

**THE SEMANTICS AND GRAMMAR OF
VIETNAMESE CLASSIFIERS**

SIM SOOK HUI

(B.Arts. (Hons.), NUS)

**A DISSERTATION SUBMITTED
FOR THE DEGREE OF MASTER OF ARTS
DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE
NATIONAL UNIVERSITY OF SINGAPORE**

2003

ACKNOWLEDGEMENTS

I am greatly indebted to Dr Umberto Ansaldo, a true mentor, who had been consistently encouraging and had inspired a lot of ideas and concepts which were incorporated in my work. During these two years, he guided me on the right track and had relentlessly watched over me making sure my research was progressing smoothly and the quality of the dissertation is maintained. He never failed to offer a solution whenever a problem cropped up, and with his knowledge and patience, numerous doubts and issues were resolved. Another area which I had received a lot of help on was the acquiring of resources for my research. A lot of material could not have been gathered without Dr Ansaldo's help and I am indeed tremendously grateful to him for everything he had done to ensure the completion of my research.

I am also grateful to Dr Mark Donohue, who had constructively suggested new lines of investigation within my research topic.

Before I express my gratitude to my Vietnamese informants, I would like to thank the Research and Graduate Studies Division (Faculty of Arts and Social Sciences) for the funding of my overseas fieldwork which was conducted August - September 2003. I was able to elicit and check Vietnamese data with native speakers and that guaranteed the quality of my data. For Vietnamese informants, I must express special thanks to Mr Thach Ngoc Minh. He was my informant in Singapore who also helped me to establish contacts in the University of Social Sciences and Humanities in Ho Chi Minh City. Miss Kim Anh,

the secretary of the Linguistics department, had been kind enough to arrange for me to meet the Head of the department, Mr Nguyen Van Hue. I am also thankful to Miss Bui Cam Ly who hosted me for the first few days when I was in Vietnam. Without her, I would be terribly lost. I am very grateful to Mr Nguyen for agreeing to be my informant and I had also benefitted from the materials and resources that he loaned to me. In addition, he introduced me to more Vietnamese informants: Miss Le Tu Anh and Mr Chien Dung. I am personally indebted to Miss Le Tu Anh because of her consistent efforts in helping me with my research and for her kindness and friendship. Other Vietnamese friends and informants that I would like to thank are Mr Phan Thai Hoang Son, Mr John Pham, Mr Ton Hung Cuong, Mr Nguyen Quoc Loc and Mr Phan Hieu Hao. I also have to thank Mr Phan Huy Thuy for being my informant and spending long periods of time going through the data with me.

Finally, I must express my appreciation for my family, especially my mother, for taking care of my needs during the busiest periods of my research. Also, special thanks to Brandon Ng for being so supportive and helpful during the course of my research.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	x
ABBREVIATIONS	xii
ABSTRACT	xiii
CHAPTER 1: LITERATURE REVIEW	
1.0 Introduction	1
1.1 Historical development of studies of classifier systems	7
1.2 Establishing typologies of classifier systems	8
1.2.1 Dixon (1986): Noun classes and noun classifiers	10
1.2.2 Allan (1977): Semantic basis of classifier systems	11
1.2.3 Adams (1989): Semantic dimensions underlying classifier systems	12
1.2.4 Becker (1975): Spatial metaphor in the Burmese classifier system	12
1.2.5 Denny (1976): Semantic functions of classifiers and interaction with the world	13
1.3 Croft (1994): Semantic universals in classifier systems	13
1.4 Other studies on the extended functions of classifiers	18
1.5 Aikhenvald (2000): Classifiers, a typology of noun categorization devices	19
1.6 Aims of the thesis and significance	20

1.7	Organization of thesis	22
1.8	Summary	24

CHAPTER 2: TYPOLOGICAL FRAMEWORK

2.0	Introduction	25
2.1	Typology in a nutshell	26
2.2	Significance of typological research	27
2.3	Issues in a functional-typological approach	31
2.4	Difficulties in typological research: Methodology	33
2.4.1	Representativeness of available languages	34
2.4.2	The problem of sample size	34
2.4.3	Genetic and areal bias	35
2.4.4	The source of data	37
2.4.5	Definition of formal features	39
2.5	Aikhenvald's framework and methodology	40
2.6	My methodology and data collection	44
2.7	Conclusion	45

CHAPTER THREE: SYNTACTIC PATTERNS

3.0	Introduction	47
3.1	The Vietnamese language and classifier constructions	48
3.2	Mensural and sortal classifiers	54
3.3	Mensural and sortal classifiers in Vietnamese	58

3.4	Absence of repeaters	73
3.5	Numeral and demonstrative constructions	75
3.5.1	Num-CL-N-Dem	76
3.5.2	Num-CL-N	78
3.5.3	Num-CL	79
3.5.4	Num-N-Dem	80
3.5.5	Num-N	80
3.5.6	CL-N-Dem	81
3.5.7	CL-N	84
3.5.8	CL-Dem	85
3.5.9	N-Dem	86
3.5.10	N-Num-CL	87
3.6	Constituent order and constituency relations in classifier constructions	89
3.6.1	[Num-CL]-N	91
3.6.2	Num-[CL-N]	93
3.6.3	Num-CL-N	94
3.7	The origin of the classifier construction in Vietnamese	95
3.8	Conclusion	97

CHAPTER FOUR: SEMANTIC CATEGORIES AND CLASSIFICATION

4.0	Introduction	99
4.1	General semantic properties of noun categorization devices	100
4.1.1	Animacy	100

4.1.2	Sex and gender	101
4.1.3	Kinship	102
4.1.4	Social status	103
4.1.5	Nature	103
4.1.6	Physical properties	104
4.1.7	Function	105
4.2	Vietnamese (numeral) classifiers – basic features	106
4.3	Semantics of numeral classifiers	109
4.4	Modes of classification of animates in Vietnamese	110
4.5	Source of human classifiers	118
4.5.1	Source of <i>người</i>	118
4.5.2	Source of kinship classifiers	122
4.6	Source of non-human animate classifiers	125
4.7	Characteristics of animate classification	126
4.8	Modes of classification of inanimates in Vietnamese	127
4.8.1	One-dimensional category	128
4.8.2	Source of one-dimensional classifiers	131
4.8.3	Two-dimensional category	132
4.8.4	Source of two-dimensional classifiers	136
4.8.5	Characteristics of two-dimensional classification	137
4.8.6	Three-dimensional category	138
4.8.7	Source of three-dimensional classifiers	142
4.8.8	Characteristics of three-dimensional classification	143

4.9	Conclusion	144
-----	------------	-----

CHAPTER FIVE: FUNCTIONAL ASPECTS

5.0	Introduction	147
5.1	Functions of Vietnamese classifiers	148
5.2	Classification, individualization, referentialization and relationalization	149
5.3	Individualization and counting	152
5.4	Nominal phrase and specificity	155
	5.4.1 Deictic expressions	161
5.5	Anaphoric function of classifier constructions	163
5.6	Lexical Function of classifiers	164
5.7	Possessive constructions	166
5.8	Conclusion	167

CHAPTER SIX: VIETNAMESE CLASSIFIERS IN AREAL PERSPECTIVE AND CONCLUSION

6.0	Introduction	168
6.1	Southeast Asia as a linguistic area	169
6.2	Word order	170
6.3	Classifier systems – Constituent order	171
6.4	Classifier systems – Semantics	174
	6.4.1 Human classification	174
	6.4.2 One-dimensional classification	177

6.4.3	Two-dimensional classification	180
6.4.4	Three-dimensional classification	182
6.5	Conclusion	183
REFERENCES		192

LIST OF TABLES

Table 1.0:	Criteria for Numeral Classifiers, Concordial Classifiers, Predicate Classifiers, and Intra-locative Classifiers	11
Table 1.1:	Synopsis of Semantic Properties	15
Table 2.0:	Solutions to Genetic and Areal Bias	36
Table 2.1:	Parameters for the Typology of Classifiers	42
Table 3.0:	Differences between Mensural and Sortal Classifiers	57
Table 3.1:	Inventory of Mensural Classifiers in Vietnamese	64
Table 3.2:	Inventory of Sortal Classifiers in Vietnamese	68
Table 3.3:	Numeral and Demonstrative Phrases	75
Table 3.4:	Unclassified Nouns in Vietnamese	76
Table A:	Vietnamese Classifiers for Animates and Some Inanimates	112
Table 4.0:	Human Classifiers	113
Table 4.1:	Kinship Classifiers in Vietnamese	124
Table B:	Vietnamese Classifiers for One-Dimensional Category	129
Table C:	Vietnamese Classifiers for Two-Dimensional Category	133
Table D:	Vietnamese Classifiers for Three-Dimensional Category	140
Table 6.0:	Word Order Typology of Southeast Asian Languages	171
Table 6.1:	Num-CL-N Classifier Constructions in Southeast Asia	172
Table 6.2:	N-Num-CL Classifier Constructions in Southeast Asia	173
Table 6.3:	Palaungic Classifiers for Deities and Humans	175
Table 6.4:	Central Bahnaric Classifiers for Human and Non-human Animates	176

Table 6.5:	Classes for One-Dimensional Objects in Khmu Yuan and Southern Khmu	178
Table 6.6:	Classes for Two-Dimensional Objects in Khmu Yuan and Southern Khmu	180
Table 6.7:	Classes for Three-Dimensional Objects in Khmu Yuan and Southern Khmu	182

LIST OF ABBREVIATIONS

1	first person
2	second person
3	third person
A	adjective
ATTRIB	attribute
CL	classifier
COMP	comparative
COP	copular
D	dimensional
DEM	demonstrative
DER	derivational
DERG	derogatory
DU	dual
EXT	existential
FEM	feminine
GEN	genitive
HAB	habitual
HM	Hmong-Mien
INT	intensifier
MASC	masculine
MK	Mon-Khmer
N	noun
NEG	negator
NP	noun phrase
O	object
PERF	perfective
PL	plural
POSS	possessive
PREP	preposition
PRES	present
PROG	progressive
REL	relativizer
RESP	respected
S	subject
SG	singular
SUFF	suffix
TB	Tibeto-Burman
Q/Quan	quantifier
V	verb
VCL	verbal classifier
VN	Vietnamese

ABSTRACT

This thesis describes Vietnamese classifiers within the integrated typological framework of classifiers proposed by Aikhenvald (2000). Her work presents a functional-typological and empirically based account of noun categorization devices across languages of the world. However, up to the present, no in-depth study of the classifiers of a single language has been based on such framework; moreover, especially in the context of Southeast Asian languages, many classifier systems are still only superficially understood. Since Aikhenvald's work promises to serve as a guide for analytical work on languages and their noun categorization systems, this study adopts such framework to clarify and consolidate our knowledge about Vietnamese classifiers; in the process, the validity of the framework – against a language that is not covered in much detail by Aikhenvald – can be tested.

In this thesis, we look at the fundamental parameters for the typology of classifiers in the classifier system of Vietnamese. These include in particular the semantic organization of the system, the morphosyntactic locus of coding, the scope of categorization, the principles of choice of classifiers as well as the interaction with other grammatical categories. These topics are covered under the main chapters of the syntax, semantics, and functional analyses of Vietnamese classifiers. I also look at some properties of classifiers from an areal perspective as well as that of genetic relatedness. Finally, I evaluate Aikhenvald's framework to accommodate the Vietnamese data and suggest possible refinements.

CHAPTER ONE

LITERATURE REVIEW

1.0 Introduction

Classifiers can be defined as morphemes with the semantic function of noun categorization, i.e. they denote certain salient inherent qualities of the noun they classify.

In Allan (1977: 285), classifiers are defined based on two criteria:

- I. They occur as morphemes in surface structure under specifiable conditions.
- II. They have meaning, in the sense that a classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer).

The first criterion looks at the syntax of classifiers while the second is concerned with the semantic function of classifiers. If we just rely on the above criteria, we find that we have not arrived at a proper or complete characterization of classifiers. Goral (1978: 1) observes that "...classifiers are usually not clearly defined. Instead, a few examples are given, the grammatical category is assumed to exist, and various lexical forms are included or excluded from the set of classifiers of a given language..." This is true to a certain extent as most of the earlier studies of classifiers like Jones (1970), Greenberg (1972), and Denny (1976) only provide a vague definition of what classifiers are. It is not an easy task trying to define what classifiers are. This is probably due to these

morphemes being used in different contexts and occurring with quantifiers, demonstratives, qualifying adjectives, or nouns in the absence of any modifiers (Greenberg 1972). Classifiers also possess different functions depending on the situation and context of usage. Adams & Conklin (1973) point out that classifiers can function as nominal substitutes, nominalizers of words in other form classes, markers of definiteness, relativizers, markers of possession, and as vocatives. Even though classifiers have such a wide range of surface realizations and functions, underlyingly they fulfill the same semantic task of categorizing nouns. According to Dixon (1986: 108), classifiers provide the means for “categorization of an object in terms of relevant parameters¹ of world-view.” Hence, the term ‘classifier’ refers to all types of noun categorization devices regardless of their form or function and this would be the working definition of ‘classifier’ in my study.

Classifiers are usually introduced in the context of numeral phrases and are especially common in East and Southeast Asian languages where their presence is taken as an areal linguistic feature (Jones 1970, Bisang 1999). Such classifiers are known as ‘Numeral Classifiers’ because they appear next to a numeral or quantifier² in expressions of quantity and most of the time, they are obligatory³. They may characterize the referent of a noun in terms of its animacy, shape, size, function or use, and other inherent properties. Similarly, we can distinguish different types of classifiers based on their syntactic environment and other properties such as their semantics, conditions for use, kinds of

¹ We will go through these parameters in Section 2.5.

² Quantifiers are words like ‘few’, ‘many’, ‘a lot’, ‘some’, ‘every’, etc.

³ Section 3.5.5 shows us that in Vietnamese, some nouns are directly enumerable without a classifier.

origin, etc. Besides numeral classifiers, another type of classifier that has undergone much linguistic scrutiny is 'Noun Classes' or 'Genders'. Noun class or gender systems are well-attested in Indo-European languages and they are also a typological feature widespread in languages of Africa (Aikhenvald 2000: 77) as well as Northern Australia and New Guinea. Noun classes are grammaticalized agreement systems which correlate with core semantic characteristics like animacy, sex, or humanness. Besides these two types, Aikhenvald (2000) lists a whole spectrum of noun categorization devices documented in various languages of the world and includes:

- A. 'Noun Classifiers' which characterize the noun by itself and co-occur with it in a noun phrase.
- B. 'Possessed Classifiers' which characterize a possessed noun in a possessive construction.
- C. 'Relational Classifiers' which characterize the way in which the referent of a possessed noun relates to that of the possessor in a possessive construction.
- D. 'Verbal Classifiers' which occur on the verb but they categorize a noun, which typically functions as an intransitive subject or a direct object, in terms of its shape, consistency, size, structure, position, and animacy.
- E. 'Locative Classifiers' which occur on locative adpositions.

F. ‘Deictic Classifiers’ which are associated with deictics and articles.

Adopting our earlier working definition of classifiers, we should not be confused and treat ‘classifiers’ just as the abbreviation for ‘numeral classifiers’ (cf. Dixon 1986, Lobel 2000). Similarly, one must be aware that some linguists use ‘noun classifiers’ to refer to ‘classifiers’ in general (cf. Denny 1976) unlike Aikhenvald’s definition of noun classifiers.

Classifiers have long been a topic of interest in typological studies. About two decades ago, the systematic typological study of classifiers began with Greenberg’s pioneering study, focusing on numeral classifiers and substantival number (1972). His paper characterizes numeral classifier languages in terms of the existence of numeral classifier constructions and embarks on a rigorous investigation of the occurrence/non-occurrence of classifiers in the context of counting. It also notes that classifiers can appear in other contexts with demonstratives, adjectives, or by itself. Some of the synchronic generalizations he made pertaining to the typology of classifiers are:

A) Only four patterns of word order are found for classifier constructions, which contain the three elements, quantifier⁴ (Q), classifier (CL), and noun (N). They are [Q-CL-N], [N-Q-CL], [CL-Q-N] and [N-CL-Q]. The two word orders where the quantifier is separated from the classifier i.e. [CL-N-Q] and [Q-N-CL] do not occur.

⁴ Quantifiers here refer to numerals, the numerical interrogative ‘how many’, and indefinite quantifiers such as ‘few’ and ‘many’.

B) A common pattern is that the relative order of the quantifier and the classifier do not change, but these elements may occur before or after the noun within a language i.e. we can have either [Q-CL]-N or N-[Q-CL], or [CL-Q]-N or N-[CL-Q]. Comparatively, the variation between [Q-CL] and [CL-Q] occurs less frequently and this variation is exemplified by three languages. In Bodo (Sino-Tibetan), the indigenous form is [CL-Q]-N while the borrowed form from Assamese (Indo-Aryan, Indo-European) is [Q-CL]-N. In Bengali (Indic), the usual order [Q-CL] may be reversed to express numerical approximation (Chatterji 1926). In most Thai languages, the [Q-CL] order generally holds but the order with the numeral 'one' is [CL-Q].

The classificatory phenomenon of possessive classificational systems is also mentioned in Greenberg (1972). It is noted that in many Oceanic and Amerind languages, the contrast between alienable and inalienable possession is elaborated through the split of the latter into classes based on the use of classifiers which are attached to the possessive affixes in place of the possessed noun. Other important hypotheses include the parallelism between classifiable nouns and collectives in their semantic non-specification of number and aversion to a direct number construction i.e. [_{NP} Q-N]/[_{NP} N-Q] or [_{NP} Num-N]/[_{NP} N-Num], and the correlation between classifiers and demonstratives such that most classifiers in quantifier constructions can also be used in deictic constructions.

Since Greenberg (1972), there has been an increase in the linguistic investigation and treatment of the classifier systems of various languages. Studies of classifiers can be grouped into three types: (1) studies that create a general typological picture, (2) studies

of individual types of classifiers and, (3) studies of specific languages. Linguists like Greenberg and Allan have provided typological generalizations and classification schemes of noun categorization devices by surveying a large set of languages. Allan (1977) examines more than fifty classifier languages around the world and Greenberg (1972) more than a hundred. Besides taking on large databases of classifier languages to rule out the problems of bias in sampling, and to make universal generalizations, there have been efforts to study classifiers from an areal typology perspective where the presence of a certain type of classifier system is taken as a typological characteristic of a linguistic area and/or of genetic relationship. For instance, Jones (1970) compares and contrasts classifier constructions in Southeast Asia and includes a wide variety of Southeast Asian languages like Burmese (Tibeto-Burman), Thai, and Vietnamese as well as Chinese, Malay, and Javanese. Adams (1989) takes on the task of describing the less well-described numeral classifier systems of Southeast Asian languages namely the Austroasiatic languages like Bahnaric (Mon-Khmer), Khasi, Khmu and Nicobarese. The above studies constitute the first type of typological study of classifiers.

Examples of the second type include Dixon (1982) and Sands (1995) where each provides detailed discussions on noun classes in Australian languages. Studies of verbal classifiers in South American Indian languages are undertaken in Derbyshire and Payne (1990); Papuan languages are considered by Lang (1975), and Brown (1981). Numeral classifiers are discussed in Greenberg (1972), Allan (1977), and Goral (1978).

The contribution of studies of classifier systems in specific languages cannot be underestimated because it feeds into forming an overall typological picture and provides

a good source of secondary data for further analytical work. This is the third type of study. Downing (1996) provides a detailed description and an in-depth discussion of the numeral classifier system in Japanese. Daley (1998) concentrates on the use of classifiers and their functions in Vietnamese narrative texts. Other language-specific accounts can be found in Nguyễn Đ. H. (1957) which looks at Vietnamese, Haas (1978a), and Beckwith (1993) which looks at Thai, Bisang (1993) which looks at Hmong (Miao-Yao), and Aikhenvald (1994) which looks at Tariana (North Arawak, Arawak).

1.1 Historical development of studies of classifier systems

According to Dixon (1982: 159), the category of gender was recognized very early in western linguistic scholarship, around the fifth century B.C., by Protagoras who divided Greek nouns into ‘feminine’, ‘masculine’, and ‘inanimate’ by correlating the orthographic endings of words with the nature of the things they represent. Subsequently, the notion of gender was extended to languages outside the Indo-European family when linguists found examples of categories in African languages which were grammatically similar to gender in the Indo-European languages. However, these gender-like systems do not have the same semantic correlation with sex and animacy as the Indo-European languages as they often do not have a masculine/feminine distinction. In addition, they are often rather large classes with up to eight distinctions or more in contrast to gender systems in Indo-European languages. According to Creider (1975), nouns in Bantu languages are classified with regards to shape i.e. whether they are extended, curved or non-extended. They are also classified according to whether they are humans, animals or things, and whether they refer to something sticky, lumpy, intangible, or in the liquid

state. Therefore, the term ‘gender’ was thought to be unsuitable for many other languages where nouns had semantic correlation with other categories besides gender, and in some cases, no semantic correlation with gender at all. Hence, the term ‘noun class’ came to be used as an umbrella term to characterize “grammaticalized agreement systems which correlate—at least in part—with certain semantic characteristics (particularly in the domain of human and animate referents)” (Aikhenvald 2000: 19). Gender is then regarded as a subset within the larger set of noun class.

When linguists first look at Southeast Asian languages, numerical classification is a construction that proves to be puzzling and becomes an area of interest to explore further. The typical lack of a singular/plural distinction in the noun phrases and the necessity of categorizing objects into groups for counting strike speakers of European languages as an unusual feature and often a frustrating semantic problem (Adams 1989). Hence, numeral classifiers, especially in Southeast Asian languages became the next noun categorization device that received much attention and prominence, besides noun classes. This can be seen from the proliferation of works on numeral classifier systems in East and Southeast Asian languages (cf. Jones 1970, Greenberg 1972, Allan 1977, Goral 1978, Haas 1978, Erbaugh 1986, Adams 1986; 1989, Downing 1996, Bisang 1993; 1999, etc).

1.2 Establishing typologies of classifier systems

Aikhenvald (2000)’s definition suggests that in order to establish the defining criteria for noun classes, we have to start with an analysis of the formal and functional features of this system. Even though noun classes have a variety of morphological and syntactic

realizations, as well as specific semantic bases, it is still possible to draw a general criterion of membership based on a thorough examination of these features in classifier languages. In fact, this form of linguistic investigation has led other linguists to come up with a typology of various classificatory phenomena like numeral classifiers, noun classifiers, possessed classifiers, verbal classifiers, etc. These noun categorization devices have already been briefly introduced in Section 1.0, which we will not repeat here. In (1), we have a numeral classifier construction in Vietnamese:

- (1) *ba con chó*
 three CL dog
 ‘three dogs’

In (2), we have a noun classifier construction in Yidiny (Australian) (Dixon 1977):

- (2) *buri birmar*
 CL:FIRE charcoal
 ‘(hot) charcoal’

Example (3) shows a possessed classifier construction in Tariana (Aikhenvald 2000: 2):

- (3) *tʃinu nu-ite*
 dog 1SG-CL:ANIMATE
 ‘my dog’

Example (4) shows a verbal classifier construction in Waris (Papuan) (Brown 1981: 96)

- (4) *sa ka-m put-ra-ho-o*
 coconut 1SG-to VCL:ROUND-GET-BENEFAC-T-IMPERATIVE
 ‘Give me a coconut’ (lit. ‘coconut to-me round.one-give’)

Although the criterion of size is often mentioned, the syntactic distribution, morphology, semantics, and functions of classifier systems have proven to be crucial in the identification of classifier systems and in shedding light on how people categorize the world through their language.

In the sections below, we will first look at some of the prominent works on classifiers that have helped to increase our knowledge and understanding of classifier systems around the world before focusing on Aikhenvald (2000), which is an in-depth study of classifier systems around the world and it is used to clarify and consolidate my knowledge on Vietnamese classifiers.

1.2.1 Dixon (1986): Noun classes and noun classifiers

Dixon (1982; 1986) distinguishes between the grammatical category of noun classes and the lexico-syntactic phenomenon of noun classification (including numeral classifiers) on morphological and syntactic grounds, recognizing the fact that noun classes always have some degree of semantic correspondence as well. He also states a correlation between language type e.g. isolating, agglutinative, and inflectional with the type of noun categorization system. Noun class systems are typically found in languages with a fusional or agglutinating profile whereas numeral classifiers are often found in languages with an isolating profile.

1.2.2 Allan (1977): Semantic basis of classifier systems

In Allan (1977), the defining criteria and characteristics for four types of classifiers: (1) numeral classifiers, (2) concordial classifiers (or noun classes), (3) predicate classifiers (or verbal classifiers), and (4) intra-locative classifiers (or locative classifiers) are given. Allan's criteria are summarized in the table below:

Table 1.0
Criteria for Numeral Classifiers, Concordial Classifiers, Predicate Classifiers, and Intra-locative Classifiers

Classifier Type	Criteria
1. Numeral Classifier	Classifier is obligatory in expressions of quantity and can occur in anaphoric or deictic expressions.
2. Concordial Classifier	Occur in languages in which classifying formatives are affixed to nouns, plus their modifiers, predicates, and proforms.
3. Predicate Classifier	Classifier that occurs on the verb which varies according to certain discernible characteristics of the 'objects conceived as participating in an event whether as actor or goal.'
4. Intra-locative Classifier	Classifier is embedded in some of the locative expressions which obligatorily accompany nouns in most environments.

Allan also asserts that the use of classifiers is often governed by semantics. Other works that focus on the semantic bases of classifiers include Adams (1989), Becker (1975), Denny (1976), and Croft (1994).

1.2.3 Adams (1989): Semantic dimensions underlying classifier systems

Adams (1989) concentrates on the detailed description of numerical classification in Austroasiatic languages which are “not generally well-described” (Adams 1989: 1). She also addresses the issue of the borrowing of classifiers in the Mon-Khmer, Nicobarese, and Aslian families by identifying the source of classifiers, either native or borrowed. In order to characterize native versus non-native aspects of the Austroasiatic classification system, Adams gives a detailed account and analysis of the semantic bases underlying numeral classifiers in the languages in these families. Adams shows recurring similarities in the types of classes that appear in different numeral classifier languages including Austroasiatic languages, for example, animacy versus inanimacy. Within the inanimate group, the classification of objects can be further distinguished by their shape, which essentially refers to the dimensionality of the object. Hence, objects are divided and classified accordingly to whether they are perceived as one-dimensional, two-dimensional, or three-dimensional.

1.2.4 Becker (1975): Spatial metaphor in the Burmese classifier system

Becker (1975) explores the Burmese classificatory system by subscribing to the notion that classifiers are part of a system for organizing experience whereby certain semantic polarities recur. According to Becker, all classifiers in Burmese have deictic implications as all things range out in relation to a conceptual center. For animate beings, the conceptual center is Buddhahood hence people (and holy things) closest to Buddhahood are classified by *pa* meaning ‘close’. For other animate beings, status is the defining

criterion for which classifier to use. In the case of inanimate things, the center is self thus objects are ranged at four distances from the self: part of self, on self, nearby self, and far from self. Furthermore, the structure underlying classification of inanimate things divides the self into head and body, associating them either with the head (metaphorically top, round) or with the body (metaphorically bottom, straight).

1.2.5 Denny (1976): Semantic functions of classifiers and interaction with the world

Denny (1976) seeks to describe the specialized semantic and syntactic properties of classifiers and to demonstrate that the semantic function of classifiers is to express three kinds of human interactions with objects and the world. The three kinds of human interactions are: (1) physical interaction such as handling, (2) functional interaction such as using an object as a vehicle, and (3) social interaction such as interacting appropriately with a human compared to an animal, or a high status person compared to a low status one. According to Denny, the advantage of a semantically-based theory of classifiers is that it enables the essential commonalities of classifiers to be perceived across the variety of classifier systems in different languages. This view is further developed in Croft (1994) which we will focus on in the section below.

1.3 Croft (1994): Semantic universals in classifier systems

Croft argues that if one distinguishes different types of classifier systems according to their grammatical and semantic-pragmatic function then it is possible to make regular

cross-linguistic generalizations about these systems. He claims that there are “substantive universal implicational hierarchies governing the semantic distinctions found among particular types of classifiers” (Croft 1994: 145). This means that there is a different hierarchy of semantic distinctions associated with each type of classifier system and these distinctions can be accounted for by the semantic/pragmatic function of the construction in which they are found. The different types of classifier systems and their functions are listed below (Croft 1994: 147):

<i>Classifier Type</i>	<i>Semantic/Pragmatic Function</i>
Noun Class	Determination (Reference)
Numeral Classifier	Enumeration
Possessive Classifier	Possession
Predicate Classifier	Spatial Predication

For example, noun class systems perform the function of tracking and distinguishing different referents in discourse so they utilize the higher levels of the taxonomic hierarchy, of which the most important semantic properties are based on animacy, and humanness. Possessive classifiers are used in indicating “the ownership of an object by a person” and “the degree of control over how that object is to be used” (Croft 1994: 164). Hence, given the types of physical objects that a person owns, the most salient distinction is between those that are edible (food) and inedible (implement). Croft claims that the primary reasons for this are because food is fuel for the body while implements are not, food is consumed and not used again unlike implements, and food is ingested and is internal to human body whereas implements are external.

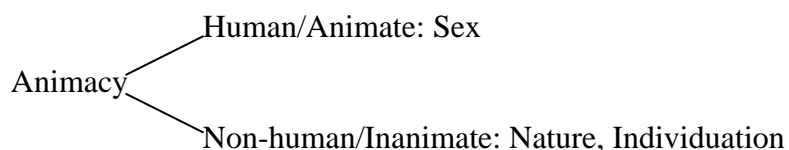
The semantic properties in Croft’s taxonomy are listed and briefly introduced in the table below (Frawley 1992). A more explicit discussion of these properties is presented in Chapter Four.

Table 1.1
Synopsis of Semantic Properties

Semantic Properties	Definition
1. Animacy (Animate)	The influence that the entity has over the execution or instantiation of an event. Criteria: life and locomotion, topicality, potency, cultural importance, discourse salience, etc.
(a) Animate: Human vs. Non-human	This category is self-explanatory. In denoting the animacy of entities, languages differentiate generally between animates and inanimates, and more specifically among humans, animals, and inanimate objects.
(b) Animate: Sex	Animate entities are characterized by biological sex i.e. male or female. Many languages restrict the male/female distinction to humans but some languages e.g. German extends this distinction to animals as well.
(c) Animate > Human: Kin	Familial relations among humans.
(d) Animate > Human: Social Status	The encoding of the non-familial, social relations of entities.
(e) Animate: Posture	Animate beings are divided according to whether they are standing, sitting or lying. Sometimes, this property also applies to inanimates which can stand upright (extended vertically), lie (extended horizontally) or ‘sit’ (non-extended).
2. Inanimate: Function	Specific uses that entities have or the kinds of actions that are performed on them.
(a) Inanimate: Nature	Natural classes such as trees, plants, birds, insects, etc. versus artifacts.
(b) Inanimate: Shape/Dimensionality	Three values for dimensionality: one-dimensional (long), two-dimensional (flat), and three-dimensional (round).

Semantic Properties	Definition
(c) Inanimate > Shape: Orientation/Direction	In addition to shape, some languages make distinctions based on the orientation of an extended object that is if it is vertically or horizontally extended.
(d) Inanimate > Shape: Rigidity/Consistency	Some languages distinguish objects according to how rigid or flexible they are.
(e) Inanimate > Function: Edibility	This category is self-explanatory. Edibles are secondarily divided into edible vs. potable or the state of food which can be raw, wild, cooked, etc. For non-edibles, the distinction is between valued vs. non-valued possession. Further distinctions are made on the possessed item's function and purpose.

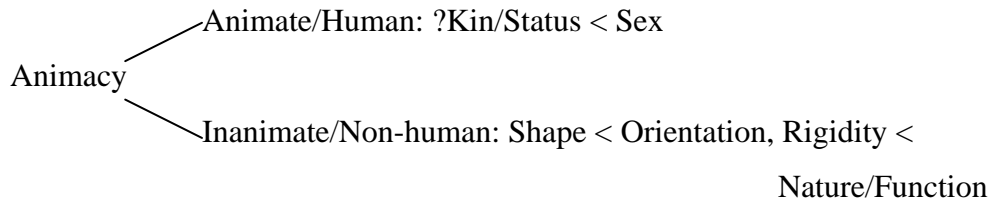
With the above semantic properties, we can look at Croft's hierarchy of semantic distinctions. For noun class systems, the human/non-human and male/female distinction, as well as the animate/inanimate distinction is used. In the human/animate class, gender is the primary distinction and in the non-human/inanimate class, nature and individuation are the primary distinctions as seen below:



Adapted from Croft (1994: 148)

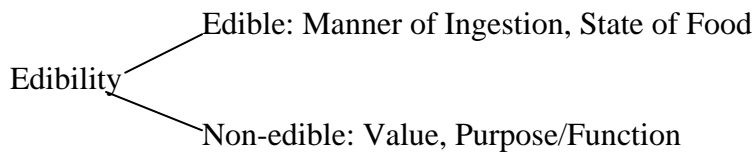
For numeral classifiers, the animate/inanimate and human/non-human distinctions are used. Further distinctions are based primarily on social status, including kinship relations (Adams and Conklin 1973) in the animate/human class. In the inanimate/non-human class, the primary distinction is shape which represents dimensionality (Adams 1989).

When talking about dimensionality, some languages also differentiate objects by orientation and rigidity. Secondary to shape, the distinction made is of nature or function.



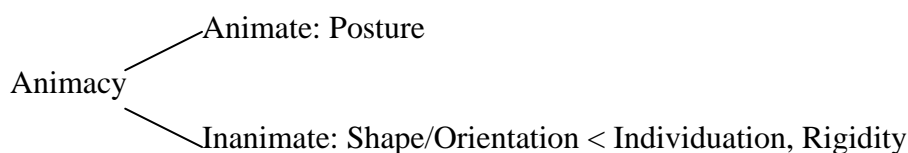
Adapted from Croft (1994: 152)

For possessive classifiers, the primary distinction is between edible and other possessed items. The secondary distinction between edibles is based on manner of ingestion and state of food. For non-edibles, the distinction is based on value, and purpose or function.



Adapted from Croft (1994: 155)

For predicate classifiers, the primary distinction is animate/inanimate. Secondary to the animate class is posture and to the inanimate class is shape or orientation. The inanimate class is further distinguished by individuation and rigidity.



Adapted from Croft (1994: 157)

1.4 Other studies on the extended functions of classifiers

A more thorough account of the functional variation among language specific classifier systems in the East and Southeast Asian languages is taken up in Bisang (1999). In addition to the basic functions of classification and individualization, Bisang claims that numeral classifier systems in some languages can expand their functional range to include relationalization or referentialization. This is made possible by the common cognitive basis which classifiers share. The range in variation of functions stems from two diachronic processes of development in the system, which he calls (1) *item oriented process* and (2) *category oriented process*. Item orientated process of development refers to “the context of counting individual items which are of particular cultural importance. The result is a new construction containing a new and clearly defined position for the classifier. The construction applied to the relatively few nouns successively spreads over to a wider range of nouns until all the count nouns of a given language become part of a classifier system” (Bisang 1999: 158-9). This system is predominant in Chinese and Japanese. Category orientated process development refers to “a categorial system already existing in the language. This system, which is part of word formation, is based on taxonomy and meronymy. In both cases, we have two positions filled by nouns. In a process of reanalysis, one of these positions i.e. the first position in the language described here, is reanalysed as a classifier” (Bisang 1999: 159). Usually, class nouns form the first step on the pathway of grammaticalization from noun to classifier and this system is predominant in Vietnamese, Thai, and Hmong. Hence, a typology based on the combination of the different functions of classifiers can be elucidated and this offers another perspective from which classifiers can be analyzed.

1.5 Aikhenvald (2000): Classifiers, a typology of noun categorization devices

Even though a large corpus of literature on different features of classifier systems exists and typological studies have been carried out by numerous linguists, Craig (1992) and Aikhenvald (2000) both point out that there is an urgent need to establish a comprehensive typology of classifiers. This is motivated by a number of factors. First and foremost, a large amount of new data on classifier systems has been generated recently. Besides the fact that these data need to be organized in a systematic way, their existence creates the opportunity of providing a typology of reasonable scope and validity.

Secondly, there exists a pervasive terminological confusion in the literature as there is a lack of an “overarching unified analysis of classifier systems” (Aikhenvald 2000: 1) in the languages of the world. This creates a lot of difficulty for the cross-linguistic comparison of noun categorization devices as well as the analysis of new data. The current literature is somewhat confusing with regards to the terms used to describe classifier systems. For instance, Denny (1976) uses ‘noun classifiers’ to describe noun categorization devices which can be misleading because Craig (1992: 284) has reserved the use of ‘noun classifiers’ for “a distinct type of classifiers that stand alone with their referent noun, independent of quantification or possession”. Both Allan (1977) and Kiyomi (1992) have used ‘concordial classifiers’ to refer to noun classes, ‘predicate classifiers’ to refer to verbal classifiers, and ‘intra-locative classifiers’ to refer to locative classifiers. This profusion of terminology creates a great amount of confusion for researchers and readers alike.

Due to the above factors, Aikhenvald has constructed “a new, integrated typological framework” to provide a comprehensive and unified approach at our present stage of knowledge about the structure and mechanisms of human languages and human cognition. It is also intended as a guide for analytical work on previously undescribed languages and their mechanisms for noun categorization. Besides presenting us with a solution to the above problems, Aikhenvald’s work is also significant because her study is based on examination of the grammars of about 500 languages representing each major language family and each linguistic area across the globe. She has looked at every language on which she could find data and which has noun categorization devices. This approach allows the typology to be as comprehensive as possible given our present knowledge. For example, recent work on South American and Papuan languages has shown that classifiers and genders **DO** co-occur and that it is possible for languages to have two distinct systems of classifiers (See Aikhenvald 2000 for more details).

1.6 Aim of the thesis and significance

In this study, I am going to describe and analyze the Vietnamese classifier system and compare it to Aikhenvald (2000), as a ‘test case’ of whether her typology and approach are robust enough to work on actual data.

There are several motivations for this research. First, we already know that the study of classifier systems in specific languages feeds into forming an overall typological picture which provides a good source of secondary data for further analytical work. Vietnamese is a Southeast Asian language that has not been explored substantially and there are not

many references available in English. Besides Emeneau (1951), Thompson (1965), and Nguyễn Đ. H. (1997), there are very few works that investigate the different phenomena in the language. Besides the lack of data, most linguists had to rely too heavily on material which are outdated and not reviewed in light of the recent theories or new data that have become available. By looking at the classifier system of Vietnamese, we are providing a possible updated source of secondary data for analytical work in a related field or other areas of grammar in Vietnamese.

In addition, the previous descriptions of the classifier system of Vietnamese are also inadequate and a holistic typological treatment has not been carried out in Vietnamese. The classifier system in Vietnamese was first formally explored more than four decades ago in Nguyễn Đ. H. (1957) even though there were detailed descriptions of it in Emeneau (1951) and Thompson (1965). Other than that, Vietnamese classifiers are analyzed mostly along with the classifier systems of other Southeast Asian languages (cf. Jones 1970, Allan 1977, Bisang 1999) or languages in the Austroasiatic family (cf. Adams 1989). Although Daley (1998) uses Vietnamese classifiers as her data, she only focuses on the use of classifiers in narrative texts and the various discourse functions that Vietnamese classifiers serve. In other words, the noun categorization system in Vietnamese has not been properly characterized nor sufficiently analyzed in the recent decades. Bearing this in mind and given my interest and prior research in the classifier system of Vietnamese, I believe it is significant and meaningful to carry out this study.

Another reason for studying the Vietnamese classifier system is the fact that Vietnamese data are not covered in great detail or extensively in Aikhenvald. There are also differences between Aikhenvald's and my treatment of Vietnamese classifiers. I am treating the grammatical and semantics aspects of Vietnamese classifiers in depth whereas Aikhenvald only briefly mentions them. Hence, this thesis provides useful insights and new perspectives on Vietnamese as a refinement of Aikhenvald's framework.

1.7 Organization of thesis

In Chapter Two, we first examine Aikhenvald's framework and parameters for the typology of classifiers. Since Aikhenvald aims to present a "functional-typological and empirically based account of noun categorization devices across languages of the world" (2000: 4), I will devote some time to a discussion of what typological research is, with reference to the principles underlying it and what it encompasses. Some of the limitations (cf. Newmeyer 1998) of embarking on a typological research will also be addressed and possible solutions suggested. Subsequently, I will describe my methodology and means of data collection.

Chapter Three is a syntactic analysis of Vietnamese classifiers. First, I will look at the morphosyntactic realization of classifiers in Vietnamese. Then, I propose a current and inventory of Vietnamese classifiers and try to elicit 'true' classifiers from measure terms which are also found in non-classifier languages like English. Finally, we look at how

Vietnamese classifiers interact with other elements in the noun phrase and the consequences of those interactions.

Chapter Four starts off by establishing the list of semantic parameters that tends to be coded in different types of classifiers. This is followed by a close examination of the semantics of numeral classifiers in Vietnamese within the general theory of the semantics of numeral classifiers proposed by other researchers and existing in the literature. The Vietnamese styles of classification are presented and we look at the various categories in which items are organized and the properties that are considered salient in the organization of objects in the Vietnamese culture. We also establish a hierarchy along the lines of Croft's (1994) 'universal implicational hierarchy' for numeral classifiers.

Chapter Five concentrates on the functions of Vietnamese classifiers in the various constructions they occur in. The semantic notions of specificity and definiteness are discussed and applied to classifier constructions in Vietnamese. We also discuss polyfunctionality and how it applies to classifiers in Vietnamese.

In Chapter Six, we look at some properties of Vietnamese classifiers in areal perspective. These properties include word order, semantics, and functions of classifiers. Finally, we summarize our findings in the light of Aikhenvald's framework, which we also evaluate if we need to revise the framework in light of our analysis of Vietnamese classifiers. One of the things we evaluate is the possibility of treating Vietnamese as a multiple classifier language which is the idea proposed by Aikhenvald. We see that this is over theory-specific and it is not a sensible way of interpreting Vietnamese classifiers. The concept of

multiple classifiers is problematic in itself and it does not help in the analysis of classifiers in general.

1.8 Summary

At the start of the chapter, I went through what classifiers are and the prevalent treatment of noun categorization devices by linguists around the world. Subsequently, we embarked on the historical development of studies of classifier systems to show how different linguists look at classifiers at different points in time. This eventually led to the various typologies of classifier systems as more studies reveal that it is possible to come up with a proper characterization of various classificatory phenomena through different criteria such as morphosyntactic properties, semantics, pragmatic function, etc. I also gave a brief review with regards to the studies mentioned. After that, I went on to explain and highlight the aims of the thesis and the significance of undertaking this research. Lastly, I gave an outline of the subsequent chapters and how I have organized the thesis.

We now proceed to Chapter Two where we will examine Aikhenvald's framework in greater detail and also look at the theory and methodology employed in typological studies.

CHAPTER TWO

TYPOLOGICAL FRAMEWORK

2.0 Introduction

In this chapter, I present a brief account of typological research which aims at positioning the contribution of this dissertation within the field of typology. I have explicitly not adopted a formal semantic or syntactic framework for my description because I believe that a broadly typological approach will be most useful in exemplifying and investigating the Vietnamese classifier data. This is because most of the literature relevant to this thesis has not been written with a formal semantic or syntactic model, and so to adopt such a model would make my work less comparable with the existing literature; and also because there is ample documentation of the benefits of a typological approach. In the next few sections, I shall detail some of those benefits, and the style of discourse usual in typological studies.

As mentioned in Section 1.7, Aikhenvald's framework on noun categorization devices is established on functional-typological grounds empirically. By an empirically-based account, we mean that the description of patterns is based on observed language data where they can be verified or proven. In order to fully appreciate what Aikhenvald sets out to do, we have to understand what exactly is meant by typological research. Even though there are significant advantages in undertaking a typological study, there are

several issues that need to be addressed before typological generalizations can be valid and trusted (Newmeyer 1998). After looking at these fundamental issues associated with any typological research, I focus on my own data collection process and how my dissertation is integrated within the above enterprise.

2.1 Typology in a nutshell

In the context of linguistics, typology is:

The classification of languages or components of languages based on shared formal characteristics. (Whaley 1997: 7)

In a nutshell, typology engages in comparisons between languages and ultimately, it has the goal of identifying cross-linguistic patterns and correlations between these patterns. In any piece of typological research, three significant propositions are present (Whaley 1997):

1. The research must involve cross-linguistic comparison.
2. A typological approach must involve classification of either components of languages or languages.
3. The classification must be based on formal features of the language.

At this point, we need to explain what we mean by *formal features*. If we base a classification in terms of the formal features of a language, it means that one should

classify languages in terms of the *forms out of which a language is composed – its sounds, morphemes, syntax, or discourse structure* (Whaley 1997). These features are the chunks of information that one finds in a phrase or a sentence. Typology's focus on the formal features of language may pose a problem when we try to look at semantics and pragmatics as there is a general debate as to whether semantics constitutes a formal feature. According to Whaley, formal features are used to convey meaning, as a result, typologists have always been concerned with semantic categories such as gender or agent, and how these categories are manifested by the formal units of language. Therefore, "the emphasis on formal features in the definition of typology should not be taken to exclude semantic considerations" (Whaley 1997: 14). However, in Newmeyer (1998), the ambiguity involved with the definition of a formal feature appears rather problematic to the justification of a typological research. We will defer discussions until the later part of the chapter (see Section 2.4.5). Going back to the three propositions, we see that for Aikhenvald (2000), these three propositions form the base of her study because her study analyses and classifies about 500 languages by examining the morpho-syntactic properties of noun categorization devices. It also takes into account the semantic properties of classifiers, which prove to be crucial and pertinent to our understanding of how language is used by people in different parts of the world to categorize humans and objects.

2.2 Significance of typological research

One might question the advantages or even validity of doing a typological research over subscribing to a grammatical framework that models how language works, such as

Functional Grammar (Dik 1989) or Cognitive Grammar (Langacker 1987). Typology, like other theories of grammar, seeks to enhance our knowledge on the most fundamental question of what language is like. However, there are specific goals and advantages of typology which cannot be found in other theories or grammatical frameworks. One such goal is the discovery of homogeneity between remotely similar languages and explaining why it exists (Greenberg 1963). Despite the fact that an immense variety of languages are spoken, there are certain core properties that all languages have in common, known as **language universals**. It is widely accepted by linguists that the underlying unity of language is far more striking than differences present in languages and understanding why this is the case enables us to know about the nature of language. For example, we refer to oral stops. If we examine the distribution of stops in the world's languages, we can derive the **absolute universal** that all languages have at least one stop. The second observation is that *almost all* languages have the stops [p], [t] or [k] which is a **universal tendency**. The third observation is that if a language has two voiceless stops, then one is a [t] and this is an **implicational universal**. In the process of trying to explain these universals, we are trying to find out how sounds are incorporated into languages and this type of enquiry is not possible if we just rely on a theoretical framework that makes use of phonological theory to explain patterns of assimilation or alternation etc. in one language or in a variety of languages (cf. Whaley 1997: 8-10).

The other goal of typology is to classify either (a) components of languages or (b) languages, based on shared features. Part (b) is what Aikhenvald (2000) does and the area that is explored is noun categorization systems. From studies in this field, we know that different languages have different noun categorization devices which may not have

seemed so apparent about three or four decades ago. Also, we have a clearer picture of which languages use numeral classifiers and which languages use other types of classifiers. This form of investigation usually goes beyond just stating the correspondence between language and type of classifier system. Typically, linguists would try to look for patterns within that classification and some research questions that arise out of studies of this nature are:

- (i) Whether there is any correlation between language type and classifier system?
- (ii) Whether it is a mere coincidence that languages in some specific areas have a particular type of classifier system? And if not,
- (iii) To what extent does the classifier system of a language get affected by the languages around it and in what manner?

In the process of classifying languages, numerous research questions are raised and these questions are important in helping us know what language is. It is difficult to rely only on formal linguistic theories to arrive at these questions.

There are many specific advantages of a typological research when we look at different aspects of a language or many languages, and talk about the specific goals of typology. On top of that, typology has the advantage of being compatible with any grammatical theory because of the methodology and results. Typology is also extremely useful as a tool in linguistic fieldwork. Most linguists never have the opportunity to organize large

samples of languages to test for universals. However, all linguists do research on individual languages and there are many ways in which knowledge in typology can dispel confusion when researching on previously undescribed languages. First, it creates an awareness on the kind of phenomenon present and what is typical about it. As Whaley (1997: 15) aptly puts it, “Many phenomena that might seem quite exotic in comparison to one’s native language may actually be typologically common. Such phenomena are not likely to cause confusion or frustration for the field-worker familiar with typology”. A typical example is that numeral classifiers used to be regarded as an unusual form of noun categorization device for linguists who had been studying Indo-European languages. It is only upon typological investigations that linguists realize that numeral classifiers are commonly used in Southeast Asia and East Asia, and many parts of the world. If linguists know what is common in language, it will be easier to recognize unexpected patterns in the language or languages being examined and to search for explanations. This can reveal important aspects of the language’s history, contact with other languages, etc. Many significant discoveries about languages are made in this way. Finally, a knowledge of typology provides access to a major corpus of linguistic literature. A large amount of data is generated by linguists working on different languages and in the process of engaging in typological research, a linguist will need to look through all these data in order to establish the range of variation among languages and to identify patterns in the languages.

2.3 Issues in a functional-typological approach

Thus far, we have talked about language typology in the broad sense and provided some justification for engaging in a typological research. In this section, we go on to elaborate the functional-typological approach. As I have mentioned earlier, linguists with typological interests are not contented with merely stating descriptive generalizations governing the distribution of the linguistic elements in the languages of the world as they usually want to explain those generalizations.

We take a look at Greenberg's (1963) seminal paper. Working with a sample of 30 languages, Greenberg presented the implicational relationships between their word order and properties like adposition ordering, adjective-noun ordering, determiner-noun ordering, and numeral-noun ordering. He found 10 languages with an SVO word order, and these languages were uniformly prepositional. On the other hand, there were 11 languages with an SOV word order and these languages were all postpositional. Greenberg also explained the predominance⁵ of noun-adjective and subject-verb order by attributing it to "a general tendency for comment to follow topic" (1963: 100). Since language is used to communicate and communication involves the conveying of information, it has been argued in Lambrecht (1987) that the ordering of grammatical elements like 'subject', 'object' and 'verb' mirrors the sequential flow of information. According to an information-flow-based explanation, information moves from the more thematic to the less. Hence, subjects which are more likely to be thematic than objects

⁵ Dryer (1988) has dispelled Greenberg's generalization on the predominance of noun-adjective and subject-verb order.

(Lyons 1977) occur before the object and this explains why the SVO, SOV and VSO word order types are very prevalent in the world's languages. What the above explanations have in common is that both subscribe to the notion that cross-linguistic patterns can be explained in terms of linguistic function, especially communicative function.

According to Newmeyer, the above example is one of the “three interconnected types of typological patterns that functionalists have attempted to account for” (1998: 301). First, it questions why grammatical elements pattern in a particular way with respect to one another without any reference to meaning. The second type involves identifying a grammatical construct as being cross-linguistically common and attempting to account for its frequency in terms of its meaning and function. For example as described in Newmeyer (1998), syntactic categories like ‘noun’, ‘verb’, and ‘adjective’ are widely used in many languages and functionalists would try to explain that frequency on the basis of their expression of the functions of reference, predication, and modification respectively. The third type involves taking a linguistic function (e.g. comparison, focusing, modification) and studying the means by which it is coded cross-linguistically and attempting to explain why some types of coding is preferred over other types. For instance, the function of noun categorization in languages examined from a functional-typological approach would yield firstly, the various types of noun categorization devices used in different languages. From there, we can elicit the specific functions of each type of classifier by studying how it interacts with other grammatical elements in the language. Following that, we can try to find out why a particular language would use a numeral classifier instead of a noun classifier or why two types of classifiers co-exist in a

particular language or how similar/different is the same type of classifier coded in different languages in terms of syntax, semantics, communicative functions, etc. This third type of functional-typological account more or less gives us a general idea of what Aikhenvald aims to accomplish. Aikhenvald's work is ambitious in that it looks at the two-tier function of noun categorization devices. The first is the general function of noun classification, and the second is the characteristics and functions of each individual type of classifier and providing some form of explanation for the patterns observed.

2.4 Difficulties in typological research: Methodology

In the area of typology, specific methodological problems have to do with data sampling and bias. This area is most susceptible to criticisms because a proper database upon which language universals can be elicited is crucial to the validity of the whole investigation and when making any cross-linguistic claims. Moreover, language universals, which are statements of fundamental properties of language, form the crux of typology. Ideally, a typologist should examine all the languages in the world to determine language universals. However, it is technically impossible to create a database of all human languages because many languages are now extinct and there is little or no record of them. Also, not all languages have been documented. These constraints force us to choose only a sample of languages which we can compare for cross-linguistic patterns. This method raises three problems: (1) representativeness of available languages, (2) sample size, and (3) genetic and areal sources of bias. The other two problems have to do with the database of the sample of languages: (4) the source of data, and (5) defining what are formal features in a language.

2.4.1 Representativeness of available languages

In linguistic typology, we look at the features and components of available language samples and try to derive some form of implicational relationships among these features. However, how can we be sure that these features or even the languages available are of any significance to understanding what language is? In other words, can we explain the origin of language, why languages differ, how they are learned, how they relay meaning through language, etc. by discovering that the predominant word order in the languages of the world is SOV? We probably would not know if there is a real significance behind all the typological research that has been carried out but if we continue to carry out solid investigations and elicit valid generalizations and implicational relationships, we can be sure that the findings are of great value and crucial to our current understanding of human language.

2.4.2 The problem of sample size

With regards to sample size, we need to ask how large a sample would have to be for us to be sure that some rare but possibly important feature has not been left out. According to Bell (1978: 143), if one percent of languages is of a particular type, the likelihood that a random sample of 100 languages containing no instance of that language type is 0.36. This probability decreases with the increase in sample size. Hence, the rule of the thumb is to obtain as large a sample as possible. For instance, in Greenberg's (1963) investigation of basic word order, his sample size is 30 and his results show that SVO is the most common word order pattern. However, in a similar investigation carried out by

Tomlin (1986) on a sample of 402 languages, SOV emerged as the most common. In this example, we see that generalizations made on samples of convenience of fewer than 50 languages must be viewed critically (Tomlin 1986).

2.4.3 Genetic and areal bias

It cannot be assumed that by having any random sample of 500 or more languages, we can safely make our generalizations and draw our conclusions about how language works. Another problem that has to be taken into account is genetic and areal bias, that is neither particular language families nor particular regions of the world should be over-represented in the sample. Since genetically related languages tend to share more typological characteristics than unrelated languages, we have to avoid taking too many languages from the same family in our sample. For instance, if we take a sample of 100 languages and 50 languages are from the same language family that has a basic word order of SVO, then our conclusion may be skewed towards languages having an SVO word order because it is the predominant word order in our sample. As suspected, SVO word order is not the most common word order. Researches have shown that the predominant word order in the languages of the world is SOV instead of SVO. Hence, this illustrates the importance of correcting for genetic bias. Typological features are spread over different areas so we also have to make sure that no area is over- or under-represented in the sample. Many factors would give rise to potential areal bias in typological studies. One of them is language contact. Features of language can be easily borrowed through long periods of contact or political dominance hence languages may look increasingly similar. For example, Comrie suggests Amharic (Semitic) developed

SOV order from an earlier VO as a result of the influence of the neighboring Cushitic languages (1989: 208). Hence, language contact can and often does obscure the differences between various languages. The other issue is that the ‘linguistic area’⁶ created may be based on geopolitical forces rather than linguistic ones. This is because in contact situations, the features of the culturally-dominant language are more likely to be borrowed than those of the less dominant.

There are three ways suggested by various linguists to overcome genetic and areal biases. They are discussed in Whaley (1997: 38-43). We summarize them briefly in the table below:

Table 2.0
Solutions to Genetic and Areal Bias

Method	Strength	Weakness
1. Each language family is represented in the sample based on the number of languages in that family. Language families with a greater number of members would receive greater representation in the sample than smaller language families. (Bell 1978, Tomlin 1986)	- It provides a general idea as to the proportion of existing languages in the world that contain a particular linguistic trait.	- The actual frequency of different language families is due not only to purely linguistic reasons but also to historical factors. Hence, numbers may not be indicative of the actual preferences that languages have.

⁶ An area in which at least one linguistic property is shared more often than elsewhere in the world to an extent which is unlikely to be due to chance, but which is probably due either to contact or remote genetic relationships (Dryer 1989: 266).

Method	Strength	Weakness
2. Gather languages that bear only very distant or no genetic relationship and are not from the same culture area. Only 1 language will be selected from each language family for the sample. (Bybee 1985, Perkins 1980; 1989)	<ul style="list-style-type: none"> - It more accurately reflects purely linguistic preferences than the first method. - The required sample is manageable in size. 	<ul style="list-style-type: none"> - It may not be possible to construct a sample of 50 languages that are distinct enough in geographic terms because of the presence of linguistic areas.
3. First, group languages into genera. Then group the genera into 6 large geographical areas. To determine if a pattern is statistically significant, the pattern must be present in the genera of each of the 6 areas. (Dryer 1989; 1992)	<ul style="list-style-type: none"> - Controls for severe genetic bias by counting genera rather than individual languages. 	<ul style="list-style-type: none"> - One is forced to decide which genus a language belongs. It may be difficult since the genetic affiliation of many languages is controversial. - One must accumulate information on an extremely large number of languages.

After a discussion of the problems associated with sampling and bias, we have to look within each language in our sample to compare the feature(s) we want to explore. At this stage, we have to consider two problems with the database and they are elaborated in the next two sections.

2.4.4 The source of data

Most typological research requires typologists to gather information on languages that they do not know first-hand. The most common method is to use published materials. Even though this is the quickest and most convenient method that yields abundant data on many languages, there are several problems. First, some of these materials only offer a brief and incomplete account of the language and do not provide us with the crucial information that is needed (Stassen 1985). The second problem is that we cannot be

confident in the accuracy and usefulness of the secondary data. In the course of writing this dissertation, I have encountered many instances where native speakers of Vietnamese have protested against the grammaticality of the data presented in some journal articles. This is a serious problem because the analysis based on the data would be seriously flawed and the conclusions cannot be valid at all. The third problem is the terminology and its cross-linguistic applicability and consistency (Newmeyer 1998). As Newmeyer has noted:

Is one grammarian's 'particle' another grammarian's 'clitic'? Is one grammarian's 'conjunction' another grammarian's 'complementizer'? The typologist has to answer such questions as well as he or she can, even though often presented with limited means for deciding (an alternative, of course, is to cancel the relevant languages from the sample). (Newmeyer 1998: 327)

The other problem is the tendency for secondary sources to use other secondary sources as primary material and errors, if they occur in any one of the sources, are perpetually passed on (Mallinson and Blake 1981). The problems mentioned above are some deficiencies resulting from the use of secondary data but since it is such a productive and convenient way of eliciting data, we cannot dismiss it totally. The only way to minimize the damage is to be mindful of the dangers in the use of secondary sources and adopt a more critical attitude and not rely solely on secondary sources for one's database.

2.4.5 Definition of formal features

Earlier on we mentioned that the classification of languages in any typological research must be based on the formal features of the languages which means that a formal analysis of the language(s) involved is a prerequisite to any typological studies. This entails two tasks that the typologist has to do:

- I. Identify grammatical elements in a single language
- II. Identify grammatical elements cross-linguistically

Some typological research applies insufficient formal analysis in identifying the grammatical elements in a particular language. Often, there is no indication of why an element is identified as a particular lexical item whereas others are not. Even if some indication is given, it is based on loose semantic and pragmatic criteria. Newmeyer (1998) criticizes the appeal to semantic criteria in both language-particular and cross-linguistic identification of grammatical elements. He claims that it is a mistake to identify grammatical items based on the semantic roles that they play or on their discourse functions because this induces circularity and we cannot hope to understand the relationship between syntactic categories and meaning. Clearly, a sufficient formal specification is desirable in order for a functional explanation to elucidate more robust generalizations. However, it is obvious that Newmeyer's radical criticism is partly motivated by ideological stances rather than empirical preoccupations. In this study, both formal as well as functional dimensions of classifiers in Vietnamese will be defined and described (see Section 2.6).

Having covered the more important issues in typological research in the earlier sections, we are now equipped with the background knowledge to take a look at Aikhenvald's framework and methodology followed by my methodology and database.

2.5 Aikhenvald's framework and methodology

As we have mentioned earlier, Aikhenvald (2000) is an attempt to present a functional-typological, empirically based account of noun categorization devices across languages of the world. It involves taking the linguistic function i.e. noun categorization, and studying the means by which it is coded cross-linguistically and attempting to explain why some types of coding are found in some languages and not in others. At the same time, Aikhenvald proposes a typology of classifiers primarily based on the morphosyntactic loci of classifier morphemes. This means that she looks at the types of classifier morphemes which acquire surface realization in natural languages. Unlike many of the earlier works (e.g. Allan 1977, Croft 1994), which only focus on distinguishing the various types of classifier systems, Aikhenvald has chosen to view classifier types as focal points on continua of different parameters of a typology of noun categorization devices. This means that the properties of different classifier types will be shown to be gradient rather than categorial. This approach is insightful because in the various types of classifier systems fuzziness exists, categories have vague boundaries and are internally organized from central focal values, the prototype, to less focal instances (Frawley 1992). We can observe this fuzziness when there are instances of classifier systems that do not seem to fit nicely into any of the types, and in the case of multiple

classifier system⁷. Different prototypes of classifiers which display all the definitional and most of the contingent properties of a type are characterized by referring to the parameters below (Aikhenvald 2000: 15-16):

- A) Morphosyntactic locus of coding
- B) Scope, or domain of categorization
- C) Principles of choice, or ‘assignment’ of noun categorization devices
- D) Kind of surface realization
- E) Agreement
- F) Markedness relations
- G) Degree of grammaticalization and lexicalization

After having established the ‘focal’ points or ‘prototypes’ on the continuum of noun categorization devices, she shows the interaction of the above types with the dimensions listed below:

- H) Interaction with other grammatical categories
- I) Semantic organization of the system
- J) Evolution and decay
- K) Language acquisition and dissolution

A brief account of these parameters is given in the table below:

⁷ This is the case where the same or almost the same set of morphemes can be used in more than one classifier environment.

Table 2.1
Parameters for the Typology of Classifiers

Dimension	Explanation
A) Morphosyntactic locus of coding	This refers to the different morphosyntactic loci where a noun categorization device can be realized.
B) Scope, or domain of categorization	This refers to the elements or constituents that noun categorization devices can refer to within different phrasal categories.
C) Principles of choice, or assignment of noun categorization devices	The choice of a classifier may depend on some semantic properties of the referent of the noun it categorizes or on some other properties of a noun (e.g. morphological or phonological).
D) Kinds of surface realization	Some noun categorization devices are realized with an affix or a clitic, while others often appear as separate words.
E) Agreement	Some noun categorization devices, especially noun class and gender system, involve agreement and some do not.
F) Markedness relations	Some noun categorization devices have a functionally and/or formally unmarked term while others tend not to.
G) Degree of grammaticalization	Some noun categorization devices are highly grammaticalized closed sets while others tend to involve a lexical choice.
H) Interaction with other grammatical categories	Different types of noun categorization show different dependencies with other grammatical categories
I) Semantic organization of the system	Noun categorization devices differ in their hierarchy of semantic distinctions. They also differ in the organization of their system and the degree of their semantic transparency. Other differences involve the syntactic and pragmatic functions they perform, and they also differ in how they respond to socio-cultural influence.
J) Evolution and decay	Different types of noun categorization devices differ in their etymological sources and in the way they develop and also fall out of use.

Dimension	Explanation
K) Language acquisition and dissolution	Distinct noun categorization devices show fundamental differences in how they are acquired by children, and what processes they undergo under language dissolution.

Aikhenvald claims to have examined the grammars of about 500. Data on the following languages have come from her own fieldwork: Tariana, Baniwa, Warekena, Bare (North Arawak, Arawak), Tucano (East Tucano, Tucano), Piratapuya (East Tucano, Tucano), Paumarí (Arawá), and Manambu (Ndu, Sepik). Aikhenvald explains that she has not restricted herself to considering just some samples of the available set of languages. Rather she has looked at every language she could find data on (i.e. samples of convenience) because she does not want to impose artificial limitations on the validity of the data dictated by sampling strategies. Also, she chooses not give any statistical counts because numbers may be unreliable at this stage where we do not have enough knowledge about all the languages in this world. This is somewhat surprising because most typological researches are concerned with sampling and statistical techniques because of their significance in affecting the results. Hence, to a certain degree, we have to recognize that even though Aikhenvald's study is comprehensive in its scope, it only serves as a suggestive guide to the statistical distribution of classifier systems throughout the world.

2.6 My methodology and data collection

In this study, I look at the language structure of only one language, Vietnamese, within the broader framework developed by Aikhenvald. My work will address issues encountered in the Vietnamese classifier system in the area of syntax, semantics, and pragmatics. It is my aim to check the validity of Aikhenvald's parameters in eliciting a proper characterization of classifiers in Vietnamese. Also, the issue of whether it is possible to classify languages according to the type of classifier constructions present in the language will be discussed. Finally, we will discuss both areal and genetic implications based on the noun categorization constructions, semantic bases, and functions of classifier systems in various languages of Southeast Asia.

My data was collected from Vietnamese informants. The variety of Vietnamese⁸ in this dissertation is typical of the educated form of Vietnamese as spoken by teachers and tertiary students living in Ho Chi Minh City. There are several dialects of Vietnamese. The main differences between the various types of dialects lie in the phonological structure and to a lesser degree, the lexicon. According to Nguyễn Đ. H. (1997), most dialects form a continuum from north to south, and urban centers such as Hanoi, Huế and Ho Chi Minh City represent rather special dialects marked by the influence of educated speakers and of more frequent contacts with the other regions.

⁸ Vietnamese is the mother tongue and the home language of the ethnic majority, người Việt or người kinh of Vietnam. This group occupies mainly the delta lowlands of Vietnam.

I have gathered Vietnamese data using published sources, such as reference grammars and journal articles, by eliciting data from informants⁹ locally and through fieldwork in Vietnam. The first two methods established the foundation of my dissertation while the third enhanced the quality of my work. Published sources raise problematic issues in their discussion of classifiers and suggest solutions to those issues. However, the data in published sources cannot be trusted completely (see Section 2.4.4), and by checking the data with my informants, I could verify if the data are correct and also applicable in modern Vietnamese. I also conducted fieldwork for a duration of three weeks in Ho Chi Minh City: this was crucial not only in order to control quality and analysis of the data but also to strengthen my analysis. In Vietnam, several tertiary students (with and without linguistics background) and a professor specializing in linguistics from the University of Social Sciences and Humanities were my informants. All data are elicited verbally and I ran through most of my previous data and crucial data in other studies on Vietnamese with my informants.

2.7 Conclusion

In this chapter, I have looked at the various propositions encompassed within a typological research and specifically at the functional-typological approach. The advantage of carrying out a typological research enables us to discover homogeneity between languages and explain why it exists. Typology also helps us to classify constructions in languages as well as languages based on shared features. It is compatible

⁹ My main informant, Mr Thạch Ngọc Minh is from Ho Chi Minh City and he is teaching Vietnamese to tertiary students in Singapore.

with any grammatical theory and is an extremely useful tool in linguistic fieldwork. However, there are difficulties and limitations in a typological approach, which need to be addressed before we can trust the results of our research. These limitations have to do with data sampling, bias, and source of data. After a general introduction on typology, I elaborate on Aikhenvald's framework and methodology and my own methodology. I also talk introduce my data collection methods.

The next chapter is a characterization of the classifiers in Vietnamese based on their morphosyntactic environment. The problem of distinguishing 'true' or sortal classifiers from measure terms, and mensural classifiers will be looked at and I will talk about how they are treated in my study. Next, I provide an up-to-date list of classifiers used commonly by native speakers. The rest of the chapter will be devoted to a close examination of the various constructions that classifiers occur in, and the constituent order and relations in those constructions.

CHAPTER THREE

SYNTACTIC PATTERNS

3.0 Introduction

This chapter provides a detailed account of the syntactic properties of classifiers in Vietnamese. The first section begins with a brief examination of the structure and grammatical behavior of Vietnamese nouns, followed by examples of Vietnamese classifier constructions. Subsequently, we look at the differences between mensural and sortal classifiers. For the purpose of eliciting the differences between these two categories, I refer to Section 3.2 which provides an insightful study of the types of nouns and how they determine the use of mensural and sortal classifiers based on Dik's (1989) Functional Grammar (FG). However, we will not go into too much detail on FG because we are not using this framework as the basis of our analysis of classifiers i.e. it is not the core framework. It is useful in terms of explaining the differences between mensural and sortal classifiers. Subsequently, we include an up-to-date inventory of classifiers commonly used by native Vietnamese speakers. In the later sections, we will look at the various combinatory patterns Vietnamese classifiers take with regard to other modifiers like demonstratives and numerals in the noun phrase. This will be followed by an analysis of the constituent order and relations in classifier constructions. Specifically, we are looking at:

- (i) The morphosyntactic locus of coding
- (ii) The interaction with other grammatical categories. (see Table 2.1)

3.1 The Vietnamese language and classifier constructions

In Comrie's (1989) index of synthesis, which focuses on the number of morphemes that tend to occur per word, a continuum from isolating languages at one extreme to highly polysynthetic languages at the other extreme is defined. A strictly isolating language is one in which every word consists of only one morpheme. Morphologically, Vietnamese is an extremely isolating language where each word is a minimal form that expresses meaning in a linguistic context (Payne 1997) and grammatical relations are primarily shown by word order. Vietnamese has the basic word order SVO. In (1a), the subject *Tôi* 'I' precedes the verb *uống* 'drink' which is followed by the object of 'drink', *cà phê* 'coffee'. Example (1b) is also grammatical even though it is not allowed by pragmatics:

- (1) a. *Tôi uống cà phê*
1SG drink coffee
'I drink coffee'
- b. *Cà phê uống tôi*
coffee drink 1SG
'The coffee drank me'

The word order in attributive constructions is Modified Item + Modifier (Mod). Hence, the dominant position of qualifying adjectives is Noun (N) + Adjective (Adj) as in (2):

- (2) **người cao**
 person tall
 ‘tall person’

Since classifiers are noun categorization devices, we take a look at the lexical category of ‘nouns’ in Vietnamese. As defined by Nguyễn Đ. H. (1997: 89), a noun is an item that can occur after a numeral (**một** ‘one’, **hai** ‘two’, **mấy/vài**¹⁰ ‘a few’), a plural morpheme (**các/những**), and/or preceding a demonstrative (**này** ‘this’, **ấy/đó** ‘that’, **kia** ‘that, but further than **ấy**’):

một/hai/mấy/vài _____ **này/ấy/đó/kia**
những/các _____

In the illustration above, a noun in Vietnamese can take the slot after numerals e.g. (3a) and (3b), and after plural morphemes e.g. (3c). In the examples below, **người** ‘person’ is the noun:

- (3) a. **hai người**
 two person
 ‘two people’
 b. **mấy người**
 few person
 ‘a few people’

¹⁰ In Nguyễn (1997), words like ‘every, each, a few, some, several, many, all’ are considered as numerals whereas in Frawley (1992), these words are considered as quantifiers.

- c. ***những người***
 PL person
 ‘people’

A noun can also occur before demonstratives e.g. (4a). In some cases, it can occur in between numerals and demonstratives as illustrated by example (4b) below:

- (4) a. ***người đó***
 person DEM
 ‘that person’
- b. ***ba người này***
 three person DEM
 ‘these three people’

However, this definition of nouns may be inadequate when we look at the noun phrase (NP) below. In (5), there are 6 lexical items between the numeral, ***năm*** ‘five’ and the demonstrative, ***đó*** ‘that/those’. In complex noun phrases like (5), how do we know which is the noun?

- (5) ***năm chiếc áo-dài lụa xanh mới may đó***
 five CL VN dress silk blue new sew DEM
 ‘those five blue silk Vietnamese dresses that were just tailored’

Hence, it is more appropriate for us to look at the structure of the NP in Vietnamese to determine which is the noun:

Structure of the Noun Phrase

-3	-2	-1	0	+1	+2	+3	+4	+5
Totality	Num/PL Morphemes	CL	N	Attrib: Material	Attrib: Colour/ Size	Other Attrib	Dem	Poss
	<i>năm</i>	<i>chiếc</i>	<i>áo-dài</i>	<i>lụa</i>	<i>xanh</i>	<i>mới may</i>	<i>đó</i>	

Adapted from Nguyễn (1997: 181)

From the structure of the NP, the noun is found in slot 0.

Nouns are classified into proper nouns and common nouns and according to Nguyễn Đ.

H. (1997: 91), “simple common nouns can take both a numeral and a classifier in the NP

structure [NUMERAL – CLASSIFIER – NOUN]: *hai quả cam* (two fruit orange) ‘two

oranges’, *ba tòa nhà* (three seat house) ‘three buildings’ ”. Most of the time, classifiers

in Vietnamese are introduced in the context of a numeral NP. They are independent

lexemes that appear contiguous to numerals in such a context. The “central element” in

the numeral NP is an “*item noun*” and we can have a demonstrative as well. A well-

formed numeral NP taken from Nguyễn (1997: 95) looks like:

NUM	N'	N	DEM	
<i>hai</i>	<i>con</i>	<i>chim</i>	<i>đó</i>	
two	animal	bird	that	‘those two birds’
<i>ba</i>	<i>cái</i>	<i>bút</i>	<i>này</i>	
three	thing	pen	this	‘these three pens’

N’ denotes classifiers in Nguyễn’s terminology (1997); *con* and *cái* are two ‘generic classifiers’¹¹ that are typically used to denote animates especially animals, and inanimates respectively¹². If the nouns that are classified are something else, we have to select appropriate classifiers to categorize the nouns. Since the purpose of this chapter is of a syntactic nature, I shall postpone proper discussions on the semantic criteria for categorization till Chapter 4.

Vietnamese has a large number of classifiers. According to some estimates, Vietnamese has about 140 classifiers (Adams 1989); Nguyễn P. P. (2002) lists 195. However, the actual number of classifiers in common usage may be much less than that, some being replaced by generic classifiers like *cái* or *con*. For example, ‘house’ can be classified by two specific classifiers, which are *ngôi* and *căn*. However, it is more common to hear speakers of Vietnamese classifying ‘house’ with *cái* nowadays.

- (6) a. ***Cái nhà rất đẹp***
 CL house INT beautiful
 ‘The house is very beautiful’ (common)
- b. ***Ngôi/căn nhà rất đẹp***
 CL house INT beautiful
 ‘The house is very beautiful’ (less common)

¹¹ A generic classifier is a classifier that can be used with any or almost any noun, replacing other more specific classifiers.

¹² *Con* can be used to denote a small group of inanimates and *cái* can be used to denote some animates (see Table A in Section 4.4)

Also, if we take a closer look at this inventory of classifiers, we find that it includes units of measure like *cân* ‘pound’, *gáo* ‘scoop/dipper’, etc. Lobel has shown that, at first glance, there is no formal distinction between these units of measure and classifiers like *con* ‘classifier for non-human’ as example (7) (2000: 261) illustrates:

- (7) a. *một cân cá*
one pound fish
‘a pound of fish’
- b. *một con cá*
one CL fish
‘a fish’

However, there are distinct differences between these two categories which I will elaborate further in the next two sections. There are various terms used for these two categories, Units of measure like *cân* ‘pound’ are also known as **mensural classifiers**¹³ (see Lyons 1977, Craig 1992), measure terms, and quantifiers (see Greenberg 1972, Bisang 1999) while classifiers of the second type like *con* are known as **sortal classifiers**. Generally, sortal classifiers are often of more interest and considered as “true classifiers” by linguists.

¹³ In this dissertation, we will use mensural classifiers to refer to units of measure e.g. *cân* ‘pound’ and *lít* ‘litre’.

3.2 Mensural and sortal classifiers

Mensural classifiers are defined in Lyons (1977: 463) as a type of classifier that “individuates in terms of quantity” whereas a sortal classifier “individuates whatever it refers to in terms of the kind of entity that it is”. Mensural classifiers are used for measuring units of both mass and count nouns and are a fairly large open class. They correspond to the measure terms of non-classifier languages like ‘a *pound* of meat’, ‘a *cup* of water’, ‘a *sheet* of paper’, etc. According to Bisang (1999: 120), “The operation of measuring puts a given entity into a unit of measure which makes it countable or enumerable i.e. discrete.” The process of measuring is obligatory for the quantification of semantically unbounded mass nouns like water, sugar, salt, etc. because these entities cannot be enumerated directly by a numeral as they possess the properties of being homogenous, continuous and expandable. However, semantically bounded count nouns like books and men can also take mensural classifiers e.g. a *stack* of books and a *group* of *men* to show the arrangement of books and the distribution of men to be counted. The semantic boundaries of these count nouns are re-created by the use of mensural classifiers such that ‘a stack of books’ becomes an individuated entity which we can enumerate e.g. two stacks of books, three groups of men. However, mensural classifiers are optional for count nouns whereas they are obligatory for mass nouns.

According to Craig (1992), one of the major characteristics of mensural classifiers is that they co-exist in a language with the other type of classifiers, the sortal ones, in complex and heterogeneous classifier systems. Within the framework of Functional Grammar developed by Dik (1987, 1989), *entities* are mental constructs and *terms* are the

instruments that refer to these entities. The structure of terms includes the following types of elements: a Head of the term phrase, restrictors which modify that head (attributive phrases and clauses) and operators (grammatical elements). Several types of term operators are identified: definiteness/indefiniteness, specificity/genericity, demonstratives, quantifying expression, and special operators for questioned, relativized, and anaphorical terms. Classifiers are first introduced in the discussion of quantifying operators. Mensural classifiers are said to be term operators which specify units in terms of which the referent of the head noun (mass or count) can be counted as shown in Craig (1992: 280-1):

Number	Mensural classifier	Head nominal	Type of nominal
two	pound/cup of	sugar/butter	mass
three	pound/bag of	oranges/nails	count

Sortal classifiers do not have a direct equivalent in non-classifier languages. They are morphemes that individuate units (not quantity) by designating their semantic boundaries such as animacy, shape, and consistency, so that the referent of the head noun can be counted and they can also be used in contexts other than enumeration (Craig 1992). For instance, they can also be used in individualization, referentialization and relationalization (Bisang 1999). While the function of mensural classifiers is also individualization and making an entity countable just like sortal classifiers, there is a significant difference between the two. According to Croft (1994: 162), the purpose of mensural classifiers is to **create** the unit to be counted whereas sortal classifiers **actualize** the semantic boundaries which already belong to the concept of a given noun. When we use mensural classifiers to measure, it is based on properties external to the entity and

“the choice of the classifier is often determined by the temporary state of an object (its quantity, or the arrangement it occurs in)” (Aikhenvald 2000: 115). So, there may be more freedom in choosing a mensural classifier. However, when we use sortal classifiers to count something, it is based on the inherent properties of the entity. There can only be one sortal classifier that is used with a particular type of entity or an entity. As can be seen in the figure below, Dik (1989) proposes that there is a category of *ensemble nouns* in classifier languages where nominal predicates can be used to refer to ensembles without distinguishing whether the entity is a set or a mass. This is possible with the use of sortal classifiers (1989: 125), “Nouns that take sortal classifiers are ensemble nouns, which can be used to refer to ensembles and neutral with respect to the set/mass distinction, as well as with respect to individuation and quantification.”

Numeral	Classifier	Ensemble Noun
a. three	animal object	of the elephant kind/idea
b. three	flat object	of the blanket kind/idea

(Adapted from Craig 1992: 279)

Sortal classifiers often appear to be semantically redundant, expressing one or more inherent semantic characteristics of the head noun (Craig 1992). They categorize the nouns in terms of their inherent properties such as animacy, shape, consistency, etc. For example, they may refer to the nature of the object, as in ‘one CL:ANIMAL bear’ or ‘one CL:PLANT tree’. They can also refer to the shape of the object: ‘two CL:STICK-LIKE pencils’ or ‘two CL:LEAF-LIKE cards’. The function of an object is also reflected by the sortal classifier: ‘two CL:TRANSPORTATION cars’ or ‘two CL:EDIBLE fruits’. For humans, sortal classifiers can also refer to the social status or kinship relation: ‘one

CL:RESPECTED monk’ or ‘one CL:DEGRADED prostitute’. The table below summarizes some of the differences between mensural and sortal classifiers:

Table 3.0
Differences between mensural and sortal classifiers

Mensural Classifier	Sortal Classifier
<ul style="list-style-type: none"> • Individuates in terms of quantity (measuring) • Conditioned by the measure of an entity, and the temporary physical properties • More freedom in the choice of a classifier • Applied to semantically unbounded nouns (mass nouns) as well as semantically bounded nouns (count nouns) • Usually do not have pragmatic uses • Cannot be omitted for mass nouns • Connection with noun is not strong • Occur in all languages under different terms • Applied to a wide variety of nouns 	<ul style="list-style-type: none"> • Individuates whatever it refers to in terms of the kind of entity it is (counting) • Conditioned by the inherent properties of an entity • More restricted in the choice of a classifier • Applied to semantically bounded nouns (count nouns) specifically • Can be used to mark definiteness or referentiality of the noun • Can be omitted under certain context • Strong connection with noun • Occur only in classifier languages • Applied for a limited and specific group of nouns

Other differences that are not listed in the table include different preferences for the lexical sources of mensural and sortal classifiers, and others have to be established based on language internal criteria.

Table 3.0, which is compiled from different sources, is useful because it highlights to us the general properties of mensural and sortal classifiers. Even though we cannot tell if *cân* in (7a) or *con* in (7b) (example (7) is reproduced here for ease of reference) has any pragmatic uses or if the connection with the noun is strong, we can make use of the properties to carry out tests to confirm whether *cân* is a mensural classifier or a sortal classifier.

- (7) a. *một cân cá*
one pound fish
'a pound of fish'
- b. *một con cá*
one CL fish
'a fish'

3.3 Mensural and sortal classifiers in Vietnamese

In Vietnamese, both mensural and sortal classifiers exist and there some researchers who argue that this distinction does not exist in Vietnamese (c.f. Cao X. H. 1988) because there is “no grammatical distinction between object nouns (entity-denoting nouns) and stuff nouns (mass-denoting nouns). They are said to behave both like ‘mass nouns’...” (Lobel 2000: 266). According to Thompson (1965: 192), there are several subtypes of

nouns, “depending on certain other special features of distribution”. He names four of them: (1) Relator nouns, (2) Mass nouns, (3) Indefinite nouns, and (4) Item nouns. Of these subtypes, mass nouns commonly appear with mensural classifiers or with quantifiers, while item (count) nouns occur with all kinds of numerators, including quantifiers, plural markers, numerals and classifiers:

- (8) Mass noun: *sữa* ‘milk’
- | | |
|-------------------|-------------------------|
| <i>sữa ấy</i> | ‘that milk’ |
| <i>nhều sữa</i> | ‘a lot of milk’ |
| <i>ba cốc sữa</i> | ‘three glasses of milk’ |
- (9) Item noun: *bàn* ‘table’
- | | |
|----------------------|----------------|
| <i>bàn ấy</i> | ‘that table’ |
| <i>nhều cái bàn</i> | ‘many tables’ |
| <i>các/những bàn</i> | ‘tables’ |
| <i>mỗi bàn</i> | ‘every table’ |
| <i>ba cái bàn</i> | ‘three tables’ |

In (8), *ba cốc sữa* is different from *ba cái bàn* in (9) because of the mass/item noun difference. There are some differences in the classifiers *cốc* and *cái* as well. In Table 3.0, we say that a mensural classifier is conditioned by the measure of an entity and it is obligatory for mass nouns. *Cốc* appears to be a mensural classifier under these two conditions. On the other hand, *cái* seems to be different from *cốc* since it individuates the table in terms of the kind of entity it is. It is also applied to a count noun specifically. Another difference we observe is that even though *cốc* and *cái* occur in the same slot in a noun phrase, we see that *cái* cannot be an independent lexical item whereas *cốc* is an

independent lexical item and it refers to the entity ‘glass’. Hence, (10a) can be interpreted in two ways whereas (10b) has only one interpretation that must be read anaphorically where the object being classified by *cái* is already known like (10c):

- (10) a. ***Cho tôi ba cốc***
 give 1SG three glass/CL
 ‘Give me three glasses’/ ‘Give me three glasses (of milk)’
- b. ***Cho tôi ba cái***
 give 1SG three CL
 ‘Give me three (of something)’
- c. ***Ông có mấy cái bàn? Cho tôi ba cái***
 2SG POSS QUAN CL table give 1SG three CL
 ‘How many tables do you have? Give me three (tables)’

In Vietnamese, we might run into problems if we just rely on the criterium of an item being independent to distinguish whether it is a mensural or a sortal classifier. This is because classifiers often come from lexical sources like nouns and verb and this extends to both mensural and sortal classifiers alike. As we will see in the next chapter, *lá* ‘leaf’ is a noun and is also used as a classifier to classify plant species and classificatorily leaf-like items like flags, sails and business cards. *Lá* has the characteristics of a sortal classifier because it individuates whatever it refers to in terms of the kind of entity it is even though it can be an independent lexical item. From the above discussion, we can see that the mass/count noun distinction is important in helping us to evaluate if a classifier is a mensural or sortal classifier. We will look at some of the ways to elicit the count-mass distinction in Vietnamese nouns.

Lobel (2000) claims that the count-mass distinction correlates with the potential of being used or not in a given syntactic position [+particularized] and other ontological criteria.

Some of these criteria include:

A) Nouns denoting discrete entities i.e. count nouns, such as people, cars and fruits can be combined with collective nouns such as ‘a group’, ‘a collection’ and ‘a bunch’ to form a collective construction. This is not possible with mass nouns. For example:

(11) a. **một chục chuối**
one dozen banana
‘approximately a dozen of bananas’

*b. **một chục cát**
one dozen sand

B) Count nouns can be modified by stative verbs denoting size, shape, length, etc. like **nhỏ** ‘small’, **dài** ‘long’ and **ngắn** ‘short’. If these verbs are combined with mass nouns, they only allow for a qualitative reading such as (12a). However, if these verbs are combined with entity-denoting nouns, then it is possible to get both the qualitative reading e.g. (13a) and dimensional reading e.g. (13b).

(12) **cát lớn**
sand to be big
‘sand which is coarse-grained’

(13) a. **áo dài**
shirt to be long
‘dress which is long’

- b. **áo dài**
 shirt long
 ‘long dress/gown’

C) The connection between mensural classifiers and their respective mass/count nouns is less than that of sortal classifiers and their head nouns. This can be seen from instances where stative verbs such as *đầy* ‘to be full of’ or quantifiers such as *rưỡi* ‘half’ can come between the mensural classifier and the noun e.g. (14a) whereas this is not possible with sortal classifiers where these verbs can only follow the entire phrase e.g. (14b).

- (14) a. **một cân rưỡi chó**
 one pound half dog
 ‘one and a half pounds of dog (meat)’
- b. **một con chó rưỡi**
 one CL dog half
 ‘one and a half dog’

In Vietnamese, the notion of a ‘discrete entity’ is highly relevant from a syntactic point of view. I will not go into a detailed discussion since this topic is elaborated in depth in Lobel (2000). So far, I have established that there are two types of nouns, i.e. mass and count nouns; there are also ways to distinguish between mensural and sortal classifiers which serve the purpose of measuring and counting respectively. Within the group of mensural classifiers, we have several subtypes corresponding to measures, collectives, and kind (c.f. T’sou 1976, Bisang 1999):

I. Measures – [+exact, -entity]

The first subtype corresponds to measure words like ‘pounds’, ‘litres’ and ‘metres’ where the measure is exact even though they refer to no discrete physical entity.

II. Collectives – [-exact, +entity]

The second subtype corresponds to collective words like ‘a group’, ‘a bunch’ and ‘a string’ where the quantity is not exact but there is a definite sense of well-defined discrete entity or entities.

III. Kind –[-exact, -entity]

The third subtype corresponds vaguely to words like ‘species’, ‘type’ and ‘a fit of’ that characterizes mainly abstract nouns where the measure is neither exact nor does it refer to a discrete physical entity.

Sortal classifiers [+exact, +entity] denote an exact quantity (unless if a non-exact quantifier is used in the numeral NP) and refer to discrete physical entities (T’sou 1976: 1217-1218). In this study, I will focus more on sortal classifiers since they are unique to classifier languages and exhibit more interesting behavior in comparison to mensural classifiers. They are also more culturally rooted because the choice of a sortal classifier depends wholly on how an entity is perceived by the community of speakers and hence it is culture-specific.

The tables below are collections of the inventory of classifiers from grammars of Vietnamese, descriptions of Vietnamese language and my informants. Table 3.1 includes mensural classifiers and their subtypes including those that are used to classify mass entities with the parameter [+exact, -entity], collection of entities with the parameter [-exact, +entity], and abstract entities with the parameter [-exact, -entity]. Table 3.2 includes sortal classifiers used to classify discrete physical entities with the parameter [+exact, +entity].

Table 3.1
Inventory of Mensural Classifiers in Vietnamese

I. Measures

1. **bao** ‘packet, box, sack, bag, pack’ – items that come in one of the above forms: cigarettes, rice, cement, matches.
2. **cân** ‘kilogram’ – entities that can be weighed.
3. **gáo** ‘scoop’ – fluid entities.
4. **giọt** ‘drop’ – tears, water, oil.
5. **hạt/hột** ‘drop, speck’ – diamonds, pearls, rice, rain.
6. **hộp** ‘box’ – matches.
7. **lít** ‘litre’ – fluid entities.

8. **ly** ‘glass’ – fluid entities.
9. **ve** ‘flask’ – fluid entities.
10. **vò** ‘jar’ – fluid entities.

II. Collectives

11. **bộ** ‘set, series, pack, collection’ – composite entities: tables and chairs, living room furniture, clothes, playing cards, books, a trio (with similar characteristics), machines, teeth, pair of beautiful legs (informal), skeletons, giblets, innards, intestines, brains, chest, moustache, books.
12. **bụi** ‘clump, grove, bush’ – grass, bamboo, shrubs.
13. **buồng** ‘bunch, room, compartment’ – bananas, areca nuts, livers, (a pair of) lungs.
14. **cặp** ‘couple, pair’ – items occurring in sets of two, naturally connected or inseparable: eyes, glasses, elephant’s tusks, married couples, hinges.
15. **chòm** ‘tuft, clump, bunch, group’ – collection for items like beard, hair, stars, bees, trees, flowers.
16. **chùm** ‘cluster, bunch’ – balloons, grapes, keys.
17. **chuỗi/tràng** ‘a string of’ – pearls, coins, laughter.
18. **cụm** ‘branch, tuft, bushes, bundle’ – trees, onions, flowers, clouds, forests.

19. **dải** ‘band, range’ (literary use) – mountains, land, rivers.
20. **dãy** ‘chain, row, line’ – houses, mountains, rooms.
21. **dòng** ‘current, stream, course’ (literary use) – water, tears, (**con/cái**) rivers.
22. **đám** ‘mass, patch’ – mass or compact items: clouds of dust, clouds, heavy showers, silver, fire, conflagration, fields, meadows, plots of (cultivable) land, jungles, forests, gardens, funeral processions, festivals, holidays, crowds of people.
23. **đoạn** ‘section, length’ – roads, writings.
24. **khối** ‘block, mass, bulk’ – big 3-dimensional objects: cement, stones, brains.
25. **mảnh** ‘bit, piece’ – paper, wood, feelings.
26. **múi** ‘section, segment’ – fruits and vegetable: garlic, pomelos, oranges.
27. **nhát/lát** ‘slice’ – ginger, meat.
28. **phường** ‘a band’ – robbers, bandits, musicians.
29. **rặng** ‘row, line, chain’ – mountains, bamboo
30. **súc** ‘log, roll’ – wood, paper, meat, cloth.
31. **tảng** ‘slab, block’ – big 3-dimensional objects: stones, meat, glacier.

32. **tập** ‘volume, stack’ – paper, books.

III. Kind

33. **bài** ‘lesson, text’ – music, songs, (**cái**) titles, poems.

34. **căn** – (**cái/cơn**) sickness.

35. **cơn** ‘outburst, fit, attack’ – sudden violent and negative states, with weather, emotions, and state of health: typhoon, gust of wind, heavy wind storm, squall and shower, burst of sunlight on a dark day, sickness, madness, sudden fever, fit of anger.

36. **cuộc** ‘party, match, game’ – with events involving competitive interaction or struggles: ball games, chess games, wars, committee meeting, revolution, election, strikes, investigations, military operations, peace, life.

37. **chuyến** ‘trip, journey, voyage’ – transport: trains, airplanes, ferry boats, goods-carrying vehicles.

38. **hơi** ‘breath’ – wind, breeze, breath.

39. **khoa/kỳ** ‘ - (**cuộc**) exams.

40. **làn** ‘wave’ – rivers, wind, public opinion.

41. **loại** ‘kind, sort, type’ – animals, everything

42. **món** ‘item’ – food, money, goods, debts, presents

- 43. **nét** – feelings, handwriting, beauty, sadness, happiness.
- 44. **niềm** – positive emotions: happiness, love, fortune.
- 45. **trận** ‘period’ – soccer matches, storms, laughter, pain, wind, rain, heat.

Table 3.2
Inventory of Sortal Classifiers in Vietnamese

- 1. **ả** ‘girl, woman, she’ (often with derogatory meaning) – secretaries, lovers, mistresses.
- 2. **bóng** ‘shadow, silhouette’ (literary use) – (**con**) ghosts.
- 3. **bản** ‘edition, copy’ – music, treaties, contracts, epics.
- 4. **bức** – relatively thin and flat objects of fairly large size: large photographs, pictures, curtains, window shades, bamboo blinds, walls of various materials, (**cái**) paintings, statues, screens, embroidery, (**cái/lá**) letters.
- 5. **cái** – generic classifier for inanimates: tables, chairs, paintings,
- 6. **cánh** ‘wing, petal’ – fields for cultivation, sails, duckweeds, panels of doors, bows.
- 7. **cảnh** ‘scenery, landscape, view’ – mountains, gardens.
- 8. **căn/gian** ‘house, compartment’ – (**cái**) houses, rooms, lifts, kitchens.

9. **cây** ‘tree, plant’ – (1) plant species: orange trees, lotus plants, grass.
- (2) stick-like items: pencils, pens, nails, candles, silk, paddy, straws.
10. **chiếc** – (1) items that are one of a pair: chopstick, shoe, elephant’s tusk, sock, stocking.
– (2) a large number of movable, manufactured or constructed items: bridges, small boats, junks, sampans, ships, airplanes, vehicles, bicycles, motorcycles.
– (3) other items: rings, leaves, mats.
11. **cổ** ‘ancient, antique, archaic’ – tea set, (**cái**) playing cards, (**cái**) coffins, (**cái**) cars.
12. **con** – (1) non-human animates, lower humans and supernatural beings: all animals and insects, gamblers, merchants, actors, hostages, ghosts and monsters.
– (2) items depicting images of animates: playing cards, chess pieces, postage stamps, seals, kites, dolls.
– (3) items created by humans for human use: spinning tops, triggers, oars, knives, dice, digits, figures.
– (4) items that seem to have a life of their own: roads, rivers, boats, ships, dikes.
– (5) body parts that are very close to humans: heart, eye (one), pupil (one).
13. **củ** ‘bulb, tuber’ – plants that have tubers and bulbs: carrots, onions, sweet potatoes, peanuts, garlic, ginger.
14. **cuốn** – literary material of considerable thickness: books, novels, diaries, dictionaries, notebooks, exercise books.
15. **cục** ‘lump, piece, clot’ – (**cái**) tumors, (**hòn**) bricks, blood.

16. **cửa** ‘entrance, door’ – (**cái**) pass at the border, restaurants, compartments, apartments, space in the room.
17. **chàng** ‘he, him’ – (**người**) youths, son-in-laws.
18. **chân** – accountants, servants, secretaries.
19. **đàn** – groups of domesticated animals: cows, sheep, horses.
20. **đoàn** – groups of humans: students, workers, people.
21. **đứa** – children and low status people.
22. **gã** ‘guy, chap’ – male youths.
23. **gốc** ‘source, origin’ – most trees like **cây**
24. **hòn** – stones, stone-like objects and islands: islands (in rivers, lakes and the sea), stones, pebbles, marbles, diamonds, (**viên**) precious stones, ice, blood clots, broken bricks.
25. **lá** ‘leaf’ – (1) leaves of plant species: tea, banana, betel.
– (2) leaf-like items: (**cây/con**) card in games, (**cánh**) sails, flags, liver, spleen, lung, (**cái/tấm**) business card, (**cái** or **bức**) letters.
26. **lời/câu** ‘word/sentence’ – (**cái**) greetings, poems, songs, music, writing.
27. **mái** ‘roof’ – (**cái**) sampans, (**cái/căn**) houses, hair.

28. **mũi** – arrows, knives, mallets
29. **nấm** ‘mushroom’ – graves, small heaps of land.
30. **nải** ‘hand’ – with bananas.
31. **ngọn** ‘summit, peak’ – peak-shaped items: (**cái or trái**) hills, mountains, treetops, a flame, current of air or breeze, current of water, thunderstorms, fire, candles.
32. **ngôi** ‘kingship, throne’ – (**cái**) graves/tombs, stars, houses.
33. **người** ‘person, man’ – adult human beings: chefs, friends, husbands, wives, actors, actresses, relatives.
34. **nhà** ‘house’ – knowledgeable people: scientists, travellers, teachers, artists, explorers, writers.
35. **nốt** – music notes, moles, scars, gall.
36. **ổ** ‘nest’ – (**bộ**) machines and engines, silk cocoons, (**cái**) loaves of bread, (**cái**) door knobs.
37. **ống** ‘hose, tube’ – 1-dimensional items: pumps, water pipes, troughs, flutes.
38. **sợi** ‘thread, fibre’ – 1-dimensional, flexible items: thread, cord, wick, (**cái**) tie, rope, wire, chain, belt, hair.

39. **tấm** – 2-dimensional, flat pieces of material: placards, blackboards (also classified with **cái**), cloth, boards, hide, maps, mirrors (also classified with **cái**), tapestry, photographs, drawings or paintings of a person, panes of glass, nets, screens, silk, cotton, business cards (also classified with **cái** or **lá**), planks, flat pieces of timber.
40. **tên** – derogatory for humans: prisoners, alcoholics, robbers, enemies.
41. **thanh** ‘slat, bar’ – 1-dimensional thin, rigid pieces of metal including weapons: swords, iron, metal.
42. **tờ** – loose sheets of paper or documents: paper, currency notes, newspaper, journals, contracts, curriculum vitae, written orders, treaties, pacts.
43. **trái** ‘fruit’ – (1) fruits of plant species: fruits, melons, cucumbers.
– (2) small 3-dimensional objects: bombs, hills, mountains, hearts, kidney, the Earth, globes.
44. **que** – matches.
45. **viên** ‘pill, to be round’ – 3-dimensional small artificially-shaped items: meatballs, precious stones, diamonds, marbles, pills, bullets.

The tables above serve as a rough guide to the inventory of both mensural and sortal classifiers in Vietnamese. It also gives us some idea of the number of classifiers and how entities are classified in Vietnamese culture. Even though we are dealing with the syntax

of Vietnamese classifiers in this chapter, we find it necessary to define a set of ‘true’ (i.e. sortal) classifiers to serve as data for our analysis.

3.4 Absence of repeaters

Repeater constructions can be the source for large systems of noun categorization devices and this phenomenon is widespread in Southeast Asian languages. A repeater appears when any – or almost any – noun can be used in the classifier slot to classify the same noun, or semantically related nouns. An example of a repeater construction from Burmese is given below (Becker 1975: 113):

- (15) ○ {kA vL} ○ {kA
 river one CL
 ‘one river’

Repeaters are not found in Vietnamese and classifiers are not an open class. Even though a number of classifiers in Vietnamese are derived from nouns, for example *cây* ‘tree, plant’, *người* ‘man, person’, *trái* ‘fruit’ and *lá* ‘leaf’, these classifiers are not repeaters because they do not classify themselves, unlike in Thai and Burmese. This simply means that the examples below are ungrammatical. Instead, the nouns ‘leaf’ and ‘person’ would mostly be unclassified or they would take a generic classifier like *cái* (for inanimates) or *con* (for animates).

- (16) *a. ***một lá lá***
 one CL leaf
- b. ***một cái lá***
 one CL leaf
 ‘a leaf’
- (17) *a. ***một người người***
 one CL person
- b. ***một con người***
 one CL person
 ‘a person’

Languages with N-[Num-CL] constituent order appear to use repeaters more than other languages with a different constituent order. Thai, Mal (Mon-Khmer), Lao (Tai) and Burmese are typical examples of languages with repeaters and they all have N-[Num-CL] constituent order (Aikhenvald 2000). On the other hand, Vietnamese has [Num-CL]-N constituent order just like Mandarin and repeaters do not occur in either language. The absence of repeaters seems to be related to the word order of languages and there could be two possibilities why this is so:

- I) As shown in (10-11), the repeater strategy with the order [Num-CL]-N would produce two adjacent identical words, which would be disfavoured by haplology;
- II) Repeaters might arise as a result of in-line production: in a N-[Num-CL] language, one may produce a noun, and even a numeral, before deciding on the appropriate classifiers, leaving no option but to repeat the noun as classifier.

Another point about derived classifiers in Vietnamese is that even though this is a potentially productive process, in actual fact, not many classifiers are derived from nouns or verbs (Refer to Table 3.2 for source of classifiers). Derived classifiers co-exist with classifiers with no nominal or verbal source denoting measures, collectives, kinds, etc.

3.5 Numeral and demonstrative constructions

Numerals and classifiers are closely related because classifiers are usually introduced in the context of numeral phrases. Classifiers also combine with nouns and demonstratives in various ways to form deictic expressions. Greenberg (1972) points out that the other context where a classifier is frequently used is in a deictic construction. All the different constructions involving counting and reference are listed in the table below:

Table 3.3
Numeral and Demonstrative Phrases

Construction Phrases	Examples
Num-CL-N-Dem	<i>hai-cái-cổng-ấy</i> ‘those two gates’
Num-CL-N	<i>ba-cây-dù</i> ‘three umbrellas’
Num-CL	<i>sáu-tám (thiếp)</i> ‘six (cards)’
Num-N-Dem	<i>ba-màu-đỏ</i> ‘those three colours’
Num-N	<i>nam-giờ</i> ‘five hours’
CL-N-Dem	<i>con-bò-này</i> ‘this cow’
CL-N	<i>chiếc-xe hơi</i> ‘car’
CL-Dem	<i>con-ấy</i> ‘that one’

N-Dem	<i>chỗ-này</i> ‘this place’, <i>bàn-kia</i> ‘that table’
N-Num-CL (restricted use)	<i>nhà-ba-căn</i> ‘a three-compartment house’

In the following sections, we will look at all the constructions in Vietnamese consecutively.

3.5.1 Num-CL-N-Dem

In Vietnamese, classifiers are not always obligatory in the context of counting or quantification as we shall see in Section 3.5.1 – 3.5.5. The deciding criterion lies on the head noun that is being classified. Some nouns are never classified under any context and a classifier is not necessary. Some of these unclassified nouns are given in Emeneau (1951: 100):

Table 3.4
Unclassified Nouns in Vietnamese

buổi	half a day
bữa	a meal
câu	a sentence
châu	a continent
chỗ	a place
chuyện	a story
đình	official residence of a high official
đêm	night

giá	price
giờ	hour
họ	extended family, family name
làng	village
lẽ	reason, cause
lễ	religious ceremony, religious festival

When a person wants to count something and reference needs to be established, Example (18) is the most conventional one:

- (18) ***hai trái đào này***
two CL peach DEM
‘these two peaches’ (*trái* 1. fruit, 2. CL: fruits and 3-D items)

When a certain attribute of the peach needs to be included or introduced in the construction above, the attribute will be incorporated before the demonstrative like:

- (19) ***hai trái đào đẹp này***
two CL peach pretty DEM
‘these two beautiful peaches’

The [Num-CL-N-DEM] order only applies to nouns which can be classified. According to Emeneau (1951: 94), “classified nouns are numerable only when preceded by a classifier”. A typical example of a classified noun would be ***sách*** ‘book’. It is ungrammatical to have:

*(20) **hai sách này**
two book DEM

The choice of classifier is determined by the semantic properties of the entity (see Chapter 4) and the cultural values of speakers of the language. In Vietnamese, books belong to the larger category of literary material and because of the thickness and form, they are classified by *cuốn* ‘to roll, wrap’. If we have a thin magazine, it would be classified by *tờ* ‘sheet, piece of’. Sometimes, the choice of the classifier also reveals a person’s perception of the entity they are talking about.

3.5.2 Num-CL-N

This is the most common and widely used construction in Vietnamese to count entities that can be classified. Typical examples are:

(21) **hai chiếc máy bay**
two CL airplane
‘two airplanes’ (*chiếc* 1. CL for machinery)

(22) **ba tờ báo**
three CL newspaper
‘three newspapers’ (*tờ* 1. CL for thin literary material)

It is possible to include attributes to form a more descriptive construction like:

(23) **hai chiếc máy bay lớn và mới**
two CL airplane big and new
‘two big and new airplanes’

[Num-CL-N-Dem] and [Num-CL-N] are common constructions that we usually come across in formal contexts and grammars of Vietnamese.

3.5.3 Num-CL

This construction where the noun is omitted may appear quite often especially in a context where the noun has already been mentioned and constitutes old information. For example:

- (24) *Tôi có hai con chó, sáng này một con mới chết*
1SG POSS two CL dog morning DEM one CL new die
'I have two dogs, this morning one of them just died'

In (24), it is possible to omit *chó* 'dog' in the second clause because we already know that the entity we are referring to are dogs and it is unambiguous that *một con* 'one CL' must be classifying the same dog mentioned in the first clause. Another example is given in Emeneau (1951: 84):

- (25) *Cha tôi mua ba con trâu. Con thứ nhất đen, con thứ nhì trắng và con thứ ba đỏ đỏ*
father 1SG buy three CL buffalo CL first black CL second
white and CL third reddish
'My father bought three buffaloes. The first one is black, the second is white and the third is reddish'

A typical situation where the noun of a classifier construction is often dropped is when we are buying something that is physically present. More often than not, we would point

and just specify the quantity (Num) with the appropriate classifier. If we are unsure of the appropriate classifier for that entity then the generic classifier *cái* is often used.

3.5.4 Num-N-Dem

From Section 3.5.1, we see that not all nouns need to be classified in Vietnamese and unclassified nouns are directly numerable without any need for classifiers. Numeral classifier languages almost always include in their inventories nouns which do not take a classifier (Dixon 1986: 106). Even though we are concerned with classifier constructions and this type of demonstrative phrase does not involve any classifiers at all, nonclassified nouns have always been the topic of interest in studies of classifier systems. Hence, we will include some constructions of this type in the following two subsections.

(26) *một ngày kia*
 one day DEM
 ‘one of those days’

(27) *Tôi thích ba màu này*
 1SG like three colours DEM
 ‘I like these three colours’

3.5.5 Num-N

This subsection contains nonclassified nouns that do not take a classifier in the context of counting and other contexts. They are directly numerable in the examples below:

- (28) **hai năm**
 two year
 ‘two years’
- (29) **hai chuyện**
 two story
 ‘two stories’
- (30) **hai tháng**
 two month
 ‘two months’

We have seen more examples of unclassified nouns in the earlier section (cf. Emeneau 1951: 100) like **bữa** ‘a meal’, **chỗ** ‘a place’, **làng** ‘a village’, **đêm** ‘night’ and **mùi** ‘smell’. A nonclassified noun cannot be omitted in the [Num-N] construction unlike classified nouns in Section 3.5.3 [Num-CL(-N)], otherwise we would just be counting numbers.

3.5.6 CL-N-Dem

Classifiers are also found in referring constructions combining with demonstratives like **này** ‘this, these’, **đó/ ấy** ‘that, those’ and **kia** ‘that, those (further than **ấy**)’ and **nọ** ‘other, that, those’, in the absence of numerals. This is a very common construction which can be found in simple sentences and interrogatives. In most referring expressions, classified nouns would be preceded by their appropriate classifiers. For instance:

- (31) **Căn/cái nhà kia rất cũ**
 CL house DEM INT old
 ‘That house over there is very old’ (house can be classified by **căn** or **cái**)

Classifiers can also be found with interrogative substantives like *gì* ‘what’, *nào* ‘which’, *mấy* ‘how many’ (for numbers less than 10) and *bao nhiêu* ‘how many’ (for numbers more than 10) to form interrogative structures as in (32) and (33) below:

(32) *Anh đọc cuốn sách nào?*
2SG.MASC read CL book which
‘Which book are you reading?’

(33) *Anh mua mấy con cá?*
2SG.MASC buy how many CL fish
‘How many fishes did you buy?’

Obviously, there are other interrogative substantives like *ai* ‘who’, *đâu* ‘where’, *bao giờ* ‘when’, *bao lâu* ‘how long’ and *sao* ‘why’ but we see that these interrogatives do not need an entity they can refer to as oppose to the earlier ones. Hence, before using interrogative substantives like *nào* ‘which’, *mấy* ‘how many’ and *bao nhiêu* ‘how many’, we need to determine some entity of reference like *núi* ‘mountain’ in (34a) and *sao* ‘star’ in (35a). Then the interrogative serves to supplement additional information on the entity we have established. It is ungrammatical otherwise.

(34) a. *Anh đi hòn núi nào?*
2SG.MASC go CL mountain which
‘Which mountain did you go?’

*b. *Anh đi nào?*
2SG.MASC go which

- (35) a. ***Bầu trời có bao nhiêu ngôi sao?***
 sky POSS how many CL star
 ‘How many stars are there in the sky?’
- *b. ***Bầu trời có bao nhiêu?***
 sky POSS how many

For interrogative substantives like *ai* ‘who’, *đâu* ‘where’, *bao giờ* ‘when’, *bao lâu* ‘how long’ and *sao* ‘why’, we cannot establish any reference because that is the answer we are expecting.

- (36) a. ***Anh đi đâu?***
 2SG.MASC go where
 ‘Where did you go?’
- *b. ***Anh đi con đường đâu?***
 2SG.MASC go CL road where
- (37) a. ***Anh gặp ai?***
 2SG.MASC meet who
 ‘Who did you meet?’
- *b. ***Anh gặp người bạn ai?***
 2SG.MASC meet CL friend who

Gì ‘what’ is a special instance whereby it can be used as an interrogative substantive with or without establishing a reference as in (38) and (39):

- (38) ***Anh ăn gì?***
 2SG.MASC eat what
 ‘What did you eat?’

- (39) **Anh tìm cuốn sách gì?**
 2SG.MASC look for CL book what
 ‘What book are you looking for?’

Another construction where classifiers appear is in interrogative structures with the confirmation-seeking tag **phải không**. Sometimes, it is just reduced to **không** appearing at the end of an interrogative like:

- (40) **Anh có chiếc xe hơi (phải) không?**
 2SG.MASC POSS CL car true not
 ‘You have a car, don’t you?’

The purpose of the above discussion is to illustrate that classifiers are frequently tied up with referring expressions and in situations where reference is already established. They are not restricted to numerated constructions only.

3.5.7 CL-N

This order does not appear as frequently as the [CL-N-Dem] order. The difference when the demonstrative is not present is that we usually take [CL-N] as denoting something very general as no reference is established¹⁴. For example:

- (41) **Con trâu là một con vật không thể thiếu được ở nhà nông dân**
 CL buffalo COP one animal indispensable in house
 farmer
 ‘Buffalos are indispensable in a farmer’s family’

¹⁴ The generic reading of [CL N] in Vietnamese contrasts with Cantonese where [CL N] is typically specific (Matthews and Yip 1994).

In (41), *con trâu* ‘CL buffalo’ does not have specific reference and just refer to buffaloes in general.

3.5.8 CL-Dem

The last type of construction is usually used in situations where the noun with its classifier has already been mentioned recently and then the noun is omitted in the subsequent utterance leaving the classifier and demonstrative behind. The classifier thus possesses an anaphoric function by being co-referential with the noun.

(42) Q: *Anh muốn cuốn sách nào?*
2SG.MASC want CL book which
‘Which book do you want?’

A: *Cuốn kia*
CL DEM
‘That one’

(43) *Cái bàn này dài gấp tư cái kia*
CL table DEM long more than four CL DEM
‘This table is four times longer than that one’

The noun can also be omitted when the classifier construction occurs in predications where the species that would be indicated by the noun is still unknown but its “class” as indicated by the semantics of a classifier can be given (Emeneau 1951: 84):

(44) Q: *Con ấy là con gì?*
CL DEM COP CL what
‘What is that?’

A: **Con ấy là con vịt**
 CL DEM COP CL duck
 ‘That is a duck’

In the above example, the classifier **con** can also occur with the interrogative substantive **gì** ‘what’.

3.5.9 N-Dem

There are two types of N-Dem constructions. The first is when the noun is nonclassified such as **làng này** ‘this village’. The second type is more interesting and relevant to this research because it is found when a classified noun can occur just with the demonstrative without the classifier. For example:

(45) **Tôi muốn ngồi bàn kia**
 1SG want sit table DEM
 ‘I want to sit at that table’

(46) **Trong giỏ này toàn cua lớn**
 in basket DEM all crab big
 ‘All that is in this basket is big crabs’

In (45), we would expect **bàn** ‘table’ to be classified by **cái** (generic classifier) but it is perfectly grammatical without the classifier. Similarly in (46), **giỏ** ‘basket’ should be classified by **cái** and **cua** ‘crab’ should take the classifier **con** (classifier for non-human animates) but the use of classifiers doesn’t seem necessary here. Emeneau (1951: 99) has

categorized this phenomenon as “exceptional” and “rare”, and claims it occurs only with demonstratives. However, there are many instances where classified nouns can occur without classifier. We will try to explain this in Chapter 5 when we look at the functions of classifiers.

3.5.10 N-Num-CL

According to one of my informants, it is possible to have this construction in the context of counting¹⁵. However, this use is very restricted and only limited to the informal listing of items such as:

- (47) ***chó một con, xe đạp một chiếc, bút ba cây, sách hai cuốn...***
 dog one CL bicycle one CL pen three CL book two CL
 ‘one dog, one bicycle, three pens, two books...’

However, in sentences, it is not possible to have this construction and we always have to use the construction [Num-CL-N] as below:

- (48) *a. ***Tôi muốn ăn bò ba con***
 1SG want eat cow three CL
 b. ***Tôi muốn ăn ba con bò***
 1SG want eat three CL cow
 ‘I want to eat three cows’

The other context where we encounter the [N-Num-CL] construction is when we have:

¹⁵ The counting construction [N-Num-CL] exists in Cantonese too (Matthews and Yip 1994).

(49) *nhà ba căn*
house three ?CL/compartment
'a three-compartment house'

(50) *râu ba chòm*
beard three ?CL/tuff
'a three-tuff beard'

The examples above are not entirely similar to (47) albeit syntactically alike. In (47), the numeral and classifier are engaged in counting, however, in (49) and (50), *căn* 'compartment, CL for houses' and *chòm* 'tuff', both mensural classifiers of collective nature, appear to modify the head nouns, *nhà* 'house' and *râu* 'beard' attributively. Also, at this juncture, we start to suspect that *căn* is not a classifier but a noun. Hence, *ba căn* 'three compartments' tells us that there are three rooms or compartments in the house and it can also be translated into the sentence below:

(51) *nhà có ba căn*
house POSS three compartment
'house with three compartments'

The fact that a lot of classifiers are derived from nouns or have nominal uses poses a rather problematic area in Vietnamese grammar. For instance in (52), we might expect that *cành* 'branch' is a classifier for tree but in fact it should be interpreted as the head noun where it is modified by *cây* 'tree'. Similarly, Nguyễn P. P. (2002) has described *búp* 'bud' as a classifier when it should be a noun modified by *hoa* 'flower' as in (53).

(52) *một cành cây*
 one branch tree
 ‘one tree branch’

(53) *hai búp hoa*
 two bud flower
 ‘two flower buds’

Examples like (49-53) show us that we cannot hope to distinguish classifiers from other lexical items based solely on the syntactic order or by looking at the lexical word itself.

The word *cây* can be a noun denoting trees and plants, or a classifier for trees and rigid two-dimensional objects, depending on the context. There is no single method or aspect that allows us to fully understand the Vietnamese classifier system and that is why we have to examine the system by looking at the syntax, the semantics, and the functions of these classifiers to achieve an integrated way of understanding how the system works.

We have looked at all the possible combinations involving numerals, classifiers, and demonstratives and discussed the context in which they occur. In the next section, we will look at the relationship between these different constituents and the significance behind it.

3.6 Constituent order and constituency relations in classifier constructions

Thus far we have seen that Vietnamese NPs have [Num-CL-N] constituent order as, for example, in Chinese and Hmong. However, under certain restrictive contexts as exemplified in the previous section, it is possible to have [N-Num-CL] (Nguyễn Đ. H.

1990) which is also present in Cantonese. If we look at the general syntactic rules of Vietnamese as introduced in Section 3.1, it seems significant that Vietnamese has the particular order [Num-CL-N]. In Vietnamese, the noun phrase consists of a head noun, which may be followed by other constituents for example, another modifying noun, a pronoun, an ordinal number, a verb, a demonstrative, and a relative clause.

- (54) *phòng khách*
 room guest
 ‘guest room’
- (55) *sách của tôi*
 book POSS 1SG
 ‘my book’
- (56) *lớp nhất*
 grade first
 ‘top grade’
- (57) *người lười biếng*
 person lazy
 ‘lazy person’
- (58) *đường về*
 road to return
 ‘the way back’
- (59) *cô này*
 young lady DEM
 ‘this young lady’
- (60) *Nhà mà chú tôi vừa mua năm ngoái*
 house REL uncle 1SG just buy year last
 ‘The house that my uncle just bought last year’

If the classifier just serves to modify the head noun, we would predict that it would follow the head noun and resemble the examples above. However, we find that this is not the case. In Vietnamese, the numeral and classifier always come before the noun:

- (61) *hai con mèo*
 two CL cat
 ‘two cats’

There are several ways to analyze the classifier construction [Num-CL-N], as we shall see in Sections 3.6.1 – 3.6.3.

3.6.1 [Num-CL]-N

The first way is to say that the numeral and the classifier form a single constituent and the classifier is the head of the NP in a classifier construction. Therefore, in (62), [Num-CL] takes the initial position and the noun *mèo* ‘cat’ modifies this constituent by specifying the exact animal being classified by *con*. In this instance, we can say that *mèo* is not really obligatory since it is not the head and can be omitted when the preceding context, verbal or others, has already specified the exact entity being classified. This yields the [Num-CL] construction (see Section 3.5.3). For example:

- (62) *Tôi có hai con mèo, một con mập, một con ốm*
 1SG POSS two CL cat one CL fat one CL thin
 ‘I have two cats, one is fat and one is thin’

In (62), we can omit *mèo* ‘cat’ in the subsequent utterance because we already know that we are referring to cats and it is unambiguous that *một con* ‘one CL’ must be classifying the same cats mentioned in the first clause. Another construction that attenuates the importance of the noun is when we have the [CL-Dem] construction (see Section 3.5.8). We see that the classifier behaves like the head in the example below and the classifier is used for reference instead of the noun:

(63) Q: *Anh thích bức tranh nào?*
 2SG.MASC like CL painting which
 ‘Which painting do you like?’

A: *Bức này*
 CL DEM
 ‘This one’

(64) Q: *Cái ấy là cái gì?*
 CL DEM COP CL what
 ‘What is that?’

A: *Cái ấy là cuốn sách*
 CL DEM COP CL book
 ‘That is a book’

From the examples above, we see that the classifier is always used in conjunction with the numeral or the demonstrative in contrast to the classified noun which gets omitted under certain contexts. We can thus argue that the numeral or demonstrative, and the classifier form a single constituent which is more important than the noun, which acts like a modifier.

3.6.2 Num-[CL-N]

The second way to analyze the classifier construction in Vietnamese is that the numeral stands alone while the classifier and the noun forms a single constituent i.e. Num-[CL-N]. This may be plausible if we look at nominal compounds comprising of CL-N and referring expressions i.e. [CL-N-Dem] and [CL-N] constructions (see Sections 3.5.6 and 3.5.7). In the examples below, it seems that the classifier is closely related to the noun:

- (65) *người bạn tôi*
 CL friend 1SG
 ‘my friend’
- (66) *Con chó rất dễ thương*
 CL dog INT cute
 ‘Dogs are very cute’
- (67) *Cuốn sách này hay*
 CL book DEM interesting
 ‘This book is interesting’
- (68) *Cái bàn đó hư rồi*
 CL table DEM spoil PERF
 ‘That table is spoilt already’

The examples above show that classifiers occur with nouns in many constructions even without the presence of a numeral. Another motivation for saying that the classifier and the noun form a single constituent is consistency in the structure of the Vietnamese NP. According to Emeneau (1951: 85), numerated constructions have the schema below:

Numeral	Classifier	Classified noun	\pm Attribute(s)	Demonstrative numerator
	Nonclassified noun			

Adapted from Emeneau (1951: 85)

This schema suggests that the numeral is a separate category from the classifier and the classified/unclassified noun, and the classifier and the classified noun cannot be separated. This is the fixed word order for NPs and we can have a representation of classifier constructions as Num-[(CL)-N] where the classifier is only needed when the noun is a classified noun. The classified noun is numerable only when preceded by a classifier and the nonclassified noun is directly numerable. If we say that the numeral and the classifier form a single constituent i.e. [Num-CL]-N as in Section 3.6.1, then we will have to include Num-N for nonclassified nouns. We will thus have a more complicated schema to explain numerated constructions if we treat [Num-CL] as a single constituent:

Numerator	Classifier	Classified noun	\pm Attribute(s)	Demonstrative numerator
Numerator		Nonclassified noun		

3.6.3 Num-CL-N

The third approach to interpreting classifier constructions is to treat the numeral, the classifier, and the noun as separate constituents i.e. Num-CL-N. This would entail a ‘flat’ structure with ternary branching NPs. This approach is particularly useful in the case of

Vietnamese because “...in isolating languages it is often difficult to work out syntactic criteria for heads, especially since either a classifier or a noun can be omitted under specifiable discourse conditions.” (Aikhenvald 2000: 105). This is especially applicable to Vietnamese as we have seen in the above data that either the numeral, the classifier, or the noun can be omitted under the right conditions. Moreover, Vietnamese is a typical isolating language with independent lexemes and agreement. Hence, it is difficult to establish the syntactic head of the NP and the constituency relations. This constitutes a strong argument for not considering constituency relations or head and modifier relations in the classifier constructions in Vietnamese. Hence, in this study, we will treat the numeral NP as a flat structure without constituency relations as shown from the above data.

3.7 The origin of the classifier construction in Vietnamese

Notwithstanding how we analyse the classifier construction in Vietnamese, we observe that the [Num-CL-N] construction is exactly the same as Chinese. There is a possibility that this particular order i.e. [Num-CL-N] is borrowed from Chinese. Greenberg (1972) raised the example of Bodo, a Sino-Tibetan language, where there are two constituent orders: the ‘indigenous’ one has the order CL-Num, while the one borrowed from Assamese has Num-CL. Vietnamese could have undergone the same process where the ‘indigenous’ construction has the order N-Num-CL (which has a very restricted use) while the one borrowed from Chinese has the order Num-CL-N. There are several reasons for hypothesizing this. Firstly, Vietnamese was under Chinese influence for a long time. Through the ten centuries of Chinese political domination, Chinese served as

the medium of education and official communication, at least among the educated classes of scholars and officials. From the early days of Chinese rule (111BC – AD939), the Chinese rulers taught the natives not only Chinese calligraphy, but also the texts of Chinese history, philosophy and literature. A large number of loanwords from Chinese were also accepted into Vietnamese. A classifier which seems to be of Chinese origin is *cuốn* (Southern Vietnam), or *quyển* (Northern Vietnam). This classifier means ‘to roll, to wrap’ in Vietnamese so we can roughly say *một cuốn/quyển sách* is equivalent to ‘one roll book’. In Ancient China, books were in the form of scrolls and they were also classified by what we know in modern Mandarin as *juàn* ‘to roll’ or *gǔn* ‘roll’ in Yue. Even though books are classified by *běn* in Modern Chinese, Vietnamese has retained the classifier *cuốn/quyển* for books till today. In fact, *quyển* and *juàn* are also phonologically similar suggesting a strong possibility that it is indeed a loan from Chinese. In addition, word order is rather easily borrowed especially as a result of areal contact (cf. e.g. Comrie 1987, Matthews 1997). Hence, through the long periods of contact, Vietnamese could have borrowed some classifiers along with the construction of [Num-CL-N] from Chinese.

Besides looking at language external reasons, if we were to look at the structure of Vietnamese noun phrases again (refer to Section 3.6), we would see that within the noun phrase, the head noun always precedes the modifiers. How is it that only in a numeral noun phrase, the numeral and classifier precedes the noun? We would expect Vietnamese to have [N-Num-CL] structure which appears relatively rare and obsolete in today’s usage.

3.8 Conclusion

After the whole discussion about the syntactic properties of classifiers in Vietnamese, we can briefly summarize:

- (I) Numerals and classifiers are related to a large extent but the same classifiers may appear outside numeral noun phrases and expressions of quantity, in nominal compounds and referring expressions.
- (II) It is not obligatory for numeral classifiers to appear contiguous to numerals because of the presence of nonclassified nouns and certain specifiable contexts.
- (III) Classifiers are especially versatile and varied in their usage and a large part of their usage is governed by pragmatics.
- (IV) The constituency relationship between the numeral, classifier and the noun is not straightforward and cannot be easily determined. However, the data presented suggest a flat structure without constituency relations.
- (V) Likelihood of a borrowed word order for numeral noun phrases which supersedes the use of the indigenous order and becomes the acknowledged form.

In chapter 5, we will look more closely at point (III), which we have summarized above. In any study of classifiers, it is always important to understand the functions of classifiers

in a particular language or different languages. This understanding not only facilitates the usage of classifiers by learners of that language, it also shed some light on the different grammatical choices a speaker can make to achieve a certain linguistic function. For now, we will look at the semantics of Vietnamese classifiers in the next chapter. We analyze the nature of the semantic dimensions underlying classifier systems and how the Vietnamese classifier system adhere or not adhere to it. The semantic bases of some important classifiers are discussed and analyzed. The styles of classifying nouns are also elicited and presented in tabular form.

CHAPTER FOUR

SEMANTIC CATEGORIES AND CLASSIFICATION

4.0 Introduction

This chapter deals with aspects of the semantics of noun categorization devices in Vietnamese. The semantic features encoded in noun categorization reflect principles of human cognition and how humans perceive the world (Aikhenvald 2000). Not only have classifiers been argued to offer a unique perspective on how humans construct representations of the world and encode them into their language (Lakoff 1986), they have also been thought to reflect how humans interact with the world (Denny 1976). We will review some general theories of the semantics of numeral classifier systems focusing in particular on Adams (1989), Frawley (1992), Croft (1994) and Aikhenvald (2000) before we examine Vietnamese classifiers and the styles of classification of animates and inanimates in Vietnamese. There are several semantic properties that we will come across when we look at the styles of classification for entities in Vietnamese culture. For example, the distinction [+animate] and [-animate] is is salient, as well as, [+human] and [-human]. For inanimate items, the properties of shape or dimensionality i.e. one-dimensional, two-dimensional and three-dimensional; orientation, flexibility, nature, and function are important in the system of categorization. The source of classifiers and the characteristics of the individual categories will also be included in the respective sections. The chapter ends with a summary highlighting the characteristics of the semantic

organization of the Vietnamese classifier system. Specifically, we are looking at Dimensions (C), (G), (I) and (J) of Table 2.1 in Section 2.5. They are:

- (i) Principles of choice, or assignment of noun categorization devices
- (ii) Degree of grammaticalization
- (iii) Semantic organization of the system
- (iv) Evolution and decay

4.1 General semantic properties of noun categorization devices

Classifiers are particular morphological means to signal the semantic classes that noun instantiate. In this section, we look at the specific properties that tend to be encoded by classifiers to determine the meaning of the nouns they classify. A number of basic semantic parameters tends to be encoded in different types of classifiers and we shall look at them one by one. These parameters fall in the classes of **Animacy, Sex and Gender, Kinship, Social Status, Nature, Physical Properties, and Function.**

4.1.1 Animacy

There are several properties that languages employ to determine linguistic animacy. Besides the straightforward biological criterion where things that are alive should be considered as animate, other properties that determine the relative animacy of an entity are influence over the execution or instantiation of an event, topicality, potency, cultural importance, and discourse salience. All languages apparently make a distinction between

animate and inanimate entities (See Allan 1977: 299 and Frawley 1992: 89), making the animate/inanimate division primary and universal (Adams 1986: 248). However, more often than not, we find that the distinction between humans and non-humans is more pervasive and this is reflected by Comrie (1989)'s universal scale of the encoding of animacy where encoding preference is from left to right:

Human > Animal > Inanimate

Languages such as Khasi (Mon-Khmer) (Adams 1986: 248), Tzeltal (Mayan) and Yurok (Ritwan, Algic) (Allan 1977: 299), and Indonesian (Macdonald 1976: 82-3, 87) show that a three-way distinction is widely used. Also, as Adams (1986) observes, a word for a body part is frequently used to encode the whole class of animates, especially of animals. The other interesting observation is that roundness has a universal role in the denotation of objects by shape and dimension (Frawley 1992). Within the human category, further distinctions are made in terms of sex, age, social status and kin relation. At times, certain humans and animals are more likely to be treated (morphologically) as more animate than other humans. This is usually a function of additional semantic properties. Humans or animals that are specific, proximal, and of high social status are usually more animate.

4.1.2 Sex and gender

Animate entities are characterized by biological sex which is a semantic property. Gender is a formal or coding property which does not necessarily have any connection with sex because it is purely a structural device with its own peculiar rules of operation (cf.

Chapter 1, Section 1.1). Sex is associated with animacy and humanness and many languages restrict the male/female distinction to the category of humans, leaving non-humans undifferentiated. In some languages, females are classed with non-human animates because they are perceived as less culturally significant entities. Children, adolescents, or individuals for whom sex is irrelevant or non-salient belong to the neuter class. This is well-reflected by Frawley (1992: 102), “a language can encode sex and does not do so completely across the class of human entities”.

4.1.3 Kinship

The next semantic property normally associated with animates is kinship: the familial relations among humans. Familial relations are intrinsic to humans thus only humans are differentiated in terms of kinship. Biological sex and age also come into play when considering kinship relations. Kinship is closely tied to social and cultural organization and the use of kin terms often extends to the pronominal system and classifier system and is unique among the Mon-Khmer systems (Adams 1989: 84). For instance, in Vietnamese, kin terms extend to non-kin in the pronominal system and classifier system. However, one should note that only a subset of kin terms can be used as pronouns and an even smaller subset used as classifiers. Aikhenvald (2000) also observes the use of kinship terms as a source of classifiers, especially noun classifiers, for Mayan languages (Kanjobalan branch), Australian, Austroasiatic, Austronesian, and a few South American languages e.g. Dâw (Makú).

4.1.4 Social status

Social status encodes the non-familial, social relations of entities and it is essentially a relation between two entities. It is not restricted to humans and can apply to animals, inanimates and even to situations. In all languages, there are ways of signaling relative social rank; one of which is through honorifics, which are direct grammatical encodings of relative social status. Social status markers are related to other semantic properties like age, sex, and nearness of kin. For instance, an older person (relative to the speaker) may trigger an honorific compared to a younger person. Social status can also be related to occupation where a doctor or a politician may command more respect than an ordinary blue-collared worker. Social status can have a pragmatic function to accord respect for a person. An honorific would be used to achieve the desired effect.

4.1.5 Nature

In the classification of entities, we sometimes distinguish them according to whether they occur naturally or they are artificially manufactured i.e. we are looking at the nature of these entities. Some of the more prominent natural classes are trees, plants, birds, insects, artifacts, weapons, etc. We see that these classes are significant by their ability to occur in individual classes because they are culturally important.

4.1.6 Physical properties

The properties of inanimate entities usually involve physical properties: (a) **Extendedness**, which subcategorizes dimensionality, shape and directionality (b) **Interiority**, (c) **Size**, and (d) **Consistency**.

Extendedness has to do with whether some entity occupies a physical position or not. It is a more basic quality which states that space is occupied and the way it is occupied is described by dimensionality and shape. There are three values for dimensionality i.e. one-dimensional, two-dimensional, and three-dimensional. The use of dimensionality here (and later in the thesis) calls for some qualification. It is virtually impossible for real-world objects to be one-dimensional (only a straight line would qualify, geometrically speaking). However, things having one salient dimension seem to be classified together. What this suggests is that dimensionality involves *cognitively salient* features, as opposed to ontological facts about the classified items in the real world (see Lee 1988: 228 for some discussion of classification from this perspective). Dimensionality is closely related to shape as extended one-dimensional entities are typically long, extended two-dimensional entities are typically flat, and extended three-dimensional entities are typically round or curved. However, there are times where dimensionality and shape diverge, giving rise to examples like ‘a long piece of cloth’ where a flat piece of cloth is extended one-dimensionally. In this example, the salient feature is the length of the cloth rather than how flat it is. The second subcategory of extendedness is direction where languages make a distinction between entities that are vertically extended and those that are horizontally extended.

The second physical property of inanimates is interioricity. This feature represents “the containedness of an entity or the way an entity differentiates its inside from its outside.” (Frawley 1992: 125). We will not go through this here because this property is not relevant to Vietnamese. (cf. Frawley 1992 for a detailed discussion).

Size is the third physical property and we distinguish only two values: large and small. A large number of languages distinguish between large and small entities, and often we find that size is encoded with shape i.e. small round objects.

The last major physical property that we look at is consistency. This property is characterized by two values i.e. flexibility and rigidity. Sometimes, consistency interacts with other properties like dimensionality to distinguish between flexible one-dimensional objects, rigid one-dimensional objects, and flexible two-dimensional objects etc.

4.1.7 Function

The final property to be considered in the classification of inanimate entities is function. Many languages have ways of eliciting the specific uses that entities have or the kinds of actions performed on them. Many of the function-based classes are language-specific and reflect the cultural values of the society and its speakers. Some common functional properties include: vehicular transport, housing, literary material etc.

4.2 Vietnamese (numeral) classifiers – basic features

Numeral classifiers are perhaps the most commonly recognized type of noun categorization device. According to Aikhenvald's typology of classifiers, some contingent properties of numeral classifier systems are (2000: 98):

- (I) The choice of a numeral classifier is predominantly semantic.
- (II) Numeral classifier systems differ in the extent to which they are grammaticalized and they can be an open lexical class.
- (III) In some numeral classifier languages, not every noun can be associated with a numeral classifier. Some nouns take no classifier at all; other nouns may have alternative choices of classifier, depending on which property of the noun is in focus.

In the previous chapter, we have seen how classifiers in Vietnamese operate within an attributive noun phrase (NP) and how they interact with other grammatical categories like number and demonstratives. We have also shown that classifiers in Vietnamese are obligatory in enumeration. From the typical numeral noun phrase in Vietnamese, we see the exemplification of various characteristics of a numeral classifier:

- (1) *hai con gà*
 two CL:ANIMAL chicken
 'two chickens'

Some examples of quantifying phrases in Vietnamese:

- (2) **nhều cuốn sách**
many CL:BOOK book
'many books'
- (3) **một số con chó**
few CL:ANIMAL dog
'a few dogs'
- (4) **mấy con mèo**
several CL:ANIMAL cat
'several cats'

Classifiers in Vietnamese often appear in referring noun phrases:

- (5) **con gà đó**
CL:ANIMAL chicken DEM
'that chicken'
- (6) **con bò này**
CL:ANIMAL cow DEM
'this cow'
- (7) **hai con gà đó**
two CL:ANIMAL chicken DEM
'those two chickens'

Another typical trait of numeral classifier systems that is evident in Vietnamese is that quite a number of classifiers are derived from nouns. This is acknowledged in Craig (1986: 6), "classifiers denote attributes of the referents of the nouns they are derived from, either directly or in a metaphorical manner". As this is a rather productive process, the inventory of classifiers in Vietnamese is quite large. Some of the most common

objects used metaphorically for the classification of objects by shape are plants, plant parts, and body parts (see Adams 1986 and Mithun 1986). An example of this would be *cây* in Vietnamese. *Cây* refers to ‘a tree’ or ‘a plant’ in Vietnamese and it is also used as a classifier to classify plant species and one-dimensional, rigid objects. For example:

- (8) *cây táo*
CL: TREE apple
‘apple tree’
- (9) *cây viết chì*
CL: 1-D pencil
‘pencil’
- (10) *cây đinh*
CL: 1-D nail
‘nail’

Drawing on our observations, we can be sure that there is a numeral classifier system within Vietnamese even though Vietnamese classifiers can occur in other contexts besides a numeral phrase. Another observation made by Greenberg (1972) about numeral classifier languages is that they typically lack obligatory plural marking. This is the case in Vietnamese. Greenberg (1972: 26) suggests that numeral classifiers are required in these languages because classifiable nouns without plural markings are like “collectives in their semantic non-specification of number and in their avoidance of a direct number construction”. Thus, Greenberg (1972) claims that the numeral classifier is an individualizer performing the same function as a “singulative derivational affix in languages with the singular/collective opposition”. However, this explanation may be insufficient to account for languages which have no plural markings for nouns and, at the

same time, are also lacking in classifiers. Also, as we have seen, certain nouns in a transnumeral¹⁶ language like Vietnamese, can be unclassified, for semantic or morphological reasons, thus having a direct number construction such as:

(11) *hai năm*
 two year
 ‘two years’

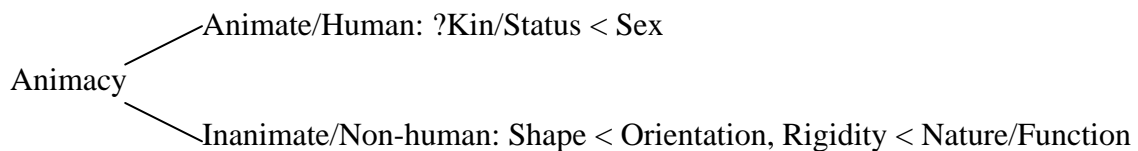
(12) *một chuyện*
 one story
 ‘a story’

4.3 Semantics of numeral classifiers

Since the choice of a numeral classifier is predominantly semantic, the classifier that is employed depends on how the entity is perceived. For example, in Section 4.1.7, we mentioned that classification by function is culture-specific. Similarly, for other types of classification, we depend on the speakers’ perception of the object being classified. Croft (1994) suggests that there is a different hierarchy of semantic distinctions associated with each type of classifier system and these distinctions can be accounted for by the function of the construction in which they are found. For numeral classifiers, the animate/inanimate, and human/non-human distinction is used. Further distinctions in the animate/human class are based primarily on age, sex, social status, and kinship relations (Adams and Conklin 1973). In the inanimate/non-human class, the primary distinction is shape, which is closely related to dimensionality. Shape and dimensionality should not be taken to mean the same thing because in some languages, there is a divergence between

¹⁶ This term denotes non-obligatory expression of nominal plurality.

shape and dimensionality. It appears that all numeral classifier systems that utilize dimensionality make a three-way distinction: one-dimensional (long or stick-like), two-dimensional (flat) and three dimensional (round). When talking about shape and dimensionality, some languages also differentiate objects by orientation or direction (vertical vs. horizontal). Secondary to shape, we have consistency (rigid vs. flexible) and finally, the distinction made is of nature or function.



Adapted from Croft (1994: 152)

Aikhenvald (2000: 306) lists the typical semantics of numeral classifiers where the properties of animacy, social status, kinship, directionality and orientation, physical properties, nature, quanta, arrangement, and functional properties are always important. Having reviewed in detail the basic features and theory of the semantics of numeral classifiers, we will look at the various modes of classification of animates and inanimates in Vietnamese and the source of classifiers used.

4.4 Modes of classification of animates in Vietnamese

As we can see from Table A, animates and inanimates are distinctively classified into different categories in Vietnamese with the exception of a few problematic categories which we will elaborate on later. Under animates, there is a further distinction between

humans and non-humans. Humans are usually classified by *người* ‘human, person’, *đứa*, *ông* ‘grandfather, gentleman, mister’ *bà* ‘grandmother, madam, missus’, and *ngài* ‘gentleman, mister’. Some examples of these classifiers are given below:

- (13) *hai người đầu bếp*
two CL:HUMAN chef
‘two chefs/cooks’
- (14) *hai đứa em*
two CL:LOW HUMAN younger sibling
‘two younger siblings’
- (15) *hai ông giáo sư*
two CL:HUMAN.MASC professor
‘two male professors’
- (16) *hai bà bác sĩ*
two CL:HUMAN.FEM doctor
‘two female doctors’

Non-human animates like insects, fishes, birds, mammals etc. are classified by *con* ‘child, young one’. However, the distinction blurs when we see that *con* can also be used for a certain group of humans and inanimates.

Table A
Vietnamese Classifiers for Animates and Some Inanimates

	Human: being friendly and polite	Human: young and/or less repected	Human: senior and/or respected + divine beings	Human: very repected	Human: disrespected + lower divine beings	Animals	Inanimates: possess life + bearing images of animates	Inanimates : body parts reflecting life	Inanimates: created for human use
Classifier	<i>Người</i> 'person, man'	<i>Đứa</i>	<i>Ông</i> 'grandfather' (masculine) <i>Bà</i> 'grandmother' (feminine)	<i>Ngài</i> 'gentleman, mister'	<i>Con</i> 'child, younger one'				
Examples	~ <i>bà</i> (grandmother) ~ <i>bạn</i> (friend) ~ <i>chồng</i> (husband) ~ <i>người diễn viên</i> (actor/actress) ~ <i>ông</i> (a man)	~ <i>con</i> (child) ~ <i>trẻ</i> (young person) ~ <i>cháu</i> (grandchild/ nephew/niec e) ~ <i>kẻ cắp</i> (pickpocket)	~ <i>chủ</i> (master/boss) ~ <i>bác học</i> (scientist) ~ <i>giáo sư</i> (professor) ~ <i>sư</i> (monk/priest) ~ <i>tổ</i> (ancestor) ~ <i>thần</i> (deity/divine being)	~ <i>tổng thống</i> (president) ~ <i>bộ trưởng</i> (minister)	~ <i>hát</i> (actor/actress) ~ <i>đĩ</i> (prostitute) ~ <i>tin</i> (hostage) ~ <i>hầu</i> (maid/servant) ~ <i>chiên</i> (Christian believer) ~ <i>con buôn</i> (merchant) ~ <i>ma</i> (ghost) ~ <i>quỷ</i> (devil/evil spirit) ~ <i>quỷ sứ</i> (demon) ~ <i>quái vật</i> (monster)	~ <i>bò</i> (bull/cow/ox) ~ <i>nhện</i> (spider) ~ <i>rắn</i> (snake) ~ <i>cá</i> (fish) ~ <i>giun</i> (earthworm) ~ <i>ếch</i> (frog) ~ <i>cóc</i> (toad) ~ <i>cua</i> (crab)	~ / <i>cái/chiếc</i> <i>tàu</i> (boat) ~ <i>sông</i> (river) ~ <i>suối</i> (stream) ~ <i>nước</i> (tide) ~ <i>đường</i> (road) ~ / <i>cái/chiếc</i> <i>diều</i> (kite) ~ <i>búp bê</i> (doll) ~ <i>rối</i> (puppet) ~ <i>quay</i> (spinning top) ~ <i>lắc</i> (pendulum)	~ <i>mắt</i> (eye) ~ <i>ngươi</i> (pupil of the eye) ~ / <i>trái/cái</i> <i>tim</i> (heart)	~ <i>bài</i> (a card in games) ~ <i>tem</i> (stamp) ~ / <i>cái dao</i> (knife) ~ <i>trỏ chuột</i> (pc mouse) ~ <i>dấu</i> (seal)

Within the class of humans, there are further distinctions made in terms of status, age, sex and other socially important factors. Hence, besides the generic human classifier *người* ‘human, person’, we have *đứa* which is used to classify human beings with low social status, and *ông* ‘grandfather, gentleman, mister’ or *bà* ‘grandmother, madam, missus’ for distinguishing people according to age and according respect. In the Vietnamese society, age is usually associated with social status. Hence, young children are always grouped together with people of low social status. This can be seen from the use of the classifier ‘*đứa*’ for both groups of people. Also, an old man is always accorded with respect by the use of *ông* ‘grandfather, gentleman’. However, these two features do not always coincide because we can use *ông* for someone young whom we respect e.g. bosses:

Table 4.0
Human Classifiers

Classifier	Classificatory Features
1. <i>người</i> ‘human, person’	[+human]
2. <i>đứa</i>	[+human, -status]
3. <i>con</i> ‘child, younger one’	[+human, -status]
4. <i>ngài</i> ‘gentleman, mister’	[+human, +high status, +masc]
5. <i>ông</i> ‘grandfather, gentleman, mister’	[+human, +status, +masc, (+age)]
6. <i>bà</i> ‘grandmother, madam, missus’	[+human, +status, +fem, (+age)]

There are other human classifiers but we will just provide a few of them in Table 4.0 and we will discuss the rest in Table 4.2. Most of the human classifiers are not marked for

gender with the exception of most kinship classifiers. The term is self-explanatory; it refers to classifiers that are derived from kin terms such as *ông* ‘grandfather’, *bà* ‘grandmother’, *cô* ‘aunt’, *chị* ‘elder sister’, *em* ‘younger sibling’, etc. which we will look at later.

Within the class of non-human animates, we see only the use of one classifier, *con* ‘child, young one’. For example:

(17) *hai* *con* *cua*
two CL:NONHUMAN ANIMATE crab
‘two crabs’

(18) *hai* *con* *ma*
two CL:NONHUMAN ANIMATE ghost
‘two ghosts’

This is a particularly interesting classifier that is not easy to characterize. We observe that *con* can also classify humans possessing some characteristics that are viewed as degrading according to the society’s value system, such as gamblers, prostitutes, drug addicts, merchants, etc. (see Table A). It is not difficult to see that due to semantic extension, a classifier used to classify animals and ghosts is used to classify humans for derogatory purposes such as:

(19) *một* *con* *đĩ*
one CL:DERG prostitute
‘a prostitute’

Con can also classify certain important body parts like *tim* ‘heart’¹⁷, *mắt* ‘eye’ and *ngươi* ‘pupil of the eye’ because of their closeness to human beings. These body parts can also be seen as more animate because they reflect certain aspects of the human soul.

- (20) **Con tim của tôi buồn**
CL heart POSS 1SG sad
‘My heart is sad’

Con can also classify items which are or depict images of animates such as kites (image of a bird), dolls (image of a human), chess pieces (carvings of different animals) and seals. Transport pathways like road, street, dike, river, and vehicles traversing waterways such as boats or sampans are classified by **con**. These entities appear to have a life of their own and behave like animates and this may be the reason why they are classified by an animate classifier even though they are inanimates.

- (21) **Em tôi có hai con búp bê**
Younger sibling 1SG POSS two CL doll
‘My younger sibling has two dolls’

- (22) **Con đường này rất ồn ào**
CL road DEM INT noisy
‘This road is very noisy’

A set of related items like digits, figures, playing cards, postage stamps, etc are also classified by **con** and this group of entities has the common property of being created by

¹⁷ Literary use. *Tim* ‘heart’ is more often classified by *trái* ‘fruit’ because of its shape (see Section 4.8.6-7).

humans for human use (but also they often picture humans/animates). The heterogeneity that we observe for *con* can be explained by looking at the various processes of semantic extension. In terms of prototypicality, some members in a given class are perceived as more salient and cognitively more central (Lakoff 1986, 1987). The peripheral members are associated with that class because they share at least one feature with one, or more, prototypical members i.e. chaining. According to Lakoff (1986: 17-8), systems of classification tend to be structured around these cognitive principles:

- 1) Centrality: The basic members of the category are central.
- 2) Chaining: Complex categories are structured by chaining: central members are linked to other members, which are linked to other members, and so on.
- 3) Experiential Domains: There are basic domains of experience, which may be culture-specific. These can characterize links in category chains.
- 4) Idealized Models: There are idealized models of the world – myths and beliefs among them – that can characterize links in category chains.
- 5) Specific Knowledge: Specific knowledge overrides general knowledge.
- 6) The Other: Conceptual systems can have an “everything else” category. It, of course, does not have central members, chaining, etc.

7) No Common Properties: Categories on the whole need not be defined by common properties.

8) Motivations: The general principles given make sense of a certain classification, but they do not predict exactly what the categories will be.

“In general, as the number of items counted by a classifier increases, the number of semantic connections between the classifier and the items grow fewer” (Craig 1986: 7).

Now, we look at the classifier **con** and try to come up with the semantic structure of **con**

with reference to the prototype approach and other extension principles listed above. **Con**

appears to have two central groups of members. The first covers all kinds of non-human animates i.e. reptiles, mammals, insects, birds, etc. This group is very salient because it encompasses anything that possesses a life, moves and yet non-human. In fact, lower supernatural beings are also counted by **con** because mythically they move but they are

no longer human. **Con** is not used for higher supernatural beings because they possess

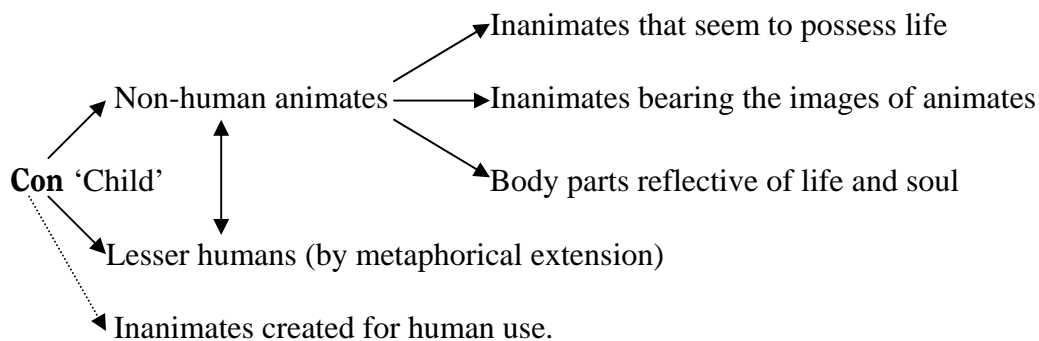
derogative meaning. The second group covers lesser humans. This can be argued as a metaphorical extension because **con** refers to ‘child’ or ‘younger one’ in nominal usage.

Even though we do not know the etymological source of the classifier, we can deduce

that since **con** refers to children, who are considered as less than adult humans and also low in social status, it is used to characterize people who are low in social status. Also,

the first central group serves to reinforce this principle since lesser humans are often compared to animals. Subsequently, we can extend the first central group to inanimates

that appear to have a life and/or have a likeness of animates. This would include members like transport pathways, kites, dolls, chess pieces, etc. This central group can also be extended to body parts that reflect certain aspects of the human soul and thus, are considered more animate. There is a group of peripheral objects that cannot be linked to the central groups. They are items like playing cards, postage stamps, figures, etc. Finally the semantic network would look something like this:



4.5 Source of human classifiers

In this section, we will look at the sources of some commonly used human classifiers such as *người* ‘human, person’, *ông* ‘grandfather, gentleman’, *bà* ‘grandmother, lady’, *em* ‘younger sibling’, *chị* ‘elder sister’, etc.

4.5.1 Source of *người*

The generic classifier for classifying ordinary people regardless of gender differences, and certain occupations such as worker, artisan, cook, rickshaw man, etc. (Nguyễn 1957)

is *người*. According to Adams (1989), the origin of this word is *ngươi* ‘eye’ and it is used today to refer to ‘human, person’. It is usually used when there is no reference or emphasis on the sex and age of the person one is referring to. This form may be used to refer to a man or a woman:

- (23) *Nhà anh có mấy người*
 house 2SG.MASC POSS several person
 ‘How many people are there in your family?’

We see in (23) that *người* in nominal usage is not classified by the generic classifier because the generic classifier is derived from *người* itself and the use of repeater constructions¹⁸ is not seen in Vietnamese. Hence, *người* is an unclassified noun in unmarked cases, as in (23) and (24a). However, one may argue that *người* in these two examples are classifiers instead of unclassified nouns and that the noun is the missing element because a derived classifier can, through metaphorical extension, guide us to the correct referent. We see that this is not a plausible analysis because if we replace *người* in (24a) with *đũa* (which only functions as a classifier), the sentence is ungrammatical as in (24b). This illustrates that *người* must be analyzed as a noun instead of a classifier.

¹⁸ A repeater appears when the specific object itself (or part of it) is used as a numerative, other terms are: self-classifiers, identical classifiers, auto-classifiers (Pe 1965: 166, Benton 1968: 115). An example from Thai (Hundius & Kolver 1983: 190):

a *prathêet sǎam prathêet*
 land three land
 ‘three countries’

- (24) a. ***Tôi là một người vui tính***
 1SG COP one person cheerful
 ‘I am a cheerful person’
- *b. ***Tôi là một đứa vui tính***
 1SG COP one person cheerful
- *c. ***Tôi là một người người vui tính***
 1SG COP one CL person cheerful

Also, if *người* is classified by another classifier, it is to emphasize certain qualities of the person we are talking about as in (26), (27), and (28). The literal meaning for (28a) and (28b) are the same but in (28a), the addition of the classifier *con* and *người* ‘person’ makes the sentence more formal and emphasizes the selfish quality of the person more so than (28b). The next most frequently used classifier is the generic classifier for animates, *con*.

- (25) ***Con người là một thực thể của xã hội***
 human being COP one entity POSS society
 ‘Human beings are an entity of society’
- (26) ***Tôi không muốn nói về (con) người đó***
 1SG NEG want say about CL person DEM
 ‘I don’t want to talk about that person!’
- (27) ***Anh ta là một (con) người xấu xa***
 3SG.MASC COP one CL person bad
 ‘He is a bad person’
- (28) a. ***Con người anh¹⁹ ta rất ích kỉ***
 CL person 3SG.MASC INT selfish
 ‘He is very selfish’

¹⁹ ‘*Con người anh ta*’ is a resumptive subject.

- b. *Anh ta rất ích kỉ*
 3SG.MASC INT selfish
 ‘He is very selfish’

The use of *con* with *người* is generally optional and depends much on the pragmatic functions of the utterance. For instance:

- (29) a. *Anh của tôi là một con người yêu nước*
 elder brother POSS 1SG COP one CL person patriotic
 ‘My elder brother is a patriotic person’
- b. *Anh của tôi là một người yêu nước*
 elder brother POSS 1SG COP one person patriotic
 ‘My elder brother is a patriotic person’

One of my informant mentions that (29a) is more formal and occurs more in written genre. Also, the quality possessed by the referent must be of significance and something great. Another informant says that (29a) is a marked form and we need a certain context to utter it. For instance, the elder brother is not very close to the family. It may be that if one wishes to emphasize the undesirability of a person or tries to distance oneself from the person he is referring to then the classifier *con* is used. This can be viewed as iconicity at work, i.e. a case where a real-world property of what is being denoted is reflected by the form of some linguistic expressions (Haiman 1980). In (29a), when there is emotional distance between the elder brother and the speaker, the linguistic form (which is a classifier in this case) is added to emphasize the emotional distance between the speaker and his brother.

Besides functioning as a noun and classifier, *người* is often used as a nominaliser whereby a verb or a so-called adjective, or a noun is combined with it to produce a noun.

Some examples of this:

(30) *người giúp việc*
 person help work
 ‘servant/assistant’

(31) *người bán hàng*
 person sell good
 ‘salesman’

(32) *người lạ*
 person strange
 ‘stranger’

4.5.2 Source of kinship classifiers

In Vietnamese, numerous kin terms function as classifiers. Adams (1989: 84) notes that:

“...their usage in the classifier system is related to their usage in the pronominal system in which kin terms refer to non-kin. The family in Vietnamese society is the major social unit whose structure is also used to order other kinds of social relationships.”

If we observe the use of kin terms in pronominal and classifier systems, we see that some of these kin terms are more frequently used than others. The set of kin based classifiers is smaller than kin based pronouns and two instances are already shown in Section 4.4. *Ông*

and *bà* are kin terms meaning ‘grandfather’ and ‘grandmother’ respectively. They are also used to refer to respected figures and elderly people in the pronominal system.

- (33) *Bà* *tôi nam nay tám mươi tuổi*
 grandmother 1SG year DEM eighty age
 ‘My grandmother is eighty years old’
- (34) *Chúng tôi cảm ơn ông* *về mọi sự giúp*
 1PL thank 2SG.MASC.RESP about every matter help
đỡ của ông
 support POSS 2SG.MASC.RESP
 ‘We thank him for all his help’

This semantic base is extended to that of the classifier system and hence *ông* and *bà* are used to classify a respected male and female individual respectively.

- (35) *(Ông) bác sĩ này rất tốt*
 CL doctor DEM INT good
 ‘This doctor is very good’

Besides that, *ông* is also used for male persons of a senior age and likewise for *bà*. It would be very strange to use *ông* to classify a very young but well-respected man.

Vietnamese speakers would usually use *chú* ‘father’s younger brother’ or *cậu* ‘mother’s brother’ because the element of respect is present without forcing senior age on the young person. This demonstrates that when a classifier is derived from a certain source, the semantic basis is usually retained and maintained closely within the classification system.

This can be explicated by taking *ông* as an example. When *ông* is used as a classifier,

just the family feature is dropped and since that feature is irrelevant then the feature of generation is also redundant. This produces our semantic base for *ông* and we see that every other feature is maintained quite closely:

Noun: **Ông**
 Family: 1sg
 Generation: +2
 Respect: +
 Sex: Masc

Classifier: **Ông**
 Respect: +
 Sex: Masc

The range of kinship classifiers is exemplified in the table below:

Table 4.1
Kinship Classifiers in Vietnamese

Kinship Classifiers	Gloss	Entities Classified
Ông	grandfather	thầy ‘master/teacher’, tổ ‘ancestor’
Bà	grandmother	bác sĩ ‘doctor’, giáo sư ‘professor’
Bác	father’s elder brother	công nhân ‘worker’, nông dân ‘farmer’
Chú	father’s younger brother	bộ đội ‘soldier’
Cô	father’s sister	ca sĩ ‘singer’
Cậu	mother’s brother	sinh viên ‘student’, học sinh ‘student’
Anh	elder brother	lính ‘soldier’
Chị	elder sister	giáo viên ‘teacher’
Em	younger sibling	sinh viên ‘student’
Con	child	người ‘person’, chó ‘dog’

4.6 Source of non-human animate classifiers

As mentioned in Section 4.4 above, *con* is used as a classifier for non-human animates and also beings of lower class, natural or supernatural. According to Adams (1989), the latter usage has the function of marking the group as “less than an adult” or “less than a healthy human adult who is not entirely in control/not morally competent” since *con* refers to ‘child’ in Vietnamese. However, from the list of inanimate entities that can also be classified by *con*, such as knife, stamps, and digits, it seems we have to be careful to claim that the classifier *con* is derived from the noun *con* ‘child’ even though there could be distant relations here such as ‘knife’ has some degree of animacy since it can kill; and stamps often portray faces or animals. It could be that *con* (CL) and *con* (N) both developed from the same source and the relation is lost in transition or we have missed out some important attributes that allows us to link these inanimates to the prototypical members in the group. It could also be that *con* (CL) and *con* (N) have different etymology. At this point, we do not have the answer but from our earlier discussion of the semantic structure of *con* in Section 4.4, we can claim that *con* (CL) is a metaphorical extension of *con* ‘child, young one’.

4.7 Characteristics of animate classification

The feature [+human] is very salient in Vietnamese and human classifiers such as *người*, *ông* and *bà* only classify humans and nothing else. Even so, not every human is eligible to be classified by the human classifiers as illustrated in the above sections and in the examples below:

- (36) a. ***con nghiện***
CL addict
'drug addict'
- *b. ***người nghiện***
CL addict
- (37) a. ***con mụ***²⁰
CL old hag
'shrew/old hag'
- *b. ***người/bà mụ***
CL old hag

Also, the use of classifiers is culturally dependent and hinges very much on the social values upheld within the society at a particular point in time. For instance, Emeneau (1951) includes actors and actresses as being classified by *con* i.e. *con hát*. This reflects the view that actors and actresses are considered as lowly occupations. However, in today's society, this is hardly the case as actors and actresses are accorded more respect by the use of the human classifier *người* and the name of the occupation has been

²⁰ When a noun denoting a human has an inherent derogative meaning, it cannot be classified by any of the human classifiers.

changed to *diễn viên* i.e. *người diễn viên* ‘actor/actress’. Within the [+human] group, there is a further distinction according to age and social status. An adult is distinguished from a child, and a president is classified differently from a hawker. This feature is found in other unrelated Southeast Asian languages like Thai and Burmese (see Becker 1975 and Chapter 6). Burmese examples (38) and (39) are taken from Daw (2000: 98):

- (38) *thanga nga: pa:*
 monk five CL
 ‘five monks’
- (39) *lu ta jau’*
 person one CL
 ‘a person’

In Burmese, monks, nuns, queens, kings, parents, spirits, etc. are classified by *pa:* whereas normal people of other kinds are classified by *jau’* (Goral 1978).

4.8 Modes of classification of inanimates in Vietnamese

In Vietnamese, as Croft’s (1994) semantic hierarchy predicts, inanimates are classified primarily according to dimensionality. Every concrete entity consists of three dimensions: length, width and height. Although theoretically every concrete object has three dimensions, some entities or some aspects of an entity are visualized as one, two, or three-dimensional due to human cognition and saliency of features (Frawley 1992). In addition to the one, two, and three-dimensional classifiers used in the classification of inanimates, we often see the use of a generic classifier *cái* which can sometimes be substituted for the above dimensional-salient classifiers without a change in meaning.

4.8.1 One-dimensional category

The category one-dimensional is taken to mean objects in which a particular dimension is salient for the purpose of classification. This saliency is not determined by absolute values but by the proportion of the related dimensions like length versus width and width versus height etc. As we can see from Table B, besides the primary cognitive basis, dimensionality, for classifying inanimates, we also see the interference of secondary parameters like rigidity (*cây* vs. *sợi*), shape (*ngọn*), nature (*gốc*), etc. in the organization of one-dimensional items in Vietnamese. One-dimensional items that we find in Vietnamese are usually salient in either length or height and they are classified by a number of different classifiers:

The central members of *cây* ‘tree, plant’ are plants and trees, and long and rigid stick-like items like pillars, pencils, pens and candles:

- (40) *một cây cột*
 one CL pillar
 ‘a pillar’

The central members of *sợi* ‘thread, fibre’ are long and flexible thread-like items like belts, strands of hair, ties and chains:

- (41) *một sợi dây lưng*
 one CL belt
 ‘a belt’

Table B
Vietnamese Classifiers for One-Dimensional Category

	Shape: peak	Nature: trees + Consistency: rigid	Consistency: flexible	Nature: trees	Nature: metal pieces	General
Classifier	<i>Ngọn</i> ‘top, peak, crest, summit’	<i>Cây</i> ‘tree, plant’	<i>Sợi</i> ‘thread, fibre’	<i>Gốc</i> ‘foot (of a tree), root’	<i>Thanh</i> ‘slat, bar’	<i>Cái</i>
Examples	tall trees and other tall objects with a peak shape. ~ <i>đồi</i> (hill) ~ <i>núi</i> (mountain) ~ <i>giáo</i> (spear)	plants and trees. ~ <i>táo</i> (apple tree) ~ <i>cam</i> (orange tree) ~ <i>viết chì</i> (pencil) ~ <i>kem</i> (stick of icecream) ~ <i>đinh</i> (nail) ~ <i>đèn</i> (lamp post) ~ <i>gươm</i> (sword) ~ <i>gỗ</i> (log of wood) ~ <i>cột cờ</i> (flagpole) ~ <i>cột</i> (pillar) ~ <i>chèo</i> (oar) ~ <i>đá</i> (ice block) ~ <i>thước kẻ</i> (ruler)	~ <i>râu</i> (hair on face) ~ <i>lông</i> (hair on body) ~ <i>bấc</i> (wick) ~ <i>dây</i> (thread/rope/wire/chain) ~ <i>tơ</i> (silk thread) ~ <i>dây lưng</i> (belt) ~ <i>chuỗi hạt</i> (necklace)	bushy trees especially fruit-bearing ones. ~ <i>cam</i> (orange tree) ~ <i>chuối</i> (banana tree) ~ <i>dừa</i> (coconut tree) ~ <i>xoài</i> (mango tree)	long, thin pieces of rigid metal including weapons. ~ <i>sắt</i> (iron) ~ <i>gươm</i> (sword) ~ <i>kim loại</i> (metal)	~ <i>cột</i> (pillar) ~ <i>chèo</i> (oar) ~ <i>dù</i> (umbrella) ~ <i>ống hút</i> (straw) ~ <i>đinh</i> (nail) ~ <i>riêng</i> (ax) ~ <i>gươm</i> (sword) ~ <i>chiếc cà vạt</i> (necktie) ~ <i>chiếc dây lưng</i> (belt) ~ <i>chuỗi hạt</i> (necklace) ~ <i>gân</i> (veins)

The central members of *ngọn* ‘top, peak, summit’ are tall objects shaped like the peak of a mountain like hills, mountains and spears:

- (42) *một ngọn núi*
 one CL mountain
 ‘a mountain’

The central members of *gốc* ‘root, source’ are bushy trees like orange trees, banana trees and mango trees:

- (43) *một gốc chanh*
 one CL lemon
 ‘a lemon tree’

The central members of *thanh* ‘slat, bar’ are long and thin pieces of rigid metal like iron and swords:

- (44) *một thanh gươm*
 one CL sword
 ‘a sword’

Cái is the generic classifier for inanimates so it appear in all categories of one, two, and three-dimensional objects. Since the use of *cái* is so pervasive, the items classified by semantically-specific classifiers can often be classified by *cái*. The specific classifier can even be replaced by *cái* in informal narratives (cf. Erbaugh 1986) or when the speaker is

not sure which is the correct specific classifier for an object. Some examples of one-dimensional objects classified by *cái* are axes, swords, oars and nails.

- (45) *một cái/cây chèo*
 one CL oar
 ‘an oar’

4.8.2 Source of one-dimensional classifiers

As we have seen in the previous sections, numerous classifiers in Vietnamese are derived from nouns. Examples of one-dimensional classifiers that are derived from nouns are *cây* ‘tree, plant’, *gốc* ‘foot (of a tree), *ngọn* ‘top, peak, crest, summit’ and *sợi* ‘thread, fibre’.

The examples below illustrates the nominal usage:

- (46) *Trong vườn có nhiều cây*
 PREP garden POSS many tree
 ‘There are many trees in the garden’
- (47) *Gốc cây cam có nhiều con kiến*
 foot tree orange POSS many CL ant
 ‘There are many ants at the foot of the orange tree’
- (48) *Trên ngọn/đỉnh núi đó có một ngôi chùa*
 PREP summit mountain DEM POSS one CL temple
 ‘There is a temple at the summit of that mountain’
- (49) *Áo này làm bằng sợi nhập khẩu từ Anh*
 shirt DEM make by thread import PREP England
 ‘This shirt is made from thread that is imported from England’

Since we have established that repeater constructions do not occur in Vietnamese, nouns from which classifiers are derived are usually unclassified and if they are classified, they are classified by another classifier that is semantically close to it, for example:

- (50) a. **một gốc cây**
one CL tree
'a tree'
- *b. **một cây cây**
one CL tree
- c. **một cây**
one tree
'a tree'

4.8.3 Two-dimensional category

This section focuses on the classification of two-dimensional inanimates: those that are salient in both length and breadth but not in depth or height. Again, the emphasis is that this saliency is not determined by absolute values but by the proportion of the related dimensions, length and width. In addition, secondary features like function (*bức*), nature (*tờ, cuốn, bản*), etc. also serve to further distinguish items in this category.

Looking at Table C, under two-dimensional classifiers, we have *tấm, bức, tờ, cuốn, bản, lá, cánh, đám* and *mảnh*. Their organization principles are captured in Table C and examples are given below:

Table C
Vietnamese Classifiers for Two-dimensional Category

	Shape: flat	Shape: leaf-like	Shape: wing or petal-like	Function: screening	Nature: natural landscapes and people	Nature: literate culture		
						Thin copies, sheets of paper	Thick bound copies, volumes	Printed copies
Classifier	<i>Tấm</i>	<i>Lá</i> 'leaf'	<i>Cánh</i> 'wing, petal'	<i>Bức</i>	<i>Đám</i> 'mass, patch' (large) <i>Mảnh</i> 'piece' (small land)	<i>Tờ</i>	<i>Cuốn</i> 'to roll up'	<i>Bản</i> 'copy, piece'
Examples	~ <i>lụa</i> (bolt/piece of silk) ~ <i>vải</i> (bolt/piece of fabric) ~ <i>thảm</i> (carpet) ~ <i>da</i> (piece of leather) ~ <i>màn</i> (curtain) ~ <i>sáo</i> (bamboo blind) ~ <i>gương</i> (mirror) ~ <i>gỗ</i> (flat piece of timber) ~ <i>bảng</i> (blackboard) ~ <i>ván</i> (plank/board) ~ <i>hình</i> (photo) ~ <i>thiệp</i> (card) ~ <i>vé</i> (ticket)	~ <i>bài</i> (a card in games) ~ <i>thiệp</i> (card) ~ <i>thư</i> (letter) ~ <i>cờ</i> (flag) ~ <i>chắn</i> (shield) ~ <i>gan</i> (liver) ~ <i>lách</i> (spleen) ~ <i>phổi</i> (lungs)	~ <i>cửa</i> (wing of a door) ~ <i>thư</i> (letter) ~ <i>buồm</i> (sail) ~ <i>chân vịt</i> (propeller blade) ~ <i>tay</i> (hand) *~ <i>đàn ông</i> (a group of man) *~ <i>đồng</i> (a cultivable area, piece of land)	~ <i>màn</i> (curtain) ~ <i>sáo</i> (bamboo blind) ~ <i>phên</i> (wall of bamboo/leaves) ~ <i>tường</i> (wall of brick/stone) ~ <i>vách</i> (wall of mud/wood) ~ <i>bình phong</i> (screen) ~ <i>thêu</i> (embroidery) *~ <i>ảnh/hình</i> (photo) *~ <i>thư</i> (letter) *~ <i>điện</i> (telegram)	~ <i>cỏ</i> (field of grass) ~ <i>rừng</i> (forest/jungle) ~ <i>vườn</i> (garden) ~ <i>đất</i> (piece of land) ~ <i>đồng</i> (cultivable area) ~ <i>bụi</i> (cloud of dust) ~ <i>mây</i> (cloud) ~ <i>sinh viên</i> (group of students) ~ <i>cưới</i> (wedding) *~ <i>vải</i> (bolt/piece of fabric)	~ <i>báo</i> (newspaper) ~ <i>tạp chí</i> (magazine) ~ <i>giao kèo</i> (contract) ~ <i>lý-lich</i> (curriculum vitae) ~ <i>hiệp-ước</i> (treaty) ~ <i>di chúc</i> (will) ~ <i>khai</i> (declaration/statement) ~ <i>trình</i> (a report) ~ <i>chép</i> (transcript)	~ <i>sách</i> (book) ~ <i>chuyện</i> (novel) ~ <i>nhật ký</i> (diary) ~ <i>sổ</i> (record book)	~ <i>báo cáo</i> (report) ~ <i>chép</i> (transcript) ~ <i>đàn/nhạc</i> (music score)

The central members of **tấm** include items that are flat regardless of the material e.g. cloth, leather, carpets, mirrors, boards, photos, cards, etc:

- (51) **một tấm kính**
one CL mirror
'a mirror'

The central members of **bức** include items that have the function of acting as a screen such as blinds, curtains and walls. However, it also classifies items like letters, photos and telegrams:

- (52) **một bức ảnh**
one CL photo
'a photo'

The central members of the next three classifiers **tờ**, **cuốn** 'to roll up', and **bản** 'copy, piece' include items that belong to literature culture. **Tờ** classifies thin, flat copies of documents perceived as a single sheet. **Cuốn** classifies thick, bound volumes of books, and **bản** classifies printed copies of documents usually with many sheets:

- (53) **một tờ giấy**
one CL paper
'a sheet of paper'

- (54) **một cuốn tự điển**
one CL dictionary
'a dictionary'

- (55) **một bản địa đồ**
 one CL map
 ‘a map’

The central members of **lá** ‘leaf’ are items that resemble the shape of leaves e.g. cards, flags and shields. Body parts that look like leaves e.g. lungs and livers are also classified by **lá**:

- (56) **một lá thiệp**
 one CL card
 ‘a card’

The other shape classifier is **cánh** ‘petal, wing’ and the central members are items that resemble the shape of petals and/or wings. They are sails, letters and hands. It is also used as a mensural classifier for collections of people.

- (57) **một cánh buồm**
 one CL sail
 ‘a sail’

The central members of **đám** ‘mass, patch’ include wide and vast natural landscapes e.g. grass fields, jungles, gardens and land. For small patches, we can use the classifier **mảnh** ‘piece, plot’ instead. **Đám** can also refer to groups of people at certain ceremonies like weddings or funerals. **Mảnh** can also refer to items like paper and cloth, or even feelings.

- (58) *một đám cỏ*
 one CL grass
 ‘a field of grass’
- (59) *một mảnh đất*
 one CL earth
 ‘a plot of land’

4.8.4 Source of two-dimensional classifiers

Some of the forms in Vietnamese are Sino-Vietnamese and ultimately of Chinese origin. An instance of this would be *cuốn*, the classifier for thick bound books. According to Gage (1979: 4), this form refers to a roll of books. This form of books is associated with Chinese culture where scrolls were common²¹. Considering the amount of literary influence Chinese has on Vietnamese, this form for books is most likely to be of Chinese origin. Besides forms of Chinese origin, we find forms of Mon-Khmer origin in all sections of the classificatory system of Vietnamese. For two-dimensional objects, one Mon-Khmer cognate is *lá*. According to Adams (1989), forms which mean ‘leaf’ are found throughout Mon-Khmer as classifiers and in many of these languages, the classifier subsumes the same kinds of items for example, leaves, papers and plates. We will look at this in more detail in Chapter 6.

²¹ In classical Chinese, books are referred to as *yi juàn shu* ‘one roll book’.

4.8.5 Characteristics of two-dimensional classification

In Vietnamese, the classificatory principles are rather transparent and straightforward. When a two-dimensional object is mentioned, the classifier for that object can usually be predicted based on the physical characteristics alone. For instance, *tấm* can be used with any material that has the feature [+flat] like silk, cloth, leather, wood, glass or paper because *tấm* denotes something like a flat piece of material. Some two-dimensional items can also be classified by the generic classifier *cái* without a change in meaning. For example, a blackboard and a mirror can be classified by *cái* instead of *tấm*; a card can be classified by either *cái* or *lá*. However, when *hình* ‘photograph, drawing, painting’ is classified with *cái* instead of *tấm*, *cái hình* denotes ‘a likeness, appearance, image’ in contrast to *tấm hình*, which denotes a photograph. *Bức* is a classifier for things that function like a wall for example a curtain, a wall or a blind, or things hanging on the wall like a painting or photograph. However, it is interesting to note that a blackboard is not classified by *bức*. This could be due to the fact that a blackboard was not originally hung on the wall and, in ancient times, it usually stood on the ground.

Under two-dimensional inanimates, some objects are further classified by the property of shape and these are categorized by the classifiers *lá* and *cánh*. *Lá* which is derived from ‘leaf’ can classify any two-dimensional items that resembles the shape, thickness or texture of a leaf, including body parts.

- (60) *Tôi đang viết la thư cho mẹ tôi*
1SG PROG write CL letter give mother 1SG
'I am writing a letter to my mother'

Cánh which is derived from 'wing' or 'petal' has a similar function.

- (61) *Tàu này có hai cánh buồm*
Boat DEM POSS two CL sail
'This boat has two sails'

The things that *đám* classifies are quite different from the rest of the classifiers in this class because it classifies extensive natural landscapes instead of individual entities. The use of plant parts as classifiers for inanimates is a common feature within the languages in the Mon-Khmer family (see Adams 1989).

Lastly, there are specific classifiers for lands, fields and gardens in Vietnamese which show that these entities are central in the Vietnamese society. For example, lands, fields and gardens are classified by *đám*. If their surface area is small, then *đám* can be replaced with *mảnh* as in (59) above.

4.8.6 Three-dimensional category

The use of the term three-dimensional needs explication since we have seen in the preceding sections that there are a number of objects that are classified as in terms of only one or two dimensions although they have length, breadth and height. In our analysis of

the classification of three-dimensional inanimates, this term is taken to apply to round or spherical items as well as bulky items including cube shapes. In other words, objects which are visualized as extending roughly equally in length, width and height are considered here as three-dimensional (i.e. there is no clear perception of one of the dimensions being salient, which is typically what otherwise determines 1/2-dimensional classifiers).

Under three-dimensional classifiers, we have *hòn*, *viên*, *cục*, *củ* and *trái*. Their organization of principles are captured in Table D and examples are given below:

The central items of *hòn* ‘ball, piece’ are items that belong to the class of islands and stones. For instance, pebbles, marbles, rubies, jade, and brick are classified by *hòn*. It also classifies items that are perceived to be almost round:

- (62) *một hòn bi*
 one CL marble
 ‘a marble’

The central items of *viên* ‘pill, to be round’ are items that are very smooth and rounded. To a large extent, this category of items includes artificial objects or artificially-shaped objects like meatballs, pills, bullets and fishballs since these items can be made so perfectly smooth or rounded.

Table D
Vietnamese Classifiers for Three-Dimensional Category

	Nature: islands/stones + Shape: round	Shape: artificially rounded + Size: small	Nature: fruits + Shape: fruit- like	Size: small	Size: big	Size: big	Nature: underground roots/fruits	General
Classifier	<i>Hòn</i> ‘ball piece’	<i>Viên</i> ‘pill, tablet’	<i>Trái</i> ‘fruit’	<i>Cục</i> ‘lump, piece’	<i>Tảng</i> ‘slab, block’	<i>Khối</i> ‘block, mass, bulk’	<i>Củ</i> ‘tuber/bulb’	<i>Cái</i>
Examples	~ <i>cù lao</i> (island) ~ <i>núi</i> (mountain) ~ <i>đá</i> (small stone) ~ <i>nước đá</i> (ice) ~ <i>ngọc</i> (precious stone) ~ <i>sỏi</i> (pebble) ~ <i>bi</i> (marble) ~ <i>gạch</i> (piece of broken brick) *~ <i>máu</i> (clot of blood) ~ <i>dái</i> (testicle)	~ <i>viên chả</i> (minced meat) ~ <i>thịt băm</i> (meatball/hamburger) ~ <i>viên gạch</i> (brick) ~ <i>đá</i> (small ice/stone) ~ <i>viên ngói</i> (tile) ~ <i>viên mực</i> (a tablet of Chinese ink) ~ <i>ngọc</i> (precious stone) ~ <i>kim cương</i> (diamond) ~ <i>bi</i> (marble) ~ <i>thuốc</i> (pill) ~ <i>cá viên</i> (fishball) ~ <i>đạn</i> (bullet) ~ <i>trụ</i> (cylinder)	~ <i>bưởi</i> (pomelo) ~ <i>dưa</i> (cucumber) ~ <i>dứa</i> (pineapple) ~ <i>dừa</i> (coconut) ~ <i>banh</i> (ball) ~ <i>cầu</i> (shuttlecock) ~ <i>bom</i> (bomb) ~ <i>cật</i> (kidney) ~ <i>tim</i> (heart) ~ <i>cật</i> (kidney) ~ <i>vú</i> (breast) ~ <i>cổ</i> (voice box) *~ <i>trái núi</i> (mountain) *~ <i>đất</i> (the globe, the Earth) *~ <i>pháo</i> (cannon)	~ <i>đá</i> (medium-sized ice/stone) ~ <i>đất</i> (lump of earth) ~ <i>gạch</i> (piece of brick) ~ <i>huyết/máu</i> (clot of blood) ~ <i>tẩy</i> (eraser)	~ <i>đá</i> (big stone) ~ <i>thịt</i> (meat) ~ <i>băng</i> (glacier)	~ <i>óc</i> (brain) ~ <i>bê tông</i> (cement) ~ <i>đá</i> (big stone) *~ <i>tình</i> (immense feeling – literary)	~ <i>cà rốt</i> (carrot) ~ <i>hành</i> (onion) ~ <i>khoai lang</i> (sweet potato) ~ <i>khoai tây</i> (potato) ~ <i>lạc</i> (peanut) ~ <i>tỏi</i> (garlic) ~ <i>từ</i> (yam) ~ <i>cải</i> (white radish) ~ <i>cải trắng</i> (turnip) ~ <i>chuối</i> (bulb of banana tree)	~ <i>khinh khí cầu</i> (hot-air balloon) ~ <i>bóng đèn</i> (light bulb) ~ <i>chuông</i> (bell) ~ <i>nốt</i> (pimple) ~ <i>nút</i> (button)

- (63) ***một viên kim cương***
one CL diamond
'a diamond'

The central items of ***cục*** 'lump, piece' are items that are rather small and yet not too round. Hence, quite a wide range of objects can be classified by this classifier such as stones, blood clots, erasers and bricks:

- (64) ***một cục đường***
one CL sugar
'a cube of sugar'

The central items of ***củ*** 'tuber, bulb' are underground roots or fruits. For example, carrots, onions, potatoes, peanuts, garlic and yam:

- (65) ***một củ gừng***
one CL ginger
'a piece of ginger'

The central items of ***trái*** 'fruit' are fruits or items shaped like a fruit. This includes all kinds of fruits, and some gourds and melons. Items shaped like fruits include balls, bombs and shuttlecocks. It also includes various body parts like the heart, kidney and breasts. The only item which seems odd in the category is mountains. According to one of my informants, in the Vietnamese society, mountains are perceived as having a fruit-like shape from afar.

- (66) **một trái bom**
one CL bomb
'a bomb'

The central items of **tảng** 'slab, block' are items that are big and not too round such as stones, meat and glaciers:

- (67) **một tảng đá**
one CL stone
'a block of stone'

The central items of **khối** 'block, mass' are items that are similar to those classified by **tảng**. The only difference is that it can also classify brains and feelings (literary use).

- (68) a. **một khối đá**
one CL stone
'a block of stone'

4.8.7 Source of three-dimensional classifiers

Vietnamese has two classifiers derived from plant parts, **củ** and **trái**. **Củ** refers to the tuber or bulb of a plant, and in the classifier usage, it can only classify plants that have tubers or bulbs like carrots, potatoes, onions, ginger etc. **Trái** refers to fruits in general and it has a wider classificatory scope that includes fruits, things and body parts that

resemble a fruit, and mountains and hills²². It is interesting to note that the use of plant parts as classifiers is a pervasive phenomenon in Mon-Khmer languages (Adams 1989).

4.8.8 Characteristics of three-dimensional classification

The classifiers for three-dimensional items show some overlapping and the items classified by *hòn*, *viên* and *cục* are rather similar. For example, stones, ice and precious stones can be classified by *hòn*, *viên* and *cục* with minimal changes in meaning. According to Võ P. (1988), there are many classifiers for stone. For instance, stone can be classified by *cục* (CL: small 3-D things), *hòn*, *viên*, *tảng* (CL: big 3-D things), *khối* (CL: big 3-D things). However, not all items can be classified by different classifiers without a change in meaning. When a brick is classified by *hòn*, it depicts a piece of broken brick. If it is classified by *viên*, it means a brick as a whole (Emeneau 1951). This suggests that *hòn* can be used for a part of a thing whereas *viên* is used to classify the whole object, especially for a small rounded object. The most important characteristic of the objects that *viên* classifies is that the objects must be whole, mostly round, and small (*viên* is derived from ‘pill’ or ‘tablet’ in Vietnamese but it is of Chinese origin meaning ‘to be round’) and it must be artificially shaped.

²² It is not clear as to why hills and mountains are perceived as fruits and counted by *trái*.

Cục is quite similar to *viên* in the things that it classifies, however the most distinctive difference between the two is the feature [+round]. *Cục* has the meaning of ‘piece, small lump’ which indicates irregular shape. Hence, while it classifies small three-dimensional items like *viên* does, the objects are mostly of an irregular shape and not round. Both *tảng* and *khôi* have very limited usage being used only for stones and for *khôi*, brains as well.

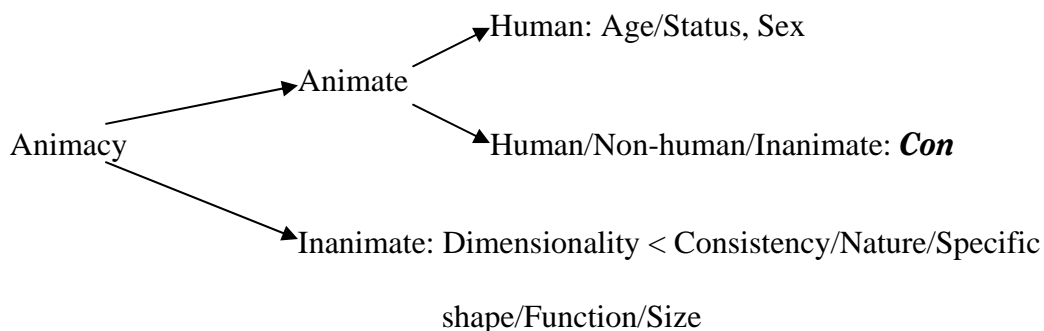
4.9 Conclusion

From the above semantic account of the Vietnamese classifier system, we can summarize several characteristics:

- (I) There is a very sophisticated animate classification where different degrees of differentiation according to age and status occur within the class of human beings.
- (II) Kin relations are used widely in the system.
- (III) The semantic tie between the lexical source and the classifier is often preserved. Hence, the origin of the classifier is often transparent.
- (IV) Many classifiers are derived from a nominal source and we do not have any instance where a classifier is derived from a verb. This suggests that the classifier denotes the objects being classified.

- (V) Plant parts are used as classifiers and this occurs repeatedly within the one, two, and three-dimensional inanimate classification. This characteristic is noted in Adams (1989) where it is prevalent in languages of the Mon-Khmer subbranch.
- (VI) The classificatory principles for most animates and inanimates are quite straightforward and we can characterize the items classified according to the saliency of physical dimensions.

To a large extent, the Vietnamese numeral classifier system adheres to the implicational hierarchy of semantic distinctions developed by Croft (1994). As mentioned briefly in Section 4.3, we can explain this hierarchy of semantic distinctions by referring to the function of numeral classifiers. Numeral classifiers are used in enumeration which involves the cognitive processes of individualization of units and identification of those units as being of the same kind. It is argued that the individualization, categorization, reference tracking and anaphoric functions of numeral classifiers interact in some ways to give rise to various significant parameters that are consistently found in most numeral classifier systems. These various functions will be looked at in greater detail in Chapter 5. In the case of Vietnamese specifically, our hierarchy of semantic distinctions might look like:



If we look at Tables A, B, C, and D again, we see that two semantic properties can be combined to characterize the items being classified by a particular classifier. For instance in the classification of inanimates, we find that the property, *Nature*, occurs with *Shape* for the classifiers *hòn* ‘ball, piece’ and *trái* ‘fruit’, and with *Consistency* for the classifier *cây* ‘tree’. When Vietnamese speakers are asked to classify an item, they look at the primary shape which is actually *Dimensionality*. If the item is one-dimensional and is a plant, they can either classify it using *cây* or *gốc*. If the item is not a tree but it is rigid, we can also use *cây*. The property *Shape* also occurs with *Size* for the classifier *viên* ‘pill, to be round’. We go on to look at the functions of Vietnamese classifiers in the next chapter.

CHAPTER FIVE

FUNCTIONAL ASPECTS

5.0 Introduction

In this chapter, we look at the functions of Vietnamese classifiers in the various constructions (as discussed in Chapter Three) they occur in. We will highlight several important points relating to the functions of classification, individualization, referentialization, and relationalization, as discussed in Bisang (1999) and we also discuss how these principles can be applied to Vietnamese. Subsequently, the semantic notions of specificity and definiteness are discussed with reference to classifier constructions in Vietnamese. We will also look at how the anaphoric and lexical functions of classifiers manifest themselves. Polyfunctionality is discussed in Ansaldo (1999) and Enfield (2003), as languages in Southeast Asia appear to have lexical items that display numerous functions. Enfield (2003) looks at the verb ‘acquire’ in Southeast Asian languages and Ansaldo (1999) looks at comparatives in Sinitic and Southeast Asian languages. The notion of polyfunctionality can be applied to classifiers, especially within the Southeast Asian region, as we see that this grammatical category fulfill a variety of functions e.g. referentialization, relationalization, marking specificity, etc. in the Sinitic and Southeast languages studied and they are not redundant (cf. Pacioni 1996, Matthews & Pacioni 1998, Bisang 1999).

5.1 Functions of Vietnamese classifiers

According to Bisang (1999), there is a high degree of indeterminateness in nouns of many Southeast Asian languages, including Vietnamese. This means that nouns in these languages express only a mere concept of objects which can be further specified for different interpretations. For instance, a noun like *chó* can mean ‘dog’ in (1) or ‘dogs’ in (2) according to the context that we specify:

- (1) *Đây là con chó của tôi*
DEM COP CL dog POSS 1SG
‘This is my dog’
- (2) *Hai con chó của tôi dễ thương*
two CL dog POSS 1SG cute
‘My two dogs are cute’

This example illustrates that Vietnamese does not have “compulsory expression of nominal plurality” (Greenberg 1972: 25) and is a case of transnumerality. According to Greenberg (1972), a noun in a transnumeral language cannot occur in immediate combination with a numeral so a classifier is first needed to individualize the noun before we can count it. As a direct consequence of this observation, the core function of classifiers has always been “to make nouns enumerable by individualizing and classifying them”. However, other studies have shown that this is not the only function that classifiers serve to fulfill in many languages (see Aikhenvald 2000, Bisang 1999, Lobel 2000, Matthews & Pacioni 1998). For instance, Bisang (1999) suggests that there is a wide range of functional variation among language specific classifier systems in languages such as Japanese, Chinese, Vietnamese, Thai, Cantonese and Hmong. Bisang

shows us that it is possible for classifiers to be maximally used for the functions of individualization, classification, referentialization, and relationalization.

5.2 Classification, individualization, referentialization and relationalization

As we have mentioned in the earlier sections, individualization and classification are the core functions of classifiers and they apply to all classifier languages. In order to carry out classification, we need to find a set of properties which constitute a certain class and from there we can select an object or a multitude of objects to be included in that class. The operation of classification can be used for two purposes as pointed out in Bisang (1999: 115):

“Classification can be employed to compare one particular sensory perception and its properties to the properties of other sensory perceptions in order to identify that particular perception by subsuming it under a certain concept or it can be employed to establish a sensory perception as an individuum by actualizing the inherent properties which constitute its conceptual unity.”

The first purpose of classification can be stated simply as identification and the second purpose is individualization. The difference between identification and individualization is that, in the former, it is possible to identify a sensual perception without explicitly referring to its inherent properties whereas, in the latter, one has to refer to the inherent properties of a sensual perception and identify it before individualizing it (Bisang 1999).

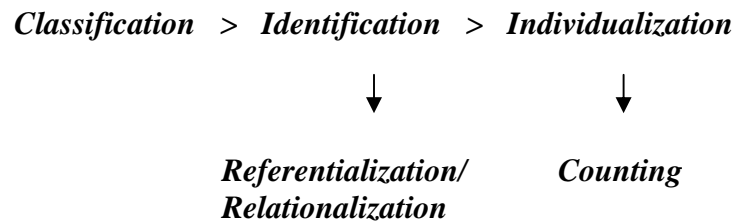
Since the operation of classification is required before performing identification and individualization, Bisang (1999:116) has established the following hierarchy:

Classification > Identification > Individualization

Classification is seen as “a pre-requisite to identification” (Croft 1994: 161) and a sensory percept has to be identified before it can be individualized. Thus, classification must first be established before identification and individualization. When we classify entities, we have to elicit the properties which determine the semantic criteria of inclusion in the classification system. Some of these properties are gender, animacy, dimensionality, functionality, etc. These properties form a hierarchy of semantic distinctions where the first distinctions made in classifier systems are almost universally between animates and inanimates, and humans and non-humans (Adams & Conklin 1973, see Chapter Four). The semantic criteria of classification and hierarchy of semantic distinctions underlying the Vietnamese classifier system has already been covered in Chapter Four.

It is pointed out in Bisang (1999) that the functional expansion of classifiers into the marking of reference or the marking of possession and relative clauses can be seen in the context of identification. If reference is defined as “an act of identifying some entity that the speaker intends to talk about” (Bisang 1999: 116 from Croft 1991: 110), then the function of identification marks the point where classifiers take on the function of referentialization. As for relationalization, one can understand that we have to first identify the head noun before it can be modified by a possessor or a relative clause.

Hence, the function of relationalization is also subsumed under the process of identification. The functional expansion of classifiers is thus schematically represented in Bisang (1999:116) as follows:



With these functions, Bisang (1999:117) suggests the following typology where there are four different types of classifier languages grouped according to their functions and some examples of these languages are given:

I. Classification & individualization

- Japanese (classifiers occur only with numerals)
- Chinese (classifiers occur with numerals and demonstratives)
- Vietnamese (individualization, but not necessarily in the context of counting)

II. Classification & individualization & referentialization

- Thai (secondary function in combination with adjectives [stative verbs] in the sequence N-CL-ADJ)

III. Classification & individualization & relationalization

Cantonese (classifiers can be used in possessive and relative constructions)

IV. Classification & individualization & referentialization & relationalization

Hmong (with referentialization being a secondary function)
Weining Miao

This typology offers a different perspective to other typologies where classifiers are distinguished according to morphosyntactic properties, semantic properties, etc. (e.g. Dixon 1982; 1986, Allan 1977, Croft 1994, Aikhenvald 2000). Also, we see that Bisang (1999) has classified Vietnamese as having the functions of classification and individualization only. This does not entirely capture the wide range of functions that Vietnamese classifiers have as we will see below.

5.3 Individualization and counting

In the earlier section, we have seen that in classifier languages, classifiers are often needed to individualize a noun before it is enumerable. This is due to the high degree of indeterminateness of nouns in these languages. Classifiers “individualize a given count noun by designating its semantic boundaries” (Bisang 1999: 120) or “by designating its natural unit” (Croft 1994: 163) before we can count it. Classifiers only occur with counting whereas quantifiers occur with measuring and apply to all languages of the

world. Even though both types of numeratives²³ have the function of individualization, there is a difference semantically. The primary purpose of quantifiers is to put a given entity into a unit of measure to be counted whereas “classifiers actualize the semantic boundaries which already belong to the concept of a given noun” (Bisang 1999: 121). Bisang (1999) calls the former *creative individualization* which is based on external units of measurement, and the latter *actualizing individualization* which is based on inherent properties of the noun. To illustrate this, we have the following example from Vietnamese:

- (3) ***Tôi có hai ly nước***
1.SG have two Q:glass water
'I have two glasses of water'
- (4) ***Tôi có hai con mèo***
1.SG have two CL cat
'I have two cats' (***con*** 1. CL for non-human animates)

In (3), the unit of measurement created for water is ‘glass’ hence we are able to count the number of glasses of water. However in (4), we do not need to create a unit to enable us to count cats. Instead, we identify that a cat is a type of animal and we actualize the semantic properties of an animal i.e. [+animate, -human] by the use of ***con***. In (3), we see that we can change the unit of measurement because we decide what to measure water with or what sort of container water come in. For instance,

²³ Cover term for both classifiers and quantifiers (Pe 1965).

- (5) *Tôi có hai lít nước*
 1.SG have two Q:litre water
 'I have two litres of water'
- (6) *Tôi có hai xô nước*
 1.SG have two Q:bucket water
 'I have two buckets of water'
- (7) *Tôi có hai chai nước*
 1.SG have two Q:bottle water
 'I have two bottles of water'

However, it is impossible for us to use another classifier to replace *con* in (3) because the salient inherent properties i.e. [+animate, -human] of a cat do not change. Therefore, the use of *con* can be extended to other entities that possess the same inherent qualities as a cat which basically covers all animals.

In the context of enumeration, we have to first establish a set of properties which constitute a certain class i.e. classification, then we identify an element as belonging to that class i.e. identification, and finally we establish that that element is relevant for counting based on its inherent properties i.e. individualizing. These operations involved in counting constitute the core functions of classifiers. To illustrate these operations in Vietnamese, we have the context where oranges are being counted. The class we have is [+fruit], we identify that an orange is a type of fruit because it is bore from a tree and we include it when we are counting the elements belonging to the class [+fruit]. So we would have to use the classifier *trái* because the semantic criterion [+fruit] matches the semantic property of an orange:

- (8) *Mẹ tôi mua ba trái cam*
 mother 1.SG buy three CL orange
 ‘My mother buy three oranges’ (*trái* 1. fruit, 2. CL for fruits & 3-D round objects)

However, there are many instances where classifiers are also required in Vietnamese for purposes other than counting. For example, classifiers can be used with or as demonstratives, with possessives or in nominal phrases. In the following sections, we will look at the various functions Vietnamese classifiers display under these different contexts and syntactic constructions.

5.4 Nominal phrases and specificity

Let us look at the following examples from Chapter Three, Sections 3.5.6 and 3.5.7:

- (9) *Căn/cái nhà²⁴ rất cũ*
 CL house INT old
 ‘The/a particular house is very old’

In the above examples, we see that Vietnamese classifiers can precede a noun outside the context of enumeration i.e. [CL N]. As a result, many grammars have tried to equate classifiers in Vietnamese to definite or indefinite articles in English (cf. Võ P. 1988: 27). However, when we look closer at the data, we will see that Vietnamese is one of the languages which does not distinctively differentiate definites from indefinites but marks the specific/nonspecific distinction instead. The function of classifiers in such nominal

²⁴ A house can be classified by either *căn* or *cái*

phrases is to indicate specificity rather than definiteness as noted in Frawley (1992), Matthews & Pacioni (1998), Pacioni (1996). For instance, (9) can be interpreted as definite or indefinite but it must be specific. Specificity is an independent semantic property which can be marked by tense, mood, definiteness, context of utterance etc. According to Frawley (1992: 70), “Specificity implies that an entity is uniquely determined” and Givón (1984) points out that specificity is a matter of the degree of referential accessibility of an entity in a projected world. Let us look at the following example which is taken from Bisang (1996: 541):

?(10) ***Tôi mua quả cam***
 1SG buy CL orange
 ‘I buy the orange/I buy an orange’

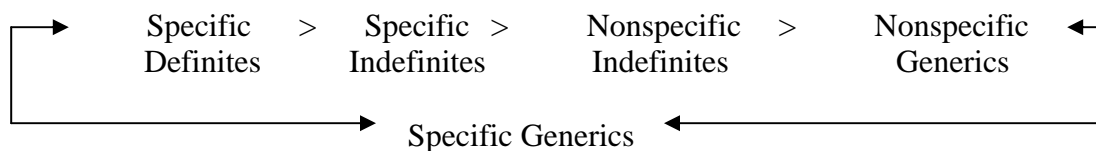
Example (10) is claimed to be “systematically ambiguous with regard to definiteness vs. indefiniteness” (Bisang 1999: 146) which would support the argument that Vietnamese classifiers do not mark definiteness. However, I found that the example above is considered unclear and uncommon by speakers of Vietnamese. When we want to buy oranges as opposed to apples or bananas, we will not include the classifier because we are talking about a type of fruit in general i.e. generic, and we do not know the number we are buying (see example 11a). It is only when we need to refer to certain aspects of the orange such as the number or the position in respect to the speaker that we need to include the classifier (see example 11b):

- (11) a. ***Tôi mua cam***
 1SG buy orange
 ‘I buy orange/s’
- b. ***Tôi mua một quả cam***
 1SG buy one CL orange
 ‘I buy an orange’ (*quả* is the same as *trái*)
- c. ***Tôi mua quả cam này***
 1SG buy CL orange DEM
 ‘I buy this orange’

Even though definiteness and specificity are two distinctive semantic properties, they are related in the sense that both subscribe to referential accessibility²⁵:

“The unifying factor in all this is that definites tend to be *known*, and if they are known, they tend to be *referentially accessible*; if they are referentially accessible, they tend to be *specific*; if they are specific, they tend to be *definite*.” (Frawley 1992: 76)

Frawley (1992: 77) proposes a modified version of Givón (1984)’s account of the coding of nouns with regard to their definiteness and specificity in the form of a scale given below:



Hence, (11a) is nonspecific generic, (11b) is specific indefinite, and (11c) is specific

²⁵ Maybe this is why the notions of definiteness and specificity are thought to be the same thing.

definite. (11b) is specific because when the numeral ‘one’ is used, the entity is “itemized, counted, delimited, and specific” (Frawley 1992: 76). The same goes for (11c). If a demonstrative is chosen, the entity is “singled out contextually” hence attributing specific reference (Frawley 1992: 76). Therefore in Vietnamese, specifics are coded similarly by the use of classifiers and nonspecifics are less likely to occur with a classifier.

We will look at some more examples below:

- (12) **Tôi muốn mua con gà**
 1SG want buy CL chicken
 ‘I want to buy chickens’
- (13) **Con khỉ rất thông minh**
 CL monkey INT clever
 ‘Monkeys are very clever’

In the above examples, we see that even though we are talking about nonspecific generics, we have to include the classifier **con** before the noun. This happens regardless of the position of the classifier phrase, whether it is in the subject or object position. It seems to indicate that different classifiers have varying degrees of closeness with the nouns they classify. The semantic link between **con** and its referents is very strong and it is only in special contexts that a **con**-entity occurs without **con**. At the other extreme, we have the generic classifier **cái** which has a rather weak semantic link with its referents. For instance in (14) and (15), **ghế** ‘chair’ can be optionally classified by **cái** but in (16a), **chó** ‘dog’ must take the classifier **con** in order for the utterance to be grammatical. According to one of my informant, when **ghế** ‘chair’ takes the classifier **cái** in (14), there

seems to be a certain emphasis on the total number of chairs, i.e. the number is being specified. This means that there is a certain focus when we include the classifier, that draws the attention of people to the number or the noun. However, there are others who are not sensitive to this subtle difference.

(14) ***Trong nhà tôi có hai (cái) ghế thấp***
 in house 1SG POSS two CL chair short
 ‘There are two short chairs in my house’

(15) ***(Cái) ghế thứ nhất hư rồi***
 CL chair first spoil PERF
 ‘The first chair is spoilt’

(16) a. ***Con chó thứ nhất bị bệnh rồi***
 CL dog first sick PERF
 ‘The first dog is sick’

*b. ***chó thứ nhất bị bệnh rồi***
 dog first sick PERF

The observation that different classifiers have varying degrees of closeness with the nouns they classify also has other implications. Bisang (1999) points out that classifiers are not compulsory in the context of counting and gives example (17)²⁶ by Lobel (2000: 296) to show that classifiers occur if the noun has to be syntactically referentialized i.e. if it has to be individualized for being further modified as in (17b):

(17) a. ***Trong nhà hát kia có mười bảy ghế***
 in theatre DEM POSS ten seven chair
 ‘There are seventeen chairs in that theatre’
 Lit.: ‘That theatre is seventeen-chaired’

²⁶ I have made some corrections to the data but the content has not been altered in any way.

- b. ***Trong nhà hát kia có mười bảy cái ghế làm bằng gỗ tốt***
 in theatre DEM POSS ten seven CL chair make by wood good
 ‘There are seventeen wooden chairs in that theatre’
 Lit.: ‘There are seventeen chairs in that theatre which are
 made of precious wood’

However, speakers of Vietnamese would say that the classifier ***cái*** is optional in both (17a) and (17b) just like in (14) and (15). When the classifier is included, there is an emphasis on the number of chairs. We look at another example below:

- (18) a. ***Cuốn sách thứ hai rất hay***
 CL book second INT interesting
 ‘The second book is very interesting’
- b. ***Cuốn thứ hai rất hay***
 CL second INT interesting
 ‘The second book is very interesting’
- *c. ***Sách thứ hai rất hay***
 book second INT interesting
- (19) (*a. ***Trong chuồng này có 15 bò***
 in stable DEM POSS 15 cow
 ‘There are 15 head of cattle in this stable’
 (Lit.: ‘This is a stable made for 15 head of cattle’) (adapted from
 Lobel 2000: 296)
- b. ***Trong chuồng này có 15 con bò***
 in stable DEM POSS 15 CL cow
 ‘There are 15 head of cattle in this stable’

My own example, (18), shows that the classifier ***cuốn*** cannot be optional in the context of counting. Also, Lobel’s example, (19), cannot be grammatical because the classifier ***con*** is omitted which is not accepted by Vietnamese speakers. Hence, it would be erroneous

to say that the occurrence of a classifier is not confined to any lexical selection as the data above does not support this analysis.

5.4.1 Deictic expressions

Classifiers in Vietnamese also appear in referring expressions and occur with the deictic element, demonstratives, such as *này* ‘this’, *ấy* ‘that’, *kia* ‘that’ (distance is further than *ấy*) etc. to indicate the position of the referent with respect to the speaker. Let us look at some examples below:

- (20) a. Q: *Anh muốn cuốn sách nào?*
2.SG.MASC want CL book which
‘Which book do you want?’
- b. A: *Cuốn (sách) kia*
CL book DEM
‘That one/book’
- *c. A: *Sách kia*
book DEM
- (21) *Cái bàn này dài gấp tư cái kia*
CL table DEM long more than four CL DEM
‘This table is four times longer than that one’

In the answer to question (20a), we have the classifier *cuốn* followed by the demonstrative *kia* ‘that’ and the head noun is optional. In this example, we can see that the classifier has an anaphoric function which is why we can omit the head noun in the answer and just use the classifier with the appropriate demonstrative. The head noun is

seen as redundant and most speakers would just use the [CL DEM] construction if reference has already been established. This tells us about the relationship between the classifier and the demonstrative: classifiers occur rather frequently with demonstratives and the classifier is seldom omitted unless the head noun is a nonclassified noun for example *chỗ này* ‘this place’. Other examples where the strong link between the classifier and demonstrative can be seen are:

- (22) *Cuốn sách đó đâu?*
CL book DEM where
‘Where is that book?’
- (23) *(Cuốn) sách đâu?*
CL book where
‘Where is/are the book/s?’

The above examples also complement and reinforce our previous argument that the function of classifiers is to indicate specificity. Example (22) has a more specific reference to a particular book whereas in (23), the degree of specificity is lower because the book/s is not singled out contextually like (22). One reason given by Frawley (1992: 77) is that the remote demonstrative, the one that expresses a position away from the speaker, tends to be more specific because remote or distant objects are more likely to be perceived in their totality, and hence individuated, delimited and bound. Another example that illustrates this is another interrogative construction. In (24), the person is asking for a specific book whereas in (25), the person is asking for any book:

(24) **Anh cần cuốn sách nào?**
2SG.MASC need CL book which
'Which book do you need?'

(25) **anh cần sách gì?**
2SG.MASC need book what
'What book do you need?'

The interrogative 'which' is more specific than 'what' because it limits the selection to a few books that are either visually accessible or which can be actively retrieved from memory whereas 'what' presents an unlimited choice of books which the person may or may not know what he needs. Hence, the classifier is used in (24) and not in (25).

5.5 Anaphoric function of classifier constructions

Already mentioned before, Vietnamese classifiers are used with anaphoric function in many different sentence structures especially in answers to a question. For instance:

(26) a. Q: **Anh mua xe đạp chưa?**
2SG.MASC buy bicycle yet
'Have you bought a bicycle yet?'

b. A: **Dạ, tôi mua một chiếc mới rồi**
Yes 1SG buy one CL new PERF
'Yes, I have already bought a new one'

The person who answers the question tends to avoid repeating the head noun **xe đạp** 'bicycle' so the classifier **chiếc** is employed anaphorically to replace the head noun.

Another example is:

- (27) a. Q: **Anh thích bức tranh nào?**
 2SG.MASC like CL painting which
 ‘Which painting do you like?’
- b. A: **Bức kia**
 CL DEM
 ‘That one (over there)’

In both formal and informal conversations, classifiers are used for anaphoric purposes because of their ability to be co-referent with the head noun.

In Vietnamese, the use of noun classifiers depends on whether the referent is contextually established. To a large extent, noun classifiers are obligatory when the head noun is mentioned the first time and constitutes as new information. The head noun can be omitted if the object is physically present. Similarly to what Reid (1997: 167) observed for Ngan’gityemerri (Australian), “there is a preference, once a specific noun is contextually established, to refer to the noun with a classifier.”

5.6 Lexical function of classifiers

In the previous section, we saw that classifiers in Vietnamese mark specificity syntactically. Classifiers can also be used to specify the meaning of a noun. For example, we have the noun **xoài** ‘mango’:

- (28) a. **cây xoài**
 CL mango
 ‘mango tree’ (**cây** 1. tree, plant, 2. CL for trees & 2-D rigid objects)

- b. ***gốc xoài***
 CL mango
 ‘a short & bushy mango tree’ (***gốc*** 1. foot of a tree, 2. CL for trees)
- c. ***trái xoài***
 CL mango
 ‘mango fruit’ (***trái*** 1. fruit, 2. CL for fruits & 3-D round objects)

In the above examples, the same noun ***xoài*** is used with different classifiers to specify which aspect of the mango we are talking about. Another interesting example is the noun ***đá*** ‘stone, ice’:

- (29) a. ***cục đá***
 CL stone/ice
 ‘small piece of stone/ice’ (***cục*** 1. lump, clot, piece, 2. CL for 3-D lumpy objects)
- b. ***hòn đá***
 CL stone/ice
 ‘average-sized stone/ice’ (***hòn*** 1. ball, piece, mass, 2. CL for 3-D objects)
- c. ***viên đá***
 CL stone/ice
 ‘small round piece of stone/ice’ (***viên*** 1. CL for 3-D objects of regular shape)
- d. ***tảng/khối đá***
 CL stone/ice
 ‘big piece of rock/ice’ (***tảng*** 1. block; ***khối*** 1. piece, block)

In the above example, the same noun ***đá*** is used with different classifiers to indicate the size. Hence, we see that classifiers in Vietnamese are **NOT** redundant because they help

us to identify the noun more specifically and by the use of the various distinct classifiers, the speaker conveys his thoughts and perception of the entity he is talking about. At the same time, he tries to evoke the desired imagery in the listener's mind.

5.7 Possessive constructions

Classifiers can occur in possessive constructions and can have anaphoric function just like those in deictic constructions. It seems that classifiers are also as readily used as substitutes for the possessed noun as in a deictic construction. We can have a deictic classifier occurring with the demonstrative without the head noun if the reference has been established previously or if it is well-understood what the speaker is talking about:

- (30) *Cái áo này đẹp hơn cái ấy*
CL shirt DEM pretty COMP CL DEM
'This shirt is nicer than that one'

A classifier in a possessive construction can perform the same function as long as reference is established prior to the classifier and, under that context it is understood what the speaker meant:

- (31) a. *Cái áo của tôi đẹp hơn cái của chị ấy*
CL shirt POSS 1SG pretty COMP CL POSS 2SG.FEM
'My shirt is nicer than yours'
- b. *Cái áo của tôi đẹp hơn cái áo của chị ấy*
CL shirt POSS 1SG pretty COMP CL shirt POSS 2SG.FEM
'My shirt is nicer than yours'

- c. *Cái áo của tôi đẹp hơn (áo) của chị ấy*
CL shirt POSS 1SG pretty COMP shirt POSS 2SG.FEM
'My shirt is nicer than yours'

From (31a), we see that Vietnamese speakers would refrain from repeating the head noun in a simple sentence especially if the head noun has just been mentioned. Hence, the classifier is used in place of the head noun. We can even have zero anaphora like (31c) where neither classifier nor head noun is essential for the sentence to make sense.

5.8 Conclusion

From the discussion of the data and observations above, we see that classifiers in Vietnamese are polyfunctional and they possess the functions of classification, individualization, referentialization, and relationalization. In Bisang's typology (1999), Vietnamese is claimed to possess only the functions of classification and individualization. However, we see that it can also take on other functions such as referentialization and relationalization. Besides those, Vietnamese classifiers also have a pragmatic use where the semantic property of specificity is the main focus in contrast to definiteness. In the next chapter, we will look at classifiers in Southeast Asian languages and discuss them in the light of areal typology. Genetic relatedness is also relevant in our discussion of areal typology. Chapter Six is also a concluding chapter for this study where we will highlight the main points of the thesis and discuss the significance of this research.

CHAPTER SIX

VIETNAMESE CLASSIFIERS IN AREAL PERSPECTIVE AND CONCLUSION

6.0 Introduction

In this chapter, we examine some properties of Vietnamese classifiers in areal perspective²⁷. I focus on three aspects particularly interesting from areal perspective to reflect on issues of genetic relation and areal patterns. They are:

- (a) Word Order
- (b) Constituency Order
- (c) Semantic Classification

Southeast Asia has been characterized as a linguistic area because of the abundance of linguistic features shared here (cf. Ansaldo 1999). In the first few sections, we look at some of the similarities among Southeast Asian languages, with main focus on classifier constructions. We will also discuss the possible reasons for the patterns uncovered in our study. Areal phenomena cannot be studied separately from genetic relations since areal phenomena may be easily mistaken for genetic relatedness (Comrie 1987). Hence, we also look at the genetic affiliations of some languages especially that of Vietnamese. In

²⁷ A comprehensive areal study of classifiers is beyond the scope of this research.

the second half of the chapter, we will review chapters one to five. The main points and consequences of certain findings in the respective chapters will be highlighted. We will also discuss and comment on Aikhenvald (2000) in terms of its usefulness and adequacy in studies of classifier systems at the end of the chapter.

6.1 Southeast Asia as a linguistic area

In Southeast Asia, “multilingualism is a way of life, and as a result there has occurred much remarkable linguistic convergence.” (Matthews 1997: 56). In a multilingual society, languages come into contact with each other and particularly, borrowing occurs. Borrowing affects primarily vocabulary but it could have deeper effects on the structure of the language if contact is extensive. Hence, it is natural that foreign loan words are adapted and added into the dominant language over time, and as a result, the foreign language and dominant language would look increasingly similar due to the lexical and structural borrowings. For example, heavy contact between the Chinese and the Vietnamese took place politically, socially, and culturally when Vietnam was dominated by the Han people in 111 B.C. The Han tried to spread their culture to the colonies and there existed an upper class of Vietnamese who absorbed the Han language and these people helped to spread Chinese culture and language. As a result of this prolonged and heavy contact between Chinese and Vietnamese languages, almost one-third of the Vietnamese vocabulary is of Chinese origin (Nguyễn V. H. 1995). Chinese borrowings permeate all levels of Vietnamese from ‘affixes’ to simple words to whole expressions (Nguyễn Đ. H. 1990). A new language, Sino-Vietnamese, is also a result of heavy contact

between the two languages where the sources are of Chinese origin and the words are adapted to the Vietnamese pronunciation system. As observed in Comrie (1987: 10): “In some cases, the contact can be so intense among the languages in a given area that they come to share a significant number of common features, setting this area off from adjacent languages, even languages that may happen to be more closely related genetically to languages within the area.” Indeed, though there are numerous language families with different genetic affiliations (Austroasiatic, Austronesian, Tai-Kadai, Hmong-Mien and Sino-Tibetan) (Matthews 1997) in a relatively small linguistic area, they are less diverse and often share linguistic features due to contact with other speakers over several generations. Some examples of common features are the use of comparatives (cf. Ansaldo 1999), and the verb ‘acquire’ (cf. Enfield 2003).

6.2 Word order

The languages of Southeast Asia share many typological features. Morphologically, they are prototypical isolating, and analytical, with no case-marking, and no verb agreement inflection, with the exception of Tibeto-Burman languages which demonstrate some inflectional properties (Enfield, m.s.). The table below is adapted from Enfield (m.s.: 11) which shows the word-typology of Southeast Asian languages:

Table 6.0
Word Order Typology of Southeast Asian Languages

	Tai	MK	HM	Sinitic	TB
V-O	+	+	+	+	-
N-A	+	+	+	-	+/-
N-Gen	+	+	-	-	-
N-Rel	+	+	+	-	-
Prep-N	+	+	+	+/-	-

From Table 6.0, we see that Tai and Mon-Khmer (MK) are identical to each other in terms of their word order where they are both strongly left-headed (Enfield m.s.). In fact, there are various works suggesting some affiliation between Vietnamese, which is argued strongly to be of MK stock, and Tai languages (Maspéro 1912; 1927, Luo 1996, Vũ Đ. N. 1996) due to shared lexical items. However, there are several problems in the identification of systematic correspondences in the grammatical structure of these languages such as cases of weak lexical or phonological link and onomatopoeia forms. It was only recently that linguists have established that the similarities in these two language families occur as a result of areal contact (Comrie 1987). Hence, similarities in languages are not necessary nor sufficient to establish genetic relatedness, as similarities may be more reflective of language contact.

6.3 Classifier systems – Constituent order

Another linguistic feature that linguists observe in Southeast Asian languages is the occurrence of classifiers, especially numeral classifiers. As noted in Emeneau (1951), Jones (1970), and Haas (1978b), virtually all the languages of Asia use classifiers, though

to a different extent. According to Jones (1970), languages towards the northwest portion of Southeast Asia, including Thai, Burmese, Vietnamese, Chinese and non-Chinese dialects of South China, make extensive use of classifiers. Classifiers occur in these languages in relatively stable syntactic constructions with only minor variations for special uses. In Khmer and Malay, the incidence of classifiers is relatively high. In an attempt to categorize these languages of Southeast Asia according to the structure of noun phrases in classifier constructions, Jones (1970: 3) has come up with the two tables below. The first table includes languages that have the widest geographic distribution and have the Num-CL-N word order:

Table 6.1
Num-CL-N Classifier Constructions in Southeast Asia

	<i>Num.</i>	<i>Clf.</i>	<i>Noun</i>	
1. Amoy	sā	nui	huê	Two-clf-flower
2. Vietnamese	hai	con	chó	Two-clf-dog
3. Yao	pyei	taub	Juq	Four-clf-dog
4. Meo, Blue	plaub	lub	tsev	Four-clf-house
5. Nung (Tai)	slám	ǎn	bòk	Three-clf-flower
6. White Tai	hả	tô	pa	Five-clf-fish
7. Black Tai	song	tô	mà	Two-clf-horse
8. Brôu ²⁸	tapoât	lám	alic	Six-clf-pig
9. Katu	pe	nak	ayi	Three-clf (persons)-we
10. Sedang	pê'a	ngê'	kúan-kojrai	Two-clf-girl
11. Cham	pak	dray	kan	Four-clf-fish
12. Malay	tiga	buah	rumah	Three-clf-house
13. Indonesian	dua	orang	anak	Two-clf-children
14. Cebuano	upát-(ka)	buquk	bátaq	Four-(linker)-clf-children

²⁸ Brôu, Katu and Sedang are Mon-Khmer languages of Vietnam.

From the table above, we see that Vietnamese has the same type of constituent order as Chinese (Sinitic), Tai languages like Nung, Western Austronesian languages like Malay and Indonesian, and other Mon-Khmer languages. Hence, we see that we cannot establish any clear genetic relation just by looking at constituent order and areal influence appears to be the only explanation for the patterns observed.

The second table includes languages that have the N-Num-CL-N word order:

Table 6.2
N-Num-CL Classifier Constructions in Southeast Asia

	<i>Noun</i>	<i>Num</i>	<i>Clf</i>	
1. Burmese	khwéi	θòun	gáun	Dog-three-clf
2. Lolo	ts'ò	seū	joù	Man-three-clf
3. Lahu	g'a^	suh	hkeh	Chicken-seven-clf
4. Lisu	a ¹ na ⁵	sa ³	ma ³	Dog-three-clf
5. Maru	myaw	pit	tau	Horse-four-clf
6. Răwang Nung	yit	hti	gung	Wild dog-one-clf
7. Palaung	kā	ū	tō	Fish-one-clf
8. Cambodian	monuh	bɛy	neaq	Person-three-clf
9. Khmu?	knēʔ	mooy	too	Rat-one-clf
10. Karen, Sgaw	pɣa	lwì	ɣa	Person-four-clf
11. Thai	măa	s—ŋ	tua	Dog-two-clf
12. Shan	mə	hah	tō	Horse-five-clf
13. Javanese	kertas	sepuluh	lembar	Paper-ten-clf (sheets)

Generally, from the above tables, we can see that there is an indication of areal spread or influence even though we cannot ascertain the source of influence without further studies.

However, Jones (1970) concludes that the single pattern Num-CL-N found among all the languages of Vietnam is the result of spread of the pattern over the whole area, most likely of Chinese influence. This point supports our hypothesis that the word order of the Vietnamese numeral classifier construction is likely to be borrowed from Chinese (see Section 3.6.3).

6.4 Classifier systems – Semantics

In this section, we look at the nature and semantics of classifier systems for languages in the subfamilies of Austroasiatic: Mon-Khmer, Nicobarese, and Aslian.

6.4.1 Human classification

According to Adams (1989), it is common in the Austroasiatic branch to have an elaborated mode of classification for humans. This means that humans are classified differently according to their social status and other cultural specific qualities. We find that this is quite a prevalent phenomenon not only within the Austroasiatic family but also in languages like Thai, Laos, and Burmese. This may be a likely case of areal influence, which cannot be used as a means to establish genetic relatedness. We look at some of these classifications in the tables below to get a clearer idea of the types of distinction that can appear in these languages:

**Table 6.3
Palaungic Classifiers for Deities and Humans**

	Images of Buddha	Monk	King	Man (person)
I. Angku Angku (Scott 1900) Amok (Scott 1900) U (Svantesson 1983) Monglwe (Scott 1900)				to i ʔi kwin
II. Rumai Gold Palaung (Milne 1921)	sū,* chū*		p`ā	ku**
III. Waic Samtau Tai Loi (Scott 1900)				Q
Wa-Lawa-La La (Davies 1909) En (Scott 1900) Son (Scott 1900) Wa-Kengtung (Scott1900)				Q (pi) kau gao gao
Wa-Lawa Drage's Wa (Drage 1907) Wa Proper Davies' Wa (Davies 1909) Antisdell's Wa (Antisdell 1911)	hsu (pagoda)	pa (priest)		kau ka kau
Kawa (Diffloth 1977)				
Lawa Bo Luang (Mitani 1966; 1972) Umphai (Mitani 1972) Mae Sariang (Mitani 1972)		t-n t-n t-n		pui phui pui

* halo, bank of river, pagoda

** also doll

Adapted from Adams (1989: 48-9)

From the table above, we see that there is a distinction made between ordinary people, royalty, monks, and images of Buddha. The emphasis on distinguishing between religious people and ordinary people is rather unique to this group of languages within the

Austroasiatic family. Classifiers for deities, members of religious order, and royalty are believed to be borrowed from sources like Shan (Tai) and/or Burmese because of the close proximity and influence from their language and culture (Adams 1989). This is a rather obvious case of areal contact as Adams puts it: “Buddhism, widespread in Southeast Asia, has replaced a native animistic religion which some other Mon-Khmer and Aslian culture have retained” (1989: 51).

Table 6.4
Central Bahnaric Classifiers for Human and Non-human Animates

	Honorific for person	Non-honor person	Children	Bonded souls	Non-human animates	Things
Central Bahnaric Bahnaric proper (Guillemin et and Alberty 1959; 1963) All dialects	măt			kol*	(valuable animals)	(valuables)
	←	(deprecating) ←	(poetic) ←		← (fish)	tông (valuables) găi** (small spherical objects)
			←		← (fish, young dogs and cats)	kōnāng (flat things)

* gôl in Rôngao (West)

** ger in Alakong

Adapted from Adams (1989: 69)

From the table above, we see that Central Bahnaric differentiates between adults and children. Also, for adults, there is a further distinction in terms of respect and status. This is similar to Vietnamese (see Table A) even though Vietnamese possesses a much more elaborated system for classifying humans.

Even though the classification of humans is widely prevalent in Southeast Asia and more indicative of areal influence, the source of classifiers and the form of the classifier can also be taken to indicate some form of areal and genetic relations in Austroasiatic languages. For instance, the majority of the forms for classifiers for humans in general are derived from nouns meaning ‘body’, ‘person’, ‘individual’, and ‘man’. The classifier for people in Amok and U means ‘man, human being’ while in Waic and Gold Palaung, it means ‘man, body’. In Mon, it means ‘person, body, self’, and in La, it means ‘person’. The **ngaay* form in many North Bahnaric languages and Palaungic and *người* in Vietnamese may be related since they originally refer to eye and now they are used to refer to ‘person, man’ (Adams 1989). The source of classification seems unique to these languages since it is not used elsewhere.

6.4.2 One-dimensional classification

In this section, we compare the systems of classification of one-dimensional objects between Vietnamese, Khmu, and Khmer.

In the classification of one-dimensional objects, Khmu Yuan and Khmu distinguish unfelled trees from felled trees, and long rigid objects from long flexible objects. Weapons, artifacts, and things made of bamboo have specific classifiers which could be due to their cultural and survival importance:

Table 6.5
Classes for One-Dimensional Objects in Khmu Yuan and Southern Khmu

Khmu Yuan		Khmu	
Form	Items classified	Form	Items classified
tuut'	unfelled trees	tuut	unfelled trees
kr—ng`/tIngok`	stalks of things		
lem`	long objects	hlem	long rigid objects
lam`	boats		
sen`	long flexible objects	sen	long flexible objects
thian	knives	?nuang	knives, crossbows
an`	artifacts	?an	artifacts
b—k`	guns		
laang`	objects made of bamboo	hlaang	objects made of bamboo
slung`	body of water	hrong	body of water
rngtaang`	connecting paths		

Adapted from Adams (1989: 133)

In Khmu, *tuut* is used wiith plants in their natural state including trees, grass, and bamboo. The main function of *hlem* in Khmu seems to be to classify long, rigid artifacts of varying size, including felled trees. Some examples are pencils, logs, needles, drinking straws, dibble sticks, and canoes. Besides that, it is also used for body parts like fingers and bones. This class is very similar to what we see in Vietnamese for the classifier *cây* ‘tree, plant’, the classifier for trees and long, rigid objects (see Table B). The only difference is that Vietnamese does not make a distinction between felled and unfelled trees. Also, canoes are not classified as one-dimensional objects.

The Khmuic dialects also have the classifier, *sen*, for long, flexible objects. *Sen* counts things like chains, strings, strands of hair, thread, ropes and intestines which is similar to Vietnamese classifier *sợi* ‘thread, fibre’, with the exception of intestines. The main difference in Khmuic dialects is that roads and rivers are counted as one-dimensional objects whereas in Vietnamese, they are counted with the non-human animate classifier, *con* ‘child, young one’ because of their ability to move.

The Khmer system for one-dimensional items distinguishes rigid items from flexible items. Khmer has a classifier *daɛm*, meaning ‘trunk’ which classifies trees, plants, and long rigid artifacts. In addition, it is also used to count items such as pencils and cigarettes. Ehrman (1972) also cites it as classifying larger inanimate items like rifles and trains (trains can also be classified by *kriɛng*, a classifier for motorized things). Even though *cây* in Vietnamese classifies many items which in Khmer are unclassified, and some items in the *daɛm* class do not overlap with those in the *cây* class, there is some similarity in the semantic criteria of classification in Vietnamese and Khmer. Both languages rely on the use of the features: rigidity and shape. Moreover, we see one of the features of Mon-Khmer languages at work here: inanimate classifiers are often derived from plant parts through metaphorical extensions. In Khmer, *daɛm* means ‘trunk’ and in Vietnamese, *cây* means ‘tree, plant’. The classifier for string-like things in Khmer is *sɛay* meaning ‘blood vessel, nerve, fibre’. This is also similar to the Vietnamese classifier *sợi* which also means ‘thread, fibre’ and classifies long, flexible objects. The

common lexical roots that these languages share are indicative of their genetic relatedness.

6.4.3 Two-dimensional classification

The classes for two-dimensional objects in the Mon-Khmer subfamilies are not easy to characterize and the problem is complicated by the fact that classes for two-dimensional objects often have more than one possible organising principle. Let us look at the Khmuic Khmer languages again. Within the Khmuic system, there are classifiers for a great variety of objects. In all of the languages, the classes are basically for the same kinds of objects. There are special classes for fibrous coverings, flat, rigid items, places, fields, leaves, nets, traps, and books (Adams 1989). Also, many of these same types of classes can be found in other Mon-Khmer branches as well:

Table 6.6
Classes for Two-Dimensional Objects in Khmu Yuan and Southern Khmu

Khmu Yuan		Khmu	
Form	Items classified	Form	Items classified
phǖn`	clothes	phǖn	cloth, blanket, clothes
pheen`	sheet of paper	piang	plates, plank, board
taar`	flat surface, sheet of paper	hla	leaf and paper
daang`	net	s'map	letter, fascicules
raø	network, meshes	ram	letter
tǖ	place, market or plantation	?m—n	place, garden, ulcer, river, bottom, corner of box
lang	field allotment		
trl—ng`	trap	hlaang	traps of bamboo
laang`	trap		

Adapted from Adams (1989: 159)

From the table, we see that in Khmu, there is a classifier, *hla* ‘leaf’, that classifies leaves and papers. According to Adams (1989), forms which mean ‘leaf’ are found throughout Mon-Khmer and in many of these languages, leaf forms subsume the same kinds of items. For example, leaves, papers, and plates. In Vietnamese, this classifier is *lá* ‘leaf’ and classifies things like cards, sails, letters, and even body parts.

In Khmer, only a few two-dimensional objects are classified and many are related to paper and items written or drawn on. For example, *s-nlɔk* classifies leaves as well as paper although it might have a wider usage because Jenner and Pou (1981) cites it as a classifier for flat objects. This form also means ‘leaf’ which is similar to *lá* in Vietnamese. The only difference is that *lá* is not the exclusive classifier for paper in Vietnamese because shape and metaphorical extension are the organizing principles of *lá* and paper has a more specific classifier which is *tò*, the classifier for thin pieces or stacks of literate material. There are many classifiers in Khmer for books and papers. For books, paper, and writing in general, *cbap* is used. For a more detailed classification based on the kinds of binding, we have *krap* (Maspéro 1915) meaning ‘covering’ which is used for paper-bound books. For cloth books, there is *kàmpi* and for those tied by string, *khsae*. *Kan* and *kba:l* are used for volumes of books. *Rù:p* is used for pictures or something drawn on.

6.4.4 Three-dimensional items classification

The three-dimensional classification in Mon-Khmer languages serve to characterize items like:

- I. Food: fruits, vegetables, tubers, eggs, grains, nuts, meat.
- II. Naturally occurring items: stones, gems, hills, mountains, sun, moon, stars, earth, body parts.
- III. Cultural artifacts: pills, bricks, buttons, bullets, bombs.

All of the Khmuic languages have three-dimensional classifiers based on shape and the mode of classification is briefly summarized in the table below:

Table 6.7
Classes for Three-Dimensional Objects in Khmu Yuan and Southern Khmu

Khmu Yuan		Southern Khmu	
Form	Items classified	Form	Items classified
nuay`/luuk` 'fruit'	- fruits, grains, eggs, rice - breasts, womb, heart, pimples, swollen sores, eyes, kidneys - pots, baskets, hats, mortars, bottles, pails - houses, barns	hnuey	- fruits, grains - breasts, womb, heart, pimples, swollen sores - pots, baskets, hats
k—n 'a small part of something'	ingots, bricks	k—n	pills, eyeballs, pill-shaped objects
kmlò	meat, bones	mak	fruit

Adapted from Adams (1989: 107)

In the Khmuic languages, there is a morpheme meaning ‘fruit’ which classifies fruits in Southern Khmuic, including foodstuffs, body parts, household items and shelters in Khmuic Yuan. Once again, we see that it is quite common for plant parts to serve as classifiers in the Mon-Khmer languages that is indicative of some genetic relatedness. We see that the three-dimensional items classified in Khmuic and Vietnamese overlap in many ways (see Table D). For instance, hearts and kidneys are both classified by a classifier derived from a fruit. In Khmuic Yuan, there is a ‘lump’ classifier, *kmlò*, for meat and bones which is similar to the Vietnamese classifier *cục* in its classificatory principles.

In the above sections (6.4.1-6.4.4), we have looked at the various features of semantic classification in various languages of the Austroasiatic family. There are many similarities elicited from the discussion. The styles of classification can be spread to other areas quite easily but the source of classifiers and the similar morphemes used across different languages are often indicative of some underlying genetic affiliations.

6.5 Conclusion

This study has addressed the phenomenon of noun categorization, from a general perspective to a closer examination at how this function is encoded in Vietnamese grammatically and semantically. In order to understand what classifiers are and the functions encoded by the use of classifiers, I situate my study of Vietnamese classifiers within the typology of classifiers developed by Aikhenvald (2000). Aikhenvald’s

framework is based on a number of parameters which serves to characterize the different prototypes of noun categorization devices. These parameters are:

- A) Morphosyntactic locus of coding
- B) Scope, or domain of categorization
- C) Principles of choice, or assignment of noun categorization devices
- D) Kinds of surface realization
- E) Agreement
- F) Markedness relations
- G) Degree of grammaticalization and lexicalization
- H) Interaction with other grammatical categories
- I) Semantic organization of the system
- J) Evolution and decay
- K) Language acquisition and dissolution

Some of these parameters are not applicable to Vietnamese so I have looked at most of these parameters in the analysis of the Vietnamese classifier system only if they serve to enhance our understanding about Vietnamese classifiers. In the syntactic analysis of Vietnamese classifiers (see Chapter Three), I have discussed Parameters A and H by examining all the Vietnamese constructions that involve noun categorization devices, i.e. how classifiers are used in quantifier/numeral constructions (Num-CL-N) as well as deictic (CL-N-Dem) and certain interrogative constructions. Besides looking at the different morphosyntactic loci of Vietnamese classifiers, I have also looked at the relation these grammatical items have with other grammatical categories. After a close

examination of the data, it is apparent that within my study, we cannot make any claims regarding constituency relations in classifier constructions in Vietnamese because there is not enough evidence in this dissertation nor within the scope of this study to do so. Thus, I have maintained the position that the various grammatical items should be treated as separate entities instead of trying to impose constituency relations on them.

I have discussed Parameters C, G, I and J by dealing with the various aspects of the semantics of Vietnamese classifiers (see Chapter Four). In the discussion of the semantic organization of the Vietnamese classifier system, we looked at the various properties that tend to be encoded in different types of classifiers e.g. animacy, sex and gender, kinship, social, status, nature, physical properties, and function. These properties are crucial in Vietnamese as we saw that animate and inanimates are categorized according to these properties and therefore the choice of a classifier depends on these semantic properties. In Vietnamese, humans are distinguished from animals and inanimates. Humans are further categorized according to age and/or respect, and sex. We also see the use of kin terms as classifiers in Vietnamese. Even though humans are often separated from animals by the use of different classifiers, this is not absolute because there is a special classifier *con*, which can be used to classify lower human beings. Similarly, even though animals are often separated from inanimates, certain inanimates reflecting motion, life, and animacy are categorized by the same classifier as animals, *con*. Vietnamese makes a three-way distinction between inanimates based on dimensionality or shape. In the classification of one-dimensional or long objects, the secondary properties of shape, nature, and rigidity/flexibility are important. In the classification of two-dimensional or flat objects,

the secondary properties of shape, function, and nature are utilized. In the classification of three-dimensional or round objects, the secondary properties of nature, shape, and size are utilized. After dealing with the organization of the system, we discussed the lexical sources of Vietnamese classifiers.

Even though the pragmatic function of noun categorization devices is not one of the parameters in Aikhenvald's typology, we see that this is a significant area of consideration because it helps us to understand when classifiers are needed and when they are not. In Vietnamese, there is a high degree of indeterminateness in nouns. In order to specify the meaning of nouns, we often have to resort to the use of some grammatical categories to fulfill the function of specification in addition to the other core functions they serve. This results in the polyfunctionality of lexical items. In Vietnamese, classifiers are not redundant because they fulfill functions like classification, referentialization, relationalization, and specification.

Since Vietnamese is situated in a typologically interesting linguistic area, Southeast Asia, where an abundance of linguistic features are shared, we examine some of the properties of classifiers discussed in the prior chapters in areal perspective. In the earlier part of this chapter, we looked at word order, constituency order, and the semantic organization of classifiers in various Southeast Asian languages. We established that there are certain features that tend to be easily borrowed through areal contact such as word order and the mode of classification. However, there are other features that tend to be less likely borrowed and can be indications of a genetic affiliation between languages.

The parameters suggested in Aikhenvald's typology are useful in a proper investigation of classifiers even though not all of them are applicable to every language in the world. However, there seems to be a problem when we try to determine the type of classifier system in Vietnamese (numeral classifier, noun classifier, deictic classifier, etc.) just by looking at these parameters²⁹. For instance, these three types of classifiers have very similar properties with the exception of occurring in different syntactic environments. So, it seems redundant to suggest that Vietnamese has three types of classifier systems based merely on their environments. Another aspect that does not seem to be useful in Aikhenvald's typology is the concept of multiple classifier systems, which is a relatively new area of research. **Multiple classifier systems** are so-labeled because several distinct classifier types may coexist in one language i.e. multiple classifier language. For an extreme case, Nasioi (Papuan of Southern Bougainville: Hurd 1977) has possessed classifiers, numeral classifiers, deictic classifiers, verbal classifiers, and classifiers used to modify adjectives (Aikhenvald 2000: 219). These are illustrated in the examples below respectively:

- (1) *n-ee-ka-na-va*
us-DU-INTENSE-DER.SUFF-CL:house
'our house' (Hurd 1977: 155)
- (2) *nto-na-ru'* *bee-ru'-pi*
water-DER.SUFF-CL:unit of liquid three-CL:unit of liquid-PL
a-ru'-daang
this-CL: unit of liquid-inland
'These three inland lakes' (Aikhenvald 2000: 219)

²⁹ I have not done so in the earlier chapters because I find it unnecessary.

- (3) *Teni toire' tareuri-ma-O-i bau'uri-ma-O-i*
 the.FEM children care.for-PRES.HAB-DER.SUFF-CL:FEM feed-
 PRES.PROGR-DER.SUFF-CL:FEM
 'The lady who cares for the children (is) the one who feeds them' (Hurd
 1977: 144)
- (4) *tamp-a-u'*
 good-DER.SUFF-CL: taro
 'a good (taro)' (Hurd 1977: 132)

According to Aikhenvald (2000), not much attempt has been made to consider multiple classifier systems in a cross-linguistic perspective. She acknowledges that there are individual descriptions of multiple classifier systems like Worsley (1954) on Anindilyakwa (Australian), Goncalves (1987) on Mundurukú (Tupí), Bisang (1993) on Hmong (Miao-Yao), Onishi (1994) on Motuna (Papuan of Bougainville), Shepard (1997) on Machiguenga (Campa, Peruvian Arawak, Arawak) etc. and that the number of studies on multiple classifier languages has increased. However, this area still needs to be looked at in greater detail from a typological viewpoint. Multiple classifier systems are also claimed to be significant because they constitute "a strong argument in favour of the proposed typology of classifiers based on the morphosyntactic locus of coding of noun categorization devices" (Aikhenvald 2000: 184).

The morphemes used in distinct classifier environments may partly overlap in their form and/or semantics. Sometimes, the semantics of the morpheme may be the same despite a different form and vice versa. At times, the same set of morphemes can be used in more than one classifier environment. These morphemes may have different grammatical properties, or be more or less obligatory depending on what classifier environment they are in. According to Aikhenvald (2000), Vietnamese is an example of a language that

utilizes one morpheme for different environments and would be an example of a multiple classifier language:

- (I) Noun Classifiers: Vietnamese has noun classifiers because classifiers are closely related to nouns, being able to co-occur with them in a noun phrase without the presence of a numeral.
- (II) Numeral Classifiers: Vietnamese has numeral classifiers because classifiers occur in numeral noun phrases and are obligatory in the context of counting.
- (III) Deictic Classifiers: Vietnamese has deictic classifiers because classifiers obligatorily occur with deictic elements such as articles and demonstratives.
- (IV) Possessed Classifiers: In Vietnamese possessive constructions, we see that possessed classifiers characterize the possessed noun in terms of its animacy, shape, function, etc.

Some examples are given below:

- (5) ***Có một ông bác sĩ làm việc ở trường***
EXT one CL doctor do work in school
'There is a doctor working in school' [Numeral]
- (6) ***Chúng ta nên kính trọng ông bác sĩ***
We should respect CL doctor
'We should respect doctors' [Simple NP]

- (7) ***Ông bác sĩ này rất tốt***
CL doctor DEM INT good
'This doctor is very good' [Deictic]
- (8) ***Ông bác sĩ của tôi rất buồn***
CL doctor POSS 1SG INT sad
'My doctor is very sad' [Possessive]

Looking at the different morphosyntactic environments that allow for the presence of “different” types of classifiers, the questions to ask are: Is it meaningful or useful to arrive at this taxonomy of classifiers based on the environments they occur in and to call them a system? Are we complicating matters and overlooking important similarities that occur within the Vietnamese classifier system? Besides the problem that the same set of classifiers is used with no difference syntactically and semantically, the bigger problem is the task of trying to characterize the classifier type when we have the four environments lapsing into one sentence, for example:

- (9) ***Hai cuốn sách này của tôi hay***
two CL book DEM POSS 1SG interesting
'These two books of mine are interesting'

In (9), should we treat the classifier as a numeral classifier, deictic classifier or a possessive classifier? In view of all these questions which we cannot answer in Vietnamese, I have chosen not to treat Vietnamese as a multiple classifier language because the concept is problematic in itself. I propose that we do not need this concept for Vietnamese because the “different” types of classifiers are almost identical except their syntactic environments and we can hardly call that a system. In my view, classifiers in Vietnamese are items that fulfill the semantic role of categorizing nouns, are organized

semantically, and possess certain syntactic and pragmatic functions. The different syntactic environments they occur in just determine the function of the classifier and should not be seen as something that has to be categorized.

Aikhenvald's framework has helped us to describe and analyze the Vietnamese classifier system with the parameters in Section 6.5. However, the typology of classifiers, which is based on the morphosyntactic environments of these noun categorization devices is over theory-specific and inadequate for Vietnamese and possibly other Southeast Asian languages. We still cannot determine how the typology is going to accommodate Vietnamese even though we have done a thorough examination of its classifier system. Aikhenvald may need to modify the typology and reconsider the idea about multiple classifier systems.

REFERENCES

- Adams, K. L. (1986). 'Numeral classifiers in Austroasiatic'. In C. Craig (ed.) *Noun Classes and Categorization*. Amsterdam: John Benjamins Publishing Company, 241-262.
- _____. (1989). *Systems of Numeral Classification in the Mon-Khmer, Nicobarese and Aslian Subfamilies of Austroasiatic*. Canberra: Pacific Linguistics.
- _____. and Conklin, N. F. (1973). 'Towards a theory of natural classification'. *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 9: 1-10.
- Aikenvald, A. Y. (1994a). 'Classifiers in Tariana'. *Anthropological Linguistics* 34: 407-65
- _____. (1994b). 'Grammatical relations in Tariana'. *Nordic Journal of Linguistics* 7: 201-18.
- _____. (2000). *Classifiers: A Typology of Noun Categorization Devices*. New York: Oxford University Press.
- Allan, K. (1977). 'Classifiers'. *Language* 53: 285-311.
- Ansaldo, U. R. (1999). 'Comparative constructions in Sinitic. Areal typology and patterns of grammaticalization'. PhD thesis, Stockholm University.
- Antisdell, C. B. (1911). 'Elementary studies in Lahoo, Akha (Kaw) and Wa languages'. *Journal of the Burma Research Society* 1: 41-64.
- Becker, A. J. (1975). 'A linguistic image of nature: The Burmese numerative classifier system'. *Linguistics* 165: 109-121.
- Beckwith, C. I. (1993). 'Class nouns and classifiers in Thai'. *Papers from the Third Annual Meeting of the Southeast Asian Linguistics Society*: 11-26.
- Bell, A. (1978). 'Language samples'. In J. H. Greenberg, C. A. Ferguson, E. A. Moravcsik (eds.) *Universals in Human Language, Vol. 1: Method and Theory*. Stanford: Stanford University Press, 123-156.
- Benton, R. A. (1968). 'Numeral and attributive classifiers in Truquese'. *Oceanic Linguistics* 7: 104-46.

- Bisang, W. (1993). 'Classifiers, quantifiers and class nouns in Hmong'. *Studies in Language* 17(1): 1-51.
- _____. (1996). 'Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages'. *Studies in Language* 20(3): 519-97.
- _____. (1999). 'Classifiers in East and Southeast Asian languages: Counting and beyond'. In J. Gvozdanovic (ed.) *Numeral Types and Changes Worldwide*. Berlin: Mouton de Guyter, 113-185.
- Brown, R. (1981). 'Semantic aspects of some Waris predications'. In K. J. Franklin (ed.) *Syntax and Semantics in Papua New Guinea Languages*. Ukarumpa: Summer Institute of Linguistics, 98-103.
- Bybee, J. (1985). *Morphology*. Amsterdam: John Benjamins.
- Cao X. H. (1988). 'The count/mass distinction in Vietnamese and the concept of classifier'. *Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41: 38-47.
- Chatterji, S. K. (1926). *The Origin and Development of the Bengali Language*. Calcutta.
- Comrie, B. (1987). 'Introduction'. In B. Comrie (ed.) *The Major Languages of East and South-east Asia*. London: Routledge, 1-18.
- _____. (1989). *Language Universals and Linguistic Typology*. Second Edition. Chicago: University of Chicago Press.
- Craig, C. G. (1986). 'Introduction'. In C. G. Craig (ed.) *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam: John Benjamins, 1-10.
- _____. (1992). 'Classifiers in a functional perspective'. In M. Fortescue, P. Harder and L. Kristoffersen (eds.) *Layered Structure and Reference in a Functional Perspective*. Amsterdam: John Benjamins, 277-301.
- Creider, C. A. (1975). 'The semantic system of noun classes in Proto-Bantu'. *Anthropological Linguistics* 17: 127-38.
- _____. (1994). 'Semantic universals in classifier systems'. *Word* 45(2): 145-171.
- Daley, K. A. (1998). *Vietnamese Classifiers in Narrative Texts*. Texas: Summer Institute of Linguistics & The University of Texas at Arlington.

- Daw, K. K. (2000). 'A study of dimensions of interlanguage: A case of Myanmar speakers of Malay as a second language'. PhD thesis. University of Brunei.
- Davies, H. R. (1909). *Yun-nan: The Link Between India and the Yangtze*. London: Cambridge University Press.
- Denny, P. J. (1976). 'What are noun classifiers good for?' *Papers from the Annual Regional Meeting of the Chicago Linguistics Society* 12: 122-132.
- Derbyshire, D. C. and Payne, D. L. (1990). 'Noun classification systems of Amazonian languages'. In D. L. Payne (ed.) *Amazonian Linguistics: Studies in Lowland South American Indian Languages*. Austin: University of Texas Press, 243-272.
- Dik, S. C. (1987). 'A typology of entities'. In Auwera and Goossens (eds.) *Ins and Outs of the Predication*. Foris, 1-20.
- _____. (1989). *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Dordrecht: Foris Publications.
- Diffloth, G. (1977). 'Proto-Waic and the effects of register on vowel gliding'. *Paper presented at Tenth International Conference on Sino-Tibetan Languages and Linguistics*, Washington, D.C.
- Dixon, R. M. W. (1977). *A Grammar of Yidin*. Cambridge: Cambridge University Press.
- _____. (1982). *Where Have All the Adjectives Gone? and other essays in Semantics and Syntax*. Berlin: Mouton Publishers.
- _____. (1986). 'Noun classes and noun classification in typological perspective'. In C. G. Craig (ed.) *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam: John Benjamins, 105-112.
- Downing, P. (1996). *Numeral Classifier Systems: The Case of Japanese*. Amsterdam: John Benjamins.
- Drage, G. (1907). *A Few Notes on Wa*. Rangoon, Burma: Superintendent, Government Printing.
- Dryer, M. S. (1988). 'Object-verb order and adjective-noun order: Dispelling a myth'. *Lingua* 74: 185-217.
- _____. (1989). 'Large linguistic areas and language sampling'. *Studies in Language* 13: 257-292.
- _____. (1992). 'The Greenbergian word order correlations.' *Language* 68: 81-138.

- Ehrman, M. E. (1972). *Contemporary Cambodian: Grammatical Sketch*. Washington, D.C.: Foreign Service Institute, Department of State.
- Emeneau, M. B. (1951). *Studies in Vietnamese (Annamese) Grammar*. Berkeley: University of California Press.
- Enfield, N. J. (m.s.). 'Potential and achievement in Southeast Asian languages: An areal study of syntactic and semantic variation'. Work in progress for PhD, June 1997.
- _____. (2003). *Linguistic Epidemiology. Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge Curzon.
- Erbaugh, M. S. (1986). 'Taking stock: The development of Chinese noun classifiers historically and in young children'. In C. Craig (ed.) *Noun Classes and Categorization*. Amsterdam: John Benjamins Publishing Company, 399-436.
- Frawley, W. (1992). *Linguistic Semantics*. London: Lawrence Erlbaum.
- Gage, W. W. (1979). 'Infiltration of grammatical morphemes into Vietnamese from Chinese'. *Paper presented at The Twelfth International Conference on Sino-Tibetan Languages and Linguistic Studies*.
- Greenberg, J. H. (1963). 'Some universals of language with special reference to the order of meaningful elements'. In J. Greenberg (ed.) *Universals of Language*. Cambridge: MIT Press, 73-113.
- _____. (1972). 'Numeral classifiers and substantive number: Problems in the genesis of a linguistic type.' *Working Papers on Language Universals 9*. Stanford: Stanford University Press, 1-39.
- Goral, D. R. (1978) 'Numeral classifier systems: A Southeast Asian cross-linguistic analysis'. *Linguistics of the Tibeto-Burman Area* 4(1): 1-72.
- Goncalves, C. H. R. C. (1987). *Concordância em Mundurukú*. Campinas: Editora da Unicamp.
- Guilleminet, P. & Alberty, J. (1959). *Dictionnaire Bahnar-Français*. Vol. 1 Publications de l'École Française d'Extrême-Orient.
- _____. (1963). *Dictionnaire Bahnar-Français*. Vol. 2. Publications de l'École Française d'Extrême-Orient.
- Haas, M. R. (1978a). 'The use of numeral classifiers in Thai'. In S. D. Anwar (ed.) *Language, Culture and History*. Stanford: Stanford University Press, 58-64.

- _____. (1978b). 'The use of numeral classifiers in Burmese'. In S. D. Anwar (ed.) *Language, Culture and History*. Stanford: Stanford University Press, 65-81.
- Haiman, J. (1980). 'The iconicity of grammar: Isomorphism and motivation.' *Language* 54: 565-589
- Hundius, H. & Kolver, U. (1983). 'Syntax and semantics of numeral classifiers in Thai.' *Studies in Language* 7: 165-214.
- Hurd, C. (1977). 'Nasioi projectives'. *Oceanic Linguistics* 16: 111-178.
- Jenner, P. N. and Pou, S. (1981). *A Lexicon of Khmer Morphology. Mon-Khmer Studies* 9-10. Honolulu: University of Hawaii Press.
- Jones, R. B. (1970). 'Classifier constructions in Southeast Asia'. *Journal of American Oriental Society* 90: 1-12.
- Kiyomi, S. (1992). 'Animateness and shape in classifiers'. *Word* 43: 15-36.
- Lakoff, G. (1986). 'Classifiers as a reflection of mind'. In C. G. Craig (ed.) *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam: John Benjamins, 13-52.
- _____. (1987). *Women, Fire and Other Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lambrecht, K. (1987). 'On the status of SVO sentences in French discourse'. In R. Tomlin (ed.) *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins, 217-262.
- Lang, A. (1975). *The Semantics of Classificatory Verbs in Enga and Other Papua New Guinea Languages*. Canberra: Pacific Linguistics.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University.
- Lee, M. (1988). 'Language, perception and the world'. In J. Hawkins (ed.) *Explaining Language Universals*. Oxford: Blackwell, 211-246.
- Lobel, E. (2000). 'Classifiers versus genders and noun classes: A case study in Vietnamese'. In B. Unterbeck (ed.) *Gender in Grammar and Cognition*. Berlin: Mouton de Gruyter, 259-319.
- Luo, Y. X. (1996). 'Word families in Tai: A preliminary account'. *Pan-Asiatic Linguistics* 3: 850-883.

- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Macdonald, R. R. (1976). *Indonesian Reference Grammar*. Washington, DC: Georgetown University Press.
- Mallinson, G. and Blake, B. J. (1981). *Language Typology: Cross-Linguistic Studies in Syntax*. North-Holland Linguistics Series 46. Amsterdam: North Holland.
- Maspéro, H. (1912). 'Etude sur la phonétique historique de la langue Annamite: les initiales'. *Bulletin de l'Ecole Française d'Extrême-Orient* 12: 1-27.
- _____. (1927). *La Chine Antique*. Paris.
- Maspéro, G. (1915). *Grammaire de la Langue Khmère (Cambodgien)*. Paris.
- Matthews, S. (1997). 'South and Southeast Asia'. In B. Comrie, S. Matthews and M. Polinsky (eds.) *The Atlas of Languages*. London: Bloomsbury Publishing, 56-71.
- _____. and Pacioni, P. (1998). 'Specificity and genericity in Cantonese and Mandarin'. In L. Xu (ed.) *The Referential Properties of Chinese Noun Phrases*. Paris: Ecole des Hautes Etudes en Sciences Sociales, Centre de Recherches Linguistiques en Asie Orientale, 45-59.
- _____. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar*. London & New York: Routledge.
- Milne, M. L. (1921). *An Elementary Palaung Grammar*. Oxford: Clarendon Press.
- Mitani, Y. (1966). 'Descriptive study of the Lawa language (Bo Luang dialect)'. *South East Asian Studies* 4: 40-62.
- _____. (1972). 'A short vocabulary of Lawa'. *South East Asian Studies* 10: 131-168.
- Mithun, M. (1986). 'The convergence of noun classification systems'. In C. G. Craig (ed.) *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam: John Benjamins, 379-398.
- Newmeyer, F. J. (1998). *Language Form and Language Function*. Cambridge: The MIT Press.
- Nguyễn Đ. H. (1957). 'Classifiers in Vietnamese'. *Word* 13: 124-152.
- _____. (1987). 'Vietnamese'. In B. Comrie (ed.) *Major Languages of East and South-east Asia*. London: Routledge, 49-67.

- _____. (1990). 'Graphemic borrowings from Chinese: The case of chữ nôm, Vietnam's demotic script'. *Bulletin of the Institute of History & Philology, Academia Sinica*. Taipei, Taiwan 61(2): 383-432.
- _____. (1997). *Vietnamese: Tiếng Việt Không Sơn Phấn*. Amsterdam: John Benjamins.
- Nguyễn, P. P. (2002). *Những Vấn Đề Ngữ Pháp Tiếng Việt: Loại Từ và Chỉ Thị Từ* (The Grammatical Problems of Vietnamese: Classifiers and Instruction Words). Hanoi: Nhà Xuất Bản Đại Học Quốc Gia Hà Nội.
- Nguyễn V. H. (1995). *The Ancient Civilization of Vietnam*. Hanoi: Thế Giới Publishers.
- Onishi, M. (1994). 'A grammar of Motuna (Bougainville, Papuan New Guinea)'. PhD thesis, Australian National University.
- Payne, T. E. (1997). *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.
- Pe, H. (1965). 'A re-examination of Burmese classifiers'. *Lingua* 15: 163-86.
- Perkins, R. (1980). 'The coevolution of grammar and culture'. Dissertation. State University of New York at Buffalo.
- _____. (1989). 'Statistical techniques for determining language sample size'. *Studies in Language* 13: 293-315.
- Reid, N. (1997). 'Class and classifier in Ngan'gityemerri'. In Harvey M. and N. Reid (eds.) *Nominal Classification in Aboriginal Australia*. Amsterdam: John Benjamins.
- Sands, A. K. (1995). 'Nominal classification in Australia'. *Anthropological Linguistics* 37: 247-346.
- Scott, J. G. and Hardiman, J. P. (1900). *Gazetteer of Upper Burma and the Shan States*. Part I, Vol 1. Rangoon, Burma: Superintendent, Government Printing.
- Shepard, G. Jr. (1997). 'Noun classification and ethnozoological classification in Machiguenga, an Arawakan language of the Peruvian Amazon'. *Journal of Amazonian Languages* 1: 29-57.
- Stassen, L. (1985). *Comparison and Universal Grammar*. Oxford: Blackwell.
- Svantesson, J. O. (1983). Adam's (1989) personal communication on U, October 31.

- Thompson, L. C. (1965). *A Vietnamese Reference Grammar*. Honolulu: University of Hawai'i Press. Reprinted in 1991.
- Tomlin, R. S. (1986). *Basic Word Order: Functional Principles*. London: Croom Helm.
- T'sou, B. K. (1976). 'The structure of numeral classifier systems'. In P. N. Jenner, L. C. Thompson and S. Starosta (eds.) *Austroasiatic Studies*, ii. Honolulu: University of Hawaii Press, 1215-1247.
- Whaley, L. J. (1997). *Introduction to Typology: The Unity and Diversity of Language*. California: Sage Publications.
- Worsley, P. M. (1954). 'Noun-classification in Australian and Bantu: Formal or semantic?' *Oceania* 24: 275-288.
- Võ, P. (1988). 'Tiếng Việt và người Việt ở hải ngoại (The Vietnamese language and the Vietnamese people overseas)'. In N. Tran, T. Nguyễn and L. Le (eds.) *Vietnamese Language and Culture*. 26-51.
- Vũ, Đ. N. (1996). 'Những từ có liên hệ với nhau về nghĩa lịch sử âm đầu trong tiếng Việt.' (Some words that are related in meaning historically with initials in Vietnamese). Dissertation, University of Hanoi.