

Classifier syntax in Vietnamese

Andrew Simpson¹ · Binh Ngo¹

Received: 20 March 2015 / Accepted: 16 March 2018 / Published online: 6 June 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract Vietnamese is a language with a three-way split in the appearance of numeral classifiers when nouns are counted: some nouns always require classifiers (obligatory-classifier nouns), others occur only optionally with classifiers (optional-classifier nouns), and a third group never combines with a classifier (non-classified nouns). This distribution provides potentially important information on the much debated question of whether classifiers functionally combine with numerals (Bale and Coon in *Linguist Inq* 45:695–707, 2014) or with nouns (Li in *Linguist Inq* 29 (4):693–702, 1998; Cheng and Sybesma in *Linguist Inq* 30:509–542, 1999). It also appears to challenge Chierchia’s (Nat Lang Semant 6(4):339–405, 1998) characterization of the basic semantic type of nouns found in different languages, which assumes a uniform pattern of classifier occurrence in numeral classifier languages. Having described the broad distribution of classifiers in Vietnamese and the questions this raises, the article probes the syntactic properties of classifiers with the three types of noun in the language, considering double classifier patterns, fragment answers, passive constructions, and the use of classifiers with certain compound nouns. Evidence from such phenomena is shown to support the hypothesis that a uniform syntactic structure is actually projected with nouns of all types in Vietnamese, but sometimes masked by the use of nouns to overtly lexicalize both the N and CL positions in nominal projections through N-to-CL movement.

Keywords Numeral classifiers · Typology of nouns · Vietnamese

✉ Andrew Simpson
andrew.simpson@usc.edu
Binh Ngo
binhngo@usc.edu

¹ Department of Linguistics, GFS 301, University of Southern California, Los Angeles, CA 90089-1693, USA

1 Introduction

The surface patterning of numeral classifiers in Vietnamese is in various ways more complex than that observed in other better-described classifier languages of East Asia such as Chinese, Japanese and Korean, as Vietnamese shows significant variation in the obligatory, optional and also non-occurrence of classifiers when nouns of different types occur in certain constructions. This paper sets out to provide an analysis of such complex variation and how it also bears on two general theoretical approaches to the syntactic status and role of classifiers—Bale and Coon's recent (2014) claim, following Krifka (1995), that classifiers of all types syntactically combine with numerals before modifying nouns, in order to enable numerals' counting function, and Chierchia's (1998) prominent typology of nouns across languages, in which the presence or absence of classifiers partially establishes what semantic type nominals may have in a particular language. It will be shown that the richness and apparent flexibility of classifier patterns in Vietnamese pose a natural challenge to the Bale and Coon/Krifka characterization of the function of classifiers if assumed to be universal, and also to the simple four-way typology of nominals proposed in Chierchia (1998). The paper also presents a detailed syntactic examination of the varying surface distribution of classifiers in different environments, and argues that a uniform functional structure is in fact projected with count nouns of all types in Vietnamese.

The structure of the paper is as follows. Section 2 introduces a basic three-way division in the presence/absence of classifiers with nouns in counting constructions in Vietnamese, and shows how this interacts with Bale and Coon's analysis of classifiers proposed on the basis of Mi'gmaq (Algonquian) and Chol (Mayan). Section 3 examines how Vietnamese figures in Chierchia's (1998) nominal typology of languages, and whether a single syntactic structure occurs with all nouns when combined with numerals, or whether Vietnamese might be a 'mixed' nominal language, with a classifier position only projected with certain types of count noun. Section 4 makes use of a range of tests to further probe the syntax of classifiers in Vietnamese, looking at the distribution of classifiers and nouns in the 'extra *cái*' construction (where an additional classifier *cái* occurs), noun-numeral separation patterns in passive sentences, and complex patterns found with compound nouns headed by kinship terms. The general conclusions of the paper are then presented in Sect. 5.

2 Noun-dependent variation in the presence of classifiers in Vietnamese, and Bale and Coon's (2014) approach to numeral classifiers

When numerals are combined with entity-denoting count nouns in Chinese, Japanese and Korean, the presence of a classifier is regularly required, as illustrated in (1a–c):¹

- (1) a. liang *(suo) fangzi *Mandarin* b. san *(dai no) kuruma *Japanese*
 2 CL house 3 CL Gen car
 'two houses' 'three cars'
 c. du *(kwon) chayk *Korean*
 2 CL book
 'two books'

In Vietnamese, however, there is a broader range of patterning, and not all nouns require or even tolerate classifiers when combined with numerals. Nouns in Vietnamese can be broken down into three basic types: (a) nouns that obligatorily require classifiers when combined with numerals—'obligatory-classifier nouns', (b) nouns that only optionally occur with classifiers when combined with numerals—'optional-classifier nouns', and (c) count nouns that cannot occur with any classifier—'non-classified nouns'. This three-way split in Vietnamese occurs both with mono-syllabic, mono-morphemic nouns (2–4), and nouns which are bi-/multi-syllabic compounds (5–7), as illustrated below. In (2) and (5), it is seen that a classifier is always necessary when the obligatory-classifier nouns in these examples are being counted. Examples (3) and (6) show that it is fully optional for a classifier to occur with other nouns when combined with a numeral. Finally, in (4) and (7), nouns occur which are never combined with any classifier in situations of counting.

Mono-syllabic, mono-morphemic nouns

Obligatory-classifier nouns

- (2) a. hai *(con) chó b. bốn *(cuốn) sách
 two CL dog four CL book
 'two dogs' 'three books'

¹ Examples of data from languages other than Vietnamese are labeled with the relevant language name. All other unlabeled data in the paper is from Vietnamese.

Optional-classifier nouns

- (3) a. bốn (căn) phòng. b. tám (cái) làng
 four CL room eight CL village
 ‘four rooms.’ ‘eight villages’²

Non-classified nouns

- (4) a. hai màu b. hai nước
 two color two country
 ‘two colors’ ‘two countries’
- c. hai tiếng d. hai vết
 two sound two mark
 ‘two sounds’ ‘two marks’ (e.g. on the skin, furniture)

Bi-/multi-syllabic compound nouns

Obligatory-classifier nouns

- (5) a. hai *(cái) chân vịt b. hai *(con) cá voi
 two CL leg duck two CL fish elephant
 ‘two propellers’ ‘two whales’
- c. hai *(bông) hoa hồng b. hai *(chiếc) xe đạp
 two CL flower rose two CL vehicle cycle
 ‘two roses’ ‘two bicycles’

² Löbel (2000) suggests that the use of classifiers is linked to the referentiality of a noun/NP, and classifiers must occur when a noun/NP is assumed to have a distinct reference and is presupposed to exist. However, as shown in (i) and (ii) below, presupposed NPs from the optionally-classified noun group may occur either with or without a classifier, with no effect on the interpretation of the NP. Consequently, the presence or absence of classifiers with optionally-classified nouns is not linked to any particular referential properties of the noun and is an instance where optional variation in the presentation of nouns simply seems to be available.

- (i) Chúng tôi đến thăm hai (cái) làng.
 we go visit two CL village
 Tên của chúng là làng Đông và làng Tây.
 name of they are village East and village West
 ‘We visited two villages. Their names are East village and West village.’
- (ii) Chúng tôi trang trí hai (căn) phòng: nhà bếp và phòng tắm.
 we decorate two CL room kitchen and bathroom
 ‘We decorated two rooms: the kitchen and bathroom.’

Optional-classifier nouns

- (6) a. hai (khối) thiên thạch b. hai (người) thợ sửa xe
 two CL sky stone two CL workman fix vehicle
 ‘two meteorites’ ‘two auto mechanics’
- c. hai (cái) rạp chiếu phim d. hai (cái) nhà máy
 two CL house to.show movie two (CL) house machine
 ‘two cinemas’ ‘two factories’

Non-classified nouns

- (7) a. hai vương quốc b. hai chính phủ
 two king nation two government
 ‘two kingdoms’ ‘two governments’

Such patterns raise a number of questions about the syntactic structures projected in nominal phrases in Vietnamese, and the analysis of similarities and differences in the syntactic behavior of obligatory-classifier nouns, optional-classifier nouns, and non-classified nouns will be the central focus of inquiry in Sects. 3 and 4. Before beginning to probe these issues, however, it will first be shown how the surface distribution of classifiers described above bears significantly on a recent general claim about the syntax and semantics of numeral classifiers made in Bale and Coon (2014).

While the currently most widespread and dominant analysis of numeral classifiers in languages like Chinese (Li 1998; Cheng and Sybesma 1999; Chierchia 1998) is that classifiers serve to individuate nouns and syntactically combine with nouns before classifier and noun are built together with numerals³, Bale and Coon (2014) present arguments that ‘numeral classifiers are required because of the syntactic and semantic properties of the *numeral* (as in Krifka 1995), rather than the *noun* (as in Chierchia 1998) (Bale and Coon p. 695).

Bale and Coon essentially follow Krifka (1995), who proposes that the distinction between classifier languages and languages without classifiers lies in the numeral system present in such languages. In languages without classifiers, it is suggested that numerals incorporate within themselves a measuring function, while in classifier languages, the measure function is encoded in a separate morpheme, the ‘numeral classifier’. Krifka’s numeral-centered perspective on classifiers contrasts directly with the position taken in Chierchia (1998), who argues that it is specifically differences in the nominal system which account for the presence or absence of numeral classifiers across languages. For Chierchia, nouns in languages like Chinese are suggested to be mass-like/kinds in their denotation, and require classifiers to be converted into atomic sets for the purposes of counting. Bale and Coon’s significant contribution to the Krifka–Chierchia debate is to present interesting evidence from

³ For an alternative view of Chinese, see Zhang (2013) who claims that individual and individuating classifiers combine with nouns before numerals, but other classifier types may first combine with numerals before modifying nouns. See also Li (2013) for a related position.

two different languages, Mi'gmaq (Algonquian) and Chol (Mayan), which provide empirical support for Krifka's analysis, in which it is properties of numerals, not nouns, that directly result in the presence/absence of classifiers.

In Mi'gmaq, it is noted that classifiers only occur with numerals above 5 and never with 1–5. In Chol, it is reported that native Mayan numerals require the use of classifiers to count nouns, but numerals which have been borrowed from Spanish do not (and may not occur with classifiers). Consequently, both Mi'gmaq and Chol have two sets of numerals which behave quite differently when combining with nouns, and what determines whether or not a classifier occurs is the particular type of numeral that is present. Example (8) shows the non-occurrence of a classifier with a low numeral in Mi'gmaq, and (9) shows how a classifier must occur with a higher numeral:

- | | | | | |
|--------|--|----|--|----------------|
| (8) a. | na'n-ijig ji'nm-ug
five-AGR man-PL
'five men' | b. | *na'n te's-ijig ji'nm-ug
five CL-AGR man-PL
intended: 'five men' | <i>Mi'gmaq</i> |
| (9) a. | *asugom-ijig ji'nm-ug
six-AGR man-PL
'six men' | b. | asugom te's-ijig ji'nm-ug
six CL-AGR man-PL
intended: 'six men' | <i>Mi'gmaq</i> |

Relevant patterns from Chol are illustrated in (10) and (11). In (10), it is seen that a classifier must be present with the Mayan numeral *ux* 'three', but it is ungrammatical for such an element to occur in (11) with the Spanish-borrowed numeral *nuebe* 'nine'.

- | | | | | |
|---------|---|----|--|-------------|
| (10) a. | ux-p'ej tyumuty
three-CL egg
'three eggs' | b. | *ux tyumuty
three egg
intended: 'three eggs' | <i>Chol</i> |
| (11) a. | *nuebe-p'ej tyumuty
nine-CL egg
intended: 'nine eggs' | b. | nuebe tyumuty
nine egg
'nine eggs' | <i>Chol</i> |

The data in Mi'gmaq and Chol consequently provide convincing support for Krifka's numeral-based theory, which Bale and Coon note 'predicts the possibility of a language with idiosyncratic behavior among the numerals, whereas Chierchia's theory is inconsistent with such a pattern' (p. 700). For Bale and Coon, the important conclusion is that the 'requirement for a classifier is dependent, not on the noun, but on the syntax and semantics of the numeral' (p. 702), as also emphasized in the first half of the title of their paper: "Classifiers are for numerals not for nouns".

Somewhat cautiously, Bale and Coon actually do not claim their study of Mi'gmaq and Chol is proof that a Krifka-style classifier system should necessarily

be assumed to be present in all classifier languages. Towards the end of their paper, they admit the possibility that, in theory, there might be two types of classifier language, one patterning according to Krifka's predictions, and another according to Chierchia's noun-centered approach. Bale and Coon outline what a language with a 'Chierchia-style' classifier system might be expected to show in counting structures, in order to give evidence of a noun-dependent classifier system:

'Unlike Krifka's theory, Chierchia's predicts that it should be possible to have a lexical numeral that requires a classifier when modifying one noun, yet prohibits a classifier when modifying another... Such a pattern would demonstrate that the presence or absence of a classifier depends on the noun that is being modified rather than on the numeral.' (p. 704)

Bale and Coon add that whether or not such a language exists is an empirical matter that they do not attempt to resolve.

Vietnamese can now be noted to be a language which does indeed present the empirical evidence Bale and Coon suggest would clearly identify a Chierchia-type classifier-noun system. In Vietnamese, the presence vs. absence of overt classifiers in counting constructions is very clearly determined by the type of noun that occurs, and not by the numeral in any way. As noted in (2–7), certain nouns, the obligatory-classifier nouns, always require the overt presence of classifiers when combined with numerals, and this is so regardless of the type of numeral they occur with. Other nouns, the optional-classifier nouns, optionally permit the overt use of classifiers, and this optionality is independent of the type of numeral. A third set of entity-denoting nouns in Vietnamese, the non-classified nouns, never permit the overt occurrence of a classifier, and this patterning remains constant regardless of the type of numerals used for the counting of such nouns. Whether an overt classifier occurs in Vietnamese is therefore dependent on the noun that is present, rather than on the numeral, suggesting for Bale and Coon a different syntactic alignment among classifier, noun and numeral from what they identify and argue for in Chol and Mi'gmaq. The general conclusion arising from a simple comparison of evidence available in Vietnamese, Chol and Mi'gmaq would therefore appear to be that there is empirical support for both Krifka-type and Chierchia-type classifier systems, and that classifiers may perhaps vary cross-linguistically in their mode of combination with numerals and nouns. The way that patterns in Vietnamese support a specifically Chierchian-type approach to the combination of classifiers with nouns will now be explored further from a predominantly syntactic perspective, and will ultimately lead to a refinement of the description of Vietnamese given here, which references both the surface distribution of overt classifiers and the underlying presence of a classifier projection. We begin such an investigation in Sect. 3 by asking how Vietnamese potentially fits into the typology of nominals developed in Chierchia (1998), and then, in Sect. 4, consider a range of syntactic phenomena which bear on the occurrence of classifiers. Section 5 subsequently revisits the conclusions about classifier/noun relations that can be drawn on the basis of the complex patterns found in Vietnamese.

3 Vietnamese and Chierchia's (1998) typology of nominals

In his much-cited semantic analysis of nouns across different language types, Chierchia (1998) distinguishes nouns in classifier and non-classifier languages by means of the features $[\pm \text{arg}, \pm \text{pred}]$. Nouns in classifier languages are classified as $[+\text{arg}, -\text{pred}]$ and taken to represent kinds, while nouns in non-classifier languages are suggested to encode properties and have a $+\text{pred}$ specification.⁴ The necessary occurrence of classifiers with numerals in languages such as Chinese and Japanese is argued to be a direct consequence of the semantic type of nouns in classifier languages—numerals are taken to be functions from properties into quantized properties and are not able to combine directly with kinds. The presence of classifiers is hypothesized to obviate such a mismatch, converting nouns in classifier-languages from kinds into number-seeking properties. It is emphasized that nouns are taken to be semantically uniform in classifier languages: “Classifier languages (CILs) are those in which no noun can directly combine with a numeral.” (Chierchia 2015:165).

Considering how Vietnamese would be characterized in Chierchia's system, obvious complications arise from the varied patterning of nouns that is found in the language. The fact that many nouns occur with classifiers in Vietnamese, should lead to a $[+\text{arg}, -\text{pred}]$ feature set for nouns and the assumption that nouns are mass-like/kinds in Vietnamese, on a par with Chinese in Chierchia's view. However, the significant presence of nouns that occur without classifiers when combined with numerals (the set of non-classified nouns) requires a different featural classification, in which such nouns are taken to be property-like, as in English and other non-classifier languages. Vietnamese then apparently has to be seen as a hybrid language, with a mixed inventory of nouns of different featural specifications, and it needs to be conceded that not all languages follow the pure typological categories set up in Chierchia's approach. Additionally, and making matters more complicated, the occurrence of many optionally-classified nouns in Vietnamese would seem to force the assumption that a portion of nouns in the language are lexically and semantically ambiguous, having two different available specifications for the $[\pm \text{arg}, \pm \text{pred}]$ features, being either kinds or properties. Faced with such a muddying of the picture of sharp, clean divisions of nominal taxonomy across languages presented in Chierchia, one might wonder if there could be other approaches to the variation in Vietnamese that would not attribute such fundamentally different semantic properties to nouns present in the language.

For a number of reasons, some of which are criticisms of Chierchia's approach, Borer (2005) proposes a different, unifying analysis of nouns across languages, in which the default interpretation of all nouns in all languages is mass, and NPs cross-linguistically are predicates. Borer suggests that classifiers are one instantiation of CIP, other instantiations being plural markers and indefinite articles, with all such elements functionally being involved in dividing up the mass denotation of nouns. If

⁴ English is classified as $[+\text{arg}, +\text{pred}]$, while French is viewed as $[-\text{arg}, +\text{pred}]$, due to differences in the distribution of bare nominals in argument positions—possible in English (with plurals and mass nouns), but not at all possible in French.

such a unitary analysis of nominals across languages were to be adopted for Vietnamese, it would need to assume that the same basic functional structure projected with classified nouns would also be present with other non-classified nouns, hence that a classifier or classifier-equivalent position would occur not only with obligatory-classifier nouns but also with non-classified nouns, and would underlyingly be present with optional-classifier languages too, whether an overt classifier appears or not. Concerning the actual mapping of overt morphemes to syntactic positions in such an approach, with obligatory-classifier nouns it is clear that the classifier and the noun would naturally occur in the regular classifier and noun positions, CI and N, as illustrated in (12).

- (12) [DP [NumP hai [CIP cuốn [NP sách]]]]
 2 CL book
 ‘two books’

In the case of non-classified nouns, there is more than one possibility to consider. A first possibility is that the noun occurs in the N position, and there is a phonetically empty CI position, as shown in (13).

- (13) [DP [NumP hai [CIP \emptyset [NP ngày]]]]
 2 day
 ‘two days’

A second theoretical possibility is that the ‘noun’ is actually base-generated in the CI position, and the N position is empty, as schematized in (14). This second mode of analysis would amount to the suggestion that non-classified ‘nouns’ are in reality classifier-like elements similar to the measure words used with English mass nouns (e.g. ‘two pounds of sand’), and combine with a silent noun complement that they would serve to measure out.

- (14) [DP [NumP hai [CIP ngày [NP \emptyset]]]]
 2 day
 ‘two days’

Such an approach has in fact been discussed for Vietnamese in Truong (1970, p. 285) who provides examples of non-classified nouns and suggests that an unexpressed mass noun occurs in the underlying structure. Truong suggests that the non-classified noun measures out discrete amounts of the unexpressed noun. In his analysis of the examples in (15) and (16), *đồng* ‘dollar’, and *ngày* ‘day’, measure out portions of *bạc* ‘money (lit. ‘silver’) and *trời* ‘time (lit. ‘heaven’) respectively. The latter elements could therefore be suggested to occupy the N position of nominal phrases, while the non-classified nouns occupy the CI position.

(15) hai đồng (bạc)
 2 dollar (money)
 ‘two dollars’

(16) hai ngày (trời)
 2 day (time)
 ‘two days’

Trương’s approach is similar in essence to a more recent, extended hypothesis of the presence and role of ‘silent elements/SEs’ in syntactic structures, developed in Kayne (2005, 2012 and other works). Kayne posits the occurrence of a wide range of SEs in different structures in order to explain a range of morpho-syntactic patterns in English and other languages. For example, Kayne suggests that the use of color terms such as ‘red’ ‘blue’ etc. involves the use of a silent noun ‘color’, and that color terms first combine with this SE before then being used to modify a further noun. The surface sequence ‘a red car’ is consequently analyzed as resulting from an underlying sequence ‘a red COLOR car’, in which ‘COLOR’ is a noun present in the syntactic structure, but not pronounced (Kayne 2005, ch. 8, p. 213).

A potential weakness in approaches assuming the widespread presence of silent lexical elements, which has been noted in certain critiques of Kayne’s work (for example, Her and Tsai 2014; Simpson 2012), is the fact that surface forms in which ‘silent elements’ are overtly pronounced (e.g. ‘a blue color car’) do not always have meanings that are equivalent to those which arise when SEs are not pronounced. Consequently, there is a worrisome non-equivalence between SEs and overt instantiations of SEs in various instances, weakening the motivation for assuming the former. Such concerns also arise with regard to a ‘silent element’ analysis of non-classified nouns in Vietnamese. If we consider Trương’s examples in (15) and (16) and reflect on the interpretation of forms in which the bracketed nouns (*bạc* ‘money’ and *trời* ‘time’) are pronounced overtly, such forms are not necessarily equivalent to truncated forms where only the non-classified nouns *đồng*, and *ngày* are pronounced. For example, the ‘optional’ element *trời* ‘time’ is used overtly only when the speaker considers that the amount of time referred to is long. Consequently, it is unnatural for *trời* to occur overtly in a context such as (17), although the use of *tháng* ‘month’ without *trời* would be fine in such a context. This may be taken to suggest that there is no silent element *trời* present in the structure when use of *hai tháng* ‘two months’ in (17) results in an acceptable interpretation.⁵

(17) Anh ta xây xong căn nhà chỉ trong hai tháng (??trời).
 he build finish CL house only in two month time
 ‘He finished building the house in only two months.’

⁵ At least, there is no SE *trời* that is directly equivalent to overt *trời* and, if SEs are not direct equivalents to overt words, arguments for their existence based on the possible use of covert ‘equivalents’ is clearly weakened.

Furthermore, the use of *trời* ‘time’ with nouns encoding amounts of time is restricted to certain nouns, and although it is possible to combine *trời* with *ngày* ‘day’ and *tháng* ‘month’, it cannot occur with other non-classified nouns referring to portions of time such as *tuần* ‘week’, *giờ* ‘hour’, and *phút* ‘minute’, etc. as shown in (18). This seriously questions whether a covert equivalent to *trời* is present in the structure when *tuần* ‘week’, *giờ* ‘hour’, and *phút* ‘minute’ are combined with numerals.⁶

- (18) hai giờ/tuần/ phút (*trời)
 2 hour/week/minute time
 ‘two hours/weeks/minutes’

Finally, there are also instances in which it is extremely unclear what kind of silent noun might be being measured by certain non-classified nouns, as for example with *phòng* ‘room’, *nước* ‘country’, *góc* ‘angle/corner’ and *kỳ tích* ‘unique/miraculous phenomenon’.⁷

Such concerns collectively cast much doubt on the plausibility of an SE hypothesis of non-classified nouns in Vietnamese in which such elements are base-generated in CI and select for some null mass noun complement as in (15).⁸ We will therefore continue to assume that non-classified nouns are indeed nouns which head the nominal projections in which they occur, and are hence base-generated in the N position.

We will now consider the question of whether non-classified nouns in Vietnamese also project a CIP, as suggested in (13), and occur with the same extended functional structure as other nouns. Theoretically, the assumption that CIPs are present with non-classified nouns as well as with other noun types in Vietnamese would help avoid the conclusion that nouns in Vietnamese would have dramatically different semantic properties if a Chierchia-style analysis were to be applied to Vietnamese. Such an assumption also helps address the issue raised by optional-classifier nouns in a similar way. Were it to be suggested that non-classified nouns really do instantiate a type of noun that does not combine with classifiers in order to account for the surface patterning attested, such an approach would not extend to cover simple patterns found with the set of optional-classifier nouns. Specifically, the frequent instances in which these nouns occur without a

⁶ When the element *trời* ‘time’ does legitimately occur combined with a time expression such as *ngày* ‘day’, as in (16), with the meaning ‘three long days of time’, we assume that *trời* functions as the head noun of the nominal projection and that *ngày* functions as just a classifier in CI⁰ (just as nouns in many languages can serve either as the head nouns of nominal projections, or as measure words/classifiers, as for example in the alternation in English: ‘two bags’—‘two bags of gold’).

⁷ While it would theoretically be possible to retreat to a position in which a very abstract silent element STUFF is posited in cases where no overt mass noun can be added following a non-classified noun, this is not the spirit of Kayne’s or Truong’s analysis, which justifies the existence of SEs on the basis of there being overt equivalents, the SEs simply not being phonetically expressed in the structure. It is also difficult to imagine what kind of abstract STUFF non-classified nouns such as ‘country’, and ‘angle/corner’ would be measuring out. An SE measure noun analysis is therefore not *impossible*, but not well-supported or very plausible, suggesting that one should look for alternative analyses.

⁸ A further empirical argument against an SE analysis with non-classified nouns is given in footnote 11 in Sect. 4.1. There it is noted that ellipsis patterns in the extra *cái* construction indicate rather clearly that non-classified nouns do not behave syntactically as if they are paired with an unpronounced/omitted noun.

classifier cannot be accounted for by suggesting that they are nouns that do not combine with classifiers, as clearly such elements do indeed occur with classifiers on a regular basis. Assuming that a classifier position/CIP is projected with all three classes of noun in Vietnamese, but not necessarily overtly filled with a classifier in all occurrences, appears to provide the most parsimonious theoretical approach to the variability found across the three noun types. From a cross-linguistic perspective, it may also be supported by classifier-related phenomena present in other Southeast Asian languages, as noted below.

Analyzing the distribution of classifiers and nouns in Burmese and Thai, Simpson (2008) notes that there are nouns in these languages which do not have special classifiers, and are counted by means of a repetition of the noun itself, in ‘repeater’ constructions, as illustrated in (19):

- (19) a. prathêet sǎam prathêet *Thai*
 country 3 country
 ‘three countries’
- b. cun ta cun *Burmese*
 island CL island
 ‘one island’

Where nouns in Thai and Burmese accidentally do not have classifiers, and so approximate non-classified nouns in Vietnamese, it is not the case that a classifier position is absent from the structure, and in the repeater construction, both noun and classifier position are present and lexically filled. In both Thai and Burmese, the canonical arrangement of numerals, classifiers and nouns found elsewhere occurs in the linear sequence: [noun numeral classifier]. In (19), the classifier position is therefore occupied by the second, repeated occurrence of the noun.

In addition to the repeater construction, Burmese and Thai also contain other nouns which have no associated classifier, and which surface in counting constructions without any special repetition of the noun. Significantly, when this occurs, the noun is found to be positioned to the right of numerals in the regular classifier position, as shown in (20):

- (20) a. soong pii b. *pii soong *Thai*
 2 year year 2
 ‘two years’

In both sets of ‘non-classified noun’ cases in Thai and Burmese, nouns consequently come to be associated in some way with a typical classifier position, and do not cause a special classifier-less syntactic structure to occur. In Simpson (2008), both repeater and non-repeater forms are suggested to result from a movement operation, in which the head noun raises from the N position to occupy the CI position, this head-movement sometimes leaving behind a spelt-out copy in the N position, in the special case of repeaters, such as (19).

Putting the above theoretical and cross-linguistic empirical considerations together, we believe that (13), above, emerges as the strongest front-runner among potentially available analyses for the underlying structure of Vietnamese non-classified nouns, and we will therefore go forward with such an analysis to see how well it fares as more complex empirical patterns from Vietnamese are introduced in Sect. 4. Taking account of the observations on word order in Burmese and Thai in (19) and (20), in which movement may seem to relate nouns in N to the classifier position, we will further assume that the surface position of non-classified nouns in Vietnamese might either be as indicated in (13) in N, or be structurally higher in Cl, as illustrated in (21), if string-vacuous movement has taken place. Data bearing on which of these two possibilities is more likely to be correct will be presented in Sect. 4, as we come to consider extra *cái* and a range of other syntactic adjustments to nominal projections in Vietnamese. This will also lead to conclusions about the position of optional-classifier nouns in structures in which no overt classifier occurs (examples (3) and (6)).

- (21) [DP [NumP hai [CIP ngày_k [NP ngày_k]]]]
 2 day
 ‘two days’

4 Syntactic tests for functional structure in Vietnamese nominal projections

In this section we consider how a range of syntactic patterns provides information about the classifier-related structures projected in nominal phrases in Vietnamese. We investigate the patterning of classifiers and nouns in the ‘extra *cái*’ construction, noun-numeral separation patterns in passive sentences, and a special patterning of short answer forms with certain human compound nouns.

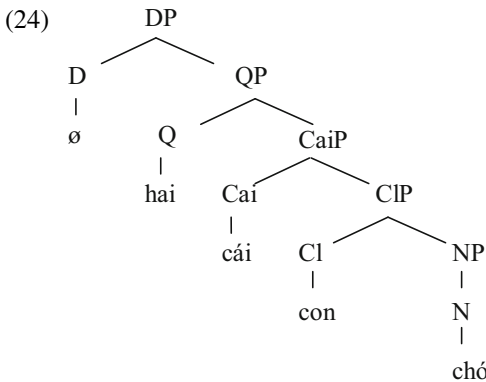
4.1 The extra *cái* phenomenon

The morpheme *cái* is a very common classifier in Vietnamese that is generally used with inanimate nouns, for example *bàn* ‘table’ in (22a), either with or without a numeral. Besides this regular classifier function, *cái* can also co-occur with other classifiers in nominal constructions as a ‘second classifier’, as shown in (22b).

- (22) a. cái bàn
 CL table
 ‘a/the table’
 b. cái con chó
 CL CL dog
 ‘the dog’

In this very special classifier doubling construction, the ‘extra *cái*’ significantly enforces interpretations of definiteness. While a simple combination of a classifier and a noun such as (22a) may be interpreted as either definite or indefinite, depending on the context (Simpson et al 2011), when extra *cái* is added as in (22b), the potential ambiguity in (in)definiteness disappears completely and *cái* forces a definite interpretation (Ngo 2012; Simpson 2008; Nguyen 2004; Nguyen 1957). In order to distinguish the two functions of *cái*, we will henceforth gloss the regular use of the classifier *cái* simply as ‘CL’ and the special, definite use of *cái* as ‘CL-DEF’. It should also be noted that although, in general, classifiers in numeral constructions do not receive stress, the special use of definite extra *cái* is always stressed. This prosodic property of extra *cái* along with its necessary definiteness are useful diagnostics for identifying which function any occurrence of *cái* has in certain situations when only one classifier occurs. As for extra *cái*’s structural position, this is below any numerals that are present, as seen in (23). We therefore conclude, following Ngo (2012), that extra *cái* occurs base-generated in a second, dedicated classifier position below numerals, which we will label simply as CaiP, to distinguish it from the regular classifier position, as illustrated in the tree in (24).⁹

- (23) hai cái con chó
 2 CL-DEF CL dog
 ‘the two dogs’



It is possible for extra *cái* to occur with all three classes of noun in Vietnamese—obligatory-classifier nouns, optional-classifier nouns, and non-classified nouns, and the interesting patterning observed leads to significant conclusions about the underlying functional structure projected with such nouns.

⁹ Note that as extra *cái* occurs between numerals and regular classifiers, it interrupts the relation between the former and the latter, indicating that it is not necessary for numerals to combine directly with the regular classifier of a noun. This, consequently, provides additional support for the assumption that classifiers have a stronger, more direct syntactic relation with nouns than they do with numerals in Vietnamese, i.e. Chierchia’s general view of the connections between nouns, classifiers and numerals is further supported over a Krifka-type perspective.

When extra *cái* occurs with obligatory-classifier nouns, it is found that the regular classifier for the noun must occur in addition to extra *cái*, and it is ungrammatical to omit the regular classifier, as shown in (25). Consequently, extra *cái* does not replace the regular classifier, rather it performs a separate function, relating to definiteness:

- (25) hai cái *(cuốn) sách
 two CL_{.DEF} CL book
 ‘the two books’

When extra *cái* occurs with optional-classifier nouns, it is again found that the regular classifier for the noun must be present as well as extra *cái*, and it is unacceptable to omit the former, although it is elsewhere optional for the regular classifier to occur with such nouns. This is illustrated in (26–28). The (a) examples show that the classifier for each noun is optional when the noun is combined with numerals in the absence of extra *cái*. The (b) examples show how omission of the regular classifier in the presence of extra *cái* is not possible. The addition of extra *cái* thus has a clear effect on optional-classifier nouns, and restricts the possibilities open to such nouns in other counting environments.

- (26) a. hai (người) nhân viên b. hai cái *(người) nhân viên
 2 CL employee 2 CL_{.DEF} CL employee
 ‘(the) two employees’ ‘the two employees’

- (27) a. hai (người) phụ huynh b. hai cái *(người) phụ huynh
 two CL parent two CL_{.DEF} CL parent
 ‘(the) two parents’ ‘the two parents’

- (28) a. hai (người) quản lý b. hai cái *(người) quản lý
 2 CL manager 2 CL_{.DEF} CL manager
 ‘(the) two managers’ ‘the two managers’

Now turning to consider non-classified nouns, when extra *cái* is combined with this class of noun, no additional classifier appears in the structure, and extra *cái* appears to be able to occur directly with the noun, as illustrated in (29) and (30) with both monosyllabic and bi-syllabic non-classified nouns:¹⁰

¹⁰ A reviewer of the paper asks whether extra *cái* may be used with other non-classified nouns such as *tuần* ‘week’, *giờ* ‘hour’, *phút* ‘minute’ and *giây* ‘second’. The answer is that extra *cái* may indeed be used with these nouns when a numeral occurs. When this is done, the interpretation is equivalent to English: ‘the two weeks’, ‘the two hours’, ‘the two minutes’ etc. Such definite noun phrases sound a little odd out of any context, and in English it is normally necessary to embed these in a special context which sets up a specific set of two minutes/hours etc that is introduced and then referred back to. Even in such enabling contexts, use of the definite determiner ‘the’ sounds a bit odd, and to make such sequences sound natural, a demonstrative is often used in place of a determiner, as in the following:

- (29) a. hai câu/màu/góc/tiếng
 2 sentence/color/corner/sound
 ‘(the) two sentences/colors/corners/sounds’
- b. hai cái câu/màu/góc/tiếng
 2 CL_{DEF} sentence/color/corner/sound
 ‘the two sentences/colors/corners/sounds’¹¹
- (30) a. hai truyền thuyết/hệ thống/xã hội/kỳ tích
 2 legend/system/society/miraculous phenomenon
 ‘(the) two legends/systems/societies/miraculous phenomena’
- b. hai cái truyền thuyết/hệ thống/xã hội/kỳ tích
 2 CL_{DEF} legend/system/society/miraculous phenomenon
 ‘the two legends/systems/societies/miraculous phenomena’

Combining the information from these patterns leads to a conclusion about the structure underlyingly present with non-classified and optionally-classified nouns when numerals and extra *cái* occur. It would seem clear that the patterning with optional-classifier nouns can only be accounted for if extra *cái* necessarily selects for a CIP complement whose head is overtly filled with a classifier—this requirement from extra *cái* eliminates the optionality seemingly present in simple count structures without extra *cái*, and necessitates the presence of an overt, regular classifier when extra *cái* is introduced (26b, 27b, 28b).

The patterning with obligatory-classifier nouns bears this conclusion out—extra *cái* does not substitute for the regular classifier that would occur with obligatory-

Footnote 10 continued

(i) ‘When I called the police for help, I was put on hold for two minutes. Those/?the two minutes seemed to last forever.’

The same discourse restrictions apply in Vietnamese, and sequences such as *hai cái tuần* ‘the two weeks’ sound a little odd out of context, but are fine once a context is created. As with English, in order for such sequences to sound fully natural, it is appropriate to add a demonstrative, as for example: *hai cái tuần đó/ấy* ‘those two weeks’. When this is done, the non-classified nouns *tuần* ‘week’, *giờ* ‘hour’, *phút* ‘minute’ and *giây* ‘second’ may all occur with extra *cái*.

¹¹ It can be noted that when extra *cái* occurs, it is never possible to elide the noun following the regular classifier, as illustrated in (i) and (ii):

- (i) hai cái con *(chó) (ii) hai cái cuốn *(sách)
 2 CL_{DEF} CL dog 2 CL_{DEF} CL book
 ‘the two dogs’ ‘the two books’

The fact that *hai cái câu* in (29b) is well-formed is therefore a further argument against Trương’s (1970) analysis of *câu* as being a classifier associated with an optional noun *nói* ‘speech’ (see the discussion of *đồng* and *ngày* in Sect. 3). If *câu* were to be a classifier in (29b), omission of its associated noun *nói* in the presence of extra *cái* should cause ungrammaticality, as with (i) and (ii) above. The observation that (29b) is fully acceptable consequently supports the analysis of *câu* as a non-classified noun, not a classifier. Similar observations can be made with other non-classified nouns.

may be taken to raise to the head of the CIP selected by extra *cái*, via movement from its base position in N, as hypothesized in Sect. 3. Vietnamese is therefore a language with overt N-to-Cl movement, as previously posited in Simpson (2008:830–833).

Such a conclusion now raises three further questions about optional-classifier and obligatory-classifier nouns. First, if N-to-Cl movement of a noun occurs with *non-classified nouns* lexicalizing the classifier position, should such movement also be assumed to occur regularly with *optional-classifier nouns* when no overt classifier is present in sequences consisting of just a numeral and a noun, for example *tám làng* ‘8 villages’ (ex. 3b)? Second, if such N-to-Cl movement is assumed to occur with optional-classifier nouns in the absence of an overt classifier, why does this movement not seem to be able to satisfy extra *cái*’s apparent requirement for overt lexicalization of the Cl head, and why must an overt classifier be selected instead in examples such as (26b/27b/28b)? Third, if N to Cl movement is taken to be a general option for nouns in Vietnamese, why should there be certain nouns that actually never allow for such an option—the class of obligatory-classifier nouns, which always require the Cl position to be filled by an overt classifier, not by movement of the noun to the Cl position? In short, once a movement operation such as N-to-Cl is posited to occur for certain nouns, what restricts this to occurring only with certain lexical items and in certain syntactic environments? What formal properties define the three noun types identified in Vietnamese and dictate their syntactic interactions in structures with and without extra *cái*?

Patterns to be considered shortly in Sect. 4.2 will suggest that N-to-Cl movement does in fact occur regularly with optional-classifier nouns when no overt classifier is present, and if there is a necessity to block such movement, an overt classifier must instead be merged in Cl, allowing the noun to remain in N. As N-to-Cl movement is therefore taken to occur both with non-classified and (often with) optional-classifier nouns, the question is what causes such movement to occur with these two classes of noun but not with the third class of Vietnamese noun, the obligatory-classifier nouns, which require the use of an overt classifier, and apparently do not permit N-to-Cl movement as an alternative way to lexicalize the Cl position?

Here we will adopt and develop a suggestion from one of the editors of the *Journal of East Asian Linguistics*, following helpful comments from two reviewers of the paper, and propose that the Cl position with non-classified nouns is not syntactically empty prior to N-to-Cl movement, but actually occupied by a classifier element which is phonetically null and affix-like, represented as $CL_{-\emptyset}$. We suggest that it is this element which triggers movement of the noun to adjoin to Cl to support the affix-like $CL_{-\emptyset}$ in a way paralleling the analysis of other instances of head-movement as involving the support of null affixes (Chomsky 1995, 2000; Radford 2004). In such a perspective, non-classified nouns would in fact be associated with classifiers, but these would simply be phonetically null. Optional-classifier nouns can be characterized as nouns that have available either overt classifier forms, or the null affix-like classifier $CL_{-\emptyset}$, so that when the latter is selected, this causes N-to-Cl movement to take place, as with ‘non-classified’ nouns.

Turning to consider the set of obligatory-classifier nouns, this group of nouns can be assumed to be lexically-specified as having available only overt classifier forms,

and not having $CL_{-\emptyset}$ as an optional classifier in their lexical listing. As a result, because N-to-Cl movement is only triggered by the presence of $CL_{-\emptyset}$, such movement will not occur with obligatory-classifier nouns, and the Cl position will only ever be lexicalized with overt classifier forms. Differences in the occurrence of N-to-Cl raising can consequently be attributed to differences in the lexical specifications of nouns in a way that is fully in line with regular assumptions about noun-classifier relations—it is well known that nouns in classifier languages are individually specified as occurring with certain classifiers (with a restricted range of classifiers often being possible with many nouns, see Zhang 2007 and Lakoff 1987 among others) and that the set of classifiers that may occur with each noun is often unpredictable. Such information must therefore be specified in the lexicon for each noun and learned by speakers on an item by item basis. In Vietnamese, only certain nouns will be listed as having available the null classifier form $CL_{-\emptyset}$, and N-to-Cl movement will only potentially be triggered with this particular set of nouns.

Now considering the patterns found with extra *cái*, what needs to be explained is why the N-to-Cl movement taken to occur with non-classified nouns appears to satisfy extra *cái*'s requirements, while that assumed to be available with optional-classifier nouns (if $CL_{-\emptyset}$ is selected) does not, making an overt classifier obligatory instead. Here we must offer a purely formal and rather mechanical modeling of this difference between non-classified and optional-classifier nouns, pending further investigation that might lead to a deeper understanding of the relevant properties of nouns of the two types in combination with extra *cái*. The technical account that we propose is as follows.

We suggest that extra *cái* does not simply require that the head of its complement, Cl, be overtly lexicalized, but that extra *cái* must additionally establish a local feature-related relation with the head of its complement, similar to the featural relation taken to exist between C and T in instances of complementizer-tense agreement, or V and C in instances where a verb selects for a particular complementizer in its CP complement.

Formally, we suggest that extra *cái* has an unvalued feature relating to the definiteness function it ultimately applies to the interpretation of the DP, a 'def-feature', and this may be valued by any element merged with a corresponding + interpretable def-feature present in Cl—intuitively, *cái* is a regular, general classifier which comes to acquire a positive specification for definiteness from some other element in the structure.¹² The elements which we propose may be merged with a valued def-feature are of two types: overt classifiers and non-classified nouns (this does, of course, seem to be a simple stipulation, but it is one which will allow us to successfully capture the patterns being considered here). Consequently, N-to-Cl movement of a non-classified noun, or the introduction of an overt classifier in Cl

¹² Expressed in other words, the property of extra *cái* we are trying to capture is that the general classifier *cái* acquires the functional role of marking definiteness due to its combination with some other element—*cái* is not inherently definite, but becomes interpreted as a marker of definiteness with, for example, optional classified nouns precisely because a second classifier is present following it. It is therefore the presence of the second classifier which results in *cái* being interpreted as definite. In featural terms, *cái* needs to be valued as definite by some other element in the structure which acts as a host for the relevant def-features (either an overt classifier or a non-classified noun).

with obligatory- and optional-classifier nouns will satisfy extra *cái*'s requirement for a def-feature, but simple N-to-Cl movement of optional-classifier nouns will not, as this will not be able to provide extra *cái* with the necessary valued def-feature.

In positing that +interpretable def-features may only be merged selectively on certain host elements, we believe there is a parallel with the selective distribution of interrogative features across different lexical host elements in questions—where a C head needs to be valued as +interrogative, such features can be supplied either by a *wh*-phrase raised to SpecCP or, in matrix yes-no questions in languages such as English, by a finite auxiliary verb raised to C, hence two distinct categories of elements are able to carry interrogative features (in English)—*wh*-phrases and auxiliary verbs. With regard to the Vietnamese extra *cái* case under consideration, we hypothesize that it is the particular lexical distribution of the valued feature that extra *cái* is in need of that causes the patterning found: in the absence of a regular classifier, non-classified nouns have acquired the ability to be merged with the definiteness-related feature that is otherwise normally only merged with an overt classifier.^{13,14}

¹³ There is, in fact, another way in which the distribution of def-features could potentially be characterized. A reviewer of the paper asks how classifiers and non-classified nouns might constitute a 'natural class' and be the only elements able to carry def-features. A possible alternative to the proposal made in the text, which would 'homogenize' the class of elements able to carry def-features in Vietnamese, is to make a rather different assumption about non-classified nouns. In place of movement of such nouns to the Cl position, it could be hypothesized that Vietnamese non-classified nouns are like the Thai and Burmese nouns which occur in repeater constructions, described in Sect. 3, example (19), and that the Cl position in Vietnamese is lexicalized by a second overt copy of the noun directly merged into Cl without any movement from the N position. A process of haplology would then trigger the deletion of the phonetic matrix of the linearly-adjacent copy in the N position, as represented in (i):

(i) [DP [NumP hai [C_{def} cái [C_{IP} nước [NP ~~nước~~]]]]]
 2 CL_{DEF} country country
 'the two countries'

If such an analysis is adopted, in which the classifier for non-classified nouns is simply homophonous to the noun itself, as in Thai/Burmese repeater constructions, this would allow for the suggestion that the only elements that can carry def-features are the set of overt classifiers. While such a reconceptualization of the way that the Cl position is lexicalized with 'non-classified nouns' is possible and would result in a potentially simpler classification of the range of elements which optionally bear def-features, we do not explore this further in the paper, as the occurrence of N-to-Cl movement seems justified by other phenomena discussed in Sects. 4.2 and 4.3, and it would not be straightforward to account for the patterns analyzed in these sections without a process of N-to-Cl raising. Additionally, as noted in the text, the distribution of features of certain types does not always occur within a clear 'natural class' of elements – auxiliary verbs in English are not obvious hosts for interrogative features and do not form a natural class with *wh*-phrases, yet both such elements are commonly assumed to carry interrogative-features (optionally, in the case of auxiliary verbs).

¹⁴ The JEAL editors ask whether the relation between a probe/head and a goal/some selected constituent (a head or a phrase) can ever reference the property of *overtness* and whether an element must have a phonetic instantiation, because the analysis of extra *cái* presented here requires the head of C_{IP} to be lexicalized either with an overt classifier or with an overt non-classified noun. Here we can note that the requirement that goals be overt is in fact assumed in a variety of analyses which deal with certain P2/V2 effects, where the Specifier of a C-domain head to which a verb has been raised must be instantiated by some overt element (and not a pro) probed by a functional head (see, for example, Holmberg 2000 on Finnish, and Manetta 2011 on Kashmiri). Hence, elsewhere probes are assumed to be able to impose a requirement that their goals must be overt.

in *bị* sentences which split apart nouns from numerals, it is found that this classifier-optionality disappears, and it is obligatory for a classifier to occur with the stranded numeral, as seen in (35).

- (35) làng (của họ) bị phá hoại tám *(cái).
 village of 3pl PASS destroy 8 CL
 ‘Eight (of their) villages were destroyed’

How is it possible to explain the obligatory occurrence of the classifier in such instances? We suggest the following account, which makes use of recent ideas on the occurrence of phases within nominal expressions in Simpson and Syed (2016) and Syed and Simpson (2017), and restrictions on movement noted in Abels (2003), Grohmann (2003) and Bošković (2014).

In Simpson and Syed (2016) and Syed and Simpson (2017), a range of syntactic phenomena from the Indo-Aryan language Bangla are argued to lead to the conclusion that nominal phrases may project an internal phase, similar to vP in clauses, in addition to a DP phase. This nominal-internal phase is identified as the projection in which numerals are base-generated and labeled QP. Elements such as AdjPs and NPs that are merged lower than QP and undergo extraction to higher positions are shown to need to pass through SpecQP as a nominal-internal escape hatch, in order to satisfy the Phase Impenetrability Condition/PIC (Chomsky 2000), which requires all inter-phasal movement to transit through the edge of phasal constituents. One consequence of such an analysis is a prediction that CIP constituents should be fully immobile and not allow for displacement, due to the interaction of the PIC and antilocality restrictions (Bošković 1994; Grohmann 2003). Elements to be extracted from phases must first raise to the edge of phases because of the PIC, but antilocality constraints prohibit the movement of the complement of a head to its specifier position. This results in the generalization, originally from Abels (2003), that:

- (36) Complements of phase heads cannot undergo movement. (Bošković 2014:32)

If QP is a phase and CIP the complement of the phase head Q, it is therefore expected that it should not be possible for CIP constituents to undergo any extraction out of QP to higher positions, though lower elements such as NP might be anticipated to permit extraction, as movement of NPs to SpecQP will not violate antilocality and will allow for satisfaction of the PIC constraint on successive cyclic movement. This prediction about the expected full immobility of CIPs seems to be correct for Chinese, Japanese, Bangla and other classifier languages, which do not permit a CIP to move and strand a numeral, as illustrated in (37) from Mandarin. The ungrammaticality of Vietnamese (34b) is similarly accounted for by the interaction of antilocality and the PIC, given the assumption that CIP is the complement of a phasal head Q, instantiated by numerals.

- (37) *_{[CIP ben shu]_k} wo mai-le _[QP san t_k] *Mandarin*
 CL book I buy-ASP 3
 Intended: ‘Books, I bought three.’

With regard to examples involving the separation of nominal phrases with optional-classifier nouns as in (35), the otherwise puzzling requirement that a classifier necessarily occur in such structures can be attributed to facilitation of the extraction of the noun/NP. Supposing it is assumed that optional-classifier nouns undergo N-to-Cl movement when an overt classifier is not made use of, in order to satisfy the affix-like properties of CL- \emptyset merged in Cl, this will restrict any movement of the noun out of the nominal expression. If the noun is moved to the Cl position, and the attempt is made to extract the containing CIP out of QP, stranding the numeral in Q, such movement should have to proceed via SpecQP in order not to violate the PIC, but movement of the CIP complement of Q to its specifier position will violate antilocality. In the absence of an overt classifier, it will thus not be possible for optional-classifier nouns to be extracted out of QP if such nouns regularly undergo raising to Cl. When an overt classifier is present, however, there will be no need for any N-to-Cl movement of the noun, and the latter will remain in N. Consequently, the containing NP may legitimately extract out of QP, as movement of the NP to the edge of the phase, SpecQP, will not violate antilocality, resulting in the acceptability of NP fronting in *bị* sentences only when an overt classifier is present.

Interestingly, such restrictions on noun/numeral separation also appear to have an effect on non-classified nouns, which is quite unexpected if N-to-Cl movement of such elements is not assumed. Although the class of non-classified nouns does not tolerate the use of distinct classifiers in simple count structures, unlike optional-classifier nouns, when non-classified nouns occur separated from numerals in *bị* sentences, a classifier does in fact appear, and must be present, as illustrated in (38) and (39).

- (38) a. Vương quốc của quân địch bị phá hoại ba *(cái).
 country of enemy PASS destroy 3 CL
 ‘Three enemy countries were destroyed.’
 b. Truyền thuyết bị phủ nhận hai *(cái).
 legend PASS disprove two CL
 ‘Two legends were disproved.’
- (39) a. Xã hội bị phân hóa hai *(cái).
 society PASS segregate two CL
 ‘Two societies were segregated.’
 b. Chính phủ bị lật đổ hai *(cái).
 government PASS overthrow two CL
 ‘Two governments were overthrown.’

This patterning can be explained in a way parallel to that with optional-classifier nouns. In nominal structures with numerals, it can be assumed that the noun regularly undergoes N-to-Cl movement. In *bi* sentences, this would result in a structure in which it would not be possible for the noun to move further, out of the QP, as the CIP complement of Q is immobile and cannot extract from QP. In order to permit extraction of the noun, it appears that an exceptional last-resort use of a classifier may occur, allowing for the noun to remain in N and undergo fronting in its containing NP constituent. Without the assumption of N-to-Cl movement, the forced introduction of classifiers with non-classified nouns and optional-classifier nouns, specifically under conditions of separation of nouns from numerals, becomes very difficult to account for. Noun-numeral separation phenomena thus provide good additional support for the idea that nouns in Vietnamese undergo movement to a classifier position when an overt classifier is not present, and that all three classes of noun project a CIP which must be lexicalized in some way.

4.3 Short answer forms consisting in the heads of human compounds

A third set of phenomena which provides potential support for the hypothesis of N-to-Cl movement with optional classifier and non-classified nouns in Vietnamese relates to the occurrence of short answer forms with certain compound nouns. Quite generally, classifier languages such as Chinese and Japanese allow for short-answer forms to questions of the type ‘How many N (are there in X/did you buy)?’ which consist of a numeral and classifier pair, as seen in (40–41). It is ungrammatical to provide simply a numeral as answer-form, and it is felt to be heavily redundant to provide the noun as well as the numeral + classifier pair.

(40) a. si *(ben) b. si ben (?shu) *Mandarin*
 4 CL 4 CL book
 ‘four books’

(41) a. yon *(satsu) b. yon satsu (?no hon) *Japanese*
 4 CL 4 CL GEN book
 ‘four books’

In Vietnamese, parallel short answer forms may consist of the pairing of a numeral and a classifier, or simply of a numeral, unlike in Chinese and Japanese. As with Chinese and Japanese, the repetition of the noun in a short answer form is felt to be heavily redundant (unlike repetition of the classifier). This is illustrated in (42):

- (42) *Có bao nhiêu cuốn sách?*
 be how many CL book
 ‘How many books are there?’
- a. *Bốn.* b. *Bốn cuốn.* c. *?Bốn cuốn sách.*
 4 4 CL 4 CL book
 ‘Four.’

Here we will consider an interesting variation of this type of short-answer form which occurs with certain compound nouns in Vietnamese, in which just the head of the compound rather than the entire noun is repeated in a short answer-form with a numeral. This pattern typically occurs with human compounds which are constructed with a kin term as the head of the compound, and also with various other human compounds which are productively created with a non-kin term head. Examples of kin term headed compounds are illustrated in (43). These items involve the kin terms *cô* ‘aunt’, *bà* ‘grandmother’, and *ông* ‘grandfather’. As nominal compounds in Vietnamese are regularly left-headed, this results in the kin term being the first syllable in the noun. As noted in the translations in (43), when elements such as *cô* ‘aunt’, *bà* ‘grandmother’, and *ông* ‘grandfather’ occur in compounds, the way they contribute semantically is to restrict the gender and (sometimes) the age of the resulting noun, and, as a result of semantic bleaching and drift, they do not retain their literal meaning ‘aunt’, ‘grandmother’ and ‘grandfather’—the nouns in (43) are regularly used to refer to people who are not related to the speaker.

- (43) a. *cô bán hàng* b. *bà giáo* c. *ông chủ nhà*
 aunt.sell.goods grandmother.teach grandfather.lord.house
 ‘female vendor’ ‘older female teacher’ ‘landlord’

Examples of relevant human compounds not headed by kin terms are given in (44), where the heads of the compounds are *thợ* ‘workman’, *người* ‘person’, and *thầy* ‘teacher’:

- (44) a. *thợ sửa xe* b. *người đánh cá* c. *thầy giáo*
 workman.fix.vehicle person.catch.fish teacher.teach
 ‘car mechanic’ ‘fisherman’ ‘teacher’

When such items occur in ‘how many’ questions, the special short answer form that is available involves a repetition of the first syllable of the compound together with the numeral, resulting in an answer form that is felt to be significantly less redundant than a full repetition of the noun. With the kin-term compounds in (43),

the short answer-forms are as shown in (45–47).¹⁶ No classifier is available with these elements (i.e. they pattern as non-classified nouns), so the use of the general human classifier *người* is not possible, and repetition of the full noun is unnatural, leaving the two common choices of either simply using a bare numeral, or adding the first syllable of the compound to the numeral, as seen in the (b) examples below. As the status of the repeated compound heads is somewhat at issue here, we will gloss these items simply as CO, BA and ONG in short answers.

- (45) Có bao nhiêu cô bán hàng ở đó?
 have how many aunt.sell.goods be there
 ‘How many female vendors are there?’
- (a) Hai. (b) Hai cô. (c) */?Hai người. (d) ??Hai cô bán hàng.
 2 2 CO 2 CL 2 aunt.sell.goods
 ‘Two.’ ‘Two.’
- (46) Có bao nhiêu ông chủ nhà ở đó?
 Have how many grandfather.lord.house be there
 ‘How many landlords are there?’
- (a) Hai. (b) Hai ông. (c) */?Hai người. (d) ??Hai ông chủ nhà.
 2 2 ONG 2 CL 2 grandfather.lord.house
 ‘Two.’ ‘Two.’
- (47) Có bao nhiêu bà giáo ở đó?
 have how many grandmother.teach be there
 ‘How many older female teachers are there?’
- (a) Hai. (b) Hai bà. (c) */?Hai người (d) ??Hai bà giáo.
 2 2 BA 2 CL 2 grandmother.teach
 ‘Two.’ ‘Two.’

With the non-kin term compounds in (44), it is again possible to use the first syllable of the compound together with a numeral as a short answer form, as seen in (48) and (49). As these are optional classifier nouns, it is also possible for a short answer to consist of a numeral and the classifier *người*, but this is less preferred than the short answer forms in (a) and (b).

¹⁶ Note that the interpretation of *cô/ông/bà* in the short answer-forms corresponds exactly to the meaning these morphemes have as heads of the compounds in (45–47), imposing general gender and age-related restrictions on the meaning of the answer-form and not signaling family membership. This indicates that the short answer forms are created by use of the compound head, not use of the separate nouns *cô/bà/ông*, which might be expected to give rise to the literal meanings of ‘two aunts/grandfathers/grandmothers’ in (45b, 46b, 47b).

- (48) Có bao nhiêu (người) thợ sửa xe trong tiệm này?
 have how many CL workman.fix.vehicle in shop this
 ‘How many car mechanics are there in this shop?’
- (a) Hai. (b) Hai thợ. (c) Hai người (d) ??Hai thợ sửa xe.
 2 2 THO 2 CL 2 workman.fix.vehicle
 ‘Two.’
- (49) Có bao nhiêu thầy giáo trong khoa này?
 have how many teacher.teach in department this
 ‘How many teachers are there in this department?’
- (a) Hai. (b) Hai thầy. (c) Hai người (d) ??Hai thầy giáo.
 2 2 THAY 2 CL 2 teacher.teach
 ‘Two.’

As short-answer forms of the type [numeral X] are regularly found to combine a numeral with an element in the classifier position, as seen in (40a/41a/42b), the patterns in (48b/49b) can be taken to suggest that there is a process in which the heads of certain compounds are able to occur raised in CI, when the residue of the compound is elided, in apparent violation of the Lexical Integrity Hypothesis (LIH), which suggests that syntactic processes may not apply to parts of words (Lapointe 1980, Selkirk 1982). Over the years, a variety of legitimate ‘violations’ of the LIH have in fact been identified, for example, the separation of parts of compound words that have non-transparent meanings in Chinese, such as (50) from Huang (1983), and similar cases in Hebrew (Borer 1988).

- (50) a. ta hen danxin zhe jian shi.
 he very carry.heart (=worry) this Cl matter
 ‘He is very worried about this matter. (Huang 1983)
- b. xin_k, wo xiang ta shi hui dan t_k de.
 Heart I think he be will carry DE
 ‘Worry, I think he will.’ (Huang 1983)

Vu (1999) demonstrates that compounds in Vietnamese may also be syntactically split apart under certain circumstances, as seen in (51), where the two-part coordinative compound *nhà cửa* lit. ‘house door’ meaning ‘houses’ is split apart by negation and the verb:

- (51) Nó chẳng thích nhà chẳng thích cửa.
 3rd NEG like house NEG like door
 ‘He doesn’t like houses.’

Many further patterns in which the LIH is violated are catalogued in Lieber and Scalise (2007), indicating that the LIH in its original formulations is over-restrictive. We suggest that the raising of the head of compounds to the classifier position in Vietnamese short answer forms is an additional example of a ‘legitimate’ LIH violation, calling for a reconceptualization of the LIH (see Lieber and Scalise 2007 for an attempt along these lines). Just as there seem to be restrictions on how the LIH may be violated in other cases (for example, the idiomatic compound-splitting operation noted by Vu (1999) in (51) may only occur in sentences with generic readings), the use of the first/leftmost syllable/morpheme of a compound as a short answer form in Vietnamese is also subject to restrictions and not available with all compounds. As illustrated in (52), it is not possible with compounds constructed with the morpheme *nhà* ‘expert’ such as *nhà báo* [lit. expert report] ‘journalist’, and it is also not possible with complex nouns such as *nhân viên* ‘employee’, *phụ huynh* ‘parent’, and *quản lý* ‘manager’.

- (52) Có bao nhiêu nhà báo trong phòng này?
 Have how many expert.report in room this
 ‘How many journalists are there in this room?’
- (a) Hai. (b) *Hai nhà. (c) Hai người. (d) ??Hai nhà báo.
- 2 CL 2 NHA 2 CL 2 expert.report
 ‘Two.’

We suggest there are two conditions governing the availability of short answer forms consisting in the first syllable of multi-morphemic nouns. First, the compound noun from which extraction and raising takes place must have its own internal hierarchical structure, and it is only the head of the compound which may raise out of the compound. As a result, other morphemes present in compounds are not legitimate targets for use as short answer forms, and a reply to (46) could not be *hai chủ* ‘two lord’ or *hai chủ nhà* ‘two lord house’, in which the second or second and third, non-head morphemes are extracted from the compound noun *ông chủ nhà* and used to construct the short answer form. Instead, it is the head of the compound *ông* which must be used in the short answer form. Second, an output condition on the movement applies, and it must be possible for the element which has been raised out of the compound (the head of the compound) to legitimately stand as an independent word, and not be lexically listed as a bound morpheme. This condition allows for elements such as *cô*, *bà*, *ông*, *thợ* and *thầy* to occur as short answer forms corresponding to compounds constructed with these elements as heads, because kin terms like *cô*, *bà*, and *ông* may all occur as independent nouns, as may the non-kin term elements *thợ* and *thầy*, as illustrated in (53).

- (53) hai người cô/bà/ông/thợ/thầy
 2 CL aunt/grandmother/grandfather/worker/teacher
 ‘two aunts/older ladies/older men/workers/teachers’

Such a filter disallows the use of bound morphemes such as *nhà* and *nhân* in short answer forms, as seen already in (52b)

While free morphemes such as *cô*, *bà*, *ông*, *thợ* and *thầy* elsewhere may occur as nouns, there is strong reason to believe that in the short answer forms they are occurring in the CI position, and not the N position—the kin term *cô*, for example, always requires a classifier when it occurs as a noun, as shown in (54) (i.e. it patterns as an obligatory-classifier noun), but appears legitimately without any supporting classifier in short answer forms, so has a different status when being used in such structures.

- (54) hai *(người) cô
 2 CL aunt
 ‘two aunts’

The patterns examined here thus suggest that the head of compounds is able to undergo raising to the classifier position, when this element is not a bound morpheme. Further patterns indicate that such movement to the CI position is not only possible with the heads of certain compound nouns, but also forced to occur. In passive-like *bị* sentences (see Sect. 4.2), there is a clear difference in patterning between compounds whose head is a kin term, and compounds whose head is a free morpheme but not a kin term, such as *thợ* and *thầy*. The latter elements allow for separation of the noun from the numeral via fronting of the former, as shown in (55). When this occurs, it is necessary for the otherwise optional classifier for such nouns to be present with the stranded numeral.

- (55) Thợ sửa xe/thầy giáo_k bị sa thải hai *(người) t_k.
 workman.fix.vehicle/teacher.teach PASS fire 2 CL
 ‘Two workmen/teachers were fired.’

However, similar fronting of the noun/NP in *bị* sentences where the noun is a compound headed by a kin term is not at all possible:

- (56) *Cô bán hàng_k bị sa thải hai t_k.
 aunt.sell.goods PASS fire 2
 Intended: ‘Two female vendors were fired.’

This difference in patterning can be accounted for straightforwardly if the head of compounds formed with a kin term regularly must undergo raising to the classifier position. This will result in the constituent being fronted in (56) being a CIP, which has already been argued to be immobile (Sect. 4.2) in virtue of being the complement of a phase. In (55), however, where the compound noun is not headed by a kin term element, it is possible (and necessary) for an independent classifier to be inserted in the classifier position, with the result that the fronted constituent is an NP and able to undergo movement. Note that it is not possible to save structures such as (56) by any similar attempt to insert a classifier (*người*) following the numeral:

- (57) *Cô bán hàng_k bị sa thải hai người t_k.
 aunt.sell.goods PASS fire 2 CL
 Intended: 'Two female vendors were fired.'

This consequently suggests that the head of kin term compounds must undergo raising to Cl, and so does not allow for the use of an independent classifier, while the heads of non-kin term compounds only optionally allow for head-raising to Cl, and do not enforce it, thus permitting the use of an independent classifier.

A reviewer of the paper asks why it might not be possible for N to raise to Cl in (56), and allow for the remnant NP to raise to SpecTP. Supposing the full noun *cô bán hàng* 'female vendor' were to raise to Cl and the remnant NP were subsequently to move to SpecTP, the output would be as in (58):

- (58) *_{[NP t_k]_m} bị sa thải <sub>[DP [QP hai _[CIP cô bán hàng_k t_m]]].
 PASS fire 2 female vendor</sub>

Alternatively, if one were to assume, as suggested here, that just the head of the compound raises to Cl with kin term compounds, the resulting derivation would be as represented in (59):

- (59) *_{[NP t_k bán hàng]_m} bị sa thải <sub>[DP [QP hai _[CIP cô_k t_m]]].
 sell goods PASS fire 2 aunt</sub>

Both such derivations can be ruled out as occurrences of illicit remnant movement, following observations in Takano (2000), who notes that 'Remnant movement of α is impossible if the head of α has moved out of α ' (Takano 2000:146). Examples such as (60) and (61) illustrate ungrammatical attempts at such 'illicit remnant movement', which fail, according to Takano, due to the remnant constituent no longer containing the features necessary to drive movement once the head has been removed.¹⁷

- (60) *_{[VP Ihr ein Buch t_i]_j} gab_i Hans t_j.
 her a book gave Hans

¹⁷ The constraint that remnant movement may not occur if the head of a constituent α has been moved out of α will also account for why an overt classifier must be used in passive sentences where an optional classifier noun is raised to the subject position, as in (35), repeated here. If the head of the NP *làng* 'village' were to be raised out of the NP to lexicalize Cl in place of merging a classifier in this position, this would make it impossible for the headless NP remnant to raise to the subject position. The obligatory use of classifiers with optional classifier nouns, and the impossibility of N-to-Cl movement in passive sentences is thus accounted for as a way to avoid illicit remnant movement.

- (35) làng (của họ) bị phá hoại tám *(cái).
 village of 3pl PASS destroy 8 CL
 'Eight (of their) villages were destroyed'

(61) *It's [_{VP} a book t_i to Mary]_j that John gave_i t_j .

The analysis of kin term compounds that has now emerged, in which the head of such nouns is regularly raised to the classifier position, leads to the expectation that kin term compounds should pattern like sequences of a classifier followed by a noun in other instances, for example, in the special 'bare classifier' construction noted to occur in Vietnamese in Daley (1998), Nguyen (2004), and others, where a classifier and a noun occur without any preceding numeral (i.e. simply as [CI N]), as illustrated in (62).¹⁸

(62) a. Cuốn sách ở đó. b. Con chó lớn lắm.
 CL book be there CL dog big very
 'The dog is there.' 'The dog is very big.'

This prediction is indeed borne out. Interestingly, there is a restriction on the use and interpretation of kin term compounds that such nouns cannot be used with generic interpretations, as illustrated in (63) and (64).

(63) Cô bán hàng Nhật lịch sự lắm.
 aunt.sell.goods Japan polite very
 Only: 'The Japanese female vendor is very polite.'
 Not possible: 'Japanese female vendors are very polite.'

(64) Tôi không thích cô bán hàng Nhật.
 I NEG like aunt.sell.goods Japan
 Only: 'I don't like the Japanese female vendor.'
 Not possible: 'I don't like Japanese female sales assistants.'

This is exactly the interpretation of bare classifier patterns elsewhere: it has often been observed that the combination of a classifier and noun without a numeral only allows for referential interpretations and not generic interpretations, as seen in (65) and (66).

¹⁸ The bare classifier construction is, incidentally, a further patterning which shows that the appearance of classifiers in Vietnamese is formally independent of numerals and conditioned only by nouns. In the bare classifier pattern, it is clearly not possible to attribute the presence of classifiers to the need for numerals to add any measuring function, as no numerals occur in such constructions. Such forms thus create a further challenge for the application of Krifka's (1995) approach to languages such as Vietnamese, Bangla, Hmong and certain varieties of Chinese where bare classifier patterns are very common (Cheng and Sybesma 2005; Simpson et al 2011; Simpson and Biswas 2016; Simpson 2017).

- (65) Cuốn sách Pháp thú vị lắm.
 CL book France interesting very
 Only: 'The French book is very interesting.'
 Not possible: 'French books are very interesting.'
- (66) 'Tôi không thích cuốn sách Pháp.'
 I NEG like CL book France
 Only: 'I don't like the French book.'
 Not possible: 'I don't like French books.'

Such a restriction on the interpretation of bare classifier forms is commonly attributed to the function that the classifier has in individualizing the noun, and is a natural explanation for the absence of generic interpretations with kin term compounds too. With such elements the classifier position is regularly lexicalized by the kin term head, which then individualizes the noun and eliminates the possibility that it can be construed generically. Note that at this stage, one might perhaps wonder whether the kin term elements which have been referred to here as the heads of the nominal compounds are actually not a part of the noun, but might actually be base-generated in the classifier position, this accounting for the restrictions on the use and interpretation of 'kin term compounds'. However, this cannot be so, as it is crucially the kin term element which nominalizes sequences such as *cô bán hàng* [lit. aunt sell goods] 'female vendor' and provides the correct meaning of such sequences when quantified by numerals. Supposing *cô* were only to be a classifier and to combine with *bán hàng* 'sell goods' as its complement, the meaning of a sequence such as '*hai cô bán hàng*' [2 CL sell goods] would be expected to be reference to two individualized instances of selling goods, rather than two female vendors. It therefore has to be concluded that the kin term in kin term compounds indeed plays a complex dual role, both nominalizing and adding lexical content to compound nouns, and also serving to individualize the noun via raising to the classifier position.

Closing this section, we have seen through all of the various Sects. 4.1–4.3 that there is empirical support of a range of types for the assumption that a uniform syntactic structure is indeed projected with all three classes of noun in Vietnamese, with a regular classifier position that is overtly lexicalized either by direct insertion of an overt classifier or via the application of movement of material from the N position, when it is hypothesized that a phonetically null classifier CL_{\emptyset} is selected and merged with optional-classifier and non-classified nouns. Such assumptions allow for a coherent account of an intricate set of patterns, and are able to explain certain otherwise puzzling restrictions on the ways that nouns and classifiers may be used and interpreted in Vietnamese. They also bear on broader, cross-linguistic issues relating to the kinds of nominal structures that may be projected in numeral classifier languages, as will now be noted in the final, conclusions section of the paper.

5 Conclusions: Vietnamese and the typology of numeral/classifier/noun relations

This paper began with an examination of Bale and Coon's recent (2014) claim that 'classifiers are for numerals not for nouns' and its support for a Krifka (1995) style analysis of numerals, classifiers and nouns over the approach proposed in Chierchia (1998). One goal of the paper was to see how the complex patterns of classifiers in Vietnamese would bear on Bale and Coon's proposals, which seem to predict syntactic structures that are quite different from those often assumed for languages such as Chinese, with numerals combining with classifiers before being built together with nouns. We also aimed to see how Vietnamese would potentially fit with Chierchia's (1998) nominal typology of languages, and whether a coherent, unified analysis of the significant variation in classifier/noun patterns could be found within Vietnamese.

In Sect. 2 it was shown that Vietnamese is a language in which the surface distribution of numerals, classifiers and nouns appear to lead to conclusions which are in fact opposite to those drawn in Bale and Coon on the basis of Chol and Mi'gmaq. Vietnamese is a language where the overt presence vs. absence of classifiers largely appears to be an idiosyncratic non-predictable property of nouns, not numerals. This provides evidence for a [numeral+[classifier+noun]] syntactic alignment in which classifiers are first combined with nouns before being built together with numerals, rather than the [[numeral+classifier]+noun] alignment Bale and Coon argue for on the basis of Chol and Mi'gmaq. The Vietnamese patterns, viewed in terms of the overt occurrence vs. non-occurrence of classifiers, thus offer support for the position that, at least in some languages, 'classifiers are for nouns, not numerals', as broadly assumed in Chierchia (1998), and that the hypothesis that 'classifiers are for numerals, not nouns', put forward in Bale and Coon (2014), cannot be universally correct.

Section 3 of the paper subsequently noted how the variegated patterning of nouns and classifiers in Vietnamese also initially appears to pose a challenge to Chierchia's typology of languages based on the properties of their nouns, with Vietnamese seeming to display mixed properties of both Chinese and English-type languages. We then set about examining whether a uniform account of Vietnamese nominals (which would accord more directly with Chierchia's typology) might result from the possibility that there is more underlying structure present with certain nouns than may initially be apparent, and that a classifier position/CIP might also be projected with 'non-classified nouns'. In Sect. 4 we considered a range of theoretical arguments and empirical patterns bearing on such an analysis, in each instance making significant use of a cross-comparison of patterns with non-classified nouns, optional-classifier nouns and obligatory-classifier nouns, in a way that is not so readily possible with other classifier languages such as Chinese and Japanese where less variation occurs. This investigation provided a variety of evidence in support of a uniform account of nominal projections in Vietnamese in which all countable nouns are embedded by means of a classifier phrase and surface differences across nouns result from whether an overt classifier occurs with a noun,

or a phonetically null affix-like form $CL_{-\emptyset}$, triggering movement of the noun to adjoin to Cl.

The assumption that count nouns in Vietnamese uniformly occur with a classifier projection leaves Chierchia's typology essentially intact, and obviates the need to assume that there are hybrid languages with mixed classifier/non-classifier properties. The suggestion that the head of a syntactically-present classifier phrase may sometimes be phonetically null also adds an extra layer of complexity to the debate on classifier syntax discussed in Bale and Coon (2014). While Bale and Coon focus their attention exclusively on overt surface forms and how lexical idiosyncrasy can be taken as evidence of a close numeral-classifier relation (or potentially a close classifier-noun relation), the consideration of broader syntactic patterns may suggest that a classifier phrase can be present even when no overt classifier actually occurs in a surface string. Vietnamese, as described and analyzed here, provides useful information on the status of classifiers within nominal projections in two relevant ways. First, if the line of argumentation presented in Bale and Coon (2014) is followed, the overt distribution of classifiers and nouns and lexical idiosyncrasy in the observed relation of overt classifiers to nouns presents evidence for the general application of a Chierchia-type alignment of numerals, (overt) classifiers and nouns in Vietnamese, with classifiers first combining with nouns before being built together with numerals. Second, the syntactic analyses of the extra *cái* construction, passive sentences, and kin term compounds suggest that such a mode of combination is present not only when there are overt classifiers, but also when a numeral is combined with a noun apparently without any classifier, and hence that, at an underlying level, a Chierchian-type syntax in fact occurs with all count nouns in Vietnamese. Classifier syntax in Vietnamese is consequently rich and informative in its complexity, and able to provide useful insights into the ways that nominal projections in classifier languages may structurally be organized, currently a topic of considerable debate.

References

- Abels, Klaus. 2003. *Successive Cyclicity, Anti-locality, and Adposition Stranding*. PhD dissertation, University of Connecticut, Storrs.
- Bale, Alan, and Jessica Coon. 2014. Classifiers are for Numerals, Not for Nouns: Consequences for the Mass/Count Distinction. *Linguistic Inquiry* 45: 695–707.
- Borer, Hagit. 1988. On the Parallelism Between Compounds and Constructs. *Yearbook of Morphology* 1: 45–66.
- Borer, Hagit. 2005. *Structuring Sense: Volume 1: In Name Only*. Oxford: Oxford University Press.
- Bošković, Željko. 2014. Now I'm a Phase, Now I'm Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45(1): 27–89.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and Not-so-Bare Nouns and the Structure of NP. *Linguistic Inquiry* 30: 509–542.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 2005. Classifiers in Four Varieties of Chinese. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard Kayne, 259–292. Oxford: Oxford University Press.
- Chierchia, Gennaro. 1998. Reference to Kinds Across Language. *Natural Language Semantics* 6(4): 339–405.

- Chierchia, Gennaro. 2015. How Universal is the Mass/Count Distinction? In *Chinese Syntax in a Cross-Linguistic Perspective*, ed. Audrey Li, Andrew Simpson, and Dylan Tsai, 147–175. Oxford: Oxford University Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 2000. Minimalist Inquiries. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Daley, Karen. 1998. *Vietnamese Classifiers in Narrative Texts*. Arlington: University of Texas, Summer Institute of Linguistics.
- Grohmann, Kleanthes. 2003. *Prolific Domains: On the Anti-locality of Movement Dependencies*. Amsterdam: John Benjamins.
- Her, One-Soon, and Hui-Chin Tsai. 2014. Color isn't Silent, Shallow isn't Deep: Two Case Studies of Evaluating Silent Elements. *Language and Linguistics* 15: 775–800.
- Holmberg, Anders. 2000. Scandinavian Stylistic-Fronting: How Any Category Can Become an Expletive. *Linguistic Inquiry* 31(3): 445–483.
- Huang, C.-T. James. 1983. Phrase Structure, Lexical Integrity, and Chinese Compounds. *Journal of the Chinese Language Teachers Association* 19: 53–78.
- Huang, C.-T. James, Yen-Hui Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Kayne, Richard. 2005. *Movement and Silence*. Oxford: Oxford University Press.
- Kayne, Richard. 2012. A Note on Grand and Its Silent Entourage. *Studies in Chinese Linguistics* 33(2): 71–85.
- Krifka, Manfred. 1995. Common Nouns: A Contrastive Analysis of English and Chinese. In *The Generic Book*, ed. Gregory Carlson and Francis Pelletier, 398–411. Chicago: University of Chicago Press.
- Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lapointe, Steven. 1980. A Theory of Grammatical Agreement. PhD dissertation, Amherst, University of Massachusetts.
- Li, Xu Ping. 2013. *Numeral Classifiers in Chinese*. Berlin: de Gruyter Mouton.
- Li, Yen-Hui Audrey. 1998. Argument Determiner Phrases and Number Phrases. *Linguistic Inquiry* 29(4): 693–702.
- Lieber, Rochelle, and Sergio Scalise. 2007. The Lexical Integrity Hypothesis in a new theoretical universe. In *On-line Proceedings of the Fifth Mediterranean Morphology Meeting*, ed. Geert Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli, and Sergio Scalise, 1–24. <http://mmm.lingue.unibo.it/>.
- Löbel, Elisabeth. 2000. Classifiers Versus Genders and Noun Classes: A Case Study in Vietnamese. In *Gender in Grammar and Cognition: I: Approaches to Gender. II: Manifestations of Gender*, ed. Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, and Mirja Saari, 259–313. Berlin: de Gruyter Mouton.
- Manetta, Emily. 2011. *Peripheries in Kashmiri and Hindi-Urdu*. Amsterdam: John Benjamins.
- Ngo, Binh. 2012. *Vietnamese Classifiers: (Non-)specificity*. MA thesis, Northeastern Illinois University.
- Nguyen, D.H. 1957. Classifiers in Vietnamese. *Word - Journal of the International Linguistic Association* 13(1): 124–152.
- Nguyen, Tüông Hùng. 2004. *The Structure of the Vietnamese Noun Phrase*. PhD dissertation, Boston University.
- Radford, Andrew. 2004. *Minimalist Syntax*. Cambridge: Cambridge University Press.
- Selkirk, Elizabeth. 1982. *The Syntax of Words*. Cambridge, MA: MIT Press.
- Simpson, Andrew. 2008. Classifiers and DP Structure in Southeast Asia. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard Kayne, 806–838. Oxford: Oxford University Press.
- Simpson, Andrew. 2012. The Silence of the Bucks: A Note on Kayne's 'Grand and Its Silent Entourage'. *Studies in Chinese Linguistics* 33(2): 87–101.
- Simpson, Andrew. 2017. Bare Classifier/Noun Alternations in the Jinyun (Wu) Variety of Chinese and the Encoding of Definiteness. *Linguistics* 55(2): 305–332.
- Simpson, Andrew, and Priyanka Biswas. 2016. Bare Nominals, Classifiers, and the Representation of Definiteness in Bangla. *Linguistic Analysis* 40(3–4): 167–198.

- Simpson, Andrew, and Tam Ho. 2013. Vietnamese and the Typology of Passive Constructions. In *Linguistics of Vietnamese*, ed. Daniel Hole and Elisabeth Löbel, 155–185. Berlin: de Gruyter Mouton.
- Simpson, Andrew, Hooi Ling Soh, and Hiroki Nomoto. 2011. Bare Classifiers and Definiteness: A Cross-Linguistic Investigation. *Studies in Language* 35: 168–193.
- Simpson, Andrew, and Saurov Syed. 2016. Blocking Effects of Higher Numerals in Bangla: A Phase-Based Analysis. *Linguistic Inquiry* 47: 754–763.
- Syed, Saurov, and Andrew Simpson. 2017. On the DP/NP Status of Nominal Projections in Bangla: Consequences for the Theory of Phases. *Glossa* 2(68): 1–24.
- Takano, Yuji. 2000. Illicit Remnant Movement: An Argument for Feature-Driven Movement. *Linguistic Inquiry* 31: 141–156.
- Truong, Văn Chính. 1970. *Structure de la Langue Vietnamiennne*. Paris: Imprimerie nationale P. Geuthner.
- Vu, Sonny. 1999. *Splitting Compounds*. Talk Presented at the Annual Meeting of the Linguistics Society of America, Los Angeles.
- Zhang, Hong. 2007. Numeral Classifiers in Mandarin Chinese. *Journal of East Asian Linguistics* 16: 43–59.
- Zhang, Niina Ning. 2013. *Classifier Structures in Mandarin Chinese*. Berlin: de Gruyter Mouton.