs are running.’ ✓compatible predicate test

The predicate test lets us make the second cut among the six nominal constructions when it comes to definiteness: between *một*-CLF-N and *hai*-CLF-N on the one hand and the bare noun, CLF-N, and *những/các*-CLF-N on the other. Only the former produces grammatical sentences since they can refer to distinct individuals. In this regard, the numerals behave like regular indefinites.

2.4 Discussion

Table 1 summarizes the behaviours of the six nominal constructions on the different definiteness diagnostics. The four diagnostics of definiteness in the first column of Table 1 can be categorized into two subgroups: the first two are uniqueness/maximality-inducing contexts (i.e., referring to unique/maximal individual(s) in the discourse and to previously introduced referent(s)), whereas the last two are anti-uniqueness/maximality contexts (referring to one or some partitive members of a set, and being held of different predicates which apply to different individuals at the same time). A definite nominal construction is expected to survive in the uniqueness/maximality contexts, but be rejected in the anti-uniqueness/maximality contexts. In Table 1, + indicates that the reading is available for that construction, and – indicates it is unavailable. Cells colored in green indicate normal behavior for definites, and yellow indicates non-definite behavior.

Table 1: Definiteness across six nominal constructions

Definiteness	<i>các</i> -CLF-N	CLF-N	<i>những</i> -CLF-N	Bare N	<i>hai</i> -CLF-N	<i>một</i> -CLF-N
Uniqueness/ maximality	+	+	+	+	+	-
Anaphoricity	+	+	+	+	+	-
Partitivity	-	-	-	-	+	+
Compatible predicates	-	-	-	-	+	+

As seen above, there are two cuts we can make among the six nominal constructions with respect to the two subgroups of definiteness-diagnosing contexts: the first cut between *một*-CLF-N and the other constructions (only the former is incompatible with uniqueness induced contexts), and the second cut between numeral-CLF-N and the other constructions (only the former is compatible with anti-uniqueness/maximality contexts).

That is to say, different nominal constructions correspond to different degrees of definiteness: *các*-CLF-N, CLF-N, *những*-CLF-N, and bare N are the most definite constructions (in surviving the uniqueness/maximality induced contexts and at the same time rejected from anti-uniqueness/maximality contexts), *một*-CLF-N is the least definite construction (in being incompatible with uniqueness contexts and compatible with anti-uniqueness contexts), and numeral(>1)-CLF-N is the hybrid construction (in surviving both uniqueness and anti-uniqueness contexts).

What else does Table 1 tell us about our three protagonists *những*, *các*, and *một*?

First, *những*-CLF-N and *các*-CLF-N belong to the same group with CLF-N and bare N, which all seem to behave as typical definite constructions in being interpreted as maximally definite as well as anaphorically definite, despite their difference in number distinction (*các* and *những* are plural, whereas CLF-N is singular, bare nouns are number-neutral). That is to say, in light of Schwarz (2013)’s distinction between weak (uniqueness-based) definite and strong (familiarity/anaphoricity-based) definite, all of these four nominal constructions can be considered strong definites. As expected, they do not survive the partitivity test nor the compatible predicate test, which force a non-uniqueness/non-maximality reading. At this point, we would like to emphasize that our investigation reveals that even bare nouns can be strong definites in Vietnamese, in contrast to what is observed by Trinh (2011: 640) who claims that bare nouns cannot be definite at all in

Vietnamese. So with regard to the debate on whether bare nouns in classifier languages can only be weak definites (as argued by Jenks 2018 for Mandarin Chinese) or can be strong definites (as defended by Dayal & Jiang 2021 for Mandarin Chinese), Vietnamese seems to provide straightforward supporting evidence for the latter claim.

Second, in light of the set of diagnostics for definiteness in Table 1, we are able to tease apart *một*-CLF-N from other numeral constructions: similar to numeral(>1)-CLF-N, *một*-CLF-N passes the two anti-uniqueness contexts, BUT unlike numeral(>1)-CLF-N, *một*-CLF-N does not survive the uniqueness and anaphoricity contexts. Furthermore, all the four diagnostics highlight one important aspect of *một* which has been previously neglected: *một*-CLF-N is consistently the mirror image of CLF-N. Both the constructions have the function of singling out one individual, but CLF-N is definite-oriented in Vietnamese (see also Bisang and Quang 2020 for a corpus-based confirmation), whereas adding *một* to the CLF-N sequence results in an indefinite reading. These points seem to suggest that *một* appears to be an indefinite article in Vietnamese.

Third, unlike *một*-CLF-N, numeral(>1)-CLF-N shows a hybrid behaviour. Like a definite, *hai*-CLF-N patterns with *các*-CLF-N, CLF-N, *những*-CLF-N, and bare N in surviving the uniqueness test and the anaphoricity test. But like an indefinite, *hai*-CLF-N also survives the partitivity and the compatible predicate test. This demonstrates that numeral(>1)-CLF-N can freely shift between definite and indefinite readings. It is also important to note that the fact that Vietnamese numeral-CLF-N (with numerals larger than one) can be anaphorically definite clearly differentiates Vietnamese from Chinese and English.

Last but not least, the four diagnostics of definiteness only tell us that *các*-CLF-N, CLF-N, *những*-CLF-N, and bare N CAN all be definite,⁶ but they do not tell us whether these constructions MUST be definite. It is time to turn to the other side of the coin in order to apply the indefiniteness diagnostics. If these constructions are inherently definite, then we should expect them not to survive the indefiniteness tests.

3 Testing for Indefiniteness

The aim of this section is to apply Dayal's (in prep.) heuristic tests for indefiniteness to the six nominal constructions under investigation. In applying these tests (including the ability to introduce first-mentioned discourse referents, together with different scope-related tests), we find that these nominal constructions also exhibit different degrees of indefiniteness.

3.1 First-mentioned discourse referents

Storytelling contexts are a classic environment which foregrounds the ability of an indefinite nominal to introduce new (first-mentioned) discourse referents. This test brings out the first cut among the six constructions when it comes to indefiniteness, between bare nouns, *một*-CLF-N, numeral (>1)-CLF-N, and *những*-CLF-N on the one hand and *các*-CLF-N and CLF-N on the other hand. The former can introduce novel referents to the discourse, whereas the latter cannot, as shown in the unavailability of the first-mentioned reading of (24e) and (24f).

- (24) a. *Hai trăm năm trước, phù thủy từng sống ở đây*
 two hundred year ago witch used.to live in here
 'Two hundred years ago, a witch/ witches used to live here.' ✓ first mentioned
- b. *Hai trăm năm trước, một cụ phù thủy từng sống ở đây*
 two hundred year ago one CLF witch used.to live in here
 'Two hundred year ago, a witch used to live here.' ✓ first mentioned
- c. *Hai trăm năm trước, hai cụ phù thủy từng sống ở đây*
 two hundred year ago two CLF witch used live in here
 'Two hundred years ago, two witches used to live here.' ✓ first mentioned

⁶ In fact, Phan & Lam (2021) further argued that there are indeed subtle differences among these definite constructions if we adopt a more fine-grained givenness hierarchy proposed by Gundel *et al* (1993) which distinguishes up to six levels of givenness.

- d. Hai trăm năm trước, **những** **mụ** **phù thủy** từng sống ở đây
 two hundred year ago PL CLF witch used.to live in here
 ‘Two hundred years ago, witches used to live here.’ ✓ first mentioned
 OR ‘Two hundred years ago, the witches used to live here.’
- e. Hai trăm năm trước, **các** **mụ** **phù thủy** từng sống ở đây
 two hundred year ago PL CLF witch used.to live in here
 *‘Two hundred years ago, witches used to live here.’ * first mentioned
 Can only meant: ‘Two hundred years ago, the witches used to live here.’
- f. Hai trăm năm trước, **mụ** **phù thủy** từng sống ở đây
 two hundred year ago CLF witch used.to live in here
 *‘Two hundred years ago, a witch used to live here.’ * first mentioned
 Can only meant: ‘Two hundred years ago, the witch used to live here.’

Two observations arise. First, the storytelling context offers us a potential way to distinguish between the two plural markers in Vietnamese: as seen in (24d) vs (24e), although both *những* and *các* can refer to previously mentioned referents (‘the witches’), only *những* can be used to introduce first-mentioned referents.

Second, on this diagnostic, bare nouns pattern with indefinite-oriented *một*-CLF-N. It is worth bearing in mind that in this storytelling context, what looks like an indefinite reading of the bare noun seems to in fact be a kind-based reading in which witches, not humans used to live here. This contrasts with the other nominal constructions that allow first mention, which do seem to have indefinite readings. It raises the question of whether bare nouns are genuine indefinites or the indefinite reading of bare nouns is derived from its kind-denoting property. The answer to this question will become clear when we take into consideration the scopal effect of bare nouns with respect to intensional verbs like want, sentential negators, and adverbial operators.

3.2 Opacity

Let us now look at the scope of these nominals with respect to intensional verbs like want, which distinguishes between the so-called opaque reading and transparent reading. For instance, say that Nam wants to meet professors during his stay at Harvard. If he is indifferent about which professor he wants to meet, such that any professor would do, we have the opaque reading (want > ∃). If Nam has a particular professor in mind (though the speaker does not know who that is because Nam hasn’t told anyone), we have the transparent reading (∃ > want).

- (25) Nam muốn gặp **giáo sư** ở Harvard
 Nam want meet professor at Harvard
 ‘Nam wants to meet a professor/professors at Harvard.’
- a. **giáo sư** nào cũng được
 professor which also ok
 ‘any professor would do.’ ✓opaque (want > ∃)
- b. **nhưng** tôi không biết **giáo sư** nào
 but I not know professor which
 ‘but I don’t know which one.’ ✓transparent (∃ > want)
- (26) Nam muốn gặp **một** **vị** **giáo sư** ở Harvard
 Nam want meet one CLF professor at Harvard
 ‘Nam wants to meet a professor at Harvard.’
- a. **vị** **giáo sư** nào cũng được
 CLF professor which also ok
 ‘any professor would do.’ ✓opaque (want > ∃)
- b. **nhưng** tôi không biết **vị** **nào**
 but I not know CLF which
 ‘but I don’t know which one.’ ✓transparent (∃ > want)

- (27) *Nam muốn gặp hai vị giáo sư ở Harvard*
 Nam want meet two CLF professor at Harvard
 ‘Nam wants to meet two professors at Harvard.’
- a. *hai vị giáo sư nào cũng được*
 two CLF professor which also ok
 ‘any two professors would do.’
 ✓opaque (want > ∃)
- b. *nhưng tôi không biết hai vị nào*
 but I not know PL CLF which
 ‘but I don’t know which ones.’
 ✓transparent (∃ > want)
- (28) *Nam muốn gặp vị giáo sư ở Harvard*
 Nam want meet CLF professor at Harvard
 ‘Nam wants to meet the professor at Harvard.’
- a. **vị giáo sư nào cũng được*
 CLF professor which also ok
 ‘any professor would do.’
 *opaque (want > ∃)
- b. *nhưng tôi không biết vị nào⁷*
 but I not know CLF which
 ‘but I don’t know which one.’
 ✓transparent (∃ > want)
- (29) *Nam muốn gặp những vị giáo sư ở Harvard*
 Nam want meet PL CLF professor at Harvard
 ‘Nam wants to meet the professors at Harvard.’
- a. **những vị giáo sư nào cũng được*
 PL CLF professor which also ok
 ‘any professor would do.’
 *opaque (want > ∃)
- b. *nhưng tôi không biết những vị nào*
 but I not know PL CLF which
 ‘but I don’t know which ones.’
 ✓transparent (∃ > want)
- (30) *Nam muốn gặp các vị giáo sư ở Harvard*
 Nam want meet PL CLF professor at Harvard
 ‘Nam wants to meet the professors at Harvard.’
- a. **các vị giáo sư nào cũng được*
 PL CLF professor which also ok
 ‘any professor would do.’
 *opaque (want > ∃)
- b. *nhưng tôi không biết các vị nào*
 but I not know PL CLF which
 ‘but I don’t know which ones.’
 ✓transparent (∃ > want)

This opacity test makes the second cut among the six constructions with respect to indefiniteness: between bare nouns, *một*-CLF-N, and numeral (>1)-CLF-N on the one hand and *các*-CLF-N, CLF-N, and *những*-CLF-N on the other hand. The former group can have either the opaque reading or the transparent reading, whereas the latter group cannot have the opaque reading, only the transparent reading. Similar to what we saw in the storytelling context above, bare nouns seem to pattern with indefinite-oriented constructions including *một*-CLF-N. Therefore, we really need to decide whether there are any deep distinctions between bare nouns and genuine indefinites, which leads us to another scopal test: the intermediate scope reading.

⁷ (26b) indicates that Vietnamese CLF-N sequence is interpreted as indefinite when it functions as the object of an intensional predicate *muốn* (‘want’). See Sudo and Trinh (2009) for more contexts of indefinite CLF-N. We thank a reviewer for discussing this point.

3.3 Intermediate scope reading

Only a true indefinite can have an intermediate scope reading in which its scope is narrower than one sentential scope-bearing element but wider than another scope-bearing element (see Farkas 1981, Ruys 1992, Abusch 1994, among others). Let us translate the classical context provided by Ruys (1992) in (31) into Vietnamese in (32).

(31) Every professor will rejoice if a student/3 students of his cheat on the exam.

(Ruys 1992)

a. Narrow scope reading: Every > If > a student of his: For every professor, the cheating on the exam by any student of his, he will rejoice.

b. Intermediate scope reading: Every > a student of his > if: For every professor, there is a specific student of his, if this student of his cheats on the exam, he will rejoice.

The ‘narrow scope reading’ is an acceptable reading for all of these constructions, whereas the ‘intermediate scope reading’ is only available with numerals-CLF-N.

(32) a. Mọi giáo viên đều sẽ vui nếu sinh viên của họ
Every teacher all FUT rejoice if student of them
gian lận trong bài thi
cheat in CLF exam

√Narrow scope reading: ‘For every professor, regarding the cheating on the exam by any student of his, he will rejoice.’

*Intermediate scope reading: ‘For every professor, there is a specific student of his, and if this student of his cheats on the exam, he will rejoice.’

b. Mọi giáo viên đều sẽ vui nếu cậu sinh viên của họ
Every teacher all FUT rejoice if CLF student of them
gian lận trong bài thi
cheat on CLF exam

√Narrow scope reading: ‘For every professor, there is an unique student of his, and if this student of his cheats on the exam, he will rejoice.’

*Intermediate scope reading: ‘For every professor, regarding the cheating on the exam by any student of his, he will rejoice.’

c. Mọi giáo viên đều sẽ vui nếu các sinh viên của họ
Every teacher all FUT rejoice if PL student of them
gian lận trong bài thi
cheat on CLF exam

√Narrow scope reading: ‘For every professor, regarding the cheating on the exam by all students of his, he will rejoice.’

*Intermediate scope reading: ‘For every professor, there are some specific students of his, and if these students of his cheat on the exam, he will rejoice.’

d. Mọi giáo viên đều sẽ vui nếu những sinh viên của họ
Every teacher all FUT rejoice if PL student of them
gian lận trong bài thi
cheat on CLF exam

√Narrow scope reading: ‘For every professor, regarding the cheating on the exam by all students of his, he will rejoice.’

*Intermediate scope reading: ‘For every professor, there are some specific students of his, and if these students of his cheat on the exam, he will rejoice.’

- e. Mọi giáo viên đều sẽ vui nếu một sinh viên của họ gian lận trong bài thi
 Every teacher all FUT rejoice if one student of them cheat on CLF exam

✓Narrow scope reading: ‘For every professor, regarding the cheating on the exam by any student of his, he will rejoice.’

✓Intermediate scope reading: ‘For every professor, there is a specific student of his, and if this student of his cheats on the exam, he will rejoice.’

- f. Mọi giáo viên đều sẽ vui nếu hai sinh viên của họ gian lận trong bài thi
 Every teacher all FUT rejoice if two student of them cheat on CLF exam

✓Narrow scope reading: ‘For every professor, regarding the cheating on the exam by any two students of his, he will rejoice.’

✓Intermediate scope reading: ‘For every professor, there is two specific students of his, and if students of his cheat on the exam, he will rejoice.’

The intermediate scope reading makes the third cut regarding indefiniteness: a fine-grained distinction between bare Ns, CLF-N, and PL-CLF-N on the one hand and numerals-CLF-N (including one and numerals larger than one) on the other hand: only the numerals allow the intermediate scope reading, as seen in (32e-f). This suggests that numerals, not bare nouns, can behave like typical indefinites in Vietnamese.

3.4 Wide scope with respect to negation

The deep division between bare nouns and genuine indefinites is further highlighted by their different scope with respect to negation. Only regular indefinites, not bare nouns, can take scope above negation.

- (33) a. Mai đã không lấy sách tôi để trên bàn
 Mai ANT not take book I leave on table

✓Narrow scope with respect to negation: ‘Mai didn’t take any books that I left on the table.’

*Wide scope with respect to negation: ‘Mai didn’t take some of the books that I left on the table.’

- b. Mai đã không lấy cuốn sách tôi để trên bàn
 Mai ANT not take CLF book I leave on table

✓Narrow scope with respect to negation: ‘Mai didn’t take the book that I left on the table.’

*Wide scope with respect to negation: ‘Mai didn’t take one of the books that I left on the table.’

- c. Mai đã không lấy những cuốn sách tôi để trên bàn
 Mai ANT not take PL CLF book I leave on table

✓Narrow scope with respect to negation: ‘Mai didn’t take the books that I left on the table.’

*Wide scope with respect to negation: ‘Mai didn’t take some of the books that I left on the table.’

- d. Mai đã không lấy các cuốn sách tôi để trên bàn
 Mai ANT not take PL CLF book I leave on table

‘Mai didn’t take the books that I left on the table.’

✓Narrow scope with respect to negation: ‘Mai didn’t take the books that I left on the table.’

*Wide scope with respect to negation: ‘Mai didn’t take some of the books that I left on the table.’

- e. *Mai đã không lấy một cuốn sách tôi để trên bàn*
 Mai ANT not take one CLF book I leave on table
 ‘Mai didn’t take one of the books that I left on the table.’
 *Narrow scope with respect to negation: ‘Mai didn’t take the book that I left on the table.’
 ✓Wide scope with respect to negation: ‘Mai didn’t take one of the books that I left on table.’
- f. *Mai đã không lấy hai cuốn sách tôi để trên bàn*
 Mai ANT not take two CLF book I leave on table
 ‘Mai didn’t take two of the books that I left on the table.’
 ✓Narrow scope with respect to negation: ‘Mai didn’t take the two books that I left on the table.’
 ✓Wide scope with respect to negation: ‘Mai didn’t take two of the books that I left on table.’

The scope with respect to negation test gives us the same cut between numerals-CLF-N (including one and numerals larger than one) on the one hand and the other nominal constructions on the other hand: only the former allows wide scope over negation. This test provides us another piece of supporting evidence for the observation that Vietnamese bare nouns are not genuine indefinites.

3.5 Differentiated Scope

In order to see what exactly makes bare nouns fundamentally different from genuine indefinites, we need to apply another scopal test, namely Differentiated Scope. This test is originally designed to highlight the so-called ‘scopelessness’ property of bare nouns (Carlson 1977, Chierchia 1995), which is in obvious contrast with that of canonical indefinites.

This test holds that only bare nouns, not regular indefinites, can take scope below certain adverbial operators such as *repeatedly*. In (34a), only bare nouns allow for different chickens to be bled (as a part of the slaughtering process), i.e., a *repeatedly* > \exists reading. This differentiated scope reading is unavailable for other nominal constructions including regular indefinites, which require the same chicken (or the same set of chickens) to undergo being bled, i.e., an \exists > *repeatedly* reading.

- (34) a. *Nam đã cắt tiết gà nhiều lần*
 Nam ANT cut blood chicken many time
 ✓ ‘Nam bled a/the chicken (the same one) many times.’ ✓ \exists > *repeatedly*
 ✓ ‘Nam bled chickens (different ones) many times.’ ✓ *repeatedly* > \exists
- b. *Nam đã cắt tiết con gà nhiều lần*
 Nam ANT cut blood CLF chicken many time
 ✓ ‘Nam bled the chicken (the same one) many times.’ ✓ \exists > *repeatedly*
 * ‘Nam bled a (possibly different) chicken many times.’ **repeatedly* > \exists
- c. *Nam đã cắt tiết một con gà nhiều lần*
 Nam ANT cut blood one CLF chicken many time
 ✓ ‘Nam bled a chicken (the same one) many times.’ ✓ \exists > *repeatedly*
 * ‘Nam bled a (possibly different) chicken many times.’ **repeatedly* > \exists
- d. *Nam đã cắt tiết những con gà nhiều lần*
 Nam ANT cut blood PL CLF chicken many time
 ✓ ‘Nam bled the chickens (of the same set) many times.’ ✓ \exists > *repeatedly*
 * ‘Nam bled the chickens (of different sets) many times.’ **repeatedly* > \exists
- e. *Nam đã cắt tiết các con gà nhiều lần*
 Nam ANT cut blood PL CLF chicken many time
 ✓ ‘Nam bled the chickens (of the same set) many times.’ ✓ \exists > *repeatedly*
 * ‘Nam bled the chickens (of different sets) many times.’ **repeatedly* > \exists

- f. *Nam đã cắt tiết hai con gà nhiều lần*
 Nam ANT cut blood PL CLF chicken many time
 ✓ ‘Nam bled the two chickens (of the same set) many times.’ ✓ $\exists > repeatedly$
 * ‘Nam bled two chickens (of different sets) many times.’ **repeatedly* $> \exists$

This test thus not only distinguishes bare nouns from genuine indefinites but also singles out bare nouns from canonical definites. So this test does not detect indefiniteness per se; what it really tells us is that bare nouns are of a different nature from the other constructions.

3.6 Kind terms and genericity

In order to see the true color of bare nouns, we must go beyond definiteness and indefiniteness to the territory of kind terms, which is to be distinguished from genericity. That is, only bare nouns can denote kinds, enabling them to co-occur with kind-level predicates which express generalizations across classes of individuals.

- (35) a. *Chó tiến hoá từ sói*
 Dog evolve from wolf
 ✓ ‘Dogs evolved from wolves.’
- b. *Con chó tiến hoá từ con sói*
 CLF dog evolve from CLF wolf
 *‘Dogs evolved from wolves.’
 only: ‘The dog evolved from the wolf.’
- c. *Những con chó tiến hoá từ những con sói*
 PL CLF dog evolve from PL CLF wolf
 * ‘Dogs evolved from wolves.’
- d. *Các con chó tiến hoá từ các con sói*
 PL CLF dog evolve from PL CLF wolf
 * ‘Dogs evolved from wolves.’
- e. *Một con chó tiến hoá từ một con sói*
 One CLF dog evolve from one CLF wolf
 * ‘Dogs evolved from wolves.’
- f. *Hai con chó tiến hoá từ hai con sói*
 two CLF dog evolve from two CLF wolf
 * ‘Dogs evolved from wolves.’

(35) tells us that among these nominal constructions, only bare nouns are compatible with kind-level predicates like *evolve*, suggesting that only bare nouns can be kind-denoting terms. On the other hand, *những*-CLF-N, *các*-CLF-N, *một*-CLF-N, and *hai*-CLF-N cannot be kind-denoting terms, and CLF-N can only have a taxonomy contrastive interpretation, as in (35b).⁸

Note that kind-level predicates should be distinguished from individual-level predicates. The latter holds true throughout the existence of an individual and is compatible with all nominal constructions.

⁸ See Dayal (2004) for how a taxonomy reading is different from genuine kind terms.

- (36) a. **Chó** *sủa* *khi* *nó/chúng* *đói*
 dog bark when it/they hungry
 ‘Dogs bark when they are hungry.’
- b. **Con chó** *sủa* *khi* *nó* *đói*
 CLF dog bark when it hungry
 ‘The dog barks when it is hungry.’
 ‘The dog (in general) usually barks when it is hungry.’
- c. **Những con chó** *thường sủa* *khi* *chúng* *đói*
 PL CLF dog usually bark when they hungry
 ‘The dogs usually bark when they are hungry.’
 ‘The dogs (in general) usually bark when they are hungry.’
- d. **Các con chó** *thường sủa* *khi* *chúng* *đói*
 PL CLF dog usually bark when they hungry
 ‘The dogs usually bark when they are hungry.’
 ‘The dogs (in general) usually bark when they are hungry.’
- e. **Một con chó** *thường sủa* *khi* *nó* *đói*
 one CLF dog usually bark when it hungry
 ‘A dog usually barks when it is hungry.’
- f. **Hai con chó** *thường đánh nhau* *khi* *bị nhốt chung*
 two CLF dog usually fight each.other when got confine together
 ‘Two dogs fight when they are confine together.’

Examples (36) show that all nominal constructions can express genericity, either on their own, or in conjunction with habitual adverbs (usually).

Overall, the distinction between kind terms and genericity, between kind-level predicates and individual-level predicates, suggests that although all the nominal constructions can express genericity, only bare nouns are able to denote genuine kind terms.

3.7 Discussion

The seven tests utilized in Section 3 can be categorized into two subgroups: the first four detect indefiniteness, while the last three are diagnostics for kind terms and genericity. Tables 2 and 3 below summarize the behaviours of the six nominal constructions with respect to the two subgroups respectively.

Similar to what is observed in the set of definiteness diagnostics, different nominal constructions also correspond to different degrees of indefiniteness, as shown in Table 2. Similar to Table 1, in Table 2, + indicates that the reading is available for that construction, and – indicates it is unavailable. Cells colored in green indicate normal behavior for definites, and yellow indicates non-definite behavior.

As illustrated above, there are three divisions among these constructions with respect to indefiniteness. The first cut is between *những*-CLF-N, numeral-CLF-N, and bare nouns versus the other constructions (only the former is able to introduce new discourse referents). The second cut separates numeral-CLF-N and bare nouns from the other constructions (only the former can obtain the opaque reading with respect to the intensional verb *want*). The third cut singles out numeral-CLF-N among these different constructions (only numeral-CLF-N can have the intermediate scope reading and wide scope above negation). That is to say, numeral-CLF-N is the most indefinite constructions (in passing all the four diagnostics of indefiniteness), *các*-CLF-N and CLF-N are the least indefinite constructions (in failing all these four diagnostics), and bare N (passing two out of four) and *những*-CLF-N (one out of four) are the hybrid constructions.

Table 2: Indefiniteness across six nominal constructions

Indefiniteness	<i>các</i> -CLF-N	CLF-N	<i>những</i> -CLF-N	Bare N	<i>hai</i> -CLF-N	<i>một</i> -CLF-N
First-mentioned discourse referent	-	-	+	+	+	+
Opacity (<i>want</i> > \exists)	-	-	-	+	+	+
Intermediate scope (<i>every</i> > \exists > <i>if</i>)	-	-	-	-	+	+
Negation (\exists > \neg)	-	-	-	-	+	+

Furthermore, the last three diagnostics enable us to go beyond definiteness and indefiniteness. Although all the nominal constructions can express genericity, only bare nouns have the special status of being able to express genuine kind terms, as summarized in Table 3. In Table 3, + indicates that the reading is available for that construction, and – indicates it is unavailable. Cells colored in blue indicate normal behavior for kind terms, and grey indicates non-kind terms, pink indicates the genericity reading.

Table 3: Kind terms & genericity across six nominal constructions

Kind terms & genericity		Bare N	CLF-N	<i>các</i> -CLF-N	<i>những</i> -CLF-N	<i>một</i> -CLF-N	<i>hai</i> -CLF-N
Kind terms	Differentiated Scope (<i>repeatedly</i> > \exists) ⁹	+	-	-	-	-	-
	Kind-level predication	+	- (taxonomy only)	-	-	-	-
Genericity		+	+	+	+	+	+

The so-called scopelessness of bare nouns (i.e., bearing narrow scope with respect to sentential adverbs like *repeatedly*) is in fact derived from its property as kind-denoting. This contrasts starkly with genuine indefinites. In this regard, Vietnamese does not support the Ambiguity Approach (Krifka 1988, Wilkinson 1991, Diesing 1992, Kratzer 1995) which argues that bare nouns are ambiguous between kind-referring and indefinites. Instead, Vietnamese data are in favour of the Neocarlsonian approach (Carlson 1977, 1989, Chierchia 1998) in which it is posited that bare nouns are kind referring at their core, and the other interpretations of bare nouns – including both definites and indefinites – can be derived from the kind reference.

4 Conclusion

In this section, we deal with the final research question: what does our multi-dimensional investigation reveal about (in)definiteness marking in Vietnamese noun phrase?

Let us now bring Tables 1 and 2 together in order to see whether *những*, *các*, *một* truly form a paradigm of lexical articles as argued by Nguyen T. C. (1975) and Nguyen H. T. (2004). Similar to Tables 1 and 2, in Table 4, + indicates that the reading is available for that construction, and – indicates it is unavailable. Cells colored in green indicate normal behavior for definites, and yellow indicates non-definite behavior.

First of all, our findings from Table 4 about *những* suggest that unlike what is claimed in the literature, *những*-CLF-N is not strictly indefinite. Rather, it can either be indefinite and definite, and in fact the definite reading is even more dominant than the indefinite one. This argues against the claim of Nguyen T. C. (1975)

⁹ Note that in Dayal (in prep.)’s original questionnaire, Differentiated Scope is classified as a diagnostic for Indefiniteness. Vietnamese data, however, suggest that it is more appropriate to locate this test in the diagnostics for a kind-denoting property.

and Nguyen H. T. (2004) that *những* is an indefinite article, instead demonstrating that at its core it is simply a plural marker, and (in)definiteness is not an inherent component of the meaning of *những*.

Table 4: (In)definiteness across six nominal constructions

	Uniqueness/ Maximality	Anaphoricity	Partitivity	Compatible predicate test	Intermediate scope	Neg	Opacity	First mentioned
CLF-N	+	+	-	-	-	-	-	-
<i>các</i> - CLF-N	+	+	-	-	-	-	-	-
<i>những</i> - CLF-N	+	+	-	-	-	-	-	+
bare N	+	+	-	-	-	-	+	+
Num> 1- CLF-N	+	+	+	+	+	+	+	+
<i>một</i> - CLF-N	-	-	+	+	+	+	+	+

Unlike *những*, *các*-CLF-N seems to be consistently definite, passing all the tests for definiteness and failing all the tests for indefiniteness. The question thus arises whether *các* can be considered a genuine definite article. Close scrutiny of Table 4 suggests that the answer to this question is a negative one. If *các* is a fully fledged article like English *the*, we should expect its presence to really make an interpretive difference in definiteness (for instance, in English, bare nouns, without the overt article *the*, can never be definite). The crucial observation is that, throughout these diagnostics, *các*-CLF-N behaves in parallel fashion with CLF-N, both as typical definites. That means that the definiteness reading is already obtained at CLF-N level, and the insertion of *các* on top of CLF-N simply adds up the plural reading.¹⁰ Furthermore, as already pointed out by Phan & Lander (2015) and Phan & Lam (2021), *các* is distributionally different from English *the* in at least two regards: (i) its optionality and (ii) its incompatibility with numerals, as illustrated in (5) and (6) repeated here as (37) and in (38), respectively:

- (37) a. *Cuốn sách rất cũ*
 CLF book very old
 ‘The book is very old.’
 b. *Các cuốn sách rất cũ*
 PL CLF book very old
 ‘The books are very old.’
- (38) a. **Các ba cuốn sách rất cũ*
 PL three CLF book very old
 ‘The three books are very old.’
 b. *Ba cuốn sách rất cũ*
 Three CLF book very old
 ‘The three books are very old.’

¹⁰ A reviewer raises the question of what happens with those Vietnamese noun phrases which do not take classifiers, such as *các sinh viên* (‘the students’) or *các chính phủ* (‘the governments’). Would the definiteness come from the noun or the pluralizer? Phan & Chierchia (submitted) argue that there is a null classifier underlyingly and a covert semantic operator which is responsible for the definiteness of such constructions.

(37a) is perfectly definite, without *các*. Therefore, the sole contribution of *các* in (37b) is plurality, and unlike English *the*, the presence of *các* is not obligatory for the definiteness of the nominal phrase. The presence of *các* can even lead to an ill-formed definite construction, as seen in (38a): *các* is obligatorily absent in the presence of a numeral. These sentences also highlight another crucial difference between Vietnamese *các* and English *the*: the former is a quantity word (containing a plurality feature) whereas the latter is not. Therefore, we can safely conclude that unlike what is argued by Nguyen T. C. (1975) and Nguyen H. T. (2004), *các* is not a genuine definite article, either.

Let us now focus on *một*. Our empirical contribution to the literature is that we not only contrast *một* with *những* and *các* but also with bare classifiers as well as with numerals larger than one. If adding *các* to the CLF-N sequence makes no difference in terms of definiteness (CLF-N and *các*-CLF-N are both definite oriented), adding *một*, on the other hand, clearly results in an indefinite reading. It can be also seen from Table 4 that unlike other numerals, *một* is unable to switch to the definite reading. That is to say, among the three article candidates listed by Nguyen T. C. (1975) and Nguyen H. T. (2004), only *một* seems to qualify as an indefinite article.

Bringing everything together, we are now able to identify the key insights about (in)definiteness marking in Vietnamese noun phrase as follows:¹¹

- i. Different nominal constructions correspond to different degrees of (in)definiteness.
- ii. Even though they are definite-oriented, *những* and *các* are not fully fledged articles since without their presence on top of the CLF-N sequence, the bare CLF-N construction can still be perfectly definite. *Những* and *các* are simply plural markers.
- iii. The presence of *một* on top of CLF-N, however, turns a definite-oriented CLF-N into an indefinite-oriented *một*-CLF-N. Furthermore, *một*-CLF-N can never switch to definite. Therefore, among the three candidates, only *một* seems to be a likely candidate for an indefinite article.¹²
- iv. Bare nouns can be interpreted not only as definite, but also as indefinite. However, different from genuine indefinites, the indefinite interpretation of bare nouns is derived from the fact that bare nouns denote kinds.
- v. Num-CLF-N in Vietnamese are not genuine indefinites, either, since in addition to the indefinite reading, Num-CLF-N can be interpreted as definite, which sets Vietnamese apart cross-linguistically.

Acknowledgments

This research has been done under the research project QG.22.32 “Definiteness in Vietnamese noun phrase” of Vietnam National University, Hanoi” to the first author, Trang Phan.

References

- Abusch, Dorit. 1994. The Scope of Indefinites. *Natural Language Semantics* 3, 88-135.
- Bisang, Walter & Quang, Kim Ngoc. 2020. (In)definiteness and Vietnamese classifiers. In: Kata Balogh, Anja Latrouite & Robert D. Van Valin, Jr. (eds.), *Nominal anchoring: Specificity, definiteness and article systems across languages*, 15–49. Berlin: Language Science Press.
- Carlson, Greg N. 1977. *Reference to kinds in English*. Doctoral Dissertation, University of Massachusetts, Amherst.
- Carlson, Greg N. 1989. The Semantic Composition of English Generic Sentences. In Gennaro Chierchia, Barbara Partee and Raymond Turner (eds.), *Properties, Types and Meaning*, 167-192. Dordrecht: Kluwer Academic Publishers.
- Chierchia, Gennaro. 1995. Individual-level predicates as inherent generics. In Gregory N. Carlson and Francis Pelletier (eds.), *The generic book*, 176-223. Chicago: University of Chicago Press.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6: 339-405.

¹¹ It goes without saying that the claims in this study need to be empirically tested with corpus driven data.

¹² Cross-linguistically, it is not unusual to find those languages which have no definite articles but indefinite articles. According to Dryer (2013), 45 out of 620 languages have no definite articles but indefinite articles.

- Clark, Herbert H. 1975. *Bridging*. In R.C. Shank and B.L. Nash Webber (eds.), *Theoretical issues in natural language processing*. New York: Association for Computer Machinery.
- Dayal, Veneeta. 2004. Number Marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27:4, 393-450.
- Dayal, Veneeta. (in prep.) *The Open Handbook of (In)definiteness: A Hitchhiker's Guide to interpreting bare arguments*. MIT Open Handbooks in Linguistics.
- Dayal, Veneeta & Li Julie Jiang. 2021. The Puzzle of Anaphoric Bare Nouns in Mandarin: A Counterpoint to Index!. *Linguistic Inquiry* December 2021:1-20.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Dryer, Matthew. (2013). Indefinite Articles. In M. S. Dryer, & M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/38>
- Đoàn, Thị Quý Ngọc, Martin B.H. Everaert & Eric J. Reuland. 2019. (In)definiteness of Vietnamese noun phrases. In Nigel Duffield, Trang Phan and Tue Trinh (eds.), *Interdisciplinary Perspectives on Vietnamese Linguistics*. Studies in Language Companion Series 211, 155–180. Amsterdam: John Benjamins.
- Emeneau, Murray Barnson. 1951. *Studies in Vietnamese (Annamese) grammar*. Berkeley: University of California
- Farkas, Donka F. 1981. Quantifier scope and syntactic islands. *Chicago Linguistics Society* 17: 59-66.
- Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274-307.
- Ihsane, Tabea & Genoveva Puskás. 2001. Specific is not Definite. *Generative Grammar in Geneva* 2. 39–54.
- Jiang, L. Julie. 2012. *Nominal arguments and language variation*. Doctoral Dissertation, Harvard University.
- Kratzer, Agelika 1995. Stage-level and Individual-level Predicates. In Gregory N. Carlson. and Francis J. Pelletier (eds.), *The generic book*, 125-175. Chicago: University of Chicago Press.
- Krifka, Manfred. 1988. The Relational Theory of Genericity. In Krifka Manfred (ed.), *Genericity in Natural Language*. Seminar für natürlich-sprachliche Systeme, Tübingen.
- Nguyen, Tai Can. 1975. *Từ loại danh từ trong tiếng Việt hiện đại* [The word class of nouns in modern Vietnamese]. Hanoi: Khoa Học Xã Hội.
- Nguyen, Dinh Hoa. 1997. *Vietnamese: Tiếng Việt không sơn phấn* [Vietnamese without veneer]. Philadelphia: John Benjamins.
- Nguyen, Kim Than. 1981. *Cơ sở ngữ pháp tiếng Việt* [Foundation of Vietnamese grammar]. Saigon: Thành phố Hồ Chí Minh.
- Nguyen, Hung Tuong. 2004. *The structure of the Vietnamese Noun Phrase*. Boston: Boston University dissertation.
- Lê, Ni-La, Forsythe, Hannah, & Schmitt, Cristina. 2018. Interactions between number and definiteness: Vietnamese children's comprehension of definite noun phrases. In Anne B. Bertolini & Maxwell J. Kaplan (eds), *Proceedings of the 42nd Annual Boston University Conference on Language Development*, 427–440. Somerville, MA: Cascadilla Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Nguyen, Hung Tuong. 2004. *The structure of the Vietnamese Noun Phrase*. Boston: Boston University dissertation.
- Phan, Trang & Lander, Eric. 2015. Vietnamese and the NP-DP parameter. *Canadian Journal of Linguistics* 60(3): 391-415.
- Phan, Trang & Lam, Quang Dong. 2021. Decomposing Definiteness in Vietnamese. *Journal of the Southeast Asian Linguistics Society* 14 (1): 1-18.
- Phan, Trang & Chierchia, Gennaro. (submitted). Putting three pieces of a puzzle into place: classifiers - plurals - articles in Vietnamese. *Taiwan Journal of Linguistics*.

- Polinsky, Maria, Soloveva, Anita & Dayal, Veneeta. 2020. *(In)definiteness in Russian: A Case Study*. Ms., Harvard University.
- Ruys, Eddy. 1992. *The Scope of Indefinites*. Doctoral Dissertation, Utrecht University.
- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7:534–559.
- Simpson, Andrew & Ngo, Binh. 2018. Classifier syntax in Vietnamese. *Journal of East Asian Linguistics* 27 (3): 211-246.
- Sudo, Yasutada & Trinh, Tue. 2009. The indefiniteness effect in Mandarin and Vietnamese. Talk, UCLA, 31 January 2009.
- Thompson, Laurence. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.
- Trinh, Tue. 2011. Nominal reference in two classifier languages. *Proceedings of Sinn und Bedeutung* 15: 629–644.
- Wilkinson, Karina. 1991. *Studies in the Semantics of Generic NP's*. Doctoral dissertation, University of Massachusetts at Amherst.

Reviewed: Received 22 September 2021, revised text accepted 14 July 2022, published 1 September 2022
Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Nathan Hill, Dr Sigrid Lew