

**Number in Classifier Languages**

**A DISSERTATION  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY**

**Hiroki Nomoto**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY**

**Advisor: Hooi Ling Soh**

**March, 2013**

**© Hiroki Nomoto 2013**  
**ALL RIGHTS RESERVED**

# Acknowledgements

Looking back at the years I spent for my doctoral studies, I cannot but realize that quite a number of people were there to play various roles. Most of them did not even know who I was, and I do not know who they were either, like those many strangers who said hi to me on the street, which by the way never happens at the places that I have lived other than in the United States. No matter how small their roles may be, they are definitely part of my memory and my life. Those who bear their collective power include the following individuals, to whom I wish to record my sincere gratitude here.

The person who deserves the first mention is no one but my advisor, Hooi Ling Soh. If I had not met her in Singapore, my life—both academic and personal—would have been totally different. I greatly appreciate that she encouraged me to pursue my graduate studies at the University of Minnesota (U of M). Though the decision making process was a painful experience, I strongly believe that I made the right decision. I have learned a lot through working with Hooi Ling. It is worth noting that my serious take on optionality and perspective towards it have developed in the course of our collaborative research on the verbal prefix *meN-* in Malay (Soh and Nomoto 2009, 2011, to appear).

Brian Reese read various drafts very carefully and gave detailed comments. His comments and questions have helped clarify my thinking and made my work more precise. I thank the other committee members, Jeanette Gundel and Michael Maratsos, for asking important questions. Addressing them, I hope, has made this study accessible to a wider readership. Besides the committee members, special thanks go to Alan Bale and Akira Watanabe, who have read drafts of parts of this dissertation and gave insightful comments and criticisms. Their comments were crucial in developing and sharpening ideas presented in this study.

Parts of this dissertation have been presented at the 2010 Annual Meeting of the Linguistic Society of America, the 29th West Coast Conference on Formal Linguistics, Workshop on

Indonesian-Type Voice Systems, Institute of Language Research Regular Colloquium and the fourth meeting of the Indonesian Languages Research Group at Tokyo University of Foreign Studies (TUFS). I would like to thank the participants at these meetings for their questions, comments and encouragement, especially Bryan James Gordon, Shinjiro Kazama, Kunio Nishiyama and Norvin Richards (in alphabetical order). In addition, I also benefited greatly from discussions with the following people: Byeong-Uk Yi, Kartini Abd. Wahab, James Stevens and Eun-Ju Noh.

This study includes data from various languages. I am grateful to the following people, who shared with me their knowledge of their languages: Kartini Abd. Wahab, Mohd Izzat Mohd Thiyahuddin, Muhammad Idham Adli bin Musa, Muhammad Nizam Md Zin (Malay); Fatemeh Gharakhani, Maryam Ataski Golestan (Persian); Bi Xi, Shu Fan, Pamela Peng (Mandarin); Nala Huiying Lee, Qizhong Chang (Singlish); Filipe Santos Kawano (Brazilian Portuguese); Iwan Setiya Budi (Indonesian); Byeong-Uk Yi, Eun-Ju Noh (Korean). I also thank the three Hmong speakers who served as consultants for Hooi Ling Soh's project supported by the College of Liberal Arts Imagine Fund Award (2009–10). In this connection, I am grateful to Andrew Simpson for giving me an opportunity to work together on a comparative study of forms and meanings of noun phrases in four classifier languages (Simpson et al. 2011). The research reported here was financially supported by the U of M and also by the JSPS Grant-in-Aid for Young Scientists (B) (#23720199), without which I could not have traveled for conferences and hired many consultants.

While the work started early in my dissertation years at the U of M, I wrote most of this dissertation while teaching at TUFS. I thus could rely on my colleagues with different expertises for information, materials and consultant recruitment: Kenji Okano (Burmese), Naotoshi Kurosawa (Portuguese), Keiko Mochizuki (Mandarin), Kyoko Niwa (Bangla) and Satoko Yoshie (Persian). Masanori Ichikawa helped me with statistics in acceptability judgment experiments that I conducted relating to the present study. I also appreciate the constant care and support from my immediate colleagues in the Malay Language Program, Naoki Soda and Faridah Mohamed. Last but not least, my former advisor, Isamu Shoho, has been always supportive, especially at important turning points in my academic life. Needless to say, those listed above are not responsible for the content of this dissertation; any remaining errors are solely mine.

Outside the academic life, I would like to thank my fellow linguistics graduate students and also to Asish Abraham for their friendship. Finally, I owe my deepest gratitude to my parents,

Masazumi and Setsuko Nomoto, for their continuous patience and support for my unusual life path. I dedicate this dissertation to them, hoping that they will understand what this dissertation means to me.

## Abstract

Classifier languages are often described as lacking genuine number morphology and treating all common nouns, including those conceptually count, as an unindividuated mass. This study argues that neither of these popular assumptions is true, and presents new generalizations and analyses gained by abandoning them.

I claim that no difference exists between classifier and non-classifier languages regarding the semantics of either nouns or numerals. Common nouns universally denote properties and are individuated, contra Chierchia (1998b). The primary evidence comes from optional classifier languages such as Malay, in which direct numeral modification of nouns without classifiers is generally possible. Moreover, upon closer examination, optional classifier use is also observed in obligatory classifier languages such as Japanese.

I propose that classifiers are a sophisticated kind of singular number marker in that they not only assert the restriction of the domain to singularities but also conventionally implicate the characteristics of the noun. Plural markers in classifier languages have the same semantic structure, and hence they are genuine plural number markers. Furthermore, classifiers and plural markers in classifier languages are subject to the same licensing condition, which involves either referential determiners or quantifiers. These similarities endorse their belonging to the same grammatical category, i.e. number. I argue that that general number, which is associated with number-neutral properties, is a universally available basic number category, along with singular and plural. Classifier languages have distinct forms for all three. Optional number marking follows from the three-way distinction number system, where the general is morphologically unmarked.

The three basic number categories can be expressed by combinations of two binary features, i.e.  $[\pm\text{Sg}]$  (atomicity) and  $[\pm\text{Pl}]$  (divisibility). They create two kinds of general numbers differing in the presence/absence of number morphology:  $\text{GN}^+$  and  $\text{GN}^-$ . Languages such as Brazilian Portuguese and Singlish indeed have two number-neutral forms, one with a plural marker and the other without.

While classifier languages distinguish all basic number categories, non-classifier languages conflate one or more of them morphologically. Languages can be classified into five types according to this criterion: (i) SG : GN : PL, (ii) SG/GN : GN/PL, (iii) SG/GN : PL, (iv) SG :

GN/PL, and (v) SG/GN/PL. Classifier languages (type (i)) do not lack number, but instead make the most fine-grained basic number distinction. The difference between classifier and non-classifier languages reduces not to semantics (Krifka 1995; Chierchia 1998b; Wilhelm 2008) or syntax (Li 1999), but to a difference in number morphology. The proposed number system and typology make it possible to account for bare “singular” kind terms in type (ii) languages (e.g. Brazilian Portuguese), a problem to Dayal’s (2004b) theory of number and definiteness marking in kind terms.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>iv</b>
<b>List of Tables</b>	<b>x</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Abbreviations</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Scope of inquiry . . . . .	1
1.2 Overview of the remaining chapters . . . . .	2
<b>2 Dissociating classifiers from counting</b>	<b>6</b>
2.1 Introduction . . . . .	6
2.2 The definition of ‘classifiers’ . . . . .	8
2.3 Classifiers are not for counting . . . . .	12
2.3.1 Optional classifier languages . . . . .	12
2.3.2 Optional classifier use in obligatory classifier languages . . . . .	15
2.4 Two little-discussed facts about classifier languages . . . . .	19
2.4.1 Three ways of interpreting noun phrases . . . . .	19
2.4.2 Fact 1: Classifiers prevent reference to subkinds . . . . .	21
2.4.2.1 Optional classifier languages: Malay and Persian . . . . .	22
2.4.2.2 Obligatory classifier languages: Japanese and Thai . . . . .	24

2.4.2.3	Crosslinguistic generalization . . . . .	26
2.4.2.4	Pragmatic factors affecting the ease of object/subkind reading . . . . .	29
2.4.3	Fact 2: Plurals do not denote kinds . . . . .	30
2.5	Summary . . . . .	33
<b>3</b>	<b>Analysis of the two little-discussed facts about classifier languages</b>	<b>34</b>
3.1	Introduction . . . . .	34
3.2	Background . . . . .	36
3.2.1	The ontology of the domain of individuals . . . . .	36
3.2.2	The relation between [object] and [kind] individuals . . . . .	41
3.2.3	Kinds . . . . .	46
3.3	Fact 1: Classifiers prevent reference to subkinds . . . . .	48
3.3.1	The idea . . . . .	48
3.3.2	An implementation in Linear Optimality Theory . . . . .	51
3.4	Fact 2: Plurals do not denote kinds . . . . .	57
3.4.1	Proposal . . . . .	57
3.4.2	Alternative accounts . . . . .	62
3.4.2.1	Individuality: Reduplication in Malay/Indonesian . . . . .	62
3.4.2.2	Associativity: <i>-Tati</i> in Japanese . . . . .	64
3.5	Summary . . . . .	70
<b>4</b>	<b>The semantics of classifiers and numerals</b>	<b>71</b>
4.1	Introduction . . . . .	71
4.2	The semantics of classifiers . . . . .	72
4.2.1	Classification function . . . . .	72
4.2.2	The core meaning . . . . .	79
4.2.3	A comparison with the “unit of counting” approach to classifiers . . . . .	85
4.3	The semantics of numerals . . . . .	87
4.3.1	Two kinds of numerals . . . . .	87
4.3.2	Compositional semantics for ‘three books’ crosslinguistically . . . . .	91
4.3.2.1	Malay: Both subsective and intersective numerals . . . . .	92
4.3.2.2	Persian: Only subsective numerals . . . . .	96
4.3.2.3	English: Only intersective numerals . . . . .	98

4.3.2.4	Watanabe's (2010) syntactic analysis . . . . .	98
4.4	Summary . . . . .	100
<b>5</b>	<b>Plural markers in classifier languages</b>	<b>101</b>
5.1	Introduction . . . . .	101
5.2	Asserted and implicated meanings of plural markers . . . . .	102
5.3	Plural markers in classifier languages and definiteness . . . . .	106
5.3.1	Plural markers $\neq$ definite markers . . . . .	106
5.3.2	Reassessment of the evidence for the definiteness of plurals . . . . .	110
5.4	Licensing by modification . . . . .	119
5.4.1	The facts . . . . .	119
5.4.2	Analysis . . . . .	124
5.4.3	Licensing by contrast . . . . .	128
5.5	Summary . . . . .	132
<b>6</b>	<b>Classifier languages and the typology of number marking</b>	<b>133</b>
6.1	Introduction . . . . .	133
6.2	The number system of classifier languages . . . . .	134
6.2.1	Classifiers and plural markers in classifier languages are true number morphology . . . . .	134
6.2.2	General number as the third number . . . . .	140
6.2.3	Optionality . . . . .	142
6.2.4	The syntax of DPs in classifier languages . . . . .	146
6.2.4.1	The DP structure . . . . .	146
6.2.4.2	D-Num agreement . . . . .	151
6.3	The typology of number marking and number features . . . . .	157
6.3.1	The typology of number marking . . . . .	157
6.3.2	Number features . . . . .	161
6.3.2.1	Morphology . . . . .	162
6.3.2.2	Semantics . . . . .	166
6.3.2.3	A comparison with previous theories of number features . . . . .	169
6.4	Implications for the theory of number and definiteness marking in kind terms . . . . .	173
6.5	Summary . . . . .	180

<b>7 Conclusion</b>	<b>182</b>
7.1 Summary of the dissertation . . . . .	182
7.2 D as the interface between two number systems . . . . .	185
7.3 Ramifications for related areas . . . . .	189
7.3.1 Language and thought . . . . .	189
7.3.2 First language acquisition . . . . .	189
<b>References</b>	<b>191</b>

# List of Tables

2.1	The effect of classifiers crosslinguistically. . . . .	28
3.1	The results of Matsumoto’s (1993) acceptability judgment experiment. . . . .	56
3.2	Two kinds of pluralities and their interpretations. . . . .	65
4.1	Combinations of numerals and classifiers in Japanese . . . . .	88
4.2	Number features and plural morphology licensing in Watanabe’s (2010) system. . . . .	99
6.1	The realizations of basic number categories in Malay and Japanese. . . . .	142
6.2	Crosslinguistic variations in ways of handling the feature-head mismatch. . . . .	150
6.3	Four basic number categories and their feature compositions. . . . .	162
6.4	Crosslinguistic variations in number feature realizations. . . . .	163
6.5	Possible and impossible lexicalization patterns of $\bar{\square}$ and $\iota$ (Dayal 2004b). . . . .	174
6.6	Number and definiteness marking in ‘SG/GN : GN/PL’ languages. . . . .	177
7.1	Basic number categories in natural languages. . . . .	184
7.2	The typology of number marking systems. . . . .	185

# List of Figures

3.1	The organization of the subdomains of individuals. . . . .	37
3.2	The relation between a kind and its subkinds. . . . .	37
3.3	An example of taxonomic hierarchies (Krifka et al. 1995:76). . . . .	38
3.4	Two subdomains of [kind] individuals. . . . .	40
3.5	A taxonomic hierarchy that guarantees the linkage between Figures 3.4a and 3.4b. . . . .	40
3.6	Figure 3.3 (reproduced). . . . .	40
3.7	Derivation of kinds from properties of [object] individuals. . . . .	47
3.8	Number-neutral and plural properties. . . . .	59
4.1	Two kinds of numerals. . . . .	89
4.2	Combination of a subsective numeral and ‘CL NP’. . . . .	90
6.1	Harley and Ritter’s (2002) feature geometry. . . . .	171
7.1	Overt determiners and two different number systems. . . . .	187

# List of Abbreviations

ACC	accusative	NOM	nominative
ASP	aspect	OBJ	object
ASSOC	associative	OM	specific object marker (Persian)
CLF	classifier	PASS	passive
COP	copula	PST	past
DAT	dative	PART	particle
DEF	definite	PRF	perfect
DUR	durative	PL	plural
EQ	equative	POL	polite
F	feminine	PRED	predicative
FOC	focus	PROG	progressive
GEN	genitive	PUNC	punctative
GN	general	Q	question
INF	infinitive	RECP	reciprocal
INDF	indefinite	REL	relativizer
INST	instrumental	SBJ	subject
LINK	linker	SG	singular
LOC	locative	SFP	sentence final particle
MOD	modifier	TOP	topic
NEG	negation		

# Chapter 1

## Introduction

### 1.1 Scope of inquiry

To many linguists and philosophers, the title of this study may not make sense at all. Classifier languages such as Japanese and Mandarin are often described as lacking number. Even those who acknowledge the presence of plural marker-like elements in classifier languages tend to refrain from treating them as genuine number markers on a par with, say, the plural marker *-s* in English for various reasons such as optionality and other meanings associated with them (e.g. Chierchia 1998b; Nakanishi and Tomioka 2004). For example, the suffix *-tati* in Japanese occurs in a noun phrase referring to multiple referents as in *gakusei-tati* [student-PL] ‘students’. However, reference to multiple referents is also possible for the same noun without *-tati*. Moreover, the referents of a noun phrase with *-tati* can also include individuals associated with the noun referent. Thus, *gakusei-tati* can mean ‘students and some non-students associated with the students’ too. Reflecting such a trend, classifier languages have had virtually no role in the theorizing of crosslinguistic variations in number marking (e.g. Corbett 2000). In this study, I argue that not only do classifier languages have genuine number morphology, but they have a rather complex number marking system.

Since countability and number morphology are thought to be closely related, such a claim has a major impact on another widely held assumption that common nouns in classifier languages, including conceptually count ones, are unindividuated and mass-like. Numerals cannot modify nouns directly, and a measure-word-like expression, a classifier, is required, as illustrated by the Mandarin examples in (1).

- (1) a. \*liang shu  
two book
- b. liang ben shu  
two CLF book  
'two books'

Analyses of classifier semantics have generally taken this assumption for granted and maintained that classifiers enable counting by individuating an unindividuated mass (e.g. Quine 1968; Krifka 1995; Chierchia 1998b, 2010; Borer 2005; Wilhelm 2008) (the “classifiers for counting” thesis). This study thus challenges this popular assumption and proposes an alternative analysis of classifier semantics.

Previous studies have regarded the difference between classifier and non-classifier languages as an extremely fundamental one: common nouns are individuated only in the latter language type, but not in the former. Chierchia (1998b), for instance, proposes a semantic parameter that connects this difference to the general availability of bare arguments (Nominal Mapping Parameter). This parameter makes the nominal syntax of the two language types very different: argument noun phrases do not project DP in classifier languages whereas they do in non-classifier languages. Li (1999) and others attempt to capture the classifier vs. non-classifier distinction by syntactic parameters. The present study, conversely, argues that the difference between classifier and non-classifier languages lies in morphology, number morphology to be more specific, and all argument noun phrases are DPs (cf. Longobardi 1994, 2005). The emerging picture of the nominal syntax and semantics of natural languages is therefore more universal than is previously thought, with morphological choices of individual languages replacing the semantic or syntactic parameters suggested by previous researchers.

## 1.2 Overview of the remaining chapters

In chapter 2, I first clarify what is meant by ‘classifiers’ in this study. I only consider as ‘classifiers’ the so-called ‘sortal classifiers’, which “name the unit in which the entity denoted by the noun naturally occurs” (Cheng and Sybesma 1999:515). Measure words meaning ‘cup’, ‘box’, ‘group’ and so forth are not classifiers but a type of nouns, as they impose a new unit on entities that come with inherent units as well as those that do not (i.e. mass nouns). Words meaning ‘kind’, ‘type’ and so forth are not considered as classifiers either, as they behave like

nouns. The chapter then invalidates the “classifiers for counting” thesis, based on the fact that direct numeral modification of nouns is generally possible in many classifier languages including Malay and Persian, and in fact also possible in languages in which classifier use has been described as obligatory. The chapter also presents two little-discussed empirical facts typically observed in classifier languages. Both facts are concerned with non-object reference and do not have bearing on counting. The first fact is that a subkind reading is readily available for expressions without a classifier, but it is impossible or very difficult to obtain for those with a classifier. The second fact is that bare plurals do not denote kinds in classifier languages.

Chapter 3 provides principled accounts for these facts. In doing so, I propose a model of the domain of individuals with two parallel subdomains of [object] and [kind]. I also present an explicit theory of the relation between ordinary objects, kinds and subkinds. A subkind reading is difficult to obtain for expressions with classifiers because classifiers are concerned primarily with properties of [object] individuals, but not with those of [kind] individuals. Consequently, a subkind reading does not satisfy the conditions specified by classifiers as faithfully as its object reading competitor. Plurals in classifier languages being incapable of denoting kinds pertains to their exclusive reference to pluralities, where singularities are not included unlike plural forms in English and other similar non-classifier languages. I argue that plurals in classifier languages do not denote kinds because the [object]-to-[kind] sort/type-shifter  $\cap$  is undefined not only for singularities (Chierchia 1998b), but also for pluralities; they are only defined for number-neutral properties.

Chapter 4 discusses the semantics of classifiers and numerals. I point out that the semantics of classifiers consists of two components, i.e. the asserted meaning and conventional implicatures. The former has to do with the singular number, i.e. restriction of the domain to singularities, whereas the latter is concerned with the noun classification function. Importantly, this analysis does not assume the “classifiers for counting” thesis. Although classifiers are not for counting (i.e. direct numeral modification), the context in which classifiers are most frequently used is in numeral modification constructions. Hence, the chapter also discusses the semantics of numerals that is compatible with the proposed classifier semantics. Following Bale et al. (2011), I propose two kinds of numerals, from which languages can choose one or both: subsective and intersective numerals. These two types differ in the type of objects for which they are defined. This definitional difference together with the choice of numeral type a language makes determines possible and impossible combinations of numerals and different nominal forms.

In chapter 5, I point out that like classifiers, the semantics of (some) plural markers in classifier languages also consists of two parts, i.e. the asserted meaning concerning number and conventional implicatures concerning the characteristics of the nouns combining with them. The parallelism not only shows that plural markers in classifier languages are genuine plural markers, as Chung (2000) and others claim, but it also explains the potential presence of multiple plural markers differing in applicable nouns, as attested in Japanese. The chapter also discusses the issue of definiteness. Plural markers in classifier languages have been often claimed to be definite in the literature (e.g. Yang 1998; Kurafuji 2004). However, I show that they are not necessarily definite; instead, they are referential, unless modified by quantifiers. Two apparent counterexamples exist to this claim. Plurals can be non-referential when they are modified or used contrastively. I claim that a non-referential interpretation is available in these cases because modifiers and contrast introduce a new situation variable that mediates between the situation variables of the NP and a higher predicate.

Chapter 6 summarizes the number system of classifier languages emerging from the discussions in the preceding chapters, and presents a typology of number marking that accommodates classifier languages properly. I argue that the general number, associated with number-neutral properties, should be treated as one of the basic number categories, along with the singular and plural. Given this conception of basic number categories, classifier languages do not lack number morphology, but on the contrary, they make the most fine-grained basic number distinction in the typology of number marking. They have distinct forms for the singular, plural and general, where the singular is realized by classifiers.

Recognizing the general number as one of the basic number categories is a crucial aspect of my number marking typology. Languages vary in how they realize the basic number categories morphologically. The difference between classifier and non-classifier languages thus stems from morphology, but not syntax or semantics. Classifier languages have distinct forms for all three basic number categories, whereas non-classifier languages morphologically conflate two or more basic number categories into one. The apparently optional use of classifiers and plural markers in classifier languages is a natural consequence of this difference, and is in fact evidence for, but not against, regarding classifiers and plural markers in classifier languages as genuine number morphology.

I argue that the three basic number categories of ‘singular’, ‘plural’ and ‘general’ are mere labels and can be decomposed into two binary number features, i.e. [ $\pm$ Sg] and [ $\pm$ Pl]. In syntax,

these features occur in the Num(ber) head, which is licensed by D by means of number and referentiality/argumenthood agreement. The two features have both morphological and semantic reflections. With regard to morphology, typically the positive value indicates the presence of the relevant marking and the negative one the lack thereof. As for the semantic reflection, the values of the two features determine the denotation of the ‘Num NP’ constituent, which is subject to further modification by numerals, if any. Assuming that NPs (before undergoing number specification by Num) universally denote number-neutral properties, [+Sg] and [+Pl] concern atomicity and divisibility respectively. A combination of two features is interpreted as the union of the two. This analysis creates two combinations for the general number, i.e. ‘[+Sg], [+Pl]’ and ‘[-Sg], [-Pl]’, which differ in the presence/absence of morphological marking. The availability of two kinds of generals account for the co-occurrence of classifiers and plural markers and bare “singular” kind terms in Brazilian Portuguese and other languages belonging to the same type. The former is often considered impossible, informed by Mandarin and Armenian facts (Li 1999; Borer 2005), but in fact possible in many classifier languages. The latter is mistakenly predicted to be impossible in Dayal’s (2004b) theory of number and definiteness marking in kind terms.

Chapter 7 is the conclusion. The chapter summarizes the chief proposals of this study, and discusses their implications for the number system in general, including that at work in the verbal domain. I also make short remarks on the ramifications of this study for studies on language and thought, and language acquisition.

## Chapter 2

# Dissociating classifiers from counting

### 2.1 Introduction

Numeral classifiers (henceforth ‘classifiers’) have been assumed to exist for the purpose of counting.<sup>1</sup> I refer to this assumption as the “classifiers for counting” thesis. In well-studied classifier languages, including Japanese and Mandarin, neither nouns considered conceptually mass (“stuff”, e.g. water) nor those considered to be count (“things”, e.g. books) can be modified by numerals directly, with the latter requiring a measure-word-like expression, a classifier, as in (1a) and (2a).

(1) Japanese

- a. hon ni \*(satu)  
book two CLF  
‘two books’
- b. mizu ni \*(hai)  
water two cup  
‘two cups of water’

(2) Mandarin

- a. liang \*(ben) shu  
two CLF book  
‘two books’

---

<sup>1</sup>In this study, I only refer to sortal classifiers as ‘classifiers’ and distinguish them from measure words, which include mensural classifiers. I will justify this decision in section 2.2.

- b. liang \*(bei) shui  
 two cup water  
 ‘two cups of water’

As this property parallels what one finds with mass nouns in non-classifier languages such as English (e.g. *a \*(piece of) furniture*), many linguists and philosophers have thought that nouns in classifier languages are mass(-like) and lack individuation, and hence they cannot be directly modified by numerals (e.g. Quine 1968; Krifka 1995; Chierchia 1998b, 2010; Borer 2005; Wilhelm 2008).<sup>2</sup> For instance, Chierchia (1998b) claims that while common nouns in non-classifier languages denote properties, common nouns in classifier languages denote kinds (the Nominal Mapping Parameter Hypothesis), which, according to his theory of mass terms and kinds, come to denote mass properties as a result of applying the ‘up’ operator ( $\cup$ ), which enables quantification.<sup>3</sup> Thus, “numerals will not be able to combine directly with nouns: a classifier will be necessary to individuate an appropriate counting level” (Chierchia 1998b:353–354). Krifka (1995) and Wilhelm (2008) propose analyses whereby the classifier subsumes a counting-unit-providing operator, which is part of either nouns or numerals in non-classifier languages. As counting is impossible without this operator in their systems, they also assume the “classifier for counting” thesis.

The first aim of this chapter is to invalidate the “classifiers for counting” thesis (section 2.3). I will do so by demonstrating that classifiers can be optional in many languages, which, to the best of my knowledge, has not received due attention in theoretical discussions of nominals, with only a handful of studies discussing them (Borer 2005; Simpson 2005; Bale and Khanjian 2009; Bale and Barner 2012). In these languages, numerals can generally modify nouns either directly or by the intermediary of classifiers.

This chapter has another aim, i.e. to present two little-discussed empirical facts typically observed in classifier languages (section 2.4). Both facts are concerned with non-object reference and do not have bearing on counting. The first fact is that expressions with and without classifiers differ with regard to the ease of obtaining a subkind reading. A subkind reading

---

<sup>2</sup>To do justice to Quine, he only suggests this view as one of two possible interpretations of the phenomenon and does not say that either is correct. The other interpretation is that noun denotations do not vary between classifier and non-classifier languages, but those of numerals do so: numerals in classifier languages decline according to the entity to be counted. As presented in chapter 3, Wilhelm (2008) can be seen as a present-day incarnation of this second interpretation.

<sup>3</sup>Common nouns in classifier languages no longer amount to masses under Chierchia’s (2010) revised theory of mass terms, though he still regards them as denoting kinds.

is readily available for expressions without a classifier, but it is impossible or very difficult to obtain for those with a classifier. Krifka (1995) has noted this interpretive difference through a crosslinguistic comparison between Mandarin and English, and proposed an explanation for it. Classifier optionality, invoked to invalidate the “classifiers for counting” thesis, enables to observe the difference in a single language. The second fact is that bare plurals do not denote kinds in classifier languages. This fact has been mentioned sporadically in studies of specific languages, but not as a recurrent pattern in classifier languages in general. This chapter only describes the facts. I will propose analyses of them in chapter 3.

## 2.2 The definition of ‘classifiers’

Before discussing the “classifiers for counting” thesis, it is necessary to clarify what is meant by ‘classifiers’ in this study. The term ‘classifiers’ has been used as an umbrella label for several noun classification systems, which include numeral classifiers, noun classifiers and verbal classifiers among others. The main concern of this study is numeral classifiers, i.e. noun-classifying morphemes that typically occur with numerals.

Two types of numeral classifiers have been distinguished in the literature, i.e. sortal classifiers and mensural classifiers. Cheng and Sybesma (1999:515) characterize the two (‘count-classifiers’ and ‘mass-classifiers/massifiers’ in their terminology) as follows: sortal classifiers “name the unit in which the entity denoted by the noun naturally occurs” whereas mensural classifiers “*create* a unit of measure.” I agree with Watanabe (2006) that mensural classifiers are in fact measure words (hence nouns), occurring in the pseudo-partitive construction such as *three cups of coffee*. Sortal classifiers are only compatible with count nouns whereas mensural classifiers are compatible with both count and mass nouns (Muromatsu 1998; Cheng and Sybesma 1999).<sup>4</sup> (3) illustrates this point.

---

<sup>4</sup>As many studies have shown, classifier languages do not lack the kind of count-mass distinction found in non-classifier languages (e.g. Kang 1994; Muromatsu 1998; Cheng and Sybesma 1999; Watanabe 2006; Yi 2010). They distinguish between count and genuine mass nouns like ‘water’ and ‘air’, which exclude fake mass nouns like *furniture* and *footwear* in English. For example, only count nouns are compatible with plural morphology and sortal classifiers; different (forms of) quantifiers, if any, are used for count and mass nouns. As in non-classifier languages, numerals cannot modify mass nouns directly in classifier languages. In both language types, mass nouns exhibit count syntax when interpreted as subkinds rather than particular objects and when an implicit container/portion is assumed, e.g. *five ice creams* meaning ‘five kinds/portions of ice cream’ (pace Dalrymple and Mofu 2012).

- (3) Japanese
- a. ni mai-no kooka/ \*oturi<sup>5</sup>  
two CLF-LINK coin/ change  
'two coins/\*change'
- b. huta-hako-no kooka/ oturi  
two-box-LINK coin/ change  
'two boxes of coins/change'

Mensural classifiers syntactically behave similarly to noun. In Mandarin, for example, mensural classifiers, but not sortal classifiers, can be optionally linked to the noun by what Cheng and Sybesma call 'modification marker' *de* as in (4).

- (4) Mandarin (Cheng and Sybesma 1999:515–516)
- a. Mensural classifiers
- (i) san bang (de) rou  
three pound MOD meat  
'three pounds of meat'
- (ii) liang xiang (de) shu  
two box MOD book  
'two boxes of books'
- b. Sortal classifiers
- (i) ba tou (\*de) niu  
eight CLF MOD cow  
'eight cows'
- (ii) jiu gen (\*de) weiba  
nine CLF MOD tail  
'nine tails'

Only mensural classifiers, but not sortal classifiers, can be modified by certain adjectives as in (6).<sup>6</sup>

<sup>5</sup>Examples like (ia) below do not count as counterexamples. In examples like this, the phrase *go mai* [five CLF] modifies the preceding noun, but not the noun following it. (ib) shows the relevant constituency.

- (i) a. hyaku-en-dama go mai-no oturi  
hundred-yen-coin five CLF-LINK change  
'change consisting of five hundred-yen coins'
- b. [hyaku-en-dama [go mai]]-no oturi

<sup>6</sup>Cheng and Sybesma (1998:note 4) note some exceptions to the second generalization, which include *yi chang bu*

- (5) Mandarin (Cheng and Sybesma 1999:516)
- a. Mensural classifiers
- (i) yi da zhang zhi  
one big sheet paper  
'one large sheet of paper'
- (ii) na yi xiao xiang shu  
that one small box book  
'that one small box of books'
- b. Sortal classifiers
- (i) \*yi da zhi gou  
one big CLF dog
- (ii) \*yi da wei laoshi  
one big CLF teacher

Assuming that *de* in Mandarin is more or less comparable to *of* in English, both of these properties are also found with the pseudo-partitive construction in English: *three cups of coffee*, *three large cups of coffee*. Mensural classifiers/measure nouns thus differ from sortal classifiers in that their nature is lexical rather than functional. Therefore, in this study, I only refer to sortal classifiers as 'classifiers' and distinguish them from mensural classifiers/measure nouns.

English does not have classifiers. However, it has two types of words that resemble classifiers. The first type is the words *item* and *piece*. These words happen to name the unit inherently associated with the atoms in the denotation of a noun (e.g. *items of clothing*, *pieces of furniture*), thus satisfying the conceptual criterion suggested by Cheng and Sybesma (1999) above. However, they actually have the properties of measure words. For example, they can be modified by adjectives rather freely and can be the host of plural morphology, as in *these five beautiful/unique and unusual/individually crafted items of jewellery*. As seen above, it is measure words that have the first property. As for the second property, classifiers do not take plural morphology, as illustrated in (6) and (7).

- (6) Japanese
- a. \*san nin-tati-no gakusei  
three CLF-PL-LINK student

---

*yingpian* [one long CLF film] 'a long movie' and *yi xiao duo huar* [one small CLF flower] 'a small flower'. Moreover, in some cases the acceptability varies among speakers (Borer 2005:99).

- b. san nin-no gakusei(-tati)  
three CLF-LINK student-PL  
'three students'

## (7) Malay

- a. \*kira-kira 300 buah-buah buku  
about 300 CLF-PL book
- b. kira-kira 300 buah buku(-buku)  
about 300 CLF book-PL  
'about 300 books'

Another type of English morphemes that may be confused with classifiers are words meaning 'kind', 'type' and the like. These words are not classifiers for the same reasons, i.e. they can be modified by adjectives and take plural morphology, as in *two different/new/fine kinds of wine*.

Some previous studies refer to words meaning 'kind', 'type' and the like in classifier languages (e.g. *zhong* in Mandarin) as 'kind classifiers'. Such words appear to occupy the same position as classifiers, as in (8).

## (8) Japanese

- a. ni tou/hiki/shurui/taipu-no kuma  
two CLF/CLF/kind/type-LINK bear  
'two heads/heads/kinds/types of bears'
- b. kuma ni tou/hiki/shurui/taipu  
bear two CLF/CLF/kind/type  
'two heads/heads/kinds/types of bears'

However, there are reason to think that these words are not classifiers but nouns. First, a construction involving such words is available not only in classifier languages but also in non-classifier languages, and as we have just seen, such words behave as nouns in the latter languages. Second, within classifier languages, these words behave differently from typical classifiers. For example, in Japanese classifiers cannot be modified by demonstratives (9a), but words meaning 'kind' can (9b).

- (9) a. \*kono tou/hiki-no kuma  
this CLF/CLF-LINK bear

- b. kono shurui/taipu-no kuma  
this kind/type-LINK bear  
'this kind/type of bear'

I assume that these words are count nouns that do not require classifiers when modified by numerals (cf. section 2.3.2 (ii)).

## 2.3 Classifiers are not for counting

### 2.3.1 Optional classifier languages

We will be forced to abandon the “classifiers for counting” thesis, the assumption that classifiers exists to enable counting, if languages exist with a general classifier system in which numerals can directly modify nouns without classifiers. In fact, such languages exist. I refer to them as ‘optional classifier languages’ as opposed to ‘obligatory classifier languages’. Not only do optional classifier languages exist, but there are many. Gil (2005) examines 400 typologically diverse languages and reports 140 of them as classifier languages. Of those 140 languages, 62 languages are identified as optional classifier languages. The languages that Gil (2005) categorizes as optional classifier languages include Minangkabau, Hungarian, Chantyal, Hatam, Tongan, Haida and Armenian. In this study, Malay is often chosen as a representative optional classifier language, as it is the optional classifier language with which I am most familiar. As in (10), both expressions with and without a classifier are grammatical in Malay. (See Gil (2005), Hamedani (2011), Bale and Barner (2012) and references cited therein for similar examples in other languages.)

- (10) Malay
- a. dua buah buku  
two CLF book  
'two books'
  - b. dua buku  
two book  
'two books'

A pair like (10) shows that numerals can be used independently of classifiers. It can be safely concluded that the “classifiers for counting” thesis cannot be maintained. Classifiers

are not required for numeral modification even in classifier languages. In optional classifier languages, one can use a classifier, as in (10a), like obligatory classifier languages, but one can also combine a numeral and noun directly, as in (10b), like non-classifier languages.

One may argue that a phonologically null classifier is present in (10b) (cf. Chung 2000; Csirmaz and Dékány 2010; Hamedani 2011). However, empirical facts exist that suggest positing a null classifier is not desirable. First, at least two interpretive differences exist between expressions with and without classifiers. They are concerned with subkind reference (section 2.4.2) and referentiality/specificity (chapter 6, section 6.2.1). Expressions with classifiers prevent subkind reference and tend to be referential/specific whereas those without classifiers do not. The account I propose for the former fact in chapter 3 makes use of the noun specification information of classifiers. Positing a null classifier can account for the first fact if the null classifier has no noun specification information. The second fact, however, remains unsolved under a null classifier analysis.

The second evidence against positing a null classifier comes from Malay. The numeral ‘one’ in Malay has both free and bound forms, i.e. *satu* and *se-*, as shown in (11a). The prefixal ‘one’ attaches only to classifiers (11a), measure words (11b–c), which impose a new measure unit not inherent to the noun, and nouns meaning ‘kind’, ‘type’ and so forth (11d), but not to nouns in general (11e).

- (11) a. satu/se- ekor kucing  
           one    CLF cat  
           ‘a cat’
- b. satu/se- cawan kopi  
           one    cup    coffee  
           ‘a cup of coffee’
- c. satu/se- kumpulan lelaki  
           one    group    man  
           ‘a group of men’
- d. satu/se- jenis  
           one    kind man  
           ‘one kind’
- e. satu/\*se- kucing  
           one    cat  
           ‘a cat’

A null classifier analysis wrongly predicts that the prefixal ‘one’ is grammatical in (11e), as it analyzes the phrase as in (12), where *se-* attaches to the classifier like (11a), but not to the noun.

(12) \**se-* Ø<sub>CLF</sub> kucing

Persian provides further evidence against positing a null classifier. Persian is an optional classifier language.

(13) Persian (Hamedani 2011:153)

- a. Bist tâ sarbâz dar xiyâbân bud-and.  
twenty CLF soldier on street be-PST.3PL  
‘Twenty soldiers were on the street.’
- b. Bist sarbâz dar xiyâbân bud-and.  
twenty soldier on street be-PST.3PL  
‘Twenty soldiers were on the street.’

While the occurrence of the plural marker *-hâ* is unconstrained in the absence of numerals, in the presence of numerals it is only possible if an overt classifier co-occurs, as shown in (14) (Gebhardt 2008; Ghaniabadi 2010; Hamedani 2011).<sup>7</sup>

(14) Persian (Hamedani 2011:153)

- a. Bist tâ sarbâz-hâ dar xiyâbân bud-and.  
twenty CLF soldier-PL on street be-PST.3PL  
‘The twenty soldiers were on the street.’
- b. \*Bist sarbâz-hâ dar xiyâbân bud-and.  
twenty soldier-PL on street be-PST.3PL

If a phonologically null classifier were present in (13b) and (14b), one would not expect a contrast like this. One could equip the null classifier with some special property that prohibits the plural marker co-occurring with it. However, such an analysis is no more than an ad hoc stipulation. The conclusion thus remains that in the ‘Num NP’ pattern, numerals directly modify nouns without the intermediary of classifiers, overt or covert, and no difference exists between

<sup>7</sup>The noun phrase in (ia) receives a definite interpretation. The issue of what brings about definiteness is an important research question and I will discuss it in chapter 5. However, it does not affect the discussion of whether a null-classifier is involved in the ‘Num NP’ pattern. I will discuss why (ib) is ungrammatical in chapter 4 (section 4.3.2).

classifier and non-classifier languages in the process whereby numerals and nouns are combined.

### 2.3.2 Optional classifier use in obligatory classifier languages

One might wonder whether such a conclusion is only possible for optional classifier languages and the “classifiers for counting” thesis is still valid in obligatory classifier languages. However, some facts suggest this is not the case. Even in obligatory classifier languages, classifiers become optional in certain contexts. I am aware of four such contexts.

**(i) Certain animate nouns** In Korean, classifiers are optional for an extremely limited number of animate nouns, including *salam* ‘person’, *haksayng* ‘student’, *kay* ‘dog’ (Kang 1994; Lee and Ramsey 2000).

**(ii) Abstract nouns** In Vietnamese, classifiers are not used for certain abstract nouns, e.g. *tu-tưởng* ‘thought’, *quan-điểm* ‘viewpoint’ and such (Nguyen 1957:131–132, Nguyễn 1997:94).<sup>8</sup> Classifiers are optional for some abstract nouns in Tai languages (Conklin 1981:364) and Japanese (Downing 1996:73). Moreover, according to Asmah (1972:94), in Malay and Iban (though these languages are optional classifier languages) certain abstract nouns do not take classifiers, that is, they only occur in the ‘Num NP’ pattern but not in the ‘Num CL NP’ pattern.

**(iii) Large and non-specific numbers** Classifiers tend to be optional or unacceptable for large numbers in many classifier languages. Classifiers are optional with multiples of ten in Nung (Saul and Wilson 1980:27) and Burmese (Hla Pe 1965),<sup>9</sup> and with large numbers such as 1000 in Thai (Aikhenvald 2000:100).

(15) Nung (Saul and Wilson 1980:27)

slám pǎm (áhn) hơn  
 three hundred CLF house  
 ‘three hundred houses’

<sup>8</sup>Bisang (1999) and Löbel (2000) present more cases in which classifiers are not obligatory in Vietnamese and claim that classifiers are not required for counting in Vietnamese.

<sup>9</sup>Precisely speaking, the Burmese case is not one of simple omission, as the order of the numeral and classifier is reversed in this case, i.e. from ‘NP Num CL’ to ‘NP CL Num’.

A similar tendency is also observed in Japanese, as in (16a). Classifiers can also be optional when the numeral slot is filled with non-specific numbers, as illustrated in (16b–c). Large numbers are non-specific in a sense, as less attention is paid to the individual members as the number increases.

(16) Japanese

- a. {san \*(ko)/ kyuu ?(ko)/ juu-go (ko)} -no gengo  
 three CLF nine CLF fifteen CLF -LINK language  
 ‘three/nine/fifteen languages’
- b. ni san (nin)-no gakusei  
 two three CLF-LINK student  
 ‘two or three/a few students’
- c. juu-suu (ko)-no sima  
 ten-some CLF-LINK island  
 ‘more than a dozen islands’

Large and non-specific numbers resemble vague-quantity expressions such as *many* and *roughly 100*. According to Matthews and Yip (2011:111), classifiers are optional with such expressions (‘relative quantifiers’ in their terminology) in Cantonese.<sup>10</sup>

(17) Cantonese (Matthews and Yip 2011:111)

- hóu dō (jī) bāt  
 very many CLF pen  
 ‘many pens’

---

<sup>10</sup>Note that the non-occurrence of classifiers with vague-quantity expressions may or may not have bearing with the tendency at issue. For example, the classifier does not occur in the Japanese equivalent of (17), as in (ia). However, this is simply because *ooku* ‘many’ is an adjective. Classifiers are only compatible with numerals in Japanese.

- (i) a. ooku (\*tu)-no pen  
 many CLF-LINK pen  
 ‘many problems’
- b. mit-\*(tu)-no pen  
 three-CLF-LINK pen  
 ‘three pens’

(iv) ‘NP(-case) Num CL’ In Japanese, classifiers can be omitted in the ‘NP-case Num CL’ pattern.<sup>11</sup> To my knowledge, the classifier being optional in this pattern has received little attention in the literature (but see Zubin and Shimojo 1993). The acceptability of ‘NP-case Num’ (without a classifier) varies from ‘completely unacceptable’ to ‘perfectly fine’, depending on both the specific construction in which the pattern occurs and the speakers. The constructions with which the pattern is most congruous include existential sentences (18) and lists (19). The (a) sentences have the ‘NP-case Num CL’ pattern, whereas the (b) sentences have the ‘Num CL-linker NP-case’ pattern. While classifiers are obligatory in the latter, they can be omitted in the former, though the result sounds a little awkward in (18a).

- (18) a. Too-hoteru-ni-wa resutoran-ga {itu-tu/ ?go} ari-masu.<sup>12</sup>  
 our-hotel-LOC-TOP restaurant-NOM five-CLF five be-POL  
 ‘Our hotel has five restaurants.’
- b. Too-hoteru-ni-wa {itu-tu/ \*go} -no resutoran-ga ari-masu.  
 our-hotel-LOC-TOP five-CLF five -LINK restaurant-NOM be-POL  
 ‘Our hotel has five restaurants.’
- (19) a. Kitte-o san (mai) to hagaki-o yon (mai) kudasai.  
 stamp-ACC three CLF and postcard-ACC four CLF give  
 ‘I’d like three stamps and four postcards.’
- b. San \*(mai)-no kitte to yon \*(mai)-no hagaki-o kudasai.  
 three CLF-LINK stamp and four CLF-LINK postcard-ACC give  
 ‘I’d like three stamps and four postcards.’

Classifier optionality in lists is also observed in other classifier languages. Gil (2005) notes that classifiers are usually not used in Vietnamese when one places an order at a food stall or when the waiter repeats his order. This example can be considered an instance of counting in a list.

In sum, classifier optionality is not restricted to optional classifier languages, but it is also found, though limitedly, in obligatory classifier languages. Both types of classifier languages allow classifier absence in the context of counting. Direct noun modification by numerals is thus, in principle, possible in classifier languages. The “classifiers for counting” thesis is untenable. There is no difference between classifier and non-classifier languages, either in the denotations of nouns and numerals or in the process whereby numerals combine with nouns.

<sup>11</sup>Whether ‘NP-case’ and ‘Num CL’ in this sequence (always) form a constituent is controversial. Nakanishi (2008) provides a good comprehensive overview of the issue.

<sup>12</sup>Both (a) and (b) sentences without a classifier improve slightly when *zenbude* ‘in total’ is added.

Consequently, the accounts of classifier existence proposed in previous studies that ascribe the difference between classifier and non-classifier languages to the difference in the denotations of nouns or numerals or the process whereby numerals combine with nouns (cf. section 2.1) have lost their basis.

At least two questions arise at this point. First, if classifiers are not for counting, why do they exist in the first place? An obvious function of classifiers is, of course, to classify nouns based on inherent features, including animacy, shape and size. I do not deny the classification function of classifiers. The question is instead how classification is related to the fact that all classifier languages use classifiers, whether optionally or obligatorily, in the counting context. Bisang (1999:144) claims that cognitively, classification is a prerequisite for counting: “Without classificational criteria for identifying an element as belonging to the set of items to be counted it would be impossible to decide whether to count that element or not.” However, this does not answer our question. For separation of elements to be counted from those not to be counted based on the characteristics inherent to a noun is achieved by nouns alone, far more precisely than by classifiers. Classifiers and counting are thus not in a means-goal relation. The second question is: why are classifiers found only in classifier languages (e.g. Japanese, Malay), but not in non-classifier languages (e.g. English, Hindi)?

Previous researchers have tried to answer the first question by comparing two types of languages, i.e. non-classifier languages such as English and obligatory classifier languages such as Japanese. We do not have to take such an indirect approach. It is more straightforward to examine minimal pairs in a single optional classifier language, i.e. expressions that only differ in the presence/absence of the classifier, as in the Malay examples in (10). They should reveal the core function of the classifier. In the next section, I will present two sets of data in classifier languages that have great significance to classifier semantics but have received little attention from researchers, i.e. the correlation between the availability of a subkind reading and the absence of classifiers (section 2.4.2) and the inability of plurals in classifier languages to denote kinds (section 2.4.3). Analyses of these facts will be proposed in chapter 3, where I will also discuss the relation of the core function of classifier to the classification function. I will propose an answer to the second question above in chapter 6.

## 2.4 Two little-discussed facts about classifier languages

### 2.4.1 Three ways of interpreting noun phrases

As seen above, optional classifier languages allow both expressions with and without classifiers. Both ‘Num NP’ and ‘Num CL NP’ are grammatical. There is, however, an interpretive difference between the two forms. The difference is concerned with whether a subkind reading is available.

To begin, it is necessary to distinguish between three ways to interpret noun phrases, i.e. the object, subkind and kind readings. On an object reading, reference is made to particular specimens of a kind. The italicized noun phrases in (20) receive an object reading.

(20) Object reading

- a. *The lion/Lions* escaped yesterday from the Hellabrunn zoo.
- b. *Gold* was stolen in yesterday’s bank robbery.
- c. *A cat* was sitting on the mat when John arrived at home.

(Krifka et al. 1995:5)

The subkind reading is also known as the ‘taxonomic reading’. On this reading, reference is made to subclasses of a kind. The noun phrases in italics in (21) illustrate the subkind reading.

(21) Subkind reading

- a. The World Wildlife Organization decided to protect *a (certain) large cat*, namely the Siberian tiger.
- b. *One metal*, namely copper, went strongly up on the market yesterday.

(Krifka et al. 1995:5)

Krifka et al. (1995) treat the type-token ambiguity as in (22) as an ambiguity between subkind and object readings.

- (22) a. *This book* got wet in the rain. (= copy; object reading)
- b. *This book* sells well. (= title; subkind reading)

(Krifka et al. 1995:77)

Subkind reading should not be confused with kind reading, in which reference is made to the entire kind to which particular specimens belong. This reading is exemplified by the italicized noun phrases in (23).

- (23) Kind reading
- a. *The lion* is a predatory cat.
  - b. *Lions* are predatory cats.
  - c. *Gold* is a precious metal.

(Krifka et al. 1995:5)

It is important to bear in mind that the three readings above are concerned with noun phrases. They are basically different interpretations of noun phrase. Though I will sometimes use the terms such as ‘object reading’ and ‘subkind reading’ loosely, to refer to sentences containing noun phrases with the relevant interpretation, this is only for convenience. Kind-referring noun phrases are distinct from the sentence-level phenomenon of characterizing sentences, which also express genericity. Hence, one cannot rush to regard a nominal form occurring in any generic sentences as referring to kinds. This is because the genericity expressed by a characterizing sentence does not stem from a particular noun phrase but from the whole sentence. Characterizing sentences thus can but do not have to contain kind-referring noun phrases. Examples of characterizing sentences are given in (24).

- (24) a. John smokes a cigar after dinner.  
 b. A potato contains vitamin C, amino acids, protein and thiamine.

(Krifka et al. 1995:3)

None of *John*, *a cigar* and *a potato* refer to kinds.

Krifka et al. (1995) introduce three diagnostic tests to distinguish kind-referring noun phrases from object-referring noun phrases. The most widely applicable of the three tests is that involving kind predicates. Kind predicates only take kind- and subkind-referring noun phrases as their arguments. For example, *be extinct* and *die out* take kind-referring noun phrases as their subjects, whilst *invent* and *exterminate* take kind-referring noun phrases as their objects. (25) illustrates how this test works.

- (25) a. *The lion* will become extinct soon.  
 b. *Lions* will become extinct soon.  
 c. *Bronze* is a metal/was invented as early as 3000 B.C.  
 d. \**A lion* will become extinct soon. (non-subkind reading)  
 e. *A (certain) lion* (namely the Berber lion) will become extinct soon. (subkind reading)

(Krifka et al. 1995:10)

(25a)–(25c) show that definite singular, bare plural and bare singular forms can refer to kinds. The contrast between (25d) and (25e) shows that an indefinite singular form cannot receive a kind reading while it can receive a subkind reading. A precondition for this test is that a language under investigation has unambiguous kind predicates. The second and third tests of Krifka et al. (1995) are more limited in their results and applicability. The second test seems to pick out noun phrases referring to well-established kinds (e.g. Coke bottles as opposed to green bottles) rather than kind-referring noun phrases as a whole, though Krifka et al. (1995) remain indecisive as to this point. The third test, which examines whether monotonicity holds in upward-entailing contexts, only works when the kind-referring noun phrase is not also in a characterizing sentence.

#### 2.4.2 Fact 1: Classifiers prevent reference to subkinds

The difference between expressions with and without a classifier can be summarized as follows: an expression without a classifier ('Num NP') has both an object and a subkind reading, whilst an object reading is impossible or very difficult to obtain for an expression with a classifier ('Num CL NP'). Two caveats must be mentioned here. First, though both an object and a subkind reading are available for 'Num NP', the former is usually far more salient than the latter. Moreover, certain pragmatic factors facilitate an object reading, which include generality of the noun, animacy, numeral largeness and world knowledge (see section 2.4.2.4 for further details).

Second, as stated in section 2.2, I assume that words meaning 'kind', 'type' and the like are not classifiers but nouns. The presence of these words in the lexicon makes referring to subkinds using 'Num NP' sound unnatural to varying degrees. The contrast between 'Num NP' and 'Num CL NP' discussed below is often so subtle that it may well be easily lost when

the researcher deliberately adds a stronger alternative expression like ‘Num kind NP’ in the comparison or when it is unexpectedly added in the consultant’s mind.

#### 2.4.2.1 Optional classifier languages: Malay and Persian

Let us consider sentences that illustrate the generalization above. The (a) sentences in (26)–(27) below (Malay) elicit an object reading, while the (b) sentences elicit a subkind reading. On an object reading, English *three magazines* means ‘three copies of magazines’ while on a subkind reading, the same phrase means ‘three titles of magazines’. While ‘Num NP’ is felicitous on both readings, ‘Num CL NP’ is only felicitous on an object reading as shown by the contrast between (27a) and (27b).

(26) ‘Num NP’

- a. Masih tinggal tiga majalah dan semua majalah itu majalah Mastika.  
still left three magazine and all magazine that magazine Mastika  
‘We still have three (copies of) magazines and all of them are *Mastika*.’
- b. Masih tinggal tiga majalah, iaitu majalah Mastika, Majalah PC dan  
still left three magazine namely magazine Mastika magazine PC and  
Nona.  
Nona  
‘We still have three (titles of) magazines, namely *Mastika*, *Majalah PC* and  
*Nona*.’

(27) ‘Num CL NP’

- a. Masih tinggal tiga buah majalah dan semua majalah itu majalah  
still left three CLF magazine and all magazine that magazine  
Mastika.  
Mastika  
‘We still have three (copies) magazines and all of them are *Mastika*.’
- b. ??Masih tinggal tiga buah majalah, iaitu majalah Mastika, Majalah PC  
still left three CLF magazine namely magazine Mastika magazine PC  
dan Nona.  
and Nona  
For: ‘We still have three (titles of) magazines, namely *Mastika*, *Majalah PC* and  
*Nona*.’<sup>13</sup>

<sup>13</sup>The sentence is acceptable if exactly one copy is left for each title, i.e. *Mastika* 1, *Majalah PC* 1, *Nona* 1. The interpretation involved in this case is an object reading, but not a subkind reading.

The generalization seems to hold in another optional classifier language, Persian. The native speakers I consulted gave mutually contradicting judgments to the Persian sentences similar to (26b) and (27b). One speaker rejected ‘Num NP’ and ‘Num CL NP’ for the subkind reading, whereas another speaker accepted both. Importantly, however, both speakers admit that the expected contrast between ‘Num NP’ and ‘Num CL NP’ existed when the numeral was *bist* ‘20’ instead of *se* ‘3’ (and the second clause was omitted). Although it is not as felicitous as the alternative with the noun *now?* ‘kind’ (28c), ‘Num NP’ (28a) can be used for a subkind reading, whereas ‘Num CL NP’ (28b) cannot.

(28) Persian

- a. ?Dar injâ bist majalle vojud dêrad.  
in here 20 magazine existence have  
For: ‘We have 20 (kinds of) magazines here.’
- b. ??Dar injâ bist tâ/jeld majalle vojud dêrad.  
in here 20 CLF magazine existence have  
For: ‘We have 20 (kinds of) magazines here.’
- c. Dar injâ bist now? majalle vojud dêrad.  
in here 20 kind magazine existence have  
‘We have 20 kinds of magazines here.’

The classifier is normally used in (colloquial) Persian, though its omission does not lead to ungrammaticality (Hamedani 2011). Ghomeshi (2003) even states that classifiers are obligatory with numerals except ‘one’. Persian thus resembles obligatory classifier languages, including Japanese. Recall the general tendency that classifier absence is tolerated more readily when the numeral expresses a large number (cf. section 2.3.2 (iii)). This tendency explains why the contrast becomes clearer with the numeral ‘20’ than ‘3’. When the numeral expresses a large number, we can observe the interpretive felicitousness of the two expressions without being obscured by the difference in acceptability of their forms. (However, the numeral cannot be too large (e.g. 143), because large numbers bias towards an object reading.)

The consultants’ mutually contradicting judgments for the Persian sentences in (28a) and (28b) can be understood as follows. For both speakers, the existence of the alternative with *now?* ‘kind’ made ‘Num NP’ and ‘Num CL NP’ sound inadequate, as the alternative expression is employed exclusively for a subkind reading. The first speaker rejected both patterns for this reason. The second speaker told me that she first thought that the relevant nominal forms

referred to particular specimens of magazines, but she had to change her initial interpretation when she encountered the part of the sentence that forced a subkind reading (omitted in (28)). This self-report suggests that her judgments resulted from coercion, though one could also argue that it simply reflects the first caveat mentioned above, i.e. an object reading is far more salient than a subkind reading. The relevant coercion is a conscious process, unlike the kind of coercion often discussed in relation to expressions such as *Ali and some boys* and *I finished the book yesterday*, which is subconscious.

#### 2.4.2.2 Obligatory classifier languages: Japanese and Thai

This contrast is not peculiar to Malay or other optional classifier languages. In fact, similar contrasts are also observed in some obligatory classifier languages in contexts where classifiers become optional (cf. section 2.3.2). When the classifier can be omitted from the ‘NP-case Num CL’ pattern in Japanese, ‘NP-case Num’ without a classifier has both object and subkind readings, but ‘NP-case Num CL’ with a classifier has only an object reading. In (29)–(30), the (a) sentences elicit an object reading, where ‘nine pets’ must be nine specimens of the same kind of pet (i.e. cats here), and the (b) sentence a subkind reading, which is concerned with the kinds of pets. The pattern with a classifier is not felicitous on a subkind reading, as in (30b).

(29) ‘NP-case Num’

- a. Uti-de-wa petto-o zenbude kyuu atukatteiru keredo, zenbu neko da.  
we-at-TOP pet-ACC in.total nine deal.in and all cat COP  
‘We have nine pets in total, and all of them are cats.’
- b. Uti-de-wa petto-o zenbude kyuu atukatteiru keredo, neko-ga itiban ureteiru.  
we-at-TOP pet-ACC in.total nine deal.in and cat-NOM most sell  
‘We have nine (kinds of) pets in total, and cats sell the most.’

(30) ‘NP-case Num CL’

- a. Uti-de-wa petto-o zenbude kyuu hiki atukatteiru keredo, zenbu neko da.  
we-at-TOP pet-ACC in.total nine CLF deal.in and all cat COP  
‘We have nine pets in total, and all of them are cats.’
- b. ??Uti-de-wa petto-o zenbude kyuu hiki atukatteiru keredo, neko-ga itiban  
we-at-TOP pet-ACC in.total nine CLF deal.in and cat-NOM most  
ureteiru.  
sell  
For: ‘We have nine (kinds of) pets in total, and cats sell the most.’

Another example showing the same point is given in (31). The predicate *zetumetus(uru)* ‘to become extinct’ is a kind predicate, which is only compatible with kind-referring arguments. Only ‘NP-case Num’ without a classifier is acceptable in this sentence.

- (31) Zetumetusi-souna tora-ga sukunakutomo ni (\*hiki) iru.<sup>14</sup>  
 extinct-likely tiger-NOM at.least two CLF be  
 ‘There are at least two tigers that are likely to become extinct.’

The second example of the interpretive contrast between expressions with and without classifiers in obligatory classifier languages comes from Thai. In Thai, one cannot omit classifiers in the counting context.

- (32) nǎŋsǎm sǎwŋ \*(lêm)  
 book two CLF  
 ‘two books’

However, classifiers may be omitted in structures involving demonstratives, as in (33).<sup>15, 16</sup>

- (33) nǎŋsǎm (lêm) ní  
 book CLF this  
 ‘this book’

According to Piriyawiboon (2010), the same contrast found between ‘Num NP’/‘NP-case Num’ and ‘Num CL NP’/‘NP-case Num CL’ in Malay and Japanese is found between ‘NP Dem’ and ‘NP CL Dem’ in Thai. ‘NP Dem’ has both an object and a subkind reading, whereas ‘NP CL Dem’ has only an object reading. (34) demonstrates the relevant contrast.

- (34) a. rót ní  
 car this  
 (i) ‘this particular car’, (ii) ‘this kind of car’

<sup>14</sup>The ‘NP-case Num CL’ pattern, i.e. *tora-ga ni hiki*, is acceptable if the sentence is intended to mean that there are at least two tiger individuals, all of which are of a kind likely to become extinct. This reading is an object reading.

<sup>15</sup>The same is true in the “less precise speech” of Vietnamese (Nguyen 1957:129).

<sup>16</sup>Some authors (e.g. Croft 1994; Aikhenvald 2000) distinguish noun-classifying morphemes that do not occur with numerals from those that do by referring to them as ‘deictic classifiers’ and so on as opposed to ‘numeral classifiers’ (= ‘classifiers’ in this study). These labels imply the presence of multiple homophonous classifier morphemes in the lexicon. I do not adopt such a view. Instead, I assume that only one classifier morpheme exists that can have multiple functions depending on the syntactic or semantic contexts in which it occurs, just as a morpheme manifests itself in different guises or allomorphs in different contexts.

- b. rôt khan ní  
 car CLF this  
 (i) ‘this particular car’, (ii) \*‘this kind of car’

(Piriyawiboon 2010:85)

### 2.4.2.3 Crosslinguistic generalization

Although the (b) pattern in (34) is not available in Malay, Persian, Japanese or English, the (a) pattern in these languages shows the same interpretive possibilities as in Thai.<sup>17</sup>

- (35) a. M: kereta ini [car this]  
 P: in mâšîn [this car]  
 J: kono kuruma [this car]  
 E: this car  
 (i) ‘this particular car’, (ii) ‘this kind of car’
- b. M: \*kereta buah ini [car CLF this], \*buah kereta ini [CLF car this]  
 P: \*in tâ mâšîn [this CLF car]  
 J: \*kono dai kuruma [this CLF car], \*kono kuruma dai [this car CLF]  
 E: (unavailable)

(M: Malay; P: Persian; J: Japanese; E: English)

Conversely, Cantonese and Mandarin lack the (a) pattern in (34). The (b) pattern in these languages shows the same interpretive possibilities as in Thai.

- (36) a. C: \*lī chē [this car]  
 Mn: \*zhe che [this car]
- b. C: lī ga chē [this CLF car]  
 Mn: zhe jia che [this CLF car]  
 (i) ‘this particular car’, (ii) \*‘this kind of car’

(C: Cantonese; Mn: Mandarin)

<sup>17</sup>Thai also allows classifiers optionally in other contexts in which classifiers may not occur in Malay, Persian or Japanese, i.e. when the modifier of a noun is the interrogative modifier *náy* ‘which’, adjectives and prepositional phrases (Iwasaki and Ingkaphirom 2005:65–67). It is not clear whether the same interpretive contrast reported by Piriyawiboon (2010) is observed in these contexts as well, though the present study predicts that it is.

These facts suggest that the relevant contrast is a general one in that it is not restricted to the counting context and is observed regardless of whether classifiers are obligatory, optional or absent in the languages in consideration. I thus propose the following generalization:

(37) The semantic effect of classifiers

Crosslinguistically, the presence of a classifier prevents a subkind reading, which is otherwise available in addition to an object reading.

Table 2.1 shows that the generalization above holds true in all language types, i.e. Japanese, Cantonese (obligatory classifier languages), Malay, Persian (optional classifier languages) and English (a non-classifier language), to the extent that the relevant expressions exist. In addition to these languages, Thai is included in the table to show the interpretive possibilities of structures involving a classifier and demonstrative, which are unavailable in the former languages except Cantonese. ‘—’ indicates that the relevant expression is unavailable or ungrammatical.

Lastly, I would like to point out an exception to the generalization in (37): a general/default classifier in Japanese, namely *tu*, allows a subkind reading.<sup>18</sup> The classifier *tu* can substitute for most specific classifiers for inanimate entities (Zubin and Shimojo 1993) and can be also used when specific classifiers are blocked (Watanabe 2012). The contrast between (38a) and (38b) below shows that the expression with *tu* can refer to subkinds much more easily compared to that with the specific classifier for machines *dai*.

- (38) a. Josei-ga yoku kau kuruma-ga yot-tu aru.  
 woman-NOM often buy car-NOM four-CLF be  
 ‘There are four (types of) cars often bought by women.’
- b. ??Josei-ga yoku kau kuruma-ga yon dai aru.  
 woman-NOM often buy car-NOM four CLF be  
 For: ‘There are four (types of) cars often bought by women.’

My analysis of the generalization (37) to be proposed in chapter 3 predicts this fact. The fact is thus more a supporting fact than a problem in the long run. The case of *tu* is a language-specific fact of Japanese, but not a crosslinguistically valid general/default classifier property. The classifier *tâ* in Persian is more general than *tu* in Japanese, being compatible with both inanimate and animate entities, but expressions with *tâ* are nevertheless not felicitous for a

<sup>18</sup>I thank Akira Watanabe for pointing out this important fact to me.

Table 2.1: The effect of classifiers crosslinguistically.

	object reading	subkind reading
<b>Expressions without classifiers</b>		
(a) NUM NP/NP NUM		
Japanese ( <i>*zassi san</i> )	—	—
Cantonese ( <i>*sāam jaahpji</i> )	—	—
Malay ( <i>tiga majalah</i> )	✓	✓
Persian ( <i>bist majalle</i> )	✓	✓
English ( <i>three magazines</i> )	✓	✓
(b) DEM NP/NP DEM		
Japanese ( <i>kono kuruma</i> )	✓	✓
Cantonese ( <i>*lī chē</i> )	—	—
Thai ( <i>rót nīi</i> )	✓	✓
Malay ( <i>kereta ini</i> )	✓	✓
Persian ( <i>in mâšīn</i> )	✓	✓
English ( <i>this car</i> )	✓	✓
<b>Expressions with classifiers</b>		
(a) NUM CL NP/NP NUM CL		
Japanese ( <i>zassi sansatu</i> )	✓	*
Cantonese ( <i>sāam bún jaahpji</i> )	✓	*
Malay ( <i>tiga buah majalah</i> )	✓	*
Persian ( <i>bist tâ majale</i> )	✓	*
English (unavailable)	—	—
(b) DEM CL NP/NP CL DEM		
Japanese ( <i>*kono dai kuruma</i> )	—	—
Cantonese ( <i>lī ga chē</i> )	✓	*
Thai ( <i>rót khan nīi</i> )	✓	*
Malay ( <i>*kereta buah ini</i> )	—	—
Persian ( <i>*in tâ mâšīn</i> )	—	—
English (unavailable)	—	—

subkind reading, as presented above.

#### 2.4.2.4 Pragmatic factors affecting the ease of object/subkind reading

The ease of testing the generalization in (37) varies from sentence to sentence. This is because some extralinguistic factors exist that favor one reading over the other. In general, a subkind reading is more difficult to obtain than an object reading. This is because (i) we talk more about objects than about (sub-)kinds and (ii) a stronger alternative expression is available for a subkind reading, i.e. one containing a word meaning 'kind'. The different degrees of acceptability reported by my Persian consultants in (28) illustrates this situation. Although the sentence without a classifier was judged better for a subkind reading than the sentences with a classifier, it did not sound as good as the sentence with the noun *now?* 'kind'.

On top of this general bias towards the object reading, I am aware that the following four more specific factors favor one reading over the other.

**Generality of the noun** The more specific the noun, the more difficult a subkind reading becomes. In (39), the subject noun phrases are intended to denote the same referent, say, four Yukon Gold potatoes. The two nouns differ in terms of their generality, *vegetables* is more general than *potatoes*. Although more than three kinds of potatoes are normally available at (American) supermarkets and hence should be familiar to many people, *three potatoes* is not as easy to receive a subkind reading as *vegetables*.

(39) I planted three vegetables/potatoes today.

**Animacy of the noun** Nouns of high animacy bias towards the object reading. For instance, in (40), the ease of an object reading is higher in the following order: *person* (human) > *cat* (non-human animal) > *flower*' (plant).

(40) I rarely see this person/cat/flower in my area.

**Largeness of the numeral** Large numbers bias towards the object reading. In (41), while the object reading is easily available for both numbers, the subkind reading is more difficult to obtain for *143* than *5*. *1,000 trees* in (42) is more likely to receive an object reading, even though the number of existing tree species far exceeds 1,000 and some large arboretums do have

more than 1,000 different tree species. In contrast, the book title *Fifty Trees of Indiana* (by T. E. Shaw) is understood as fifty different kinds of trees in the state of Indiana more naturally, or at least as naturally as, fifty most unique or popular individual trees there.

(41) Our library has 5/143 linguistics journals.

(42) This arboretum has 1,000 trees from around the world.

**World knowledge** These two sentences also indicate that the ease of object/subkind reading also depends on world knowledge, as largeness is a relative concept that can vary in different contexts. For a high school library, where linguistics is not usually taught as subject, possessing even a single copy of linguistics journal is probably worth a comment. Hence, *five linguistics journals* is more likely to mean ‘five copies of a linguistics journal’. Conversely, the same phrase would more likely mean ‘five titles of linguistics journals’ for university libraries, given that university libraries usually subscribe at least one linguistics journal. Similarly, the interpretation of *1,000 trees* depends on one’s assumption about the number of trees that an average arboretum (for him/her) has. *Fifty trees* in the book title *Fifty Trees in Indiana* naturally receives a subkind reading because fifty is within the range of numbers we expect for the tree species available in a state and also because individual trees usually do not draw our attention, unlike individual restaurants (e.g. a *Top Ten Restaurants* list normally lists ten individual restaurant names rather than ten kinds of restaurants).

### 2.4.3 Fact 2: Plurals do not denote kinds

The second little-discussed fact about classifier languages is concerned with plurals. Bare plurals in classifier languages are freely available but do not denote kinds. This situation differs from languages like English and Hindi, in which bare plurals are freely available and can denote kinds. It also differs from languages like French and Italian, in which bare plurals are highly restricted and do not denote kinds. Specific language studies have mentioned this fact sporadically (Indonesian: Sneddon et al. (2010), Sato (2008); Javanese: Sato (2008); Mandarin: Iljic (1994), Yang (1998) *inter alia*; Japanese: Nakanishi and Tomioka (2004), Mizuguchi (2004); Korean: Mizuguchi (2004), Kwon and Zribi-Hertz (2004); and Burmese: Mizuguchi (2004)). The fact is also observed in Jingpo and Persian. Though no one has claimed that the phenomenon is peculiar to classifier languages, these languages all happen to be classifier languages. I am thus

inclined to take this fact as indicating that the inability of plurals to denote kinds is related to the existence of a general classifier system in a principled way. I will articulate how the two are related in chapter 3, where I will also discuss other meanings associated with plurals in classifier languages, as some researchers invoke these meanings instead of plurality to explain the inability of plurals to denote kinds in respective languages.

Let us see concrete examples from Japanese (an obligatory classifier language), Malay, Jingpo and Persian (optional classifier languages).<sup>19</sup> In Japanese, the plural is marked by suffixes *-tati*, *-ra*, *-gata* and *-domo* (see Mizuguchi (2004:90–95) for a description of these four suffixes). In Malay, the plural is expressed by full reduplication as in *harimau-harimau* ‘tigers’. In Jingpo and Persian, the plural is marked by the suffixes *-ni* and *-hâ* respectively. As shown in the (b) examples in (43)–(46), plural forms are not compatible with kind predicates.

(43) Japanese

- a. Kyooryuu-wa zetumetusi-ta.  
dinosaur-TOP become.extinct-PST  
‘Dinosaurs became extinct.’
- b. #Kyooryuu-tati-wa zetumetusi-ta.<sup>20</sup>  
dinosaur-PL-TOP become.extinct-PST

(44) Malay

- a. Telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
telephone PASS-invent by Alexander Graham Bell at year 1876  
‘The telephone was invented by Alexander Graham Bell in 1876.’
- b. #Telefon-telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
telephone.PL PASS-invent by Alexander Graham Bell at year 1876

(45) Jingpo<sup>21</sup>

- a. 1876 ning e Graham Bell gaw fon sawk gyin ai.  
1876 year in Graham Bell TOP telephone invent SFP  
‘Bell invented the telephone in 1876.’
- b. #1876 ning e Graham Bell gaw fon-ni sawk gyin ai.  
1876 year in Graham Bell TOP telephone-PL invent SFP

<sup>19</sup>Hamedani (2011) state that plurals in Persian denote kinds. However, all her examples only show that plurals can occur in characterizing sentences and no adequate example is given that shows they can denote kinds.

<sup>20</sup>The sentence is acceptable when it means ‘The dinosaur “and some other species” became extinct’, because the suffix *-tati* makes not only additive but also associative plurals (cf. section 3.4.2.2).

<sup>21</sup>I would like to thank Hideyuki Onishi for his help in collecting Jingpo data.

- (46) Persian
- a. Bel telefon râ exterâ? kard.  
Bell telephone OM invent did  
'Bell invented the telephone.'
- b. #Bel telefon-hâ râ exterâ? kard.  
Bell telephone-PL OM invent did

This situation is in contrast with non-classifier languages such as English and Italian, in which morphologically plural forms refer to kinds, as in the Italian example in (47) and its English translation.

- (47) Italian (Chierchia 1998b:342)
- I cani sono rari.  
the dogs are rare  
'Dogs are rare.'

Though a kind reading is impossible for plurals, the other two readings, i.e. object and subkind readings, are available. In (48)–(49), the (a) and (b) sentences illustrate an object reading and a subkind reading respectively. In the (b) sentences, the plural noun phrase supplies a set of dinosaur subspecies, and the subject indicated by the brackets picks out a subset of that set. As the subject contains a kind predicate meaning 'to be extinct', these sets must be ones of subkinds.

- (48) Japanese
- a. Kyooryuu-tati-wa mada neteiru.  
dinosaur-PL-TOP still be.sleeping  
'The dinosaurs are still sleeping.'
- b. Kyooryuu-tati-nonakani-wa [sono toki zetumetusi-ta mono]-mo i-ta.  
dinosaur-PL-among-TOP that time become.extinct-PST one-too be-PST  
'Among the dinosaurs there were also some (subspecies) which went extinct at that time.'
- (49) Malay
- a. Dinosaur-dinosaur masih tidur.  
dinosaur.PL still sleep  
'The dinosaurs are still sleeping.'

- b. Ada di antara dinosaur-dinosaur [yang pupus pada masa itu].  
 be at among dinosaur.PL REL extinct at time that  
 'Among the dinosaurs there were also some (subspecies) which went extinct at  
 that time.'

That an object reading is available for plurals is no surprise. However, it is quite surprising that plurals are felicitous on a subkind reading despite their incompatibility with a kind reading, given that subkinds are subclasses of a kind.

## 2.5 Summary

In this chapter, I invalidated the popular assumption that classifiers exist for the sake of counting, enabling otherwise impossible numeral modification of nouns. Released from this influential but problematic assumption, I pointed out two little-discussed facts about classifier languages that have little to do with counting. First, the presence of a classifier prevents a subkind reading, which is otherwise obtained easily in addition to an object reading. Second, plurals do not denote kinds, but they can denote subkinds and objects. The two facts presented in this chapter will be analyzed in chapter 3.

## Chapter 3

# Analysis of the two little-discussed facts about classifier languages

### 3.1 Introduction

This chapter proposes principled accounts for the two little-discussed facts about classifier languages discussed in chapter 2. To recapitulate, the first fact is that the presence of a classifier prevents a subkind reading, which is otherwise available in addition to an object reading. The Malay and Thai examples below, involving numeral and demonstrative modification respectively, show the interpretive difference arising from the presence of a classifier.

- (1) Malay
  - a. tiga majalah  
three magazine
    - (i) ‘three copies of magazines’
    - (ii) ‘three kinds of magazines’
  - b. tiga buah majalah  
three CLF magazine
    - (i) ‘three copies of magazines’
    - (ii) ??‘three kinds of magazines’

- (2) Thai (Piriyawiboon 2010:85)
- a. rôt níi  
car this  
(i) ‘this particular car’  
(ii) ‘this kind of car’
  - b. rôt khan níi  
car CLF this  
(i) ‘this particular car’  
(ii) \*‘this kind of car’

Krifka (1995) has noted this interpretive difference through a crosslinguistic comparison between Mandarin and English, and proposed an explanation for it. I will put forward a new analysis of the fact, as Krifka’s analysis is based on the problematic “classifiers for counting” thesis, an assumption that classifiers exist to enable numeral modification of nouns. Unlike Krifka (1995), my analysis does not posit a special counting unit for subkinds nor a kind classifier.

The second fact is that bare plurals do not denote kinds in classifier languages as in (3)–(4).

- (3) Malay
- a. Telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
telephone PASS-invent by Alexander Graham Bell at year 1876  
‘The telephone was invented by Alexander Graham Bell in 1876.’
  - b. #Telefon-telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
telephone.PL PASS-invent by Alexander Graham Bell at year 1876
- (4) Persian
- a. Bel telefon râ exterâ? kard.  
Bell telephone OM invent did  
‘Bell invented the telephone.’
  - b. #Bel telefon-hâ râ exterâ? kard.  
Bell telephone-PL OM invent did

Although the fact has been mentioned sporadically in studies of specific languages, no satisfactory account has been provided for this recurrent pattern. I will propose an analysis according to which the inability to denote kinds results from a property common to plurals in classifier languages, i.e. plurality, rather than other properties associated with plurals such as associativity

and definiteness.

I will first lay out the background for the analysis in section 3.2, and then present my analyses of these two facts in sections 3.3 and 3.4.

## 3.2 Background

### 3.2.1 The ontology of the domain of individuals

I follow Carlson (1977:69) and assume that individuals consist of two basic sorts, i.e. [object] individuals and [kind] individuals. I will refer to these two sorts by superscripts ‘o’ and ‘k’ respectively. The semantic types of [object] individuals and properties are thus  $e^o$  and  $\langle e^o, t \rangle$ , and those of [kind] individuals and properties are  $e^k$  and  $\langle e^k, t \rangle$ .

Following the standard view, I assume that the domain of individuals is structured as a complete atomic join-semilattice (Link 1983; Landman 1989). According to this view, the domain contains both singular and plural individuals, with the latter being sums of the former. Individuals are ordered by a ‘part of’ relation ( $\leq$ ). As [object] and [kind] individuals are both individuals, their (sub)domains have identical structures. The difference between the two subdomains lies in what the atoms in the structure represent. In the subdomain of [object] individuals, the atoms are single specimens of a kind, whereas they are subkinds of a kind in the subdomain of [kind] individuals. For notation, I use lowercase letters for [object] individuals, i.e. specimens, and uppercase letters for [kind] individuals, i.e. kinds and their subkinds. I distinguish kinds from their subkinds with a prime symbol (′) on the latter:  $A'$  is a subkind of  $A$ .<sup>1</sup> With these notations, the subdomains of [object] and [kind] individuals associated with a kind  $A$  can be represented as in Figure 3.1. In each subdomain, the row at the bottom represents singularities and the two rows above it represent pluralities. The lines connecting individual nodes capture the ‘part of’ relation:  $a_1 \sqcup a_2 \leq a_1 \sqcup a_2 \sqcup a_3$ ,  $a_1 \leq a_1 \sqcup a_2$ , etc.

Two things must be noted regarding the subdomain of [kind] individuals. First, the kind  $A$  does not exist in Figure 3.1b. The top node  $A'_1 \sqcup A'_2 \sqcup A'_3$  cannot be identified with  $A$ , as the former is a plurality whilst the latter is a singularity. However, they are related to each other. Specifically, I treat a kind as a group, in the sense of Link (1984) and Landman (1989),

<sup>1</sup>A more rigorous way of notation would use numbers: e.g.  $A^{n-1}$  is a subkind of  $A^n$ . This is because kinds are hierarchically ordered according to a taxonomic hierarchy. A kind can be a subkind of a superordinate kind, which may in turn be a subkind of a kind at a higher level. For example, the kind ‘tiger’ has the Malayan, Sumatran and Siberian tigers as its subkinds, but it is also a subkind of a superordinate kind, say, the cat or Felidae.

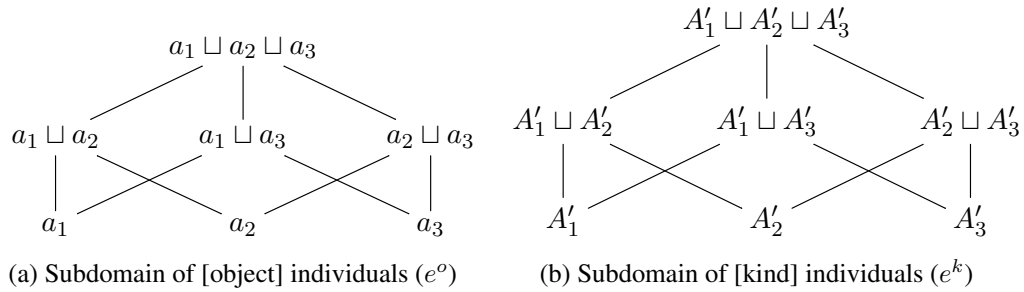


Figure 3.1: The organization of the subdomains of individuals.

formed by the sum of all its subkinds, i.e.  $A = \uparrow (A'_1 \sqcup A'_2 \sqcup A'_3)$ , where  $\uparrow$  is a group formation function that maps a sum of individuals to an atomic group individual. Groups differ from sums/pluralities in that their constituent members are not transparent and treated as an (impure) atom/singularity. Figure 3.2 shows how a kind and its subkinds are related to each other via  $\uparrow$ . Note that kind  $A$  and subkinds  $A'_1, A'_2$ , etc. are not of different nature in the sense that the former is a group whereas the latter is not. Subkinds are also groups. For example,  $A'_1$  is a group formed by the sum of all its subkinds (= “subsubkinds” for  $A$ ). If  $A'_1$  has three subkinds  $A''_1, A''_2$  and  $A''_3$ , then  $A'_1 = \uparrow (A''_1 \sqcup A''_2 \sqcup A''_3)$ . If subkinds do not have subclasses in any situation, I consider that they have themselves as their subkinds, that is to say, they have only one subkind that is identical to themselves:  $A'_1 = \uparrow A'_1$  and  $A'_1 = A''_1$ .

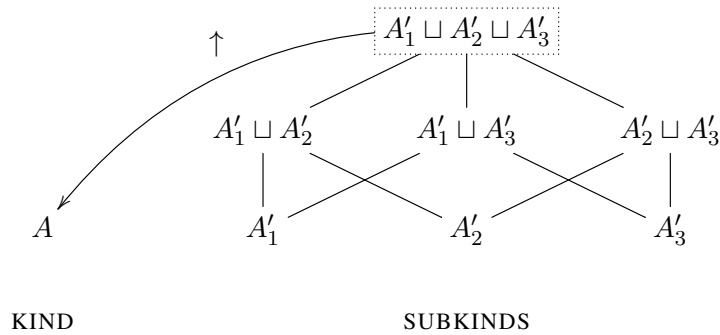


Figure 3.2: The relation between a kind and its subkinds.

Second, the subdomain of [kind] individuals must be distinguished from taxonomic hierarchies, such as the one shown in Figure 3.3. A taxonomic hierarchy is a list of names organized based on taxonomy/group relations. It differs from the subdomain of [kind] individuals in that

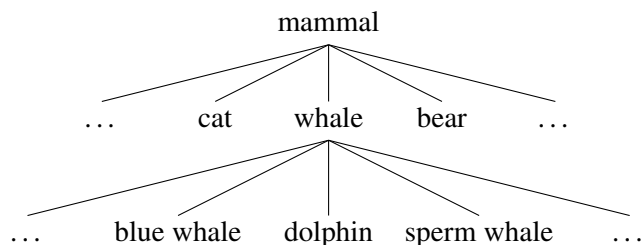


Figure 3.3: An example of taxonomic hierarchies (Krifka et al. 1995:76).

it only contains names of natural kinds and does not have a complete join-semilattice structure. As Dayal (2004b) notes, not all subkinds and their sums have a corresponding node in a taxonomy hierarchy. Elements in the [kind] subdomain are thus not necessarily natural kinds. For example, the sum of tigers and whales, which are two subkinds of mammals, do not form a natural kind and thus have no corresponding node in a taxonomic hierarchy.

Dayal (2004b) refers to the subdomain of [kind] individuals as ‘the taxonomic domain’. I find this naming misleading, as it gives the wrong impression that the two are the same thing and quantification occurs in a taxonomic hierarchy. However, neither is the case. If quantification occurred in a taxonomic hierarchy, we would have a very different mechanism for quantification over subkinds than for quantification over ordinary objects. This is because the domains of quantification would be structured differently in the two cases. The domain would have a lattice structure for ordinary objects but not for subkinds (compare Figures 3.4b and 3.5 below). However, no empirical fact suggests such a big difference. In fact, Zamparelli (1998) proposes a system where the denotations of common nouns may include taxonomic hierarchies. (5) gives an example, where the superscript ‘k’ indicates that it is a kind.

$$(5) \quad \llbracket [\text{NP animal}] \rrbracket = \{ \text{animal}^k, \text{mammals}^k, \text{tigers}^k, \text{Bengali-tigers}^k, \text{cats}^k, \text{Siamese-cats}^k, \dots \}$$

To allow numeral modification, he creates a lattice by applying a power-set operator  $\text{Pw}'$  to this set. However, the result includes a number of deviant sums, such as ‘ $\text{animal}^k \sqcup \text{tigers}^k \sqcup \text{Bengali-tigers}^k$ ’, which will never occur in my system. Zamparelli is aware of this problem and qualifies that these problematic sums are ruled out either by pragmatics or some appropriate filter incorporated into the definition of  $\text{Pw}'$ . There is a significant difference between my system and Zamparelli’s in the conception of taxonomic hierarchies. They are part of lexical meanings

and thus linguistic in Zamparelli's system. Conversely, they are extra-linguistic psychological entities referenced by language in my system. A node in a taxonomic hierarchy normally has a name. However, this is not always the case. For instance, the traditional type of mails as opposed to modern electronic ones are surely a subcategory of mails in most people's taxonomic hierarchies. They are given specialized names such as 'snail mails' for some people, but not for others. The latter group of people refer to the relevant node in a taxonomic hierarchy by describing it as in *ordinary mails*. It is because of the presence of such unnamed categories that I regard taxonomic hierarchies as psychological rather than linguistic.

I have argued above that a kind is a group formed by all its subkinds. According to this view, the function of a taxonomic hierarchy is to link one structured kind to another structured kind by identifying an atom of the former with the group formed by the totality of the subkinds of the latter. Suppose we have two subdomains of [kind] individuals, as in Figure 3.4. Given the taxonomic hierarchy in Figure 3.5, it is possible to link the two structures by identifying an atom in Figure 3.4a with the group formed by the sum of all atoms in Figure 3.4b, i.e.  $\text{Whale} = \uparrow (\text{Blue whale} \sqcup \text{Dolphin} \sqcup \text{Sperm whale})$ . This linking is not possible if our model has the taxonomic hierarchy in Figure 3.3, repeated below for convenience, because other whale subkinds exist in this list, indicated by '...' at the same level.

Taxonomic hierarchies are also referenced when evaluating expressions that can potentially receive a subkind reading. These expressions sound awkward if no taxonomic hierarchy is available that contains information about the subkinds of the relevant kind. This explains the contrast between *dinosaurs* and *dodos* in (6). The predicates in (6) require the subject to refer to subkinds of a kind. *Dodos* sounds awkward, because unlike dinosaurs, dodos are not known to have distinct subkinds, and there is no taxonomic hierarchy containing the subkinds of dodos.

- (6) a. Dinosaurs evolved/became extinct at different times.  
 b. #Dodos evolved/became extinct at different times.

It is clear from the structures shown in Figure 3.1 that subkind reference can be made in the same way as reference to particular objects, with the only difference being that the subdomain consists of [kind] individuals. For example, *two tigers* denotes the set of pluralities consisting of two atoms. If the relevant pluralities are [object] individuals (i.e.  $\{a_1 \sqcup a_2, a_1 \sqcup a_3, a_2 \sqcup a_3\}$  in Figure 3.1a), the phrase means 'two particular tigers'. If they are [kind] individuals (i.e.  $\{A'_1 \sqcup A'_2, A'_1 \sqcup A'_3, A'_2 \sqcup A'_3\}$  in Figure 3.1b), the phrase means 'two subspecies of tiger'.

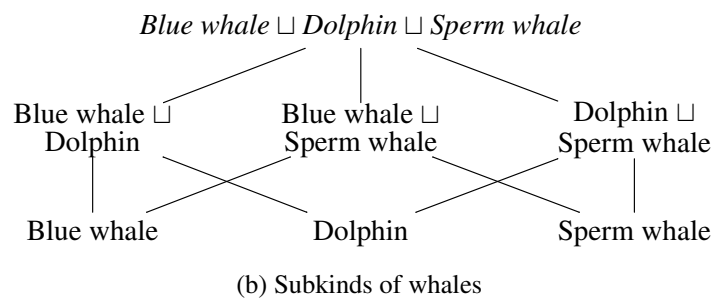
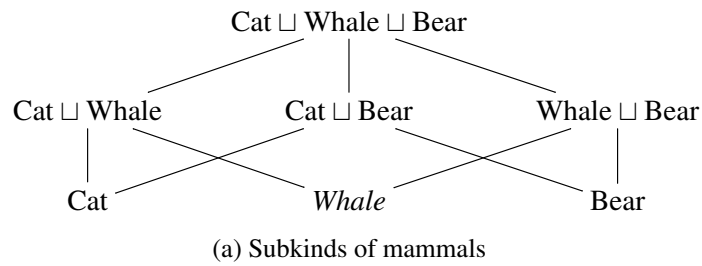


Figure 3.4: Two subdomains of [kind] individuals.

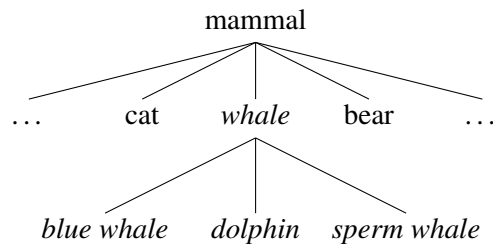


Figure 3.5: A taxonomic hierarchy that guarantees the linkage between Figures 3.4a and 3.4b.

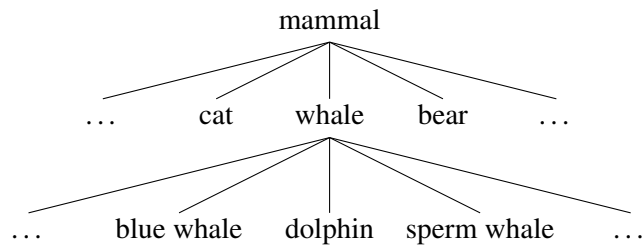


Figure 3.6: Figure 3.3 (reproduced).

### 3.2.2 The relation between [object] and [kind] individuals

I assume that [object] and [kind] individuals exist on an equal footing side-by-side. That is, it is not the case that one is basic and the other is derived from it. Most common nouns are thus ambiguous between objects and (sub-)kinds, as evidenced by the expression *two tigers* above. Dölling (1995) and Nomoto (2010) assume that [kind] is the basic sort, and derive [object] individuals by postulating a sort/type-shifting function that maps [kind] individuals to corresponding [object] properties (INST/Ins). This approach can handle the first fact about classifier languages, i.e. classifiers prevent a subkind reading. Nomoto (2010), for example, incorporates a sortal restriction to [object] in the semantics of overt classifiers. However, this approach cannot account for variations among languages in number (and definiteness) marking in kind terms (cf. Dayal 2004b). If [kind] were the basic sort, there should be no such variations and kind terms should be unmarked in number morphology in all languages.

The opposite approach, i.e. to regard [object] as the basic sort and [kind] derived from it, also faces an empirical problem. Specifically, this approach fails to account for the subkind reading available for singular forms in English such as *a/one animal* ('an/one animal species') and *this car* ('this kind of car'). As seen in Figure 3.2, subkinds are related to the relevant kind in such a way that the latter is a group (i.e. an impure atom/singularity) formed by the totality of the former. In other words, subkinds are members of a kind. Following Chierchia (1998b), I assume that [object] individuals are connected to kinds rather than subkinds. A subkind reading is thus obtained by applying a member specification function and an entity-to-property type shifter to a kind.<sup>2</sup> Given that kind- and subkind-denoting nominal forms are identical in the case of plurals as in *tigers* and *(two) tigers*, the relevant operations are not associated with overt

<sup>2</sup>The technical details involved in this process are as follows. The member specification operator  $\downarrow$  is an inverse function of the group formation operator  $\uparrow$  and turns a group, which is a singular individual of type  $e^k$ , to the sum of all its members, which is a plural individual also of type  $e^k$ :

$$(i) \quad \downarrow A = \downarrow \uparrow (A'_1 \sqcup A'_2 \sqcup A'_3) = A'_1 \sqcup A'_2 \sqcup A'_3$$

The standard type shifter from  $e$  to  $\langle e, t \rangle$  is Partee's (1987) Ident. However, a slight modification is necessary for our purpose, as our ontology of individuals is richer than that assumed by Partee, containing singular and plural individuals. The modified Ident in (ii) substitutes the 'part-of' relation  $\leq$  for  $=$ . Applying this modified Ident to (i) results in the desired property, as shown in (iii).

$$(ii) \quad \text{Ident}(x) = \lambda y[y \leq x]$$

$$(iii) \quad \text{Ident}(A'_1 \sqcup A'_2 \sqcup A'_3) = \{A'_1 \sqcup A'_2 \sqcup A'_3, A'_1 \sqcup A'_2, A'_1 \sqcup A'_3, A'_2 \sqcup A'_3, A'_1, A'_2, A'_3\}$$

morphology. Singular count nouns in English can denote kinds only when accompanied by the definite article *the* as in *the tiger*. Given these facts, it is expected that singular forms denoting subkinds must also contain *the*, contrary to the fact.

This problem does not arise if [object] individuals are connected directly to subkinds. Carlson (1977:chapter 6) proposes such a direct connection. Specifically, he postulates a lexical rule that derives for every noun a homophonous noun denoting the relevant subkinds. For example, the noun *dog* basically applies to particular dogs, and this lexical rule derives a new noun, *dog'*, which is pronounced the same as the original noun but applies to subkinds of dogs such as poodles and retrievers. This subkind-deriving lexical rule is subject to a condition: it is applicable only if the lexicon contains nouns naming the relevant subkinds. The rule can apply to *dog* because the lexicon of English contains nouns such as *poodle* and *retriever*, whereas it cannot apply to *mallard* due to the lack of specific nouns naming mallard subkinds. Carlson claims that the condition must be stated as one on the lexicon because one can talk about subkinds of mallards by means of the *kind of* construction. Consider the contrast shown in (7).

- (7) a. ?This mallard is extremely common.  
 b. This kind of mallard is extremely common.

Pointing out that abstract nouns pattern like kind terms rather than object terms, Carlson (1977:chapter 7) extends the subkind-deriving lexical rule to abstract nouns such as *virtue* and *courage*. According to his analysis, abstract nouns denote non-count [object] individuals, and the relevant lexical rule derives from them homophonous nouns denoting count [kind] individuals, i.e. (sub-)kinds. The evidence comes from the fact that some abstract nouns have corresponding plural forms whilst others do not, as in (8).

- (8) a. (i) much virtue/science  
 (ii) many virtues/sciences  
 b. (i) much courage/precision  
 (ii) \*many courages/precisions

(Carlson 1977:298)

Carlson claims that abstract nouns such as *courage* and *precision* do not have plural forms

because the lexicon does not contain nouns naming their subkinds, and hence the subkind-deriving rule, which turns a non-count [object] individual into a count [kind] individual, cannot apply to them. Moreover, as with concrete nouns (cf. (7)), it is possible to talk about subkinds by means of the *kind of* construction, even though the noun itself does not denote subkinds.

(9) They exhibited the kind of courage characteristic of all devoted hockey players.

(Carlson 1977:299)

In fact, the facts discussed by Carlson can also be accounted for by the model adopted in the present study, where the [kind] subdomain is accessible independently, without a rule deriving [kind] individuals from the corresponding [object] individuals. First, the difficulty in obtaining a subkind reading for a concrete noun as in (7a) is due to the difficulty in coming up with an existing taxonomic hierarchy containing information about the subkinds of that entity (see the discussion around (6)). This account shares the basic intuition with Carlson's account but differs from his in that it is not a lexical rule but is based on a non-linguistic construct referenced by language, i.e. the taxonomic hierarchy. The same line of account explains why some abstract nouns lack plural forms as in (8b), insofar as the plural forms are intended to denote subkinds. This qualification also applies to Carlson's account, as abstract nouns can bear a plural marker only after the subkind-deriving lexical rule has applied to them.

Next, Carlson accounts for the ability of the *kind of* construction to refer to the subkinds that cannot be referred to by simple nouns as in (7b) and (9) by analyzing *kind (of)* as a function mapping objects to subkinds (*of* is regarded semantically empty). A simplified version of Carlson's translation of *kind (of)* is given in (10), where  $R(y, x)$  means *y* realizes *x*:

(10)  $\lambda P \lambda x^k [\forall y^o [R(y^o, x^k) \rightarrow P(y^o)]]$  (cf. Carlson 1977:212)

Importantly, the semantics of *kind (of)* in (10) is not subject to the condition applicable to the subkind-deriving lexical rule.

By contrast, in my model, the noun *kind* is just one of the many abstract nouns that denote in the [kind] subdomain. Unlike Carlson's analysis, *kind* does not alter the sort of its argument from [object] to [kind]. Instead, it simply denotes [kind] individuals and the argument restricts them only to the relevant ones. The expression *kind (of)* thus translates as in (11), where *P*

stands for a [kind] property.<sup>3</sup> *Kind (of)* is semantically vacuous in the [kind] subdomain, its role being to rule out the possibility of interpreting an expression in the [object] subdomain.

$$(11) \quad \lambda P \lambda x^k . P(x^k)$$

The reason why *kind of* constructions can refer to any subkinds has to do with their structural complexity. In his recent work, Carlson (2009) assumes that concepts as meanings are limited to lexical items, i.e. words, including compounds, as opposed to phrases (cf. Murphy 2002), and claims that complex phrases such as *hungry or thirsty dog* would not take concepts as their meanings, whether the corresponding concepts such as HUNGRY-OR-THIRSTY-DOG exist or not; only lexical items such as *dog* denote concepts. Since taxonomic hierarchies are psychological constructs that capture inclusion relationships among concepts, Carlson's claim indicates that taxonomic hierarchies are not referenced in evaluating the meanings of complex expressions including *kind of* constructions. Therefore, no constraint exists as to the subkinds that *kind of* constructions can refer to, unlike those that can be referred to by simple nouns such as *mallard* or *courage*. *Kind of* constructions can refer not only subkinds that have corresponding nodes in existing taxonomy hierarchies, but also to those that do not. The latter case looks as if a new temporary node is added to the existing taxonomic hierarchy.

What is more, my model can account for some facts that were pointed out by Carlson (1977) but have not received an adequate account to date. First, a subkind reading becomes notably easier to obtain when the noun is modified, as shown by the contrast in (12).

- (12) a. ??At Seiko, they make a watch.  
 b. At Seiko, they make a watch that also serves as a juice squeezer and as a tire pump.

(Carlson 1977:299)

<sup>3</sup>Wilkinson (1995) points out that *kind of* constructions can be parsed in two ways. In one structure, *kind* and its argument form a constituent as in [*that [kind of animal]*]. In another structure, *kind* and the determiner forms a constituent and modify the common noun after *of* as in [[*that kind*] *of animal*]. She proposes that in the second structure, the transitive *kind* as in (10) and (11) combines with *that* by function composition, as shown in (i).

- (i) a.  $\llbracket \text{that}_i \rrbracket = \lambda Q [\nu y [Q(x_i) \wedge y = x_i]]$   $\langle et, e \rangle$   
 b.  $\llbracket \text{kind} \rrbracket = \lambda Q \lambda x^k . Q(x^k)$   $\langle et, et \rangle$   
 c.  $\llbracket \text{that}_i \text{kind} \rrbracket = \llbracket \text{kind} \rrbracket \circ \llbracket \text{that}_i \rrbracket = \lambda P [\nu y [P(x_i^k) \wedge y = x_i^k]]$   $\langle et, et \rangle \circ \langle et, e \rangle = \langle et, e \rangle$

The irrelevance of taxonomic hierarchies in evaluating the meanings of complex expressions is not specific to *kind of* constructions but a general semantic principle. Hence, while (12a) with an unmodified noun requires a suitable taxonomy hierarchy, (12b) with the same noun with a relative clause modifier does not. The contrast in acceptability in (12) is thought to arise because no suitable taxonomic hierarchy is available for the nouns *watch*. Carlson also points out the same contrast with regard to abstract nouns, as in (13).

- (13) a. ?Sally has a fear.  
b. Sally has a fear of small creatures.

(Carlson 1977:299)

This contrast receives the same account as given to (12), because abstract nouns can denote in the [kind] subdomain and does not have to be derived from non-count nouns in the [object] subdomain in my system.<sup>4</sup> These contrasts are puzzling in Carlson's theory. This is because the subkind-deriving lexical rule should be able to apply either to both or to neither, as the two sentences contain exactly the same noun.

My parallel subdomain model and Carlson's single primary subdomain model differ as to what is responsible for certain properties of [kind] individuals. In my model, the properties of [kind] individuals reflect the general architecture of the domain of individuals, particularly the nature of the [kind] subdomain. In Carlson's model, on the other hand, they result from particular lexical items and rules. A common noun can denote subkinds, because there exists a

<sup>4</sup> In this connection, Moltmann (2004a,b) argues that some abstract nouns, namely nominalizations, do not denote (properties of abstract) objects, unlike the explicit property-referring terms containing their corresponding adjectives. The two types of expressions receive different interpretations when they occur with evaluative predicates. Nominalizations are understood as referring to concrete particulars. (ia) thus means that friendly behavior, gestures, remarks, etc. are nice. Conversely, (ib) means that an abstract object is nice.

- (i) a. Friendliness is nice.  
b. The property of being friendly is nice.

(Moltmann 2004a:7)

The (in)compatibility with kind-predicates indicates that nominalizations denote kinds whereas the corresponding explicit property-referring terms do not.

- (ii) a. Honesty is rare.  
b. ??The property of being honest is rare.

(Moltmann 2004a:21)

homophonous (object-denoting) noun and the subkind-deriving lexical rule meets the condition of its application, which is concerned with the presence of lexical items that denote the relevant subkinds in the lexicon. *Kind of* constructions can denote subkinds because the noun *kind* functions as an [object]-to-[kind] sort-shifter as in (10). In fact, Carlson’s semantics of *kind* is much more complex than (10), with two additional conditions. These conditions ensure that only the subkinds of a kind is included, that is, the (superordinate) kind itself is not included, and that the relevant subkinds are disjoint (e.g. *two kinds of dogs* does not refer to poodles and toy poodles).<sup>5</sup> These two conditions need not be stipulated as part of the meaning of *kind (of)* in my model because they follow as direct consequences of the proposed ontology of individuals. The subdomain of [kind] individuals is structured as a complete atomic join-semilattice, which consists only of the subkinds of a kind (cf. section 3.2.1). An atom, i.e. a minimal subkind, cannot be part of another atom by definition.

To summarize, both empirical and theoretical considerations lead us to the conclusion that [object] and [kind] individuals exist in parallel, and it is not the case that one is basic and the other is derived from it.

### 3.2.3 Kinds

Although [object] and [kind] individuals exist on an equal footing, they are not completely independent of each other. Chierchia (1998b) proposes operations to connect the two subdomains. He argues that kinds can be treated as individual correlates of properties of their specimens. Given that bare NPs are initially of type  $\langle e^o, t \rangle$ , the corresponding [kind] individual is obtained by applying the nominalization operator  $\sqcap$  to the former.  $\sqcap$  is defined as in (14).<sup>6</sup>

---

<sup>5</sup>Carlson (1977:215) gives the following translation for the expression *kind of dog*, where  $S_o$  is a free variable ranging over the (sub-)kinds the derived noun *dog* may denote. The second and third lines handle these conditions respectively.

(i)  $\lambda x^k [S_o(x^k) \wedge \forall z^o \exists y^k \square [\text{dog}'(z^o) \leftrightarrow S_o(y^k) \wedge R(z^o, y^k)] \wedge$   
 $\sim \square \sim \exists w^o [\text{dog}'(w^o) \wedge \sim R(w^o, x^k)] \wedge$   
 $\sim \exists w^o \exists z^k \exists y^k [z^k \neq y^k \wedge S_o(y^k) \wedge S_o(z^k) \wedge R(w^o, z^k) \wedge R(w^o, y^k)]]$

<sup>6</sup>Not all NPs satisfying the morphosyntactic requirements for kind terms qualify as kinds (Carlson 1977; Krifka et al. 1995). For example, *people in the next room* does not denote a kind, though bare plurals in English generally denote kinds (see Carlson (1977:194–200) for a discussion of why this is so). In Chierchia’s (1998b) system,  $\sqcap P$  is undefined in such cases.

$$(14) \quad \sqcap P = \lambda w[\iota x P_w(x)]$$

(Chierchia 2010:115)

According to (14), kinds are individual concepts, specifically “functions that at any world yield the totality of the manifestations of that kind in that world” (Chierchia 2010:115). The semantic type of kinds is therefore  $\langle s, e \rangle$ . In what follows, I will substitute  $e^k$  for this complex type for simplicity and talk as if kinds obtained through  $\sqcap$  were individuals. The kind thus obtained can technically be a plurality. As our intuition is that a kind is a singular entity, the resulting plural entity is a kind “emulation” distinct from true kinds. For it to count as a true kind, it must be turned into an atom. I propose to do so by applying the group formation operator  $\uparrow$ . Figure 3.7 visually shows how kinds are derived from properties of [object] individuals.

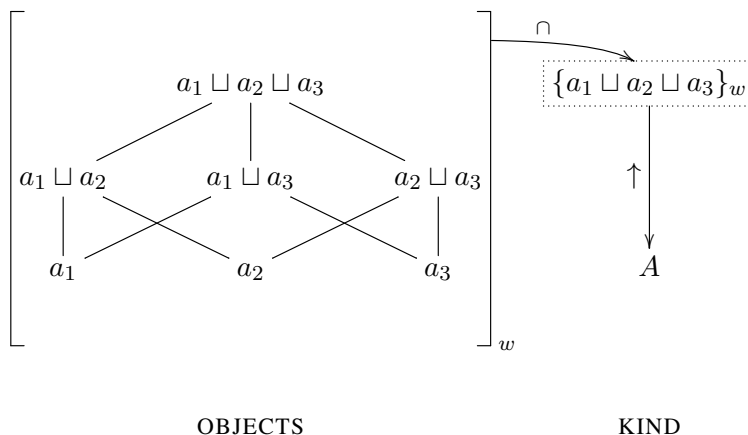


Figure 3.7: Derivation of kinds from properties of [object] individuals.

The distinction between kind emulations ( $\{a_1 \sqcup a_2 \sqcup a_3\}_w$ ; plurality) and true kinds ( $A$ ; singularity) coincides with that between the “plural” and “singular” kind terms. Dayal (2004b) demonstrates that, based on the contrast seen in (15), singular kind terms differ from plural ones in that the former do not have a transparent relation to their instantiations.

- (15) a. Rutgers professors seem to be born on weekdays.
- b. #The Rutgers professor seems to be born on weekdays.

(Edwin Williams (p.c.) cited in Dayal (2004b))

While both sums and groups have plural members, the constituent members are only transparent in the former but not the latter (Link 1984; Landman 1989). (15b) is infelicitous in the intended generic reading because *be born on weekdays* cannot be predicated of the group members, i.e. all Rutgers professors. I concur with Dayal's (2004b) claim that a singular kind term denotes in the [kind] subdomain that only contains the relevant [kind] individual ( $A$  in Figure 3.7) but not its subkinds, hence the singular morphology. To justify this claim, Dayal cites Jespersen (1927), who suggests that "the singular generic term 'denotes the kind itself' while the plural denotes 'the members of the species'," captured exactly by the system depicted in Figure 3.7. In what follows, I will not pay particular attention to the distinction between the two types and refer to them collectively as 'kinds', unless the distinction is crucial.

The processes to obtain different interpretations of nominals presented above are common to both classifier and non-classifier languages. As shown in chapter 2 (section 2.3), numerals can directly modify common nouns in classifier languages, either generally (optional classifier languages) or in limited cases (obligatory classifier languages). It is thus natural to think that there is no difference in the denotation of common nouns between classifier and non-classifier languages; in both language types, bare NPs denote properties of objects/subkinds ( $\langle e^o, t \rangle / \langle e^k, t \rangle$ ), contrary to the Nominal Mapping Parameter Hypothesis (Chierchia 1998b), according to which common nouns in classifier and non-classifier languages denote kinds ( $e^k$ ) and properties ( $\langle e^o, t \rangle$ ) respectively.

Let us now examine how the two facts about classifier languages pointed out in chapter 2 are accounted for against this background.

### 3.3 Fact 1: Classifiers prevent reference to subkinds

#### 3.3.1 The idea

The first fact is that while expressions without a classifier have both an object and a subkind reading, a subkind reading is impossible or very difficult to obtain for those with a classifier. The contrast between (16a) and (16b) illustrates this point.

- (16) Malay
- a. tiga majalah  
three magazine  
(i) ‘three copies of magazines’  
(ii) ‘three kinds of magazines’
- b. tiga buah majalah  
three CLF magazine  
(i) ‘three copies of magazines’  
(ii) ??‘three kinds of magazines’

Krifka (1995) has pointed out the same contrast between expressions without a classifier in English and expressions with a classifier in Mandarin. He ascribes the contrast to the type of counting unit encoded by numerals (English) or classifiers (Mandarin). Specifically, while the counting unit introduced by numerals in English applies to both objects and subkinds, the counting unit introduced by classifiers in Mandarin applies to either objects or subkinds.<sup>7</sup> This analysis is problematic as it assumes the “classifiers for counting” thesis, which I invalidated in chapter 2. Krifka’s analysis wrongly predicts that expressions without a classifier like (16a) are ungrammatical in a language that has expressions with a classifier as in (16b). Moreover, it regards the words meaning ‘kind’ as special classifiers for counting subkinds. I have shown that these words are in fact not classifiers but nouns in terms of their syntactic behaviors (cf. chapter 2, section 2.2).

I claim that a subkind reading is difficult for expressions with a classifier because the same expression can be interpreted more faithfully in an object reading compared to a subkind reading. Specifically, the object reading satisfies all pieces of noun specification information of the classifier, whereas the subkind reading only satisfies one of them, i.e. ‘inanimate’.

Classifiers specify the characteristics of nouns with which they combine in terms of animacy, shape and size, etc. These characteristics are primarily concerned with concrete entities,

<sup>7</sup>Specifically, Krifka analyzes *three bears* in English and *san zhi xiong* [three CLF bear] ‘three bears’ in Mandarin as in (i).  $RT_i(x, k)$  means that  $x$  is a specimen or a subspecies, or an individual sum of specimens or subspecies, of kind  $k$  in possible world  $i$ . OU and KU stand for ‘object unit’ and ‘kind unit’ respectively, and OKU is their disjunction, i.e. ‘object or kind unit’. For example,  $OU_i(\text{Ursus})(x) = 3$  in (ib) means that  $x$  consists of three individual bears.

- (i) a.  $\llbracket \text{three bears} \rrbracket = \lambda i \lambda x [RT_i(x, \text{Ursus}) \wedge OKU_i(\text{Ursus})(x) = 3]$   
b.  $\llbracket \text{san zhi xiong} \rrbracket = \lambda i \lambda x [RT_i(x, \text{Ursus}) \wedge OU_i(\text{Ursus})(x) = 3]$

which are [object] individuals, and are normally not satisfied by abstract entities including (sub-)kinds, which are [kind] individuals. The sole exception is ‘inanimate’. The noun specification information of classifiers usually consists of more than one characteristic. For instance, the classifier *buah* in (16) combines with ‘inanimate’ entities that are ‘three dimensional’ and ‘big’ such as buildings and vehicles (Khazriyati and Winskel 2009). Hence, classifiers should not be compatible with properties of [kind] individuals, as abstract entities in the [kind] subdomain have neither shape nor size, though they are ‘inanimate’.

This is not the end of the story, however. The account above predicts that no abstract nouns take classifiers, as they also generally lack characteristics specified by classifiers such as shape and size, due to their very abstractness. This prediction has some truth. There are few classifiers dedicated exclusively to abstract nouns, and abstract nouns are often modified by numerals directly (cf. chapter 2, section 2.3.2 (ii)). Yet, the prediction is still incorrect. Some classifiers are also used with abstract nouns. The Malay examples in (17) illustrate this point.

- (17) se-buah idea/analisis/cita-cita/kejayaan  
 one-CLF idea/analysis/dream/success  
 ‘a(n) idea/analysis/dream/success’

Notice that the same classifier as in (16), i.e. *buah*, is used here. The expressions are well-formed even though ideas, analyses, dreams and successes are not three dimensional or big. I suggest that some classifiers can combine with abstract nouns as a secondary effect of having ‘inanimate’ as part of their noun specification information. The other pieces of noun specification information, if any, tend to be rather non-specific and cover a wide range of entities. This enables speakers to ignore these pieces when interpreting relevant expressions. Ignoring incompatible specifications is only possible for abstract nouns, but not for concrete nouns like *majalah* ‘magazine’. This is arguably because abstract nouns always denote in the [kind] subdomain and hence do not have an object reading, in which all pieces of noun specification information of classifiers are satisfied. That is to say, a subkind reading is difficult to obtain for an expression with a classifier because a more faithful interpretation is available for the same expression, i.e. an object reading. Competition between the two interpretations does not occur for abstract nouns.

This analysis in turn predicts that a subkind reading is available for expressions with a classifier if the noun specification information of a classifier consists only of ‘inanimate’. This is

because in such a case the subkind reading is as faithful as the object reading, both satisfying the only required specification ‘inanimate’. The general/default classifier *tu* in Japanese is such a classifier. According to Zubin and Shimojo (1993), *tu* “seems not to have any core semantic representation beyond a restriction to inanimates,” unlike another default classifier *ko*, which specifies that nouns combining with it are three dimensional hand-sized manipulable solid objects. As noted at the end of section 2.4.2 in chapter 2, expressions with *tu* can refer to subkinds. The contrast between (18a) and (18b) below shows that the expression with *tu* can refer to subkinds much more easily compared to that with the specific classifier for machines *dai*.

- (18) a. Josei-ga yoku kau kuruma-ga yot-tu aru.  
 woman-NOM often buy car-NOM four-CLF be  
 ‘There are four (types of) cars often bought by women.’
- b. ??Josei-ga yoku kau kuruma-ga yon dai aru.  
 woman-NOM often buy car-NOM four CLF be  
 For: ‘There are four (types of) cars often bought by women.’

The relevance of ‘inanimate’ is supported by the fact that a subkind reading is much more difficult to obtain for nouns denoting animate entities, even with the classifier *tu*.

- (19) \*?Josei-ga yoku kau inu-ga yot-tu aru.  
 woman-NOM often buy dog-NOM four-CLF be  
 For: ‘There are four (types of) dogs often bought by women.’

An analysis like Krifka’s (1995) introduced at the beginning of this section cannot account for the contrast between (18a) and (19). Given (18a), such an analysis would analyze *tu* as encoding both object and kind units. (19) with *tu* is thus expected to be perfectly acceptable, contrary to the fact.

### 3.3.2 An implementation in Linear Optimality Theory

The idea presented above can be implemented using a constraint-based grammar as it allows one to capture the competition among potential interpretations. The specific model of constraint-based grammar that I adopt here is Linear Optimality Theory (LOT) proposed by Keller (2000, 2006). LOT is designed to account for gradient judgment data. It differs from standard Optimality Theory (OT; Prince and Smolensky 2004) in that each constraint is assigned a numeric weight and the grammaticality of a candidate is determined by the sum of the weights of the

constraints it violates. In standard OT, constraints are (partially) ordered and the grammaticality of a candidate is basically determined by the rank of the constraint that it violates. That is, a candidate is considered optimal and hence well-formed if the first constraint that it violates is ranked lower than the first constraints violated by any other candidates; additional constraint violations it incurs do not affect the grammaticality. For example, in the tableau in (20), Candidate 1 is chosen as optimal (indicated by the hand sign) because its first violation occurs at  $C_3$ , which is ranked lower than those of Candidates 2 and 3, namely  $C_1$  and  $C_2$  respectively. The two additional violations incurred by Candidate 1 at  $C_4$  are not taken into consideration.

(20)  $C_1 \gg C_2 \gg C_3 \gg C_4$

Input	$C_1$	$C_2$	$C_3$	$C_4$
☞ Candidate 1			*	**
Candidate 2		*!		
Candidate 3	*!			

In LOT, a harmony score is calculated for each candidate. The harmony score of the  $i$ -th candidate  $H_i$  is the negative of the sum of the products between the weight of each constraint  $w_j$  and the number of its violations incurred by the relevant candidate  $v_{ij}$ :

$$(21) \quad H_i = - \sum_j w_j v_{ij}$$

Suppose that the weights of the four constraints in (20) are 4, 3, 2 and 1. The harmony score of Candidate 1 is calculated as follows:

$$(22) \quad H_1 = -(4 \times 0 + 3 \times 0 + 2 \times 1 + 1 \times 2) = -(0 + 0 + 2 + 2) = -4$$

The entire LOT tableau based on the candidates and constraints in (20) and the sample constraint weights above is given in (23). Differences in harmony scores corresponds to differences in acceptability. The degree of degradedness of a suboptimal candidate compared to the optimal one is thus expressed by a numeric value. I will add in the tableaux below the standard acceptability symbols used in the literature such as ‘?’, ‘??’ and ‘\*?’ that roughly reflect the actual degradedness values.

(23)

Input	$C_1$	$C_2$	$C_3$	$C_4$	$H$
? Candidate 1	4	3	2	1	-4
☞ Candidate 2		*			-3
? Candidate 3	*				-4

Notice the two differences between the standard OT and LOT tableaux in (20) and (23). First, in the standard OT tableau, only one candidate is optimal/grammatical and the other two candidates are simply ruled out. In the LOT tableau, on the other hand, the two suboptimal candidates are not ruled out but evaluated as less acceptable. Second, although the violation pattern and the constraint ranking (in terms of domination in standard OT and weight differences in LOT) are identical, different candidates are chosen as the optimal candidate, i.e. Candidate 1 in (20) and Candidate 2 in (23). This kind of difference arises because constraint violations are cumulative only in LOT. A combined effect of violations of lower ranked constraints can be more serious than a single violation of a high ranked constraint (ganging up effects).

Let us now turn to the analysis of the fact that a subkind reading is impossible or very difficult to obtain for expressions with classifiers. The relevant example is repeated below.

- (16b) tiga buah majalah  
 three CLF magazine  
 (i) 'three copies of magazines'  
 (ii) ??'three kinds of magazines'

It is safe to assume that the constraints at work are the noun specifications carried by classifiers. The candidates are the two readings of the '(Num) CL NP' constituent, i.e. object and subkind readings. That is, the Generator, which contains the system articulated in section 3.2, takes a syntactic object, i.e. an LF representation, as the input and generates all potential interpretations for it. In the case of the classifier *buah* in Malay, the three attributes proposed by Khazriyati and Winkler (2009) are good enough to adopt as the constraints at work: INANIMATE, 3D and BIG. In Khazriyati and Winkler's classification system, the dimensionality and size parameters are dependent on the animacy parameter, that is, these parameters are only relevant to inanimate entities but not to animate entities. LOT establishes constraint rankings, i.e. numeric weights,

from gradient acceptability data based on the following equation:

$$(24) \quad (\text{Harmony score of Candidate A}) - (\text{Harmony score of Candidate B}) \\ = (\text{Acceptability of Candidate A}) - (\text{Acceptability of Candidate B})$$

However, no such data is available for the two readings of expressions with *buah*. Hence, I hypothesize working constraint weights for the three constraints above, relying on the non-numeric acceptability judgments obtained from my consultants and the dependency between the constraints. The specific weights that I hypothesize are ‘4’ for INANIMATE and ‘1’ for 3D and BIG, where the weight of INANIMATE is the square of the sum of all its dependent constraints, i.e.  $(1 + 1)^2 = 4$ . It is known that exponential constraint weighting realizes strict domination among constraints (Prince and Smolensky 2004:chapter 10). I analyze the dependent constraints 3D and BIG as strictly dominated by INANIMATE. The tableaux in (25) shows how the two interpretations are evaluated with these weights for the concrete noun *majalah* ‘magazine’. The object reading satisfies all three constraints whilst the subkind reading violates 3D and BIG, which lowers the harmonic score of the latter by 2, hence the degraded acceptability ‘??’.

(25) Concrete nouns

<i>buah majalah</i> [CLF magazine]	INANIMATE	3D	BIG	<i>H</i>
	4	1	1	
☞ (copy of) magazine <sup>o</sup> (object reading)				0
?? (kind of) magazine <sup>k</sup> (subkind reading)		*	*	-2

We have seen that unlike concrete nouns like *majalah* ‘magazine’, the classifier *buah* is perfectly compatible with abstract nouns even though they have neither dimension nor size.

(17) se-buah idea/analisis/cita-cita/kejayaan  
one-CLF idea/analysis/dream/success  
‘a(n) idea/analysis/dream/success’

I argued in the last section that this is because abstract nouns always denote in the [kind] sub-domain, and hence no competition occurs between two potential interpretations. In Optimality Theoretic terms, this means that the Generator uniquely assigns a subkind reading to abstract nouns. The tableau for *buah idea* [CLF idea] contains only one candidate as in (26). The middle row is shown only to highlight the difference between this tableau and the tableau for *buah*

*majalah* [CLF magazine] in (25), which has two candidate interpretations. This second raw is not involved in the actual form-meaning mapping evaluation. Notice that the subkind reading is optimal in (26) but not in (25) even though their harmony scores are identical.

(26) Abstract nouns

<i>buah idea</i> [CLF idea]	INANIMATE	3D	BIG	
	4	1	1	<i>H</i>
(object reading not generated)	—	—	—	—
☞ <i>idea<sup>k</sup></i> (subkind reading)		*	*	-2

I also argued that the general/default classifier *tu* in Japanese allows a subkind reading as well as an object reading as in (18a) because its noun specification consists only of ‘inanimate’.

- (18) a. Josei-ga yoku kau kuruma-ga yot-tu aru.  
 woman-NOM often buy car-NOM four-CLF be  
 ‘There are four (types of) cars often bought by women.’
- b. ??Josei-ga yoku kau kuruma-ga yon dai aru.  
 woman-NOM often buy car-NOM four CLF be  
 For: ‘There are four (types of) cars often bought by women.’

The tableaux for these two sentences are given in (27) and (28) below respectively. The four constraints for *dai* are adopted from Matsumoto (1993). The weights of each constraint are estimated based on the results of the acceptability judgment experiment conducted by Matsumoto (Table 3.1), wherein the subjects were given a sheet with sentences containing the classifier *dai* and different nouns, and judged the acceptability of those sentences on a seven-point rating scale. As before, the weight of INANIMATE is the square of the sum of all its dependent constraints. The weight thus calculated is also applied to (27) to enable inter-tableau comparison.

(27) *tu* (general classifier)

<i>kuruma-ga yot-tu</i> [car-NOM four-CLF]	INANIMATE	
	34.81	<i>H</i>
☞ four cars <sup>o</sup> (object reading)		0
☞ four (types of) cars <sup>k</sup> (subkind reading)		0

(28) *dai* (specific classifier for machines)

<i>kuruma-ga yon dai</i> [car-NOM four CLF]	INANIMATE	MECHANICAL	PLACED ON THE GROUND	DETACHED	CARRYING THINGS	<i>H</i>
四 four cars <sup>o</sup> (object reading)						0
?? four (types of) cars <sup>k</sup> (subkind reading)		*	*	*	*	-5.9

Table 3.1: The results of Matsumoto’s (1993) acceptability judgment experiment.

	MECHANICAL	PLACED ON THE GROUND	DETACHED	CARRYING THINGS	acceptability value
car	+	+	+	+	6.0
TV set	+	+	-	+	5.9
rickshaw	-?	+	+	+	5.5
escalator	+	+	+	-	5.2
radar	+	+	-	-	4.4
cable gondola	+?	-	+	+	4.1
bookcase	-	+	-	+	1.8
balloon	-	-	+	+	1.3
balloon (toy)	-	-	-	+	0.0

While INANIMATE is the only constraint for *tu*, the specific classifier for machines *dai* has four other constraints, which are dependent on INANIMATE. It is these dependent constraints that determines the availability of a subkind reading.

The two tableaux above suggest that ‘four cars’ in the object reading is equally acceptable for both *tu* and *dai*, as they have the same harmony score, i.e. 0. However, this is not actually the case. The expression with the specific classifier *dai* is much preferred compared to that with the general/default classifier *tu*. I agree with Matsumoto (1993) that this difference in acceptability is due to a general pragmatic principle, specifically the maxim of quantity, which states, “make your contribution as informative as is required for the current purpose of the exchange” (Grice

1975). When two classifiers are equally felicitous, the more informative one is preferred over the less informative one.

Finally, remember that a subkind reading is much more difficult for nouns denoting animate entities even with *tu* as in (19).

- (19) \*?Josei-ga yoku kau inu-ga yot-tu aru.  
 woman-NOM often buy dog-NOM four-CLF be  
 For: ‘There are four (types of) dogs often bought by women.’

This is so because unlike nouns denoting inanimate entities, the subkind reading is less faithful than the competing object reading, as illustrated by the tableau below:

- (29) *tu* (general classifier) and an animate noun

<i>inu-ga yot-tu</i> [dog-NOM four-CLF]	INANIMATE 34.81	<i>H</i>
☞ four dogs <sup>o</sup> (object reading)		0
*? four (types of) dogs <sup>k</sup> (subkind reading)	*	−34.81

The contrast between animate and inanimate nouns with respect to the availability of a subkind reading clearly shows that the *tu* morpheme used in subkind-reading nominals has the same meaning as the *tu* morpheme used in object-reading nominals. *Tu* being compatible with a subkind reading does not mean that the morpheme is polysemous, nor are there two homophonous *tu* morphemes.

### 3.4 Fact 2: Plurals do not denote kinds

#### 3.4.1 Proposal

To recapitulate, the second fact to be explained is that plurals in classifier languages do not denote kinds, unlike languages such as English and Italian, in which morphologically plural forms denote kinds. The relevant examples are repeated below.

- (3) Malay
- a. Telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
 telephone PASS-invent by Alexander Graham Bell at year 1876  
 ‘The telephone was invented by Alexander Graham Bell in 1876.’
- b. #Telefon-telefon di-cipta oleh Alexander Graham Bell pada tahun 1876.  
 telephone.PL PASS-invent by Alexander Graham Bell at year 1876
- (30) Italian (Chierchia 1998b:342)

I cani sono rari.  
 the dogs are rare  
 ‘Dogs are rare.’

Recall that there are two types of kinds, i.e. kind emulations and true kinds. The former is a plurality ( $\{a_1 \sqcup a_2 \sqcup a_3\}_w$  in Figure 3.7) while the latter is a singularity ( $A$  in Figure 3.7). Given the plural morphology in the examples above, the type of kind relevant here is kind emulations. The question can thus be stated more precisely as follows: why do plurals in classifier languages not denote kind emulations? In what follows, I will refer to kind emulations simply as kinds.

I argue that plurals do not denote kinds because  $\sqcap$ , which derives kinds from properties of objects, is not defined for pluralities. The definition of  $\sqcap$  is repeated below.

$$(14) \quad \sqcap P = \lambda w[\iota x P_w(x)]$$

(Chierchia 2010:115)

According to Chierchia (1998b),  $\sqcap$  is not defined for singular properties. If  $\sqcap$  is applied to a singular property,  $\iota P_w$  will necessarily be a singular individual. Such an entity is too peculiar to qualify as a kind, because it has a single instance in every world.

To see why  $\sqcap$  is not defined for plural properties either, it is crucial not to confuse pluralities and number-neutral properties, which subsume not only pluralities but also singularities. Suppose, for example, that our domain of discourse contains just three dogs, namely Fido ( $f$ ), Barky ( $b$ ) and Spotty ( $s$ ), and their sums. The number-neutral property true of the entire domain is as in (31a) while the plural property true only of the sums is as in (31b). Figure 3.8 shows the difference between these two properties visually.

$$(31) \quad \text{a.} \quad \llbracket \text{dogs} \rrbracket = \{f, b, s, f \sqcup b, b \sqcup s, f \sqcup s, f \sqcup b \sqcup s\}$$

$$\text{b. } \llbracket \text{dogs (PL)} \rrbracket = \{f \sqcup b, b \sqcup s, f \sqcup s, f \sqcup b \sqcup s\}$$

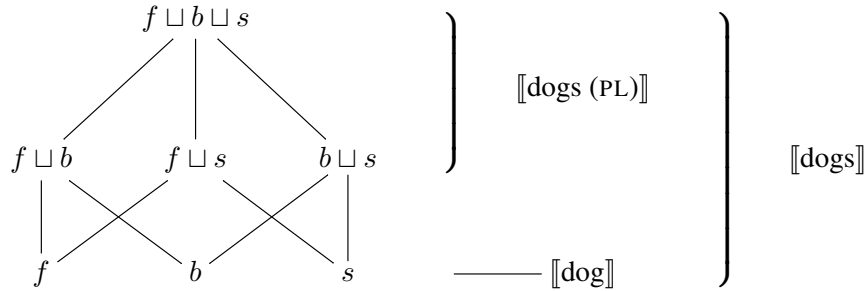


Figure 3.8: Number-neutral and plural properties.

Conceptually, our intuition about kinds is independent of the number of manifestations. Fido belongs to the dog-kind not only when he is accompanied by Barky but also when he is alone. Formally, it does not matter whether  $\cap$  is defined for pluralities. This is because the result of applying  $\cap$  to the property of pluralities (31b) is identical to that obtained by applying  $\cap$  to the number-neutral property (31a). Both result in  $f \sqcup b \sqcup s$ . By taking both perspectives into consideration, there is no positive reason to believe that  $\cap$  is defined for pluralities as long as it is defined for number-neutral properties.

In fact, the formal identity between the results of applying  $\cap$  to plural and number-neutral properties can suggest that  $\cap$  is not defined for plural properties. This becomes clear when one considers the morphological realizations of the two properties. Supposing for the moment that  $\cap$  is defined for plural properties, if one applies the inverse of  $\cap$ , i.e.  $\cup$  (defined as in (32)), to the kind obtained via  $\cap$  to recover its instantiations, only the number-neutral property (31a) is obtained, as shown in (33b) and (33c). That is, (31a) and (31b) become indistinguishable.

$$(32) \quad \text{For any kind } k, \cup k = \lambda w \lambda y [y \leq k_w]$$

(Chierchia 2010:115)

$$(33) \quad \begin{array}{l} \text{a. } \llbracket \cap \text{dogs} \rrbracket = \llbracket \cap \text{dogs (PL)} \rrbracket = \{f \sqcup b \sqcup s\} \\ \text{b. } \llbracket \cup \cap \text{dogs} \rrbracket = \{f, b, s, f \sqcup b, b \sqcup s, f \sqcup s, f \sqcup b \sqcup s\} = \llbracket \text{dogs} \rrbracket \\ \text{c. } \llbracket \cup \cap \text{dogs (PL)} \rrbracket = \{f, b, s, f \sqcup b, b \sqcup s, f \sqcup s, f \sqcup b \sqcup s\} = \llbracket \text{dogs} \rrbracket \neq \llbracket \text{dogs (PL)} \rrbracket \end{array}$$

Given the reasonable assumption that  $\cup$  does not affect the form of the NP (excluding the determiner), this coincidence does not cause any problem if the same NP form expresses the two properties. Indeed, the coincidence was considered desirable by Chierchia (1998b), as plural properties happened to be associated with the same NP form as number-neutral properties in languages such as English and Italian. However, the consequence causes a serious problem in languages that distinguish the two properties morphologically, because the recovered denotation of a plural form will incorrectly contain singularities. Classifier languages are such languages.

The so-called “plural” nouns in English with the suffix *-s* can entail the transnumeral, general number, with the plural meaning obtained pragmatically (e.g. McCawley 1968; Krifka 1989; Sauerland 2003; Rullmann and You 2006; Zweig 2009; Farkas and de Swart 2010). A question like *Do you have cats?* can be answered by *Yes, one* but not by *No, only one*. The morphologically “plural” form *children* is felicitous in the question sentence because it is associated with the general number. This explains why bare plurals in English can denote kinds. The so-called bare “plurals” are actually bare “generals.” Bare plurals denoting kinds are neither singular nor plural; hence,  $\cap$  is defined.

Unlike bare plurals in English, plurals in classifier languages are genuine plurals, associated exclusively with the plural number. To ask whether someone has a child, bare NPs are used, but not NPs with plural morphology, as illustrated in (34) and (35). While bare NPs, which are associated with the general number, are uncommitted to the number of children, NPs with plural morphology entail the existence of at least two children.<sup>8</sup>

(34) Malay

Saya tidak tahu sama ada mereka ada anak/#anak-anak.  
 I not know whether they have child/child-PL  
 ‘I do not know whether they have children.’

---

<sup>8</sup>Hosoi (2005) claims that forms with *-tati* in Japanese are compatible neither with the predicate *iru* ‘to have’ as in (35) nor with kind-predicates because these predicates are individual-level predicates but the semantics of *-tati* he proposes requires a stage-level predicate. This analysis does not work as forms with *-tati* are compatible with other types of individual-level predicates, as shown in (i).

(i) Kodomo-tati-wa kasikoi/ mozi-ga yomeru.  
 child-PL-TOP intelligent letter-NOM can.read  
 ‘(The) Children are intelligent/can read letters.’

(35) Japanese

Watasi-wa karera-ni kodomo/#kodomo-tati-ga iru-ka sira-nai.  
 I-TOP they-DAT child/child-PL-NOM be-whether know-not  
 ‘I do not know whether they have children.’

NP forms associated with plural and number-neutral properties are thus morphologically distinct in classifier languages. Therefore, it can be concluded that  $\cap$  is not defined for properties of pluralities; it is only defined for number-neutral properties.<sup>9</sup> In conclusion, plurals in classifier languages do not denote kinds because unlike plurals in English, they are genuine plurals, for which  $\cap$  is not defined.

Before leaving this section, I would like to discuss the other readings of plurals. As seen in chapter 2, plurals have object and subkind readings. Consider the Malay sentences and their translations in English in (36).

- (36) a. Dinosaur-dinosaur masih tidur.  
 dinosaur.PL still sleep  
 ‘The dinosaurs are still sleeping.’ (object reading)
- b. Ada di antara dinosaur-dinosaur [yang pupus pada masa itu].  
 be at among dinosaur.PL REL extinct at time that  
 ‘Among the dinosaurs there were also some (subspecies) which went extinct at that time.’ (subkind reading)

This fact is nothing surprising at all. The two interpretations result from plurals denoting pluralities in the [object] and [kind] subdomains respectively. Plurals having a subkind reading without a kind reading was surprising earlier, because it was implicitly assumed that kind and subkind readings were tightly connected. However, the two readings can be obtained independently in the model presented in this chapter. Nominals can denote kinds (‘kind emulations’ to be precise) when they denote number-neutral properties in the [object] subdomain. Subkinds are pluralities in the [kind] subdomain. Since the two subdomains exist on an equal footing, nominals can denote in the [kind] subdomain, i.e. have a subkind reading, whatever denotations they may have in the [object] subdomain. English happens to employ the same morphology, i.e. -s, to denote number-neutral and plural properties. “Plural” forms in English can thus have both

<sup>9</sup>In this connection, Bale and Khanjian (2009) propose the generalization that only nouns whose denotations are number-neutral obligatorily scope under negation and appear as the direct complement of measure words, and they point out a possible link between number-neutrality and kinds.

kind and subkind readings. However, other languages employ two distinct forms to differentiate these meanings, in which case one form having a subkind reading does not necessarily mean that the same form should have a kind reading. Classifier languages belong to the latter type.

### 3.4.2 Alternative accounts

In fact, some researchers are reluctant to regard the morphemes that I have seen as plural markers as genuine plural morphology because these morphemes can convey meanings that are not conveyed by more established plural markers such as *-s* in English. These researchers contend that the markers at issue indicate meanings other than plurality from which the meaning of plurality results as side effects. For instance, Mintz (2002:282) states that reduplication in Malay/Indonesian “indicate[s] individuality or variety and not plurality.” Nakanishi and Tomioka (2004); Nakanishi and Ritter (2008) claim that the suffix *-tati* in Japanese is an associative marker, of which plurality is a special case. In what follows, I will examine these two cases and show that while such views could offer alternative accounts for inability of plurals to denote kinds, they are untenable for empirical reasons. In both cases, there is no denying that the relevant markers encode plurality as their primary meaning.

#### 3.4.2.1 Individuality: Reduplication in Malay/Indonesian

According to Mintz (2002), (full) reduplication in Malay/Indonesian indicates individuality within a group. Reduplicated forms differ from the corresponding non-reduplicated forms in that multiple individuals are construed as “a number of individual units” as opposed to “a single unit.” These two meanings are not distinguished morphologically in English. The reduplicated and non-reduplicated forms meaning ‘stories’ in (37) illustrate the relevant contrast. “[T]he stories [*cerita-cerita*—HN] depicted by the wayang kulit are to be taken as separate and individual” whereas their origin, “the stories [*cerita*—HN] of Hindu heroes such as the Ramayana and the Mahabarata[,] are to be taken as a unit” (Mintz 2002:283).

(37) Malay/Indonesian (Mintz 2002:283)

Wayang kulit memainkan *cerita-cerita* dari *cerita* pahlawan Hindu seperti Ramayana  
 wayang kulit portray story.PL from story hero Hindu like Ramayana  
 dan Mahabarata.  
 and Mahabarata

‘The wayang kulit portrays (*various different*) *stories* from the *stories* of Hindu heroes  
 such as the Ramayana and the Mahabarata.’

Since individuality in this sense presumes plurality, Mintz concludes that reduplication does not encode plurality.<sup>10</sup>

While my analysis is compatible with Mintz’s analysis of reduplication, an alternative analysis is also conceivable. Specifically, the non-unitary aspect of individuality prevents reduplicated forms from referring to kinds, as kinds are normally considered as unitary entities. This line of analysis, however, is not promising. Sneddon et al. (2010:21) point out that there are sentences in which individuality or variety is of no importance. In (38), they say, “what is stressed is that there is more than one.”

(38) Indonesian (Sneddon et al. 2010:21)

Rumah-nya dekat pohon-pohon mangga itu.  
 house-3SG near tree.PL mango that

‘His house is near those mango trees.’

It is thus more accurate to describe that the semantic function of reduplication is to encode plurality and reduplicated forms often but not always convey individuality or variety as well. The second meaning is thought to arise from pragmatic inference. This inference is made possible by the contrast between reduplicated and non-reduplicated bare NPs. While both forms can be used to refer to a set of multiple individuals, the former differs from the latter in that the set consists only of pluralities, to the exclusion of singularities. The choice of the former over the latter thus highlights that every member of the set consists of multiple atoms. In addition, iconicity plays the role of reinforcing this effect.

<sup>10</sup>Alwi et al. (1998:239) and Liaw (1999:367) also recognize the meaning of individuality (their terms are ‘*keanekaan*’ [variety] and ‘diversity’ respectively) in reduplicated forms. However, for them, this meaning is an additional meaning on top of plurality. Also, they do not distinguish fully reduplicated forms from rhythmically reduplicated forms such as *sayur-mayur* ‘all sorts of vegetables’ [base: *sayur* ‘vegetable’].

### 3.4.2.2 Associativity: *-Tati* in Japanese

Forms with *-tati* and other “plural” markers in Japanese can include individuals that are associated with the host noun but are not in its extension. The referent of the host noun and the individuals associated with it are referred to as the ‘focal referent’ and its ‘associates’ (Daniel and Moravcsik 2011). *Gakusei-tati/ra/domo* [student-TATI/RA/DOMO] can thus mean either ‘(the) students’ (uniform reading) or ‘the student(s) and their non-student associates’ (non-uniform reading). Plurals with *-s* in English only have a uniform reading and lack a non-uniform reading. Uniform and non-uniform readings are often referred to as additive and associative readings. However, I would like to reserve the terms ‘additive’ and ‘associative’ to distinguish between two kinds of pluralities, which are defined as follows:

- (39) Two kinds of pluralities
- a. Additive plurality: a plurality consisting only of Xs.
  - b. Associative plurality: a plurality consisting of X(s) and X’s associates.

Morphemes that accompany nominals referring to additive pluralities are (true) plural markers whereas those that accompany nominals referring to associative pluralities are associative markers. Notice that no restriction is stipulated regarding X’s associates in (39b). This means that the associates of a noun may or may not be also X, though the associates are necessarily disjunct from X when the host noun is a pronoun or a proper name as in (40). The two kinds of plurality and their possible interpretations are summarized in Table 3.2.

- (40) a. *watasi-tati*  
I-ASSOC  
‘we (= speaker and his/her associates)’
- b. *Tanaka-tati*<sup>11</sup>  
Tanaka-ASSOC  
‘Tanaka and others’

Traditionally, the term ‘associative plural’ has been used to refer to forms denoting non-uniform associative pluralities, assuming that X’s associates are not X (e.g. Daniel and Moravcsik 2011). I do not adopt such a convention here because it is important not to conflate two distinct notions,

---

<sup>11</sup>The same string can also mean ‘more than one person named Tanaka’. In this case, *Tanaka* is used as a common noun meaning ‘bear the name Tanaka’.

Table 3.2: Two kinds of pluralities and their interpretations.

Referent	Reading	Marker	Traditional label
ADDITIVE PLURALITY	uniform	plural	additive plural
ASSOCIATIVE PLURALITY			
(i) homogeneous (associates = X)	uniform	associative	“additive plural”
(ii) heterogeneous (associates $\neq$ X)	non-uniform	associative	associative plural

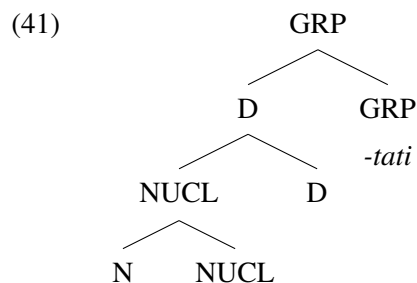
i.e. associativity and referential homogeneity/heterogeneity. An additive plurality only results in a uniform reading. By contrast, an associative plurality can result in either a uniform or non-uniform reading, depending on whether the associates are X or not. A uniform reading arises when the associates are X whereas a non-uniform reading arises when the associates are not X. Note that the reading traditionally associated with the label “additive plural” can be marked not only by plural markers but also by associative markers, as indicated by the quotation marks in Table 3.2.

Since a non-uniform reading only arises from an associative plurality, there is no denying that *-tati* and other “plural” markers in Japanese encode associativity. The question is whether they are also (true) plural markers.<sup>12</sup> If they are, the uniform reading has two sources, i.e. additive and associative pluralities. If they are not, the uniform reading has only one source, i.e. associative pluralities. In this case, inability of *-tati* nominals to denote kinds can be ascribed to potential referential heterogeneity rather than plurality, as kinds are normally considered to have homogeneous instantiations. That is to say, the form *X-tati* cannot refer to the kind X because it may contain some non-Xs in its extension.

The principle of parsimony leads one to choose the second option, provided that the empirical coverage of the two options is identical. Indeed, Nakanishi and Tomioka (2004) and Nakanishi and Ritter (2008) propose analyses of *-tati* whereby *-tati* is essentially an associative

<sup>12</sup>A similar question has been raised regarding the suffix *-men* in Mandarin, as it is obviously an associative marker (often referred to as “collective marker” in the literature on Mandarin) when it attaches to pronouns and proper names, like *-tati* in Japanese (e.g. Iljic 1994; Li 1999; Lan 2010). In fact, the question dissolves itself once the independence of associativity from referential heterogeneity is duly recognized. This is because unlike forms with *-tati*, forms with *-men* only have a uniform reading. For example, *xuesheng-men* [student-MEN] cannot be used to refer to a group containing some non-students, no matter how (the speaker thinks) these non-student individuals are related to the students. This fact indicates that *-men* attaching to common nouns is a plural marker, as Li (1999) claims. I will discuss certain peculiar properties of *-men* that cannot be accounted for by simply identifying it as either a plural marker or an associative marker in chapter 4.

marker and the uniform reading follows from it. Specifically, in Nakanishi and Tomioka (2004) the uniform reading of *otokonoko-tati* [boy-TATI] ‘boys’ is a special case of its meaning ‘a potentially non-uniform plurality that is represented by boys’, namely a case in which the plurality at issue happens to be uniform. They argue that *-tati* forms do not denote kinds and are not easily used in generic sentences because the purported generalizations are hardly precise generalizations if they allow for exceptions. Nakanishi and Ritter (2008) dissociate *-tati* from number completely. In their analysis, *-tati* heads an independent functional head GR(ou)P projected above D, NU(mber)CL(assifier) and N as shown in (41), with *-tati* given the semantics in (42).



- (42)  $\llbracket -tati \rrbracket^c = \lambda x: x \text{ is human.group}(x)(c)$   
 “*-tati* combines with a human individual  $x$  and yields a group that consists of  $x$  and his/her associate(s) with respect to the contextual parameter  $c$ ”<sup>13</sup>

In this analysis, a uniform reading is obtained if the noun and its associates happen to share the same descriptive content.

According to these analyses, it is predicted that whenever a uniform reading is available, a non-uniform reading should be also available. However, this prediction is not borne out. In many cases, a non-uniform reading is impossible or extremely difficult to obtain even if a uniform reading is available. For instance, *-tati* forms with numerals as in (43) only have a uniform reading. The relevant plurality cannot contain any non-students and its cardinality cannot exceed the number specified by the numeral.

<sup>13</sup>The term “group” in this definition is not defined explicitly in Nakanishi and Ritter (2008). I take it to stand for a context-sensitive type of plurality (distinct from  $x$  in case  $x$  is a plurality), where the elements constituting the plurality stand in some relation determined by the context. It cannot be identified with the ‘group’ in the sense of Link (1984) and Landman (1989) introduced in section 3.2, i.e. an atomic individual formed from a sum of pluralities, which triggers a collective reading. Sentences with associative subjects with *-tati* allow both collective and distributive readings. In chapter 4, I will argue that the restriction to human is not asserted or presupposed but conventionally implicated.

- (43) *san nin-no gakusei-tati*  
 three CLF-LINK student-TATI
- (i) ‘three students’ (uniform reading; no more than three people; no non-students allowed)
  - (ii) \*‘three people including at least one non-student represented by students’ (non-uniform reading based on Nakanishi and Tomioka (2004); non-students allowed)
  - (iii) \*‘a group consisting of (the) three students and their non-student associates’ (non-uniform reading based on Nakanishi and Ritter (2008); more than three people)

The empirical coverage of the two possibilities is thus not identical. Cases like (43) can be accounted for only by analyses that recognize *-tati* as a (true) plural marker besides being an associative marker.

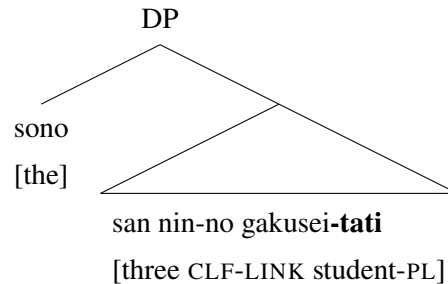
The reason why a non-uniform reading is not available in (43) is because *san nin-no gakusei* ‘three students’ is interpreted as a quantity expression rather than a referential expression. A non-uniform reading presumes a focal referent, which must be referential. A non-uniform reading becomes available with the addition of a demonstrative as in (44), though the uniform reading is still predominant. Moreover, the cardinality of the group can exceed three, as Nakanishi and Ritter’s analysis predicts.

- (44) *sono san nin-no gakusei-tati*  
 the three CLF-LINK student-TATI
- (i) ‘the three students’ (uniform reading; no more than three people)
  - (ii) ‘a group consisting of the three students and their student associates’ (uniform reading; more than three people)
  - (iii) ‘a group consisting of the three students and their non-student associates’ (non-uniform reading; more than three people)

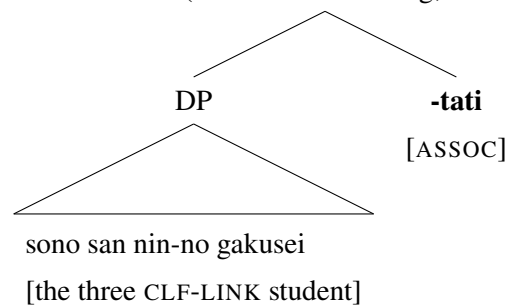
The contrast between (43) and (44) suggests that Nakanishi and Ritter’s (2008) analysis is basically on the right track, but only insofar as the nominal form to which *-tati* attaches can be interpreted as referential. An extra analysis is necessary to account for the whole range of uniform pluralities, including (43) and the first interpretation of (44). Empirical facts justify

positing two distinct *-tati* morphemes.<sup>14</sup> One *-tati* morpheme is a (true) plural marker and occurs below D. It is comparable to *-s* in English or reduplication in Malay, and is the primary source of the uniform reading. The other *-tati* morpheme is concerned with associativity and occurs above D, as proposed by Nakanishi and Ritter (2008). A uniform reading may be obtained as a special case, but only insofar as the nominal form to which it attaches can be interpreted as referential. The string in (44) thus have two possible syntactic parsings associated with different interpretations, as shown in (45).

- (45) a. The plural marker *-tati*  
 ‘the three students’ (uniform reading; no more than three people)



- b. The associative marker *-tati*
- (i) ‘a group consisting of the three students and their student associates’ (uniform reading; more than three people)
  - (ii) ‘a group consisting of the three students and their non-student associates’ (non-uniform reading; more than three people)



Notice that although both *-tati* morphemes give rise to a uniform reading, (45a) and (45b) differ in cardinality. The availability of two uniform readings with different cardinalities also supports

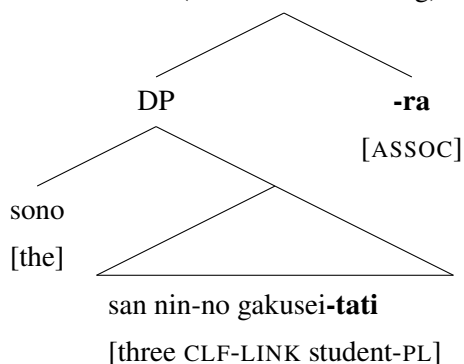
<sup>14</sup>Görgülü (2011) reaches a similar conclusion regarding *-lar* in Turkish, which involves uniform and non-uniform readings like *-tati* in Japanese.

positing the plural *-tati* in addition to the associative *-tati*.

The analysis above predicts that two “plural” markers can occur simultaneously, realized in different syntactic positions. This prediction is actually borne out. (46a) below is well-formed and is thought to have the structure shown in (46b). Two different markers, i.e. *-tati* and *-ra*, are employed in this example because repeating the same marker as in *sono san nin-no gakusei-tati-tati* and *sono san nin-no gakusei-ra-ra* sounds awkward, though not ill-formed. When two markers are used together, the preferred order between them is *-tati-ra* for common nouns and proper names as in (46a), and *-ra-tati* for pronouns, as in *kare-ra-tati* [he-PL-ASSOC] ‘they’.

- (46) a.    *sono san nin-no gakusei-tati-ra*  
           the    three CLF-LINK student-PL-ASSOC  
           (i)    ‘a group consisting of the three students and their student associates’  
                   (uniform reading)  
           (ii)    ‘a group consisting of the three students and their non-student associates’ (non-uniform reading)

b.



Additional support for positing two distinct *-tati* morphemes comes from languages in which plurality and associativity are marked by different morphemes. For example, Amharic marks these two meanings with *-och* and *innä-* respectively, as in (47).

- (47) Amharic (Gideon Goldberg (p.c.) cited in Dryer and Haspelmath 2011)
- a.    *Maryam-och*  
       Mary-PL  
       ‘more than one Mary’

- b.     innä-Maryam  
       ASSOC-Mary  
       ‘Mary with her group/followers/friends’

Unlike Amharic, the markers for the two distinct categories of plurality and associativity happen to be phonologically identical in Japanese.

To conclude, *-tati* exists as a true plural marker, which exclusively expresses additive pluralities. The same form also expresses associativity. An associative *-tati* form is capable of denoting a uniform plurality. However, this is only possible if the form to which *-tati* attaches can be interpreted as referential. Furthermore, the uniform plurality denoted by an associative *-tati* form is not identical to that denoted by the corresponding plural *-tati* form, for the former is always greater in cardinality than the latter. Therefore, it is more plausible to link inability of *-tati* forms to denote kinds to plurality, as I propose, than to link it to associativity.

### 3.5 Summary

This chapter has proposed principled accounts for the two little-discussed facts about classifier languages pointed out in chapter 2: (i) classifiers prevent a subkind reading and (ii) bare plurals do not denote kinds. The chapter has also proposed a model of the domain of individuals and the relationship between ordinary objects, kinds and subkinds. The proposed organization of the domain of individuals with two parallel subdomains of [object] and [kind] plays an important role in explaining both facts. A subkind reading is difficult to obtain for expressions with classifiers because classifiers are concerned primarily with properties of [object] individuals, but not with those of [kind] individuals. Consequently, a subkind reading does not satisfy the conditions specified by classifiers as faithfully as its object reading competitor. Crucially, my account does not rely on the problematic “classifiers for counting” thesis, which I invalidated in chapter 2. I argued that plurals in classifier languages do not denote kinds because they are pure plural forms, which do not include singularities in their denotations, and the [object]-to-[kind] sort/type-shifter  $\cap$  is not defined for them. In the next chapter, I will investigate the semantics of classifiers and plurals in classifier languages more closely.

## **Chapter 4**

# **The semantics of classifiers and numerals**

### **4.1 Introduction**

Having explained the two little-discussed facts about classifier languages, I will now discuss the semantics of classifiers in this chapter. As seen in the last chapter, classifiers specify the characteristics of the nouns with which they combine. I claim that this classification function is not the core meaning of classifiers (section 4.2.1). I argue that the core meaning has to do with number, more specifically the singular, and propose a semantics of classifiers that incorporates both the core meaning and the classification function (section 4.2.2). The proposed classifier semantics is then compared with the “unit of counting” approach to classifiers, one of the most popular views in previous studies (section 4.2.3). Although classifiers are not for counting (i.e. direct numeral modification), it is true that the context in which classifiers are most often used is in numeral modification constructions. I thus discuss the semantics of numerals that is compatible with the proposed semantics of classifier in section 4.3. I adopt Bale et al.’s (2011) analysis, wherein two kinds of numerals are available in natural language, and languages can choose one or both of these options. The choice determines possible and impossible combinations of numerals and different nominal forms.

## 4.2 The semantics of classifiers

### 4.2.1 Classification function

The most salient aspect of the meaning of classifiers is the classification function. Classifiers specify the characteristics of nouns with which they combine in terms of animacy, shape and size, etc. For instance, the classifier *buah* in Malay combines with ‘inanimate’ entities that are ‘three dimensional’ and ‘big’ such as buildings and vehicles (Khazriyati and Winskel 2009). The classifier *dai* in Japanese combines with ‘inanimate’ entities that have the following attributes: ‘mechanical’, ‘placed on the ground’, ‘detached’ and ‘carrying things’ (Matsumoto 1993). Indeed, the primary concern of many semantic studies of classifiers has been to describe the noun specification information of individual classifiers and explore if any meaningful generalizations are possible as to how nouns are grouped by classifiers in a single language or across languages.

I do not deny that the classification function is part of the meaning of classifiers. However, the analysis of the first fact proposed in chapter 3 indicates that it is not the core meaning that affects the truth conditions of the expression, but is instead something additional and can be violated.<sup>1</sup> We have seen that failure to satisfy some of the characteristics specified by the classifier causes no problem unless an alternative interpretation exists that satisfies those characteristics. This explains the difference between abstract nouns and the subkind reading of concrete nouns, with respect to the compatibility with the classifier *buah* in Malay, as shown in (1). In both cases, the noun does not satisfy the specifications ‘three dimensional’ and ‘big’. However, failure to satisfy these specifications only affects the acceptability of the subkind reading of a concrete noun but not that of an abstract noun, because competition with an alternative interpretation, i.e. an object reading, only happens for concrete nouns.

- (1) a.    tiga buah majalah  
          three CLF magazine  
          (i)    ‘three copies of magazines’  
          (ii) ??‘three kinds of magazines’
- b.    se-buah idea/analisis/cita-cita/kejayaan  
          one-CLF idea/analysis/dream/success  
          ‘a(n) idea/analysis/dream/success’

---

<sup>1</sup>The classification information being violable (in terms of Optimality Theory) only means that the relevant meanings can be ignored, but not that they can be cancelled or their negations hold true.

Moreover, as Kang (1994) and McCready (2009, 2012) point out, the mismatch of a classifier and a noun as in (2a) does not make the sentence false, but just inappropriate.<sup>2</sup> The classifier *satu* in (2a) is used with bound objects such as books and magazines; the correct classifier for humans is *nin* as in (2b).

- (2) Japanese
- a. #*Otoko-ga san satu haitteki-ta.*  
 man-NOM three CLF enter-PST  
 For: ‘Three men entered.’
- b. *Otoko-ga san nin haitteki-ta.*  
 man-NOM three CLF enter-PST  
 ‘Three men entered.’

I agree with McCready (2009, 2012), who argues that the noun specification information provided by classifiers is neither asserted nor presupposed, but is instead conventionally implicated. Conventional implicatures are a type of meaning originally introduced by Grice (1975). Grice discusses this meaning only briefly, mainly to preempt confusion with another meaning that is also neither asserted nor presupposed, namely conversational implicatures. Unlike the latter, conventional implicatures are not based on general pragmatic principles, but on particular linguistic forms. The noun specification information of classifiers is tied to particular lexical items, i.e. classifiers themselves. Hence, it is not conversationally implicated. It is Potts’s (2005) close investigation into conventional implicatures that brought about a surge of interest in them among semanticists. Potts (2005:11) characterizes conventional implicatures as speaker-oriented comments on an independent semantic core, which he refers to as ‘at-issue entailments’. He divides expressions involving conventional implicatures into two types, i.e. supplements (appositives, parentheticals) and expressives (e.g. epithets, honorifics). The italicized portions of the sentences in (3) exemplify these two types.

- (3) a. Ed’s claim, *which is based on extensive research*, is highly controversial.  
 b. I have to mow the *damn* lawn.

(Potts 2005:7)

Conventional implicatures are known to have the following properties, which distinguish

---

<sup>2</sup>I thank Brian Reese for bringing McCready’s work to my attention.

them from assertions (at-issue entailments) and presuppositions. First, unlike the latter two meanings, conventional implicatures do not affect the truth value of an entire expression. For example, numerals contribute assertive meanings, which affect the truth values of expressions containing them. The expression *three men* is false in referring to a group consisting of four men. The sentence in (4) presupposes that John has a son. If John actually does not have a son, the sentence is normally considered as lacking a truth value or undefined.

(4) John's son must be a nice guy.

The word *damn* in (3b) does not change the truth of the proposition conveyed by the sentence, nor does it make the sentence truth-value-less or undefined when the speaker does not actually view the lawn negatively.

Second, conventionally implicated meanings are scopeless, projecting from negation. The appositive construction in (5a) entails John's being a swimmer. This meaning is not lost under negation. (5b) denies John's being a nice guy, but not his being a swimmer.

- (5) a. John, a swimmer, is a nice guy.  
 b. It is false that John, a swimmer, is a nice guy.

(McCready 2009:201)

Third, conventionally implicated meanings cannot be bound. This property distinguishes conventional implicatures from presuppositions. The former cannot be bound whereas the latter can, as shown by the contrast between (6b) and (7b).

- (6) Conventional implicature  
 a. John, a swimmer, is a nice guy.  
     → John is a swimmer.  
 b. If John is a swimmer, then John, a swimmer, is a nice guy.  
     → John is a swimmer.

(McCready 2009:201–202)

- (7) Presupposition  
 a. John's son must be a nice guy.  
     → John has a son.

- b. If John has a son, then John's son must be a nice guy.  
→ no presupposition

Lastly, conventional implicatures pass unmodified through presupposition plugs such as belief contexts. This property also distinguishes conventional implicatures from presuppositions. The latter projects in a modified form. In (8a), the negative view towards Conner expressed by *that jerk* is attributed to the speaker in both the matrix and embedded clauses. By contrast, the presupposition that Conner has a son in (8b) projects as the matrix subject's belief.

- (8) a. Conventional implicature  
Sue wrongly believes that that jerk Conner got promoted. #I think Conner is a great guy.

(Potts 2005:31)

- b. Presupposition  
Sue wrongly believes that Conner's son got promoted. I think Conner has no son.

McCready (2009, 2012) concludes that the noun specification meanings of classifiers are conventional implicatures because they show the properties discussed above. I will review them using examples in Malay, as these sentences sound more natural than his examples in Japanese.

Regarding the first property, we have already seen that the use of wrong classifiers does not make an expression false. The Malay sentence in (9) contains two expressions referring to humans, i.e. *se-orang saksi* 'a witness' and *tiga ekor lelaki* 'three men'. The canonical classifier for humans is *orang*, as used in the former. However, the classifier *ekor*, which is normally used for non-human animals, is used in the second expression. In neutral contexts, the use of *ekor* for human referents is considered wrong, though it does not make the expression false. The use of *ekor* in (9), however, is felicitous, as it allows the speaker to condemn or despise the three men who robbed a woman by comparing them to non-human animals.

- (9) Malang pada pagi itu sebagaimana di-ceritakan oleh *se-orang saksi*, *tiga*  
unfortunate on morning that as PASS-report by one-CLF witness three  
*ekor lelaki* menerpa ke arah-nya, menghentam kepala pemuda tersebut  
CLF man rush to direction-3SG beat head youngster mentioned

dengan helmet.

with helmet

‘Unfortunately, in that morning, as reported by *a witness*, *three men* rushed towards her and beat the youngster’s head with their helmets.’

(*I hold it true*, Kisah peragut zaman sekarang ‘A story about modern robbers’,

<http://iholdittrue.multiply.com/journal/item/288/>

Kisah-peragut-zaman-sekarang-, accessed 11/25/2012)

Furthermore, the expression remains well-formed without any classifier as in *tiga lelaki*. Importantly, the choice of (non-)use of a particular classifier does not affect the truth conditions of the expression; the referent is identical in all three expressions. Noun specifications such as ‘human’ and ‘non-human animal’ as well as attitudes such as condemnation and despite are both concerned with how the speaker views the referent, and hence are speaker-oriented meanings, which is in accordance with Potts’ characterization of conventional implicatures.

Second, the noun specification information of classifiers is scopeless, projecting from negation. The condemnation meaning arising from the use of the classifier for non-human animals *ekor* for humans discussed above is not lost under negation, as in (10).

- (10) Adalah tidak benar bahawa tiga ekor lelaki itu meragut beg tangan-nya.  
 be not true that three CLF man that robbed bag hand-3SG  
 ‘It is not true that the three men robbed her of her handbag.’

Third, the noun specification information of classifiers cannot be bound. The condemnation meaning persists with the sentence with *ekor* in (11), regardless of the truth of the antecedent.

- (11) Kalau tiga lelaki itu berperangai seperti binatang, maka tiga ekor lelaki itu(-lah  
 if three man that behave like animal then three CLF man that-PART  
 yang) meragut beg tangan-nya.  
 that rob bag hand-3SG  
 ‘If the three men behave like animals, then (it is) the three men (who) robbed her of her handbag.’

Lastly, the noun specification information of classifiers passes unmodified through presupposition plugs such as belief contexts. Sentence (12) makes sense because it is the speaker, but not her mother, who despises the three men.

- (12) Emak saya percaya tiga ekor lelaki itu orang baik.  
 mom my believe three CLF man that person good  
 ‘My mom believes that the three men are good people.’

This sentence contrasts with a sentence in which a presupposition trigger is embedded in a belief context as in (13). In this sentence, Singapore’s having a king is a belief of the speaker’s mother, but neither one of the speaker nor an objective fact.

- (13) Emak saya percaya raja Singapura masih muda.  
 mom my believe king Singapore still young  
 ‘My mom believes that the king of Singapore is still young.’

Potts (2005) states that “CI [conventional implicature—HN] expressions are used to guide the discourse in a particular direction or to help the hearer to better understand why the at-issue content is important at that stage.” Besides expressing the speaker’s attitudes towards the asserted referent, conventional implicatures generated by classifiers facilitate interpretation by narrowing down interpretation possibilities by explicitly conveying how the speaker views the referent. As seen in chapter 2, the grammaticality of the ‘Num NP’ pattern in optional classifier languages as in (14a) shows that atoms are already accessible for counting before classifiers are introduced, at the semantic as well as syntactic levels, in classifier as well as non-classifier languages.

- (14) a. tiga majalah  
 three magazine  
 ‘three magazines’  
 b. tiga buah majalah  
 three CLF magazine  
 ‘three magazines’

However, as Rothstein (2007, 2010) and Chierchia (2010) demonstrate with count nouns such as *fence* and *mountain*, what counts as one atom is context-dependent. To quote Rothstein (2007) (“individuation” in the quotation means creating atoms):

*Fence* in comparison, is homogeneous since the same piece of fencing can be analyzed as one or several fences in the same situation under different criteria of individuation. If my house and yours adjoin each other, and both of us build a fence

between our houses and the street which meet at a certain point, we could call it “a fence” or “two fences”, depending on the context. (Is the town council charging for a permit to build fences, or giving tax deductions to those who build fences?)

A similar example is *tegami* ‘letter’ in Japanese. As the English translations of (15a) and (15b) below show, the word means either ‘a set of sheets of paper sent as a letter’ or ‘individual sheets of paper sent as a letter’. If Ken writes something on three sheets of paper, puts them in an envelope and sends it to Naomi, one can say either Ken sent Naomi a *tegami* or three *tegami*(s). There is no difference in what Ken sends, but different aspects of *tegami* are in focus, which determines the relevant atom or what counts as one *tegami*. This ambiguity concerning the atom of *tegami* is resolved by using different classifiers: *tuu* in (15a) focuses on the functional aspect of *tegami*, i.e. a medium of correspondence, and *mai* focuses on the physical aspect of it, i.e. paper (‘telic’ and ‘formal’ roles respectively, in Pustejovsky’s (1995) terminology).

(15) Japanese

- a. Ken-wa Naomi-ni tegami-o san tuu okut-ta.  
Ken-TOP Naomi-DAT letter-ACC three CLF send-PST  
‘Ken sent three letters.’
- b. Ken-wa Naomi-ni tegami-o san mai okut-ta.<sup>3</sup>  
Ken-TOP Naomi-DAT letter-ACC three CLF send-PST  
(i) ‘Ken sent a three-page letter.’ (collective reading; a single event)  
(ii) ‘Ken sent three single-page letters.’ (distributive reading; multiple events)

What counts as one can be still vague even in the presence of classifiers. This becomes most evident when the general/default classifier *tu* in Japanese is used with nouns referring to things that come as pairs such as *kutu* ‘shoe(s)’ and *hasi* ‘chopstick(s)’. Things such as shoes and chopsticks consist of two separate objects but they only function when both are used together. Consequently, either one such object or two such objects as a pair can be considered ‘one’. The classifier *tu* does not disambiguate these two possibilities as its noun specification information

<sup>3</sup>The phrase *san mai* here does not modify an implicit noun (e.g. *binsen* ‘writing paper’) as shown in (ia), but it modifies *tegami* ‘letter’. (ib) shows that the structure in (ia) is not possible.

- (i) a. Ken-wa Naomi-ni tegami-o [ $\emptyset$ <sub>NP</sub> san mai] okut-ta.  
Ken-TOP Naomi-DAT letter-ACC three CLF send-PST  
b. \*Ken-wa Naomi-ni tegami-o [binsen(-ni) san mai] okut-ta.  
Ken-TOP Naomi-DAT letter-ACC writing.paper(-DAT) three CLF send-PST

consists only of ‘inanimate’. Consider the following conversation between A and B:

(16) [At a school cafeteria.]

A: *Hasi hito-tu* totte kureru?  
chopstick one-CLF take give  
‘Can you take a *hasi* for me?’

B: Ii-yo. Hai.  
good-PART yes  
‘Sure. Here you go.’

A1: Iya, *ip-pon*-de yokat-ta nda-kedo.  
no one-CLF-with fine-PST be-but  
‘Oh, I just needed one (= a chopstick).’

A2: E? Hutuu *ni-hon* ‘sho?  
eh normally two-CLF right  
‘Wait, don’t you normally think you need two (= a pair of chopsticks) to eat?’

Both replies by A make sense, but in different situations. A1 is appropriate when B brought a pair of chopsticks, but in fact, A already had one and only needed another. A2, on the other hand, is appropriate in the situation in which B brought only one chopstick, when A needed a pair of chopsticks. Notice the use of the specific classifier *hon* (occurring in the allomorph *pon* in A1) in A’s replies. Unlike *tu*, *hon* specifies that the nouns with which it combines are inanimate entities that are saliently one-dimensional and non-circular (Matsumoto 1993). It is thus more capable of disambiguating the two possible meanings of one *hasi* than *tu*, though the vagueness concerning what counts as one *hasi* still remains slightly even when accompanied by *hon*.

#### 4.2.2 The core meaning

What is the core meaning of classifiers if the classification function is not their core meaning? I claim that the core meaning of classifiers is to restrict the noun denotation to singularities.<sup>4</sup> This claim amounts to saying that the ‘CL NP’ constituent (in ‘Num CL NP’) denotes singularities.

<sup>4</sup>McCready (2012) analyzes the classifier *zen* as restricting the domain to pairs, based on the fact that *hasi ni zen* [chopstick two CLF], which refers to two pairs of chopsticks as opposed to *hasi ni hon* [chopstick two CLF], which refers to two chopsticks. I consider that the contrast is a matter of what counts as one atom, which is context-dependent. As seen in the last section, the meaning of the noun *hasi* ‘chopsticks’ itself is vague between ‘pairs of chopsticks’ or ‘one or more chopsticks’ (cf. (16)). It is thus not the case that *zen* restricts the domain to pairs. *Zen* restricts the domain to singularities, where one singularity/atom is a pair of chopsticks.

In fact, identifying ‘CL NP’ with singular NPs is not a new idea. The very conclusion has often been drawn from the “classifiers for counting” thesis, which I have denied in chapter 2. It has been also drawn from the assumption that classifiers and plural morphology are syntactically complementary across languages (cf. Cheng and Sybesma 1999; Borer 2005; Watanabe 2010). This assumption is wrong. Classifiers and plural morphology can co-occur in many classifier languages, including Japanese, Korean, Malay, Halkomelem Salish (Wiltscko 2008), Persian (Hamedani 2011) and Yucatec Maya (Butler 2011).

(17) Japanese

- a.     gakusei-tati san nin  
        student-PL three CLF  
        ‘three students’
- b.     san nin-no gakusei-tati  
        three CLF-LINK student-PL  
        ‘three students’

(18) Yucatec Maya (Butler 2011:42)

- ka’a-túul x-ch’úupal-o’ob  
 two-CLF F-girl-PL  
 ‘two girls’

Moreover, I will argue in chapter 6 that the so-called ‘plural classifiers’ in Cantonese and Hmong are realizations of the features associated with classifiers and plural markers as a single word. In this respect, plural classifiers can be seen as co-occurrence of (the features associated with) classifiers and plural markers.

Real evidence that ‘CL NP’ denotes singularities comes from expressions with classifiers that do not involve numerals. In many languages that use classifiers in non-counting contexts, expressions with a classifier (e.g. ‘CL NP’ and ‘Dem CL NP’) are always interpreted as singular.<sup>5</sup> These languages include Mandarin (e.g. Tang 1990; Iljic 1994; Yang 1998), Cantonese (Cheng and Sybesma 1999), Thai (Bisang 1999; Piriya-wiboon 2010), Bangla, Hmong and Vietnamese (Simpson et al. 2011). In Mandarin, ‘Dem CL NP’ only refers to a singular entity, as in (19a). (19b) shows that it is not the case that the demonstrative *nei* ‘that’ selects for a singular referent. (19c) shows that ‘CL NP’ is grammatical in an object position and receives an

<sup>5</sup>I put aside ‘plural classifiers’, including *dī* (Cantonese) and *cov* (Hmong), for the moment.

indefinite singular interpretation.

- (19) a. Nei ben shu hen gui.  
 that CLF book very expensive  
 ‘That book is expensive.’/\*‘Those books are expensive.’
- b. Nei liang ben shu hen gui.  
 that two CLF book very expensive  
 ‘Those two books are expensive.’
- c. Yuehan mai-le ben shu.  
 John buy-ASP CLF book  
 ‘John bought a book/\*books.’

(Yang 1998:271)

Examples like (19a) and (19c) are often analyzed as a result of omitting the numeral *yi* ‘one’. Whether such an analysis is correct, the absence of ‘one’ makes sense if the constituent ‘CL NP’, to which the numeral ‘one’ is potentially attached, denotes singularities, which naturally licenses the absence of ‘one’. The Hmong and Bangla examples in (20)–(21) show more directly that ‘CL NP’ denotes singularities. In the (a) examples with the bare classifier construction ‘CL NP’, the referents are necessarily singular. However, when the bare classifier constructions are substituted by bare NPs without a classifier, as in the (b) examples, the referents are no longer restricted to a singular key or light.

- (20) Hmong
- a. Tus yuam sij nyob qhov twg?  
 CLF key stay where  
 ‘Where is the key?’
- b. Yuam sij nyob qhov twg?  
 key stay where  
 ‘Where is/are the key(s)?’
- (21) Bangla (Simpson et al. 2011)
- a. Tumi ki alo-Ta jele dite parbe please?  
 you Q light-CLF turn.on give can please  
 ‘Can you turn on the light, please?’
- b. Tumi ki alo jele dite parbe please?  
 you Q light turn.on give can please  
 ‘Can you turn on (the) light(s), please?’

Furthermore, Cheng and Sybesma (2005) report a fact that also points to ‘CL NP’ denoting singularities. The bare classifier construction ‘CL NP’ in Cantonese and Wenzhou cannot receive a kind reading, as in (22) and (23).

(22) Cantonese (Cheng and Sybesma 1999:511)

- a. Gáu jūngji sihk yuhk.  
dog like eat meat  
‘Dogs love to eat meat.’
- b. \*Jek gáu jūngji sihk yuhk.  
CLF dog like eat meat  
For: ‘Dogs like to eat meat.’ (Acceptable for ‘The dog likes to eat meat.’)

(23) Wenzhou (Cheng and Sybesma 2005:264–265)

- a. Kau<sup>3</sup> si<sup>3</sup>-cy<sup>1</sup> tshi<sup>7</sup> niou<sup>8</sup>.  
dog like eat meat  
‘Dogs like to eat meat.’
- b. \*Dyu<sup>8</sup> kau<sup>3</sup> si<sup>3</sup>-cy<sup>1</sup> tshi<sup>7</sup> niou<sup>8</sup>.  
CLF dog like eat meat  
For: ‘Dogs like to eat meat.’

In the theory of kinds adopted in this study, i.e. that of Chierchia’s (1998b), kinds are obtained by applying the  $\cap$  operator to properties of objects. Remember that  $\cap$  is not defined for singular properties. One can then infer that the bare classifier construction lacks a kind reading because ‘CL NP’ denotes singularities, to which  $\cap$  cannot apply.

To summarize, classifier meaning consists of two components. The first component restricts the domain to singularities, i.e. atoms. This meaning is truth-conditional and the core meaning common to all classifiers. The second component specifies the characteristics that the nouns combining with the relevant classifier have. This meaning is conventionally implicated, with its content varying from classifier to classifier. I propose the semantics of classifiers in (24). The ‘◆’ symbol indicates that the material after ◆ is conventional implicatures.<sup>6</sup>

- (24)  $[[\text{CL}]] = \lambda P \lambda x [P(x) \wedge \neg \exists y \in P [y < x]] \text{◆} \lambda P [P \subseteq \text{CLASS}]$ , where  
CLASS denotes conjoined properties that nouns with which the relevant classifier is used have, e.g. inanimate  $\wedge$  3D  $\wedge$  big (for the classifier *buah* in Malay).

<sup>6</sup>This notation is taken from McCready (2010). I assume his theory of conventional implicature computation, which is an extension of Potts’s (2005).

Classifiers are functions from number-neutral properties to properties that hold with singularities alone. Their semantic type is  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$ . Under the current view, classifiers boil down to a sophisticated type of singular number morphology, in the sense that they not only restrict the domain to singularities/atoms, but they also specify the nature of the relevant atoms. The difference between classifiers and ordinary singular morphology is whether the conventionally implicated meaning, i.e. CLASS, is included.

In the Linear Optimality Theoretic analysis proposed in chapter 3 (section 3.3.2), the contents of CLASS were treated as constraints with different numeric weights. As OT constraints, they are violable and the seriousness of their violations correlates with their associated weights. To capture such differences in importance (i.e. seriousness of violations) among the properties constituting CLASS, one can supply each property with the numeric weight associated with it. Unlike the conventionally implicated meanings, i.e. CLASS, the core truth-conditional meaning of classifiers, i.e. restriction of the domain to singularities, is not violable. Failure to satisfy it simply leads to falsity, but not varying degrees of inappropriateness. The meaning computation system involving it is thus considered to be part of the Generator. Given these ideas and the constraint weights used in chapter 3, the semantics of the specific classifier for machines *dai* in Japanese looks like (25). The constraint weights are indicated by subscripts.

$$(25) \quad \llbracket \text{dai} \rrbracket = \lambda P \lambda x [P(x) \wedge \neg \exists y \in P [y < x]] \blacklozenge \lambda P [P \subseteq \text{INANIMATE}_{34.81} \wedge \\ P \subseteq \text{MECHANICAL}_{3.9} \wedge P \subseteq \text{PLACED ON THE GROUND}_{0.5} \wedge P \subseteq \text{DETACHED}_{1.3} \wedge \\ P \subseteq \text{CARRYING THINGS}_{0.2}]$$

In terms of truth-conditional semantics, *dai* can combine with any noun as long as the noun is countable, that is, it has atoms. Gradient acceptability is observed with different nouns because different nouns satisfy the conventionally implicated meanings of *dai* to different degrees. Let us consider how *dai* combines with the following four nouns: *kuruma* ‘car’, which Matsumoto (1993) identifies as a prototypical noun used with *dai*, *hondana* ‘bookcase’, which is judged less acceptable with *dai* in Matsumoto’s experiment, *neko* ‘cat’, which is definitely unacceptable with *dai*, and *mizu* ‘water’, which is a mass noun and hence cannot combine with any classifiers including *dai* unless an implicit container is assumed.

- (26) a.    iti dai-no kuruma  
          one CLF-LINK car  
          ‘a car’

- b. ??iti dai-no honbako  
 one CLF-LINK bookcase  
 ‘a bookcase’
- c. \*iti dai-no neko  
 one CLF-LINK cat  
 ‘a cat’
- d. #iti dai-no mizu  
 one CLF-LINK water  
 #‘a water’

(27) shows the semantics of the four nouns, where the predicates in small capital represent number-neutral properties that hold of the relevant entity. As I have claimed in chapter 2, no difference exists between classifier and non-classifier languages regarding common noun denotations. They denote properties.

- (27) a.  $\llbracket \text{kuruma} \rrbracket = \lambda x. \text{CAR}(x)$   
 b.  $\llbracket \text{honbako} \rrbracket = \lambda x. \text{BOOKCASE}(x)$   
 c.  $\llbracket \text{neko} \rrbracket = \lambda x. \text{CAT}(x)$   
 d.  $\llbracket \text{mizu} \rrbracket = \lambda x. \text{WATER}(x)$

*Mizu* ‘water’ cannot combine with *dai* as in (26d) because it does not satisfy the clause ‘ $\neg \exists y \in P[y < x]$ ’ in (25). Water lacks minimal parts that cannot be divided further, at least in the same sense as cars and bookcases have minimal car or bookcase parts (cf. Link 1983; Chierchia 1998a, 2010). Therefore, (26d) ends in falsity.

Putting aside the conventional implicatures (i.e. the material after  $\blacklozenge$ ), the other three nouns can all combine with *dai*, as shown in (28). The linker *no* is ignored, as I consider it to be semantically vacuous like *of* in English.

- (28) a.  $\llbracket \text{dai kuruma} \rrbracket = \lambda x [\text{CAR}(x) \wedge \neg \exists y \in \text{CAR}[y < x]]$   
 ‘a set of car atoms’
- b.  $\llbracket \text{dai honbako} \rrbracket = \lambda x [\text{BOOKCASE}(x) \wedge \neg \exists y \in \text{BOOKCASE}[y < x]]$   
 ‘a set of bookcase atoms’
- c.  $\llbracket \text{dai neko} \rrbracket = \lambda x [\text{CAT}(x) \wedge \neg \exists y \in \text{CAT}[y < x]]$   
 ‘a set of cat atoms’

The differences in acceptability observed between the three reflect which conventionally implicated properties of *dai* are violated by these three nouns. The tableau in (29), which assembles three tableaux together for exposition, shows these differences. ‘Car’ satisfies all constraints, whereas ‘bookcase’ and ‘cat’ violate some constraints. While both of them are less acceptable than ‘car’ for this reason, ‘cat’ is much worse than ‘bookcase’ as it violates the constraint with the highest weight, i.e. ‘inanimate’.

(29)

	INANIMATE	MECHANICAL	PLACED ON THE GROUND	DETACHED	CARRYING THINGS	
<i>dai kuruma/honbako/neko</i> [CLF car/bookcase/cat]	34.81	3.9	0.5	1.3	0.2	<i>H</i>
car						0
?? bookcase		*		*		−4.2
* cat	*	*			*	−38.91

### 4.2.3 A comparison with the “unit of counting” approach to classifiers

One important aspect of the semantics of classifiers proposed in the last section (cf. (24)) is that it does not contain a number variable. This feature distinguishes my analysis from a family of analyses that contain one. The latter approach is represented by Krifka (1995) and Wilhelm (2008), among many others. I will refer to this approach as the “unit of counting” approach, as the number variable is associated with units of counting. The “unit of counting” approach obviously assumes the “classifiers for counting” thesis, which I claimed to be problematic in chapter 2.

I take Wilhelm’s (2008) analysis as an example and show problems of the “unit of counting” approach. My choice of Wilhelm (2008) rather than Krifka (1995) is because the latter involves a complication absent from the former, which makes a comparison with my analysis more complex. Specifically, Krifka assumes that bare nouns universally denote kinds and references to their specimens and subkinds are achieved through the realization and taxonomy relations respectively. His analysis thus differs from mine not only in the number variable but also with respect to these relations. Wilhelm (2008) ascribes the difference between classifier

and non-classifier languages (i.e. classifiers are used in numeral modification in the former languages but not in the latter) to a difference in the denotation of numerals between the two types of languages. She claims that numerals contain an atom-accessing function in non-classifier languages, whereas numerals in classifier languages lack one. The semantics of numerals in non-classifier languages is as in (30).

- (30) Numerals of non-classifier languages (e.g. English *three*, Dëne *taghe* ‘three’)  
 $\llbracket \text{three/taghe} \rrbracket = \lambda P \lambda x [P(x) \ \& \ \text{OU}(x) = 3]$   
 ‘a function from a set  $P$  (of atoms and sums) onto that subset of  $P$  containing the sums of three object units/atoms’  
 OU is a function which gives the number of ‘object units’ (i.e. atoms) in a plurality.

(Wilhelm 2008:55)

By contrast, the denotation of numerals in classifier languages is claimed to just denote a number. For example, in Mandarin,  $\llbracket \text{san} \rrbracket = 3$ . In classifier languages, classifiers express the atom-accessing function OU. The general classifier *ge* ‘unit’ in Mandarin thus has the following semantics:

- (31)  $\llbracket \text{ge} \rrbracket = \lambda n \lambda P \lambda x [P(x) \ \& \ \text{OU}(x) = n]$ , where  $n$  is a natural number

(Wilhelm 2008:55)

There are at least two problems with this analysis. First, some classifier languages also use classifiers in non-counting contexts (cf. (22)–(23)). Under the “unit of counting” approach, all classifiers would have two denotations in these languages, i.e. one with a number variable  $n$  (for use in the counting context) and another without it (for use in non-counting contexts).<sup>7</sup>

Second, numerals in optional classifier languages would also have two denotations, i.e. the non-classifier language type (for ‘Num NP’) and the obligatory classifier language type (for ‘Num CL NP’).<sup>8</sup> Multiple significantly different denotations for numerals must be posited in

<sup>7</sup>Given that classifiers restrict the domain to singularities, one could also characterize the two as follows: one with a number variable  $n$  (for use in the counting context) and another with the number ‘1’ (for use in non-counting contexts).

<sup>8</sup>Dalrymple and Mofu (2012) reach a similar conclusion for Indonesian, but they claim that it is not problematic for them, given a “glue” semantic approach.

a single language in well-known obligatory classifier languages too, as classifiers become optional in some contexts even in these languages (cf. chapter 2, section 2.3.2). Thus, all Sino-Japanese numerals in Japanese will be ambiguous under the “count of unit” approach, which weakens the plausibility of the approach considerably. Another type of numerals, i.e. native Japanese numerals, are limited to numbers one through ten, and are less commonly used than Sino-Japanese numerals. Unlike Sino-Japanese numerals, native Japanese numerals except *yon* ‘four’, *nana* ‘seven’ and *too* ‘ten’ almost always require classifiers.<sup>9</sup> The native Japanese numeral ‘ten’ modifies nouns directly and never takes classifiers. These facts may appear to lend support to the “unit of counting” approach, as such an approach can easily accommodate idiosyncratic behaviors in individual numerals. However, I follow the standard view that the reason why native Japanese numerals require classifiers is not semantic but morphosyntactic. Specifically, they are bound morphemes that prefix to classifiers. Like numerals, classifiers can also be divided into native and Sino-Japanese ones. Selectional relations exist between numerals and classifiers. In general, native Japanese classifiers combine with native Japanese numerals whereas Sino-Japanese classifiers combine with Sino-Japanese numerals. Some examples of numeral-classifier combinations are given in Table 4.1.

### 4.3 The semantics of numerals

#### 4.3.1 Two kinds of numerals

In the present analysis, the denotation of numerals does not vary across languages as significantly as claimed by the “unit of counting” approach. For numeral semantics, I adopt Bale et al.’s (2011) analysis. According to them, numerals are restrictive modifiers (Link 1983; Landman 2004; Ionin and Matushansky 2006), which can be subsective or intersective like adjectives. They propose the two numerals as universally available options from which individual

<sup>9</sup>The qualification “almost” is necessary because the numerals *hito*- ‘one’ and *huta*- ‘two’ (and less commonly *mi*- ‘three’) attach to some event nominals, as in (i).

- (i) *hito*-/*huta*- *aruk*-i ‘a/two walk(s)’ (cf. *take a walk*)  
*hito*-/*huta*- *huk*-i ‘a/two wipe(s)’ (cf. *give it a wipe*)  
*hito*-/*huta*- *yom*-i ‘a/two read(s)’ (cf. *give it a read*)  
*hito*-/*huta*- *nemur*-i ‘a/two sleep(s)’ (cf. *have a sleep*)

Table 4.1: Combinations of numerals and classifiers in Japanese. Native Japanese numerals/classifiers are in **bold** and Sino-Japanese numerals/classifiers are *italicized*.

	Independent use		<b>tu</b> (default)	<i>ko</i> (3D, big)	<i>dai</i> (machines)	<b>ri/nin</b> (humans)
	<b>Native</b>	<i>Sino-Jap.</i>				
1	<b>*hito-</b>	<i>iti</i>	<b>hito-tu</b>	<i>ik-ko</i>	<i>iti dai</i>	<b>hito-ri</b>
2	<b>*huta-</b>	<i>ni</i>	<b>huta-tu</b>	<i>ni ko</i>	<i>ni dai</i>	<b>huta-ri</b>
3	<b>*mi-</b>	<i>san</i>	<b>mit-tu</b>	<i>san ko</i>	<i>san dai</i>	<i>san nin</i>
4	<b>yon</b>	<i>si</i>	<b>yot-tu</b>	<b>yon ko</b>	<b>yon dai</b>	<b>yo-nin</b>
5	<b>*itu-</b>	<i>go</i>	<b>itu-tu</b>	<i>go ko</i>	<i>go dai</i>	<i>go nin</i>
6	<b>*mu-</b>	<i>roku</i>	<b>mut-tu</b>	<i>rok-ko</i>	<i>roku dai</i>	<i>roku nin</i>
7	<b>nana</b>	<i>siti</i>	<b>nana tu</b>	<b>nana ko</b>	<b>nana dai</b>	<b>nana/siti nin</b>
8	<b>*ya-</b>	<i>hati</i>	<b>yat-tu</b>	<i>hati/hak-ko</i>	<i>hati dai</i>	<i>hati nin</i>
9	<b>*kokono-</b>	<i>kyuu/ku</i>	<b>kokono-tu</b>	<i>kyuu ko</i>	<i>kyuu dai</i>	<i>kyuu/ku-nin</i>
10	<b>too</b>	<i>juu</i>	—	<i>juk-ko</i>	<i>juu dai</i>	<i>juu nin</i>

languages may choose one or both. The semantics of the two types of numerals and necessary definitions are given in (32) and (33) respectively, with some modifications made to Bale et al.'s original formulations. The two figures in Figure 4.1 visually show how the two kinds of numerals pick out the same individual in different ways when they combine with nouns denoting number-neutral properties. The subsective numeral  $3^S$  in (32a) restricts the noun denotation to a subset of it such that the cardinality is three (Figure 4.1a), like subsective adjectives such as *big* restrict the noun denotation to a subset that satisfies the adjective description (the bigness of a *big mouse* is judged in comparison with that of other mice). The intersective numeral  $3^I$  in (32b) intersects the noun denotation with the set of plural individuals whose cardinality is three (Figure 4.1b), like intersective adjectives such as *pregnant* intersect the noun denotation with the adjective denotation (a *pregnant mouse* is an individual that belongs to both the set of mice and the set of pregnant things).

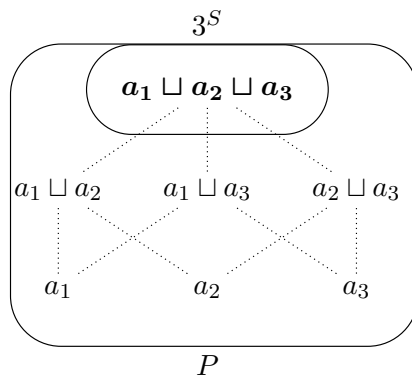
- (32) a. Subsective numerals  

$$\llbracket 3^S \rrbracket = \lambda P \lambda x. \exists Y [\Pi(Y)(x) \wedge |Y| = 3 \wedge \forall z \in Y [z \in \text{MIN}(P)]]$$
- b. Intersective numerals  

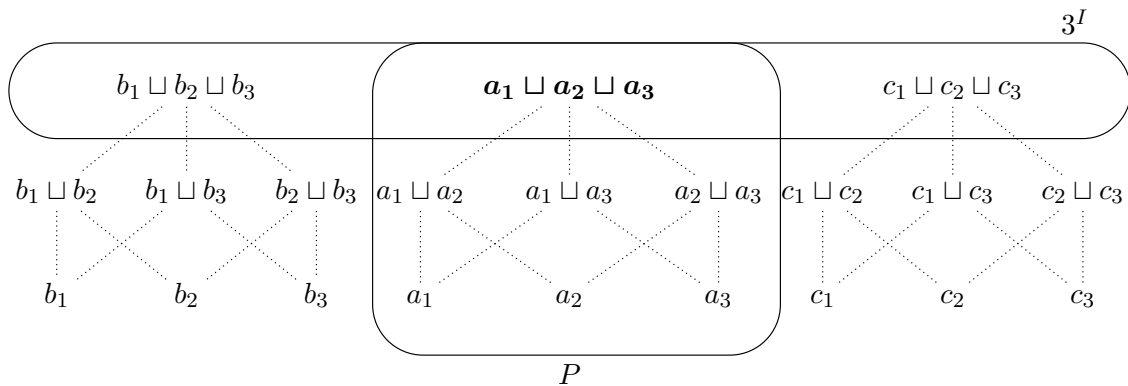
$$\llbracket 3^I \rrbracket = \lambda P. P \cap \{x : \exists Y [\Pi(Y)(x) \wedge |Y| = 3 \wedge \forall z \in Y [\text{ATOM}(z)]]\}$$
- (33) a.  $\Pi(Y)(x) = 1$  iff  $Y$  is a partition of  $x$ .  
b. A partition of a plural individual  $x$  is a set of individuals  $Z$  such that the sum of

all the elements in  $Z$  is equal to  $x$  ( $\sqcup Z = x$ ) and no two elements in  $Z$  overlap ( $\forall v, w \in Z [v = w \vee v \sqcap w = \emptyset]$ ).

- c.  $\text{MIN}(P)$  is defined iff no two smallest elements in  $P$  overlap ( $\forall x, y \in P [\neg \exists z \in P [z < x \vee z < y] \rightarrow x \sqcap y = \emptyset]$ ).  
When defined,  $\text{MIN}(P) = \{x \in P : \neg \exists z \in P [z < x]\}$ .
- d.  $\text{ATOM}(x) = 1$  iff  $x$  is an atom in the domain of the model.



(a) Subjective numerals.



(b) Intersective numerals.

Figure 4.1: Two kinds of numerals.

I will illustrate how these two types of numerals combine with different nominal forms in Malay, Persian and English shortly (section 4.3.2).

Aside from being subjective or intersective, the two numerals differ in the type of predicates with which they can combine. I assume that they are defined only for the predicates with which

they can combine. Intersective numerals greater than one are defined for plural and number-neutral properties, but not for singular ones. This is because combinations of numerals greater than one and singular properties always denote an empty set. Conversely, subsective numerals are defined for singular and number-neutral properties, but not for plural ones. This is because the MIN function in the last conjunct of (32a) is undefined for pluralities. The smallest elements of plural properties are overlapping pairs such as  $a \sqcap b$ ,  $a \sqcap c$  and  $b \sqcap c$ , and hence the first line of the definition of MIN (33c) is not satisfied.

In fact, subsective numerals defined as in (32a) are not always subsective, despite its name. Specifically, they create a new set disjoint from the original set when they combine with nouns denoting singularities, i.e. when  $P$  is a property of singularities. Bale et al. (2011) do not consider this possibility, claiming that numerals greater than one only combine with number-numeral properties. Hence, for them, subsective numerals are always subsective, and all modifiers including numerals are restrictive. The latter consequence appears favorable in view of the general semantic theory. However, such theoretical elegance makes a sacrifice of an empirical fact. Numeral modification in classifier languages obviously involves a combination of a numeral and ‘CL NP’, which I claimed to denote properties of singularities in section 4.2.2. Thus, in the Malay expression *tiga buah majalah* [three CLF magazine] ‘three magazines’, the numeral *tiga* maps the singularities denoted by *buah majalah* to the pluralities with cardinality ‘3’, as shown in Figure 4.2.

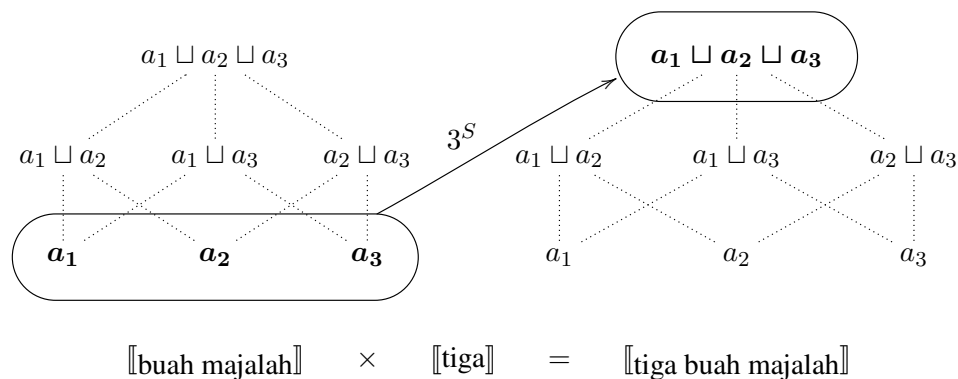


Figure 4.2: Combination of a subsective numeral and ‘CL NP’.

It may be possible to remedy this non-subsective part of subsective numerals by assuming the constituent structure as in (34a), where the classifier merges with the numeral rather than the

NP, instead of the structure that I have implicitly assumed so far as in (34b), where the classifier merges with the NP.



The idea is that the constituent ‘Num CL’ as a whole serves as a subsective numeral as if the classifier had no semantic contribution. In such an analysis, ‘Num CL’ involving a numeral can only combine with number-neutral properties, as assumed by Bale et al. (2011), and subsective numerals do not pose a problem to the generalization that all modifiers are restrictive. As we have seen in the last section, the noun specifications carried by classifiers are not asserted meanings, but conventional implicatures. Hence, classifiers will be simply identity functions on properties in the analysis under consideration:  $\llbracket \text{CL} \rrbracket = \lambda P.P$  (the conventional implicature portion is omitted). As shown in (35), the asserted meaning of ‘Num CL’ is identical to that of the numeral alone.

$$(35) \quad \llbracket 3^S \text{CL} \rrbracket = \lambda P \lambda x. \exists Y [\Pi(Y)(x) \wedge |Y| = 3 \wedge \forall z \in Y [z \in \text{MIN}(P)]] = \llbracket 3^S \rrbracket \quad (32a)$$

The most telling problem of this analysis is that the semantic contribution of classifiers, i.e. restriction to singularities, is neglected. Consequently, the phenomenon observed in several languages in which expressions that involve classifiers but not numerals obligatorily refer to singular entities (cf. section 4.2.2) will lose a natural account. Therefore, I consider the potential non-subsectivity of subsective numerals as unproblematic. In fact, a compositional analysis of complex cardinals such as *three hundred* also requires that (some) numerals be privative (Ionin and Matushansky 2006). The set of plural individuals with cardinality ‘300’ is disjoint from the set of plural individuals with cardinality ‘100’.

### 4.3.2 Compositional semantics for ‘three books’ crosslinguistically

I will now demonstrate how the two types of numerals proposed in the last section combine with different nominal forms in Malay, Persian and English. For concreteness, I discuss expressions meaning ‘three books’. It is shown that a choice of particular numeral types determines the nominal forms that can co-occur with numerals.

To recapitulate, numerals can be subsective or intersective like adjectives. Languages may

choose one or both types. Subjective numerals are defined for singular and number-neutral properties, but not for plural ones. Intersective numerals greater than one are defined for plural and number-neutral properties, but not for singular ones. The semantics of two types of numerals are repeated below.

- (32) a. Subjective numerals: Undefined for plural properties  
 $\llbracket 3^S \rrbracket = \lambda P \lambda x. \exists Y [\Pi(Y)(x) \wedge |Y| = 3 \wedge \forall z \in Y [z \in \text{MIN}(P)]]$
- b. Intersective numerals: Undefined for singular properties  
 $\llbracket 3^I \rrbracket = \lambda P. P \cap \{x: \exists Y [\Pi(Y)(x) \wedge |Y| = 3 \wedge \forall z \in Y [\text{ATOM}(z)]]\}$

#### 4.3.2.1 Malay: Both subjective and intersective numerals

In Malay, numerals can co-occur with any nominal forms, i.e. (a) NP, (b) CL NP, (c) reduplicated NP and (d) CL reduplicated NP, though co-occurrence with the latter two patterns, especially pattern (c), is uncommon (Chung 2000) and prohibited in prescriptive grammars.

- (36) ‘three books’ in Malay
- a. tiga buku  
three book
- b. tiga buah buku  
three CLF book
- c. (\*)tiga buku-buku  
three book.PL
- d. (\*?)tiga buah buku-buku  
three CLF book.PL

Patterns (a), (b) and (c) respectively denote number-neutral, singular and plural properties. As for pattern (d), where a classifier and plural morphology co-occur, I will argue in chapter 6 that classifiers and plural markers are realizations of the [+Sg] and [+Pl] features in the number head respectively, and forms with both these markers denote the union of the singularities and pluralities. Hence, pattern (d) denotes number-neutral properties.

The Malay facts follow if the language has both subjective and intersective numerals. In a context where the domain contains three book atoms ( $b_1, b_2, b_3$ ) and two pencil atoms ( $p_1, p_2$ ):

- (37) a. NP  
 $\llbracket \text{buku} \rrbracket = \{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}$
- b. CL NP  
 $\llbracket \text{buah buku} \rrbracket = \{b_1, b_2, b_3\}$
- c. NP-PL  
 $\llbracket \text{buku-buku} \rrbracket = \{b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}$
- d. CL NP-PL  
 $\llbracket \text{buah buku-buku} \rrbracket = \{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}$
- e. (the right conjunct of  $\llbracket \text{tiga}^I(3) \rrbracket$ ) =  $\{b_1 \sqcup b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup p_1, b_1 \sqcup b_2 \sqcup p_2, b_1 \sqcup b_3 \sqcup p_1, b_1 \sqcup b_3 \sqcup p_2, b_2 \sqcup b_3 \sqcup p_1, b_2 \sqcup b_3 \sqcup p_2, b_1 \sqcup p_1 \sqcup p_2, b_2 \sqcup p_1 \sqcup p_2, b_3 \sqcup p_1 \sqcup p_2\}$

(36a) results from either subsective or intersective numerals.

- (38) a.  $\llbracket \text{tiga}^S \text{buku} \rrbracket = \llbracket \text{tiga}^S \rrbracket(\llbracket \text{buku} \rrbracket)$   
 $= \llbracket \text{tiga}^S \rrbracket(\{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$
- b.  $\llbracket \text{tiga}^I \text{buku} \rrbracket = \llbracket \text{buku} \rrbracket \cap (37e)$   
 $= \{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\} \cap (37e) = \{b_1 \sqcup b_2 \sqcup b_3\}$

(36b) results from subsective numerals. Intersective numerals are undefined for ‘CL NP’, which denotes properties of singularities.

- (39) a.  $\llbracket \text{tiga}^S \text{buah buku} \rrbracket = \llbracket \text{tiga}^S \rrbracket(\llbracket \text{buah buku} \rrbracket)$   
 $= \llbracket \text{tiga}^S \rrbracket(\{b_1, b_2, b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$
- b.  $\llbracket \text{tiga}^I \text{buah buku} \rrbracket = \llbracket \text{buah buku} \rrbracket \cap (37e) = \{b_1, b_2, b_3\} \cap (37e) = \text{undefined}$

Conversely, (36c) results from intersective numerals. Subsective numerals are undefined for reduplicated NPs, which denote properties of pluralities.

- (40) a.  $\llbracket \text{tiga}^S \text{buku-buku} \rrbracket = \llbracket \text{tiga}^S \rrbracket(\llbracket \text{buku-buku} \rrbracket)$   
 $= \llbracket \text{tiga}^S \rrbracket(\{b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \text{undefined}$
- b.  $\llbracket \text{tiga}^I \text{buku-buku} \rrbracket = \llbracket \text{buku-buku} \rrbracket \cap (37e)$   
 $= \{b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\} \cap (37e) = \{b_1 \sqcup b_2 \sqcup b_3\}$

Lastly, (36d) results from either subsective or intersective numerals.

- (41) a.  $\llbracket \text{tiga}^S \text{ buah buku-buku} \rrbracket = \llbracket \text{tiga}^S \rrbracket (\llbracket \text{buah buku-buku} \rrbracket)$   
 $= \llbracket \text{tiga}^S \rrbracket (\{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$
- b.  $\llbracket \text{tiga}^I \text{ buah buku-buku} \rrbracket = \llbracket \text{buah buku-buku} \rrbracket \cap (37e)$   
 $= \{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\} \cap (37e) = \{b_1 \sqcup b_2 \sqcup b_3\}$

Two possible reasons are conceivable as to why the (c) pattern (i.e. Num reduplicated NP) is uncommon. The first possibility is that numerals are intersective for only some speakers and they are subsective for most speakers. The second possibility is that the distribution of the pattern is constrained by some yet unknown conditions. At any rate, instances of this pattern are attested in naturally occurring texts. The examples below are both from Malaysian online newspapers.

- (42) Pada Julai 2010, kerajaan mengharamkan *tujuh buku-buku kartun* beliau dengan on July 2010 government ban seven book.PL cartoon his with alasan kandungan-nya menjejaskan ketenteraman awam. excuse content-its affect peace public  
 ‘In July, 2010, the government banned seven cartoon books of his on the grounds that they would disturb the public order.’

(*Keadilan Daily*, 1/17/2012, <http://www.keadilandaily.com/?p=25585>, accessed 12/8/2012)

- (43) Pada mulanya, sebanyak *tujuh wakil* dari *enam negara-negara Islam* at first as many as seven representative from six country.PL Islam bertindak sebagai pemegang amanah . . . . act as holder trust  
 ‘At first, seven representatives from six Islamic countries acted as the trustees . . . .’

(*Utusan Malaysia*, 10/30/2012, [http://www.utusan.com.my/utusan/Bicara\\_Agama/20121030/ba\\_03/Keutuhan-Islam-di-Kota-London](http://www.utusan.com.my/utusan/Bicara_Agama/20121030/ba_03/Keutuhan-Islam-di-Kota-London), accessed 12/8/2012)

The reason why the (d) pattern (i.e. Num CL reduplicated NP) is uncommon has to do with the Gricean co-operative principle, specifically the third submaxim of the maxim of Manner: Be brief (avoid unnecessary prolixity) (Grice 1975). As stated above, both patterns (a) NP and (d) CL reduplicated NP denote number-neutral properties. The latter pattern is formally marked

compared to the former. Hence, the unmarked pattern (a) is the default option for denoting number-neutral properties while the use of the marked pattern (d) implies a marked meaning. The presence of this marked meaning restricts the distribution of the (d) pattern. In Malay, the (d) pattern tends to occur with large numbers, as demonstrated by the following examples. This is arguably due to the presence of the plural morphology.

- (44) ... ada kira-kira *tiga ratus buah buku-buku* yang telah di-jilid, ....<sup>10</sup>  
 be about three hundred CLF book.PL REL PRF PASS-bind  
 ‘... there are about three hundred bound books, ...’

(*Wikipedia ‘Hikayat Johor’,*

[http://ms.wikipedia.org/wiki/Hikayat\\_Johor](http://ms.wikipedia.org/wiki/Hikayat_Johor), accessed 12/13/2012)

- (45) ... seramai *25 orang pelajar-pelajar* dari persekitaran yang sukar dan  
 as.many.as 25 CLF student.PL from environment REL difficult and  
 mencabar ... menerima surat tawaran mereka daripada menteri.  
 challenging receive letter offer their from minister  
 ‘... 25 students from difficult and challenging environments ... received their offer  
 letters from the minister.’

(*Utusan Malaysia, 6/3/2012, http://www.utusan.com.my/info.asp?y=*  
*2012&dt=0603&pub=Utusan\_Malaysia&sec=Rencana&pg=re\_15.htm,*  
 accessed 12/13/2012)

This analysis predicts that the (d) pattern is more freely available in obligatory classifier languages, which do not generally allow direct numeral modification (i.e. the (a) pattern). In these languages, the (d) pattern is the only number-neutral form that can combine with numerals. The issue of markedness thus does not arise and the (d) pattern does not engender any marked meaning. This is indeed the case in Japanese. The (d) pattern is as natural as the (b) pattern (i.e. Num CL NP) and does not convey any special meaning. Note that the (b) pattern does not

<sup>10</sup>The phrase *tiga ratus buah buku-buku* sounds a little odd in the quoted clause alone. However, the oddity disappears when the subsequent clause in (i) is added.

- (i) tiada termasuk naskah-naskah yang belum berjilid, yang bercerai-berai, bergulung dan yang berhelai-helai  
 ‘which do not include those manuscripts which have not been bound yet, those which are torn into pieces or rolled, and those which are mere bundles of sheets.’

make the (d) pattern marked, because the constituent ‘CL NP’ denotes singular properties, but not number-neutral properties.

- (46) ‘three students’ in Japanese
- a. \*san-no gakusei  
three-LINK student
  - b. san nin-no gakusei  
three CLF-LINK student
  - c. \*san-no gakusei-tati  
three-LINK student-PL
  - d. san nin-no gakusei-tati  
three CLF-LINK student-PL

#### 4.3.2.2 Persian: Only subsective numerals

In Persian, numerals can co-occur with (a) NP, (b) CL NP and (d) CL plural NP, but not with (c) plural NP without CL.<sup>11</sup>

- (47) ‘three books’ in Persian
- a. se ketâb  
three book
  - b. se tâ ketâb  
three CLF book
  - c. \*se ketâb-hâ  
three book-PL
  - d. se tâ ketâb-hâ  
three CLF book-PL

These facts suggest that the language only has subsective numerals, because subsective numerals are not defined for plural properties.

- (48) a.  $[[se^S ketâb]] = [[se^S]]([[ketâb]])$   
 $= [[se^S]](\{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$
- b.  $[[se^S tâ ketâb]] = [[se^S]]([[tâ ketâb]]) = [[se^S]](\{b_1, b_2, b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$

<sup>11</sup>I follow Hamedani’s (2011) description here. Gebhardt (2009) describes (c) as acceptable. The variety of Persian described by Gebhardt (2009) has both subsective and intersective numerals like Malay.

- c.  $\llbracket *se^S \text{ketâb-hâ} \rrbracket = \llbracket se^S \rrbracket(\llbracket \text{ketâb-hâ} \rrbracket)$   
 $= \llbracket se^S \rrbracket(\{b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \text{undefined}$
- d.  $\llbracket se^S \text{tâ ketâb-hâ} \rrbracket = \llbracket se^S \rrbracket(\llbracket \text{tâ ketâb-hâ} \rrbracket)$   
 $= \llbracket se^S \rrbracket(\{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\}) = \{b_1 \sqcup b_2 \sqcup b_3\}$

In the discussion of Malay above, it was pointed out that the (d) pattern is a marked option conveying a marked meaning. The same is the case with Persian. The (d) pattern does not simply mean ‘three books’, but it is interpreted as definite (Gebhardt 2009; Ghaniabadi 2010; Hamedani 2011). I will discuss the relation between plurality and definiteness in the next chapter. It will be shown that the definite reading of the (d) pattern in Persian is not a product of pragmatic inference, but results from the standard syntax and semantics of definiteness that involves a null definite marker. However, it is possible that diachronically the currently rigid connection between the (d) pattern and definiteness started as a pragmatic phenomenon similar to that observed in the (d) pattern in Malay. In both Malay and Persian, the (d) pattern is subject to a restriction in terms of the number of atoms comprising the domain. The difference between the two languages is the strictness of the restriction. The restriction is loose in Malay; the number just needs to be large (relative to the contextually determined standard). In Persian, on the other hand, the number is more strictly restricted; it must be the number specified by the numeral. Unlike the (d) pattern, the (b) pattern can be either definite or indefinite.

In some languages, the co-occurrence of a classifier and a plural marker is not tolerated at all. This situation can also be understood as a result of the pressure that is brought to bear on the marked form (d) by its unmarked competitor (a). The resulting restriction is maximally strict in this case, namely a total ban on the pattern. Armenian is one of these languages.

(49) ‘two umbrellas’ in Armenian (Borer 2005:117–118)

- a. yergu hovanoc  
two umbrella
- b. yergu had hovanoc  
two CLF umbrella
- c. yergu hanvanoc-ner  
two umbrella-PL
- d. \*yergu had hanvanoc-ner  
two CLF umbrella-PL

### 4.3.2.3 English: Only intersective numerals

English is a language that only has intersective numerals. Numerals greater than one can combine with plural forms denoting plural and number-neutral properties, but not with unmarked NPs denoting singular properties. I assume that the English equivalent of classifiers is phonologically null (Sharvy 1978). Under this assumption, plural forms in English are in fact ambiguous between pattern (c), which only bears a plural marker and receives a plural interpretation, and pattern (d), which bears not only plural but also a null classifier and receives a number-neutral interpretation, as shown in (50). I will argue in chapter 6 that English lacks the (a) pattern in syntax.

- (50) b. \*three book = three [Ø book]  
 c. three books (PL) = three [book-s]  
 d. three books = three [Ø book-s]

Numerals cannot combine with unmarked NPs in English because intersective numerals are defined for plural and number-neutral properties, but not for properties of singularities.

- (51) b.  $\llbracket *three^I \text{book} \rrbracket = \llbracket \text{book} \rrbracket \cap (37e) = \{b_1, b_2, b_3\} \cap (37e) = \text{undefined}$   
 c.  $\llbracket \text{three}^I \text{books (PL)} \rrbracket = \llbracket \text{books (PL)} \rrbracket \cap (37e)$   
 $= \{b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\} \cap (37e) = \{b_1 \sqcup b_2 \sqcup b_3\}$   
 d.  $\llbracket \text{three}^I \text{books} \rrbracket = \llbracket \text{books} \rrbracket \cap (37e)$   
 $= \{b_1, b_2, b_3, b_1 \sqcup b_2, b_1 \sqcup b_3, b_2 \sqcup b_3, b_1 \sqcup b_2 \sqcup b_3\} \cap (37e) = \{b_1 \sqcup b_2 \sqcup b_3\}$

### 4.3.2.4 Watanabe's (2010) syntactic analysis

The proposed analysis offers an alternative to syntactic analyses such as Watanabe (2010). Watanabe claims that numerals are licensed only when the number head is marked for the [ $\pm$ augmented] feature, which distinguishes the singular and dual from the plural ( $\geq 3$ ) in the system he adopts, i.e. that of Harbour's (2007). According to Watanabe (2010), numerals cannot co-occur with plural nouns in Turkish as in (52) because the relevant number head only contains [ $-$ singular], but not [ $\pm$ augmented], unlike the corresponding number head in English (realized as *-s* on the noun), which contains both ([ $-$ singular,  $\pm$ augmented]).

- (52) iki öğrenci(\*-ler)  
 2 student-PL  
 'two students'

(Kornfilt 1996:119)

One problem of this analysis is that it cannot account for the fact that the plural marker *-s* in English is also associated with the general number with number-neutral interpretations as mentioned in chapter 3. In Watanabe's analysis, the plural morphology reflects [−singular], whereas a number-neutral interpretation results from either [±augmented] or [±singular, ±augmented]. Table 4.2 shows all possible combinations of the two number features (without considering redundancies) and whether they permit plural morphology in his system as well as how nominals with the relevant features are interpreted. As the last line indicates, no feature combination is left for morphologically plural forms with number-neutral (general) interpretations.

Table 4.2: Number features and plural morphology licensing in Watanabe's (2010) system.

Features	Plural morphology	Meaning
[(±singular), (±augmented)]	*	general (≥ 1)
[(±singular,) +augmented]	*	dual + plural (≥ 2)
[(±singular,) −augmented]	*	singular (= 1)
[+singular (, ±augmented)]	*	singular (= 1)
[+singular, +augmented]	*	(undefined)
[+singular, −augmented]	*	singular (= 1)
[−singular (, ±augmented)]	✓	dual + plural (≥ 2)
[−singular, +augmented]	✓	plural (≥ 3)
[−singular, −augmented]	✓	dual (= 2)
?	✓	general (≥ 1)

In the analysis proposed in this chapter, Turkish falls into the same class as Persian, but not English, and does not have intersective numerals, which can combine with forms denoting pluralities. In the feature system proposed in chapter 6, both Turkish and English plurals are '−Sg], [+Pl]'. Morphologically plural forms with *-s* in English are also associated with the feature combination '[+Sg], [+Pl]', which captures their number-neutrality.

## 4.4 Summary

This chapter discussed the semantics of classifiers and numerals. The semantics of classifiers consists of two components, i.e. the asserted meaning and conventional implicatures. The former has to do with the singular number, i.e. restriction of the domain to singularities, whereas the latter is concerned with the classification function. Most previous studies have considered the latter either as part of classifier's truth-conditional meaning or, what is worse, as the only meaning component that should be included in a description of classifier semantics. Such views will miss an important parallelism present between classifiers and plural markers that I will point out in the next chapter.

This chapter also discussed the semantics of numerals. Following Bale et al. (2011), I proposed two kinds of numerals, from which languages can choose one or both: subsective and intersective numerals. I claimed that these two types of numerals differ in the type of objects for which they are defined, and demonstrated that it is possible to derive possible and impossible nominal forms combining with numerals based on this definitional difference and the particular choice of numerals each language makes.

## Chapter 5

# Plural markers in classifier languages

### 5.1 Introduction

In the last chapter, I have argued that the core meaning of classifiers is to restrict the domain to singularities and that the classification function is not asserted but conventionally implicated. In this chapter, I point out a parallelism between classifiers and plural markers in classifier languages, taking Japanese as an example (section 5.2). The parallelism confirms the claim by Chung (2000) and many others that the existence of classifiers in a language does not necessarily preclude the possibility that the language has a genuine plural marker. On the contrary, the existence of classifiers sometimes indicates the possibility of there being (multiple) plural markers whose application is limited to certain nouns like classifiers. Section 5.3 discusses the relation between the core meaning of plural markers, i.e. restriction of the domain to pluralities, and definiteness, which has been often associated with plural markers in classifier languages in the literature. It is shown that plurals in classifier languages are not necessarily definite; instead, they are referential, unless modified by quantifiers. Section 5.4 discusses two cases in which plurals cease to be referential, i.e. modification and contrastive constructions, and proposes an analysis that explains them. The key element in explaining both the referential interpretation of plurals and the apparent loss thereof, I argue, is situation variables.

## 5.2 Asserted and implicated meanings of plural markers

In chapter 3, I have shown that plurals in classifier languages are genuine plurals, associated exclusively with the plural number, unlike bare plurals in English, which are associated with not only plural but the transnumeral, general number. Plurals in classifier languages are thus not felicitous to ask whether someone has a child, as shown in (1)–(2). Bare NPs, which are associated with the general number, are used in this context. Plurality is expressed by reduplication and the suffix *-tati* in Malay and Japanese respectively.

(1) Malay

Saya tidak tahu sama ada mereka ada anak/#anak-anak.  
 I not know whether they have child/child-PL  
 ‘I do not know whether they have children.’

(2) Japanese

Watasi-wa karera-ni kodomo/#kodomo-tati-ga iru-ka sira-nai.  
 I-TOP they-DAT child/child-PL-NOM be-whether know-not  
 ‘I do not know whether they have children.’

Restriction of the domain to pluralities is the core asserted meaning of plural markers in classifier languages. Sentences (3) and (4) are false if only one student is exercising.

(3) Malay

Pelajar-pelajar sedang bersenam di padang.  
 student.PL PROG exercise at field  
 ‘(The) students are exercising in the schoolyard.’

(4) Japanese

Gakusei-tati-ga kootei-de undoosi-teiru.  
 student-PL-NOM schoolyard-at exercise-PROG  
 ‘(The) students are exercising in the schoolyard.’

Reduplication in Malay resembles the suffix *-s* in English except for the ability of the latter to denote number-neutral properties. The suffix *-tati* in Japanese differs from these plural markers in at least two respects. First, while the plural markers in Malay and English apply to any count nouns, *-tati* typically attaches to a limited class of nouns, specifically animate nouns. Second,

while reduplication in Malay and *-s* in English are the only productive plural marker in the respective languages, Japanese has three productive plural markers other than *-tati*: *-gata*, *-domo* and *-ra*. The four plural markers differ in the level of politeness. *-Tati* is neutral. *-Gata* and *-domo* indicate that the referent is honorable and humble respectively. Some nouns take all these markers, in which case different degrees of politeness are implied by different forms. *Sensei* ‘teacher’ is one such noun. While *sensei-tati* is neutral as to whether one respects teachers, *sensei-gata* and *sensei-domo* respectively convey one’s respect and disrespect toward teachers, real or pretentious. *-Ra* expresses humbleness, but only slightly and not as strongly as *-domo*. It typically occurs with pronouns and demonstratives, as in *kare-ra* [he-PL] ‘they’ and *kore-ra* [this-PL] ‘these’.

I argue that the restriction to animate nouns and the politeness-related meanings of plural markers in Japanese are not part of their asserted meanings but are conventionally implicated. That is to say, the meanings of plural markers are structured in the same way as those of classifiers. They consist of the asserted meaning concerning number—singular for classifiers and plural for plural markers—and conventional implicatures concerning the characteristics of the noun. We can thus think of plural markers such as *-tati* as plural counterparts of classifiers. In the rest of this section, I present evidence that the meanings concerning the characteristics of the noun are not asserted but conventionally implicated.

To begin with, the restriction to animate nouns is defined in the same semantic terms employed to define nouns with which classifiers can occur. Nouns that occur with the classifier *nin* in Japanese constitute a proper subset of those to which the plural suffix *-tati* can be attached. Simplifying a bit, the former set can be defined by the semantic features ANIMATE and HUMAN and the latter by ANIMATE. These features are the content of CLASS in the semantics of classifiers proposed in the last chapter, repeated below for convenience.

- (5)  $\llbracket \text{CL} \rrbracket = \lambda P \lambda x [P(x) \wedge \neg \exists y \in P [y < x]] \blacklozenge \lambda P [P \subseteq \text{CLASS}]$ , where  
 CLASS denotes conjoined properties that nouns with which the relevant classifier is used have, e.g.  $\text{inanimate} \wedge 3\text{D} \wedge \text{big}$  (for the classifier *buah* in Malay).

The semantics of *-tati* looks like (6), where the assertion portion (i.e. the material before  $\blacklozenge$ ) restricts the domain to the set of entities that are divisible to smaller parts, i.e. the set of pluralities.

- (6)  $\llbracket \text{-tati} \rrbracket = \lambda P \lambda x [P(x) \wedge \exists y \in P [y < x]] \blacklozenge \lambda P [P \subseteq \text{ANIMATE}]^1$

This type of plurals is often found in classifier languages, because classifier languages use CLASS extensively for the singular. Because of this shared component of CLASS, both classifiers and plural markers only apply to certain semantically defined classes of nouns.

As the content of CLASS is conventionally implicated rather than asserted, the mismatch of a plural marker and a noun as in (7) does not make the sentence false, but just inappropriate.

- (7) ?Mukoogawa kara yon dai-no basu-tati-ga yatteki-ta.  
 other.side from four CLF-LINK bus-PL-NOM come-PST  
 ‘Four buses came from the other side.’

The inappropriateness is gradient, depending on how deviant the referent is from the prototypical animal, i.e. humans. For example, although *basu-tati* [bus-PL] is not appropriate, it sounds better than, say, *mise-tati* [shop-PL]. This contrast arises because buses move like humans but shops do not.

In fact, *-tati* can attach to inanimate nouns quite naturally in some cases, though the frequency is far lower than cases in which *-tati* attaches to animate nouns. When attached to inanimate nouns, *-tati* often expresses the speaker’s empathy or affection toward the entities denoted by the noun.<sup>2</sup> This property of *-tati* has been quite stable for decades and is not a result of a recent development, as evidenced by the following examples, which were written in different times.

- (8) Don Kihote-ga kare-no taisetuna kisimonogatari-no *hon-tati-o*  
 Don Quixote-NOM he-GEN important chivalric.novel-LINK book-PL-ACC  
 yak-are-ta toki, Don Kihote-wa nan-to utatte Sancho-o  
 burn-PASS-PST time Don Quixote-TOP what-that sing Sancho-ACC  
 nagusameteyat-ta-roo?  
 comfort-PST-may  
 ‘When Don Quixote got his important chivalric novel books burnt, what did he sing to  
 comfort Sancho?’ (Yuriko Miyamoto (1941) *Hondana*)

<sup>1</sup>A more rigorous formulation would include a presupposition that the NP denotation is atomistic, as *-tati* does not attach to mass nouns to convey abundance, unlike plural markers in some languages (e.g. Greek). I will return to this point in chapter 6.

<sup>2</sup>*-Domo* attaches to inanimate nouns too, and often conveys that the speaker despises or is annoyed by the entities denoted by the noun.

- (9) Hontooni kawaii *keeki-tati* bakari de dore-o taberu-ka mayottesimai-masu!  
 really cute cake-PL all and which-ACC eat-Q puzzled-POL  
 ‘Their cakes are all so cute, and I am not sure which one to eat!’

(*Machi-pita*, <http://www.machi-pita.com/station/detail.php?StationInfoID=87&reportItemID=25840>, accessed 12/15/2012)

The special connotation associated with *-tati* attached to inanimate nouns supports the claim that the general restriction to animate nouns is not asserted.

The politeness-related meanings of plural markers in Japanese exhibit the same properties of conventional implicatures that noun specifications of classifiers have (cf. chapter 4, section 4.2.1). First, different levels of politeness are tied to particular plural markers. Second, the politeness-related meaning is scopeless, projecting from negation. The meaning of respect implied by the plural marker *-gata* is not lost under negation in (10). (10) denies the teachers’ presence, but not the speaker’s respect toward them.

- (10) Sono toki sensei-gata-ga soko-ni i-ta toiu wakedewanai.  
 that time teacher-PL-NOM there-LOC be-PST that it.is.not  
 ‘It is not that the teachers were there at that time.’

Third, the politeness-related meanings of plural markers cannot be bound. The meaning of respect persists with the sentence with *-gata* in (11), regardless of the truth of the antecedent clause.

- (11) Mosi (karera-ga) sonkeis-are-teiru naraba, sensei-gata-wa tadasii koto-o  
 if they-NOM respect-PASS-ASP if teacher-PL-TOP right thing-ACC  
 it-ta hazuda.  
 say-PST should  
 ‘If they are respected, the teachers should have said a right thing.’

Lastly, the politeness-related meanings of plural markers pass unmodified through presupposition plugs such as propositional attitude verbs. In (12), the meaning of respect implied by the plural marker *-gata* is ascribed not to the matrix subject ‘parents these days’ but to the speaker.

- (12) Saikin-no oya-tati-wa sensei-gata-ga sinyoo-deki-nai to omot-teiru.  
 these.days-LINK parent-PL-TOP teacher-PL-NOM trust-can-NEG that feel-ASP  
 ‘Parents these days feel that (the) teachers cannot be trusted.’

To summarize, plural markers in Japanese have two meaning components, i.e. the asserted meaning, which restricts the domain to pluralities, and the conventional implicatures which provide information about the nouns with which they combine. As this structure is exactly the same as that of classifiers, it can be safely concluded that plural markers are plural counterparts of classifiers. One clear difference between classifiers and plural markers is that the number of plural markers available in a classifier language is far smaller than that of classifiers available in that language. Given this difference, some classifier languages including Malay having only one plural marker is not surprising.

### 5.3 Plural markers in classifier languages and definiteness

#### 5.3.1 Plural markers $\neq$ definite markers

Plural markers in classifier languages are often claimed to encode definiteness as well. I argue that such simple identification of plural markers with definite markers does not hold true. The most clear and convincing evidence comes from Persian. NPs with the plural marker *-hâ* in Persian are interpreted as definite as in (13).

- (13) Bačče-hâ-ye bâhuš unjâ bâzi mi-kard-an.<sup>3</sup>  
 child-PL-MOD clever there play DUR-do.PST-3PL  
 ‘The clever children were playing there.’

(Ghomeshi 2003:60)

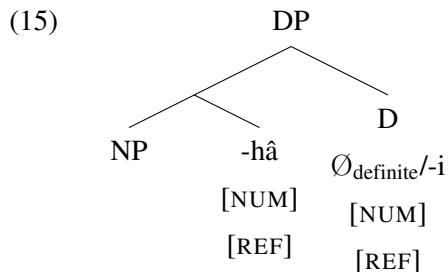
However, as Ghomeshi (2003) points out, the definite interpretation of plurals does not mean that the plural marker also encodes definiteness. She argues that the definite interpretation has to do with the syntactic licensing condition of the plural marker: the plural marker must be licensed either by the indefinite marker *-i* or a null definite marker. Plurals are interpreted as definite only if an overt marker of indefiniteness is absent as in (13). The plural marker *-hâ* can co-occur with the indefinite marker *-i* as in (14). In this case, plurals are interpreted as indefinite.

<sup>3</sup>I gloss the so-called ‘ezafe’ as MOD (modifier), as its basic function is to link a modifier and the noun modified by it. Ghomeshi’s examples represent the colloquial variety (p.c. Satoko Yoshie). However, the register difference does not affect my argument as it is mainly concerned with verbal forms.

- (14) Bačče-hâ-ye bâhuš-i unjâ bâzi mi-kard-an.  
 child-PL-MOD clever-INDF there play DUR-do.PST-3PL  
 ‘Clever children were playing there.’

(Ghameshi 2003:59)

In essence, the source of definiteness is not the plural marker but the definite determiner. Since the definite determiner is phonologically null, the plural marker appears to trigger definiteness. In a framework in which syntactic licensing is achieved by means of agreement such as Chomsky (2000, 2001), the relevant agreement takes place between the plural marker and the definite/indefinite marker with regard to number and referentiality, as shown in (15).<sup>4</sup>



This analysis is quite plausible, given that definite determiners in many languages inflect in number.

The connection between plural markers and definiteness is also observed in other classifier languages. However, in most languages, the connection is less rigid than in Persian. Forms with plural morphology are often interpreted as definite, but they can also have an indefinite interpretation. The most famous case is the suffix *-men* in Mandarin. It has been noted that NPs suffixed by *-men* are obligatorily definite and cannot be interpreted as indefinite (Iljic 1994; Yang 1998; Li 1999; Rullmann and You 2006; Huang et al. 2009). However, this generalization is in fact too strong. It is true that *-men* usually does not occur with indefinites, but indefinites with *-men* are not totally disallowed, as demonstrated by Yorifuji (1976) and Lan (2010).<sup>5</sup> One

<sup>4</sup>Ghameshi (2003) claims that the definite and indefinite markers occupy different heads, i.e. D and Q. I conflate these two categories into D, assuming that the properties that she attempted to derive from the syntactic categorical difference are captured by the lexical semantics of *-i*. I will clarify the agreement mechanism involved here in chapter 6. I will rename the referentiality feature ‘argument feature’ there.

<sup>5</sup>It is unfortunate that Yorifuji (1976) has been misrepresented in the literature on Mandarin. For example, Iljic (1994:94) writes as follows: “Next, as Rygaloff and Yorifuji notably point out, *N-men* always refers to the definite.” Yorifuji’s work should be reevaluated as one of the earliest attempts to challenge the traditional view that NPs with

of Yorifuji's examples is given in (16). The sentence was taken from a children's story collection *You Miao Ji*. It occurs as the first sentence of a story, and hence no referent has yet been established, ruling out the possibility to interpret *haizi-men* 'children' as definite.

- (16) Xiao he liushui hua hua xiang, xiang *haizi-men* zai gechang.  
 small river flowing.water rush rush sound like child-MEN PROG sing  
 'Rushing creek water sounds like children singing a song.' (Yorifuji 1976:86)

(17) is another example that shows that forms with *-men* can be indefinite. According to my consultants, *xuesheng-men* 'students' is ambiguous between definite and indefinite interpretations, with the indefinite interpretation more salient.

- (17) Xianzai yanzou de yuedui shi you *xuesheng-men* zucheng de.  
 now play MOD band is of student-MEN compose PART  
 'The band now playing is made up of (the) students.'

The plural marker *-tati* in Japanese is semantically similar to *-men* in Mandarin in many respects, including the presence of an associative use (cf. chapter 3; see Lan (2010:chapter 2) for a comparison between *-tati* and *-men*). NPs suffixed by *-tati* are usually interpreted as definite. However, Kurafuji (2004) as well as Nakanishi and Tomioka (2004) point out that there are cases where NPs suffixed by *-tati* are not interpreted as definite, though their conclusions about whether to analyze the morpheme as encoding definiteness are different. Furthermore, similar situations have been reported about plural NPs in Cantonese (Au-Yeung 2007; Matthews and Yip 2011) and Korean (Kwon and Zribi-Hertz 2004). Plural NPs in these language are often interpreted as definite, but they do in fact allow an indefinite interpretation. This is also the case in Malay, as shown in (18).

- (18) Mana-lah tahu kalau-kalau ada *anak-anak dara* kampung yang telah membuatkan  
 where-PART know maybe be virgin.girl.PL village REL PRF make  
 hati Amir cair!  
 heart Amir melt  
 'Who knows? Maybe there are some village girls who have made Amir's heart melt!'

(*Penulisan2u*, <http://www.penulisan2u.my/2011/04/novel-kemarau-kasih-9/>, accessed 8/5/2012)

---

*-men* are always definite.

These facts indicate that the connection between plurality and definiteness is not a necessary one in these classifier languages either. I thus claim that Ghomeshi's (2003) analysis for Persian also applies to these languages: the definite interpretation is not due to the plural marker itself, but to a null definite marker, which licenses the plural marker (by means of agreement). In her analysis, a null definite marker is posited when an overt indefinite marker is absent, which suggests that the definite reading is the default interpretation for plurals.

An important difference exists between Persian and these languages, however. Unlike Persian, these languages lack a dedicated indefinite marker, with indefiniteness conveyed either covertly or by words with other functions such as numeral 'one' (unaccompanied by a demonstrative). This difference makes the plural morphology in these languages appear to be linked to definiteness more directly than that in Persian. However, the definite interpretation is merely the default option and is not an unbreakable rule. An indefinite interpretation is possible insofar as the grammars of individual languages allow it. Consequently, crosslinguistic variations are expected as to how often the default definite interpretation option is chosen. The stricter the restriction against indefinites, the less likely it is for plurals to receive an indefinite interpretation. Mandarin is stricter than Japanese and Malay, as it generally prohibits indefinite NPs, including those with the plural marker *-men*, in the preverbal subject position (e.g. Huang et al. 2009:chapter 8). Lan (2010) points out this positional definiteness requirement as one of the factors that make NPs with *-men* prone to be interpreted as definite. Plurals in the subject position can be interpreted as indefinite more easily in Japanese and Malay, as illustrated by examples (3)–(4) and (18) above.

Synthesizing insights from previous studies, Huang et al. (2009:322) propose the generalization that an indefinite NP in Mandarin can occur in a sentence expressing athetic judgement in the sense of Kuroda (1992). Athetic judgment is "a direct response to the perceptual cognition of an actual situation, a perceptual intake of information about an actual situation" (Kuroda 1992:22). While athetic judgment is "a simple recognition of the existence of an actual situation", a categorical judgment is "a double judgment insofar as it involves the cognitive act of apprehending something as substance and attributing to it a certain property perceived in a situation" (Kuroda 1992:23). For example, adding the existential verb *you* 'to have' makes an indefinite subject acceptable as shown by the contrast between (19a) and (19b), because the verb changes the sentence from a simple description of someone else's experience, which cannot be directly perceived, to a description of a directly perceived situation.

- (19) a. ??Yi ge ren kan-guo ta de dianying.  
 one CLF person see-ASP he GEN movie  
 'A person has seen his movie.'
- b. You yi ge ren kan-guo ta de dianying.  
 exist one CLF person see-ASP he GEN movie  
 'A person has seen his movie.'

(Huang et al. 2009:321)

In fact, plurals with *-men* can be indefinite in the same context. *Haizi-men* 'children' in (20a) without *you* receives the default definite interpretation and refers to some known group of children (e.g. those present at the site, currently talked about). (20b) with the same form, on the other hand, can be used when one suddenly hears a distant sound of children singing, without knowing who the singing children are.

- (20) a. Haizi-men zai chang ge.  
 child-PL PROG sing song  
 'The children are singing./\*Some children are singing.'
- b. You haizi-men zai chang ge.  
 exist child-PL PROG sing song  
 'Some children are singing./\*The children are singing'

The thetic judgment account explains the indefiniteness of *haizi-men* in (16) above too.

### 5.3.2 Reassessment of the evidence for the definiteness of plurals

Now that the connection between the plural marker and definiteness turned out not to be a necessary one, the facts that have been presented as evidence for the connection must be reassessed. I will review three main supporting facts from Mandarin, as it exhibits the connection most clearly. In all cases, the same contrast is also observed between the corresponding indefinite and definite expressions in English. This coincidence, I think, has led many researchers to the conclusion that plural markers in classifier languages subsume the role played by the definite article *the* in English, i.e. definiteness marking.

First, plurals cannot occur in existential/unaccusative sentences.

- (21) a. You ren lai-le.  
 have person come-ASP  
 ‘There is somebody coming.’  
 b. \*You ren-men lai-le.  
 have person-PL come-ASP  
 ‘There are people coming.’

(Yang 1998:281)

Definite noun phrases cannot occur in this context in English either.

- (22) a. There are people coming.  
 b. \*There are the people coming.

These facts are instances of the definiteness effect.

Second, plurals cannot be used as predicates.

- (23) Tamen shi xuesheng(\*-men).  
 they be student-PL  
 ‘They are students.’

(Yang 1998:280)

The same restriction is found with definite noun phrases with *the* in English.

- (24) They are (\*the) students.

Third, while bare NPs are scopally ambiguous with respect to intensional verbs as in (25a), plurals obligatorily take wide scope as in (25b).

- (25) a. Na jia yiyuan zhengzai zhao hushi.  
 that CLF hospital currently look.for nurse  
 (i)  $\sqrt{\text{look.for}} > \text{nurse(s)}$   
 ‘That hospital is looking for a nurse/nurses (to hire).’  
 (ii)  $? \text{nurse(s)} > \text{look.for}$   
 ‘There is a nurse/are nurses that hospital is looking for.’

- b. Na jia yiyuan zhengzai zhao hushi-men.  
 that CLF hospital currently look.for nurse-PL  
 (i) \*look.for > nurses  
 ‘That hospital is looking for nurses (to hire).’  
 (ii) √nurses > look.for  
 ‘There is a group of nurses that hospital is looking for.’

(Lan 2010:19)<sup>6</sup>

Again, NPs with *-men* pattern with definite noun phrases with *the* in English.

- (26) That hospital is looking for the nurses.  
 (\*look.for > nurses, nurses > look.for)

The second and third of these three facts also hold in other classifier languages. Malay sentence (28c) is not perfectly acceptable even in the wide-scope reading because plurals without modifiers or focus normally sound incomplete in the direct verb object position in Malay.<sup>7</sup>

- (27) Japanese
- a. Karera-wa gakusei(\*-tati) da.  
 they-TOP student-PL be  
 ‘They are students.’
- b. Sono byooin-wa kangohu-o sagasi-teiru.  
 that hospital-TOP nurse-ACC look.for-PROG  
 (i) √look.for > nurse(s)  
 ‘That hospital is looking for a nurse/nurses (to hire).’  
 (ii) ??nurse(s) > look.for  
 ‘There is a nurse/are nurses that hospital is looking for.’
- c. Sono byooin-wa kangohu-tati-o sagasi-teiru.  
 that hospital-TOP nurse-PL-ACC look.for-PROG  
 (i) \*?!look.for > nurses  
 ‘That hospital is looking for nurses (to hire).’

<sup>6</sup>The judgments shown here are not Lan’s but my consultants’. Lan reports a bigger contrast between (25a-i) and (25a-ii), and a smaller contrast between (25b-i) and (25b-ii) ((25a-i): √; (25a-ii): ??; (25b-i): \*?; (25b-ii): √).

<sup>7</sup>Kester and Schmitt (2007) report a similar fact in Papiamentu. I will discuss the effects of modifiers and focus in section 5.4 below.

(ii)  $\sqrt{\text{nurses}} > \text{look.for}$

‘There is a group of nurses that hospital is looking for.’

(Nakanishi and Tomioka 2004:115)

(28) Malay

a. Mereka pelajar(\*-pelajar).

they student-PL  
‘They are students.’

b. Hospital itu sedang mencari jururawat.

hospital that PROG look.for nurse

(i)  $\sqrt{\text{look.for}} > \text{nurse(s)}$

‘That hospital is looking for a nurse/nurses (to hire).’

(ii)  $??\text{nurse(s)} > \text{look.for}$

‘There is a nurse/are nurses that hospital is looking for.’

c. (??)Hospital itu sedang mencari jururawat-jururawat.

hospital that PROG look.for nurse.PL

(i)  $*?\text{look.for} > \text{nurses}$

‘That hospital is looking for nurses (to hire).’

(ii)  $?\text{nurses} > \text{look.for}$

‘There is a group of nurses that hospital is looking for.’

As for the first fact, plurals in Japanese and Malay can occur in a context in which the definiteness effect is normally observed. According to Kishimoto (2000), the nominative phrases of possessive verbs display definiteness effects in Japanese.

(29) a. \*Taroo-ni sono/hotondo-no/subete-no/kare-no kyoodai-ga aru.

Taro-DAT that/most-LINK/all-LINK/his brother-NOM have  
‘Taro has that/most/all/his brother(s).’

b. Taroo-ni takusan-no/futa-ri-no/nan-nin-ka-no kyoodai-ga aru.

Taro-DAT many-LINK/two-CLF-LINK/some-CLF-LINK brother-NOM have  
‘Taro has many/two/some brothers.’

(Kishimoto 2000:69)

A plural form can occur as the nominative phrase in this construction.

- (30) Taro-ni (tayorininaru) kyoodai-tati-ga iru/aru (koto-wa sir-are-tei-nai).  
 Taro-DAT (reliable) brother-PL-NOM have that-TOP know-PASS-ASP-NEG  
 ‘(It is not known that) Taro has (reliable) brothers.’

Definite noun phrases are disallowed after existential/unaccusative verbs in affirmative sentences without any context in Malay (Nomoto 2009).

- (31) Ada dua naskhah majalah (\*itu) di atas meja.  
 be two CLF magazine that at on table  
 ‘There are (\*the) two magazines on the table.’

The definiteness restriction in Malay has to do with the meaning of the noun phrase rather than the syntax of it, because sentences with a definite noun phrase are acceptable when they are used in appropriate contexts, as in (32).

- (32) [A conversation over the telephone.]  
 A: Ada apa di bilik awak?  
 be what at room you  
 ‘What is in your room?’  
 B: Ingat tak, kita beli majalah di kedai buku. Ada dua naskhah majalah  
 remember not we buy magazine at store book be two CLF magazine  
 itu di atas meja.  
 that at on table  
 ‘Do you remember we bought magazines at the bookstore? The two magazines  
 are on the table.’

A plural noun phrase can occur in this construction without any contextual support.

- (33) Ada burung-burung di atas pokok.  
 be bird.PL at on tree  
 ‘There are birds on the tree.’

Given that the connection between the plural marker and definiteness is stronger in the subject position in Mandarin than in Japanese and Malay (cf. section 5.3.1), the difference between Mandarin on one hand and Japanese and Malay on the other in the context where the definiteness effect is normally observed comes as no surprise. However, three languages behaving similarly for the second and third facts requires an explanation.

I argue that the lack of a predicative nominal use and obligatory wide scope associated with plurals should be ascribed to referentiality, but not to definiteness. That is, plurals are referential. Many referential noun phrases are definite, but referential noun phrases can be indefinite as well. The following Persian data clearly shows that definiteness is irrelevant to whether an NP can be used as a predicate. As seen above, a noun phrase with the plural marker *-hâ* in Persian is interpreted as definite unless an overt indefinite marker is present. One is thus inclined to ascribe the lack of a predicative nominal use of *dânešju-hâ* ‘students’ in (34b) and the obligatory wide scope of *parastâr-hâ* ‘nurses’ in (35b) to definiteness.<sup>8</sup> However, such an analysis fails to account for (34c) and (35c), where the noun phrases are indefinite but still cannot be used as predicates and take narrow scope. In order for a noun phrase to serve as a predicate nominal or take narrow scope, it cannot take the indefinite marker *-i*, as in (34a) and (35a).

- (34) a. Ânhâ dânešju-and.  
 they student-be.3PL  
 ‘They are students.’
- b. \*Ânhâ dânešju-hâ-and.  
 they student-PL-be.3PL
- c. \*Ânhâ dânešju-hâ-i-and.<sup>9</sup>  
 they student-PL-INDF-be.3PL
- (35) a. Ân bimârestân donbâl-e parastâr mi-gard-ad.  
 that hospital sequence nurse-MOD DUR-turn-3SG  
 (i) √look.for > nurse(s)  
 ‘That hospital is looking for a nurse/nurses (to hire).’  
 (ii) √nurse(s) > look.for  
 ‘There is a nurse/are nurses that hospital is looking for.’
- b. Ân bimârestân donbâl-e parastâr-hâ mi-gard-ad.  
 that hospital sequence-NOM nurse-PL DUR-turn-3SG  
 (i) \*look.for > nurses  
 ‘That hospital is looking for nurses (to hire).’

<sup>8</sup>According to Ghomeshi (2003), sentence (34b) is difficult, but not impossible, in an equative reading, where the plural noun phrase is construed referentially.

<sup>9</sup>A glide or a glottal stop may be inserted before *-i* to prevent hiatus, though not shown here.

- (ii)  $\sqrt{\text{nurses}} > \text{look.for}$   
 ‘There are nurses that hospital is looking for.’
- c.  $\hat{\text{A}}\text{n}$  bimârestân donbâl-e      parastâr-hâ-i      mi-gard-ad.  
 that hospital      sequence-MOD nurse-PL-INDF DUR-turn-3SG  
 (i) \*look.for > nurses  
 ‘That hospital is looking for nurses (to hire).’  
 (ii)  $\sqrt{\text{nurses}} > \text{look.for}$   
 ‘There are nurses that hospital is looking for.’

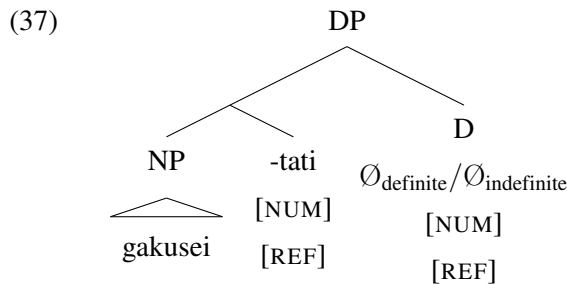
Ghameshi (2003:63) claims that “the presence of *-i* on a noun entails that the noun must be referential” and this “accounts for the fact that the resulting noun is often construed as ‘specific’.” In other words, unlike the indefinite article *a* in English, which is ambiguous between referential and non-referential, quantificational meanings (Fodor and Sag 1982), the indefinite marker *-i* in Persian unambiguously encodes referential indefiniteness.<sup>10</sup> The nominal forms and interpretations for plural referents can be summarized as follows:

- (36) a. NP: non-referential indefinite  
 b. NP-*hâ-i*: referential indefinite  
 c. NP-*hâ*: definite

Inability of a noun phrase to serve as a predicate only means that that noun phrase is referential but not that it is referential *and* definite. Likewise, the wide scope reading of plurals does not necessarily mean that plurals are definite. What it actually shows is that plurals are referential. Referential/specific indefinites are known to take the widest scope because they are indexicals, which only depend on the context of utterance and do not interact with sentence-internal operators, and resemble definites, especially demonstratives (e.g. Fodor and Sag 1982; Kratzer 1998).

I assume that other classifier languages have phonologically null equivalents of *-i* in Persian. Plural markers must be licensed either by this referential indefinite marker or by the definite marker. For example, *gakusei-tati* [student-PL] ‘students’ in Japanese have the structure shown in (37).

<sup>10</sup>But see Windfuhr (1979:34–38), who claims that *-i* in Persian is ambiguous in the same way as *a* in English is.



Given that definites are normally referential, whichever determiner licenses the plural marker, that is, whether definite or indefinite, plurals will be referential and never be non-referential. This is why they cannot be used as predicates and take wide scope with respect to intensional verbs.

The relevance of referentiality to plurals in classifier languages is not something totally unnoticed in the literature. However, referentiality has been normally thought to be part of the meaning of the plural marker itself and never been analyzed as a result of the interaction between determiners and plural markers. Hosoi (2005) correctly points out that the phenomena involving *-tati* in Japanese that are usually linked to definiteness are in fact related to “specificity,” by which he means referential indefiniteness. His analysis derives the specificity effects from *-tati* itself rather than the (in)definite marker licensing it. As a consequence, his semantics of *-tati* is unnecessarily complex. Moreover, his analysis does not extend to Persian naturally, without ignoring the distributional fact concerning the indefinite marker *-i*. Kurafuji (2004) suggests a possibility that *-tati*, which he analyzes as a definite plural marker, has the indexical use like *the* in English. Kaneko (2007) also points out the indexical nature of plurals with *-tati*. His evidence is two similarities between *-tati* and demonstratives in English.<sup>11</sup> First, both *-tati* and *that* can attach to an indefinite only if the indefinite is specific.

- (38) a. John-wa dareka(\*-tati)-o mi-ta.  
 John-TOP someone-PL-ACC see-PST  
 ‘John saw someone.’

(Kawasaki 1989:134)

- b. Tokuteino dareka-tati-e-no messeezi-o konna tokoro-ni  
 specific someone-PL-to-LINK message-ACC such place-LOC

<sup>11</sup> Another fact that he presents as evidence, namely difficulty in occurring as an appositive, is irrelevant, given that *wareware simin-tati* [we citizen-PL] ‘we the citizens’ sounds quite natural. His example, i.e. *watasi-tati simin-tati* [I-PL citizen-PL] ‘we the citizens’, sounds a little awkward arguably due to the repetition of *-tati*.

kaki-kon-desimau.  
 write-down-accomplish  
 ‘I have written down a message towards someone-TATI specific in such a place.’  
 (Kaneko 2007)

(39) a gift for that someone special

(Kaneko 2007)

The second similarity has to do with modification, which I will discuss at length in the next section. The following description about *-men* in Mandarin by Iljic (1994:94–95) is also suggestive of the indexical nature of plurals in classifier languages: NPs with *-men* refer to “a situationally anchored and defined group” and must be used in allocution. He compares NPs with *-men* to personal pronouns: “*Péngyou-men!* ‘(My dear) friends!’ functions as a qualified *you*.”

Besides referential determiners, quantifiers also license plural markers, as in (40)–(42), though these sentences sound a little awkward compared to the corresponding sentences without the plural marker. A non-referential interpretation is available in this case.

(40) Japanese

- a. Oozei/Hotondo-no gakusei-tati-ga siken-ni oti-ta.  
 many/most-LINK student-PL-NOM exam-DAT fail-PST  
 ‘Many/Most students failed the exam.’
- b. Doredake-no gakusei-tati-ga siken-ni oti-ta-no?  
 how.many-LINK student-PL-NOM exam-DAT fail-PST-Q  
 ‘How many students failed the exam?’

(41) Malay

?Ramai/Kebanyakan pelajar-pelajar gagal dalam peperiksaan.  
 many/most student.PL fail in exam  
 ‘Many/Most students failed the exam.’

(42) Mandarin (Lan 2010:63)

Rujin, xuduo fu mu-men hui miandui zheyang de wenti.<sup>12</sup>  
 today many father mother-PL will confront such MOD problem  
 ‘Today, many parents will confront such problems.’

<sup>12</sup>The is one of the 71 instances of *xuduo NP-men* that Lan found in his corpus. Though his consultants judge the sentence as unnatural, the number of attested instances is obviously too large to reject the pattern as ungrammatical.

To summarize, plural markers in classifier languages are licensed by referential determiners or quantifiers.

## 5.4 Licensing by modification

### 5.4.1 The facts

The plural forms in the examples above are unmodified. Surprisingly, plurals with modifiers behave differently. They can serve as predicate nominals and take narrow scope with respect to intensional verbs. Compare (27)–(28) above and (43)–(44) below.

(43) Japanese

- a. Karera-wa yuusyuuuuna gakusei-tati da.  
they-TOP excellent student-PL be  
'They are excellent students.'
- b. Sono byooiin-wa kodomo-no atukai-ni nareta kangohu-tati-o  
that hospital-TOP child-GEN handling-DAT be.used nurse-PL-ACC  
sagasi-teiru.  
look.for-PROG  
√look.for > nurse(s)  
'That hospital is looking for nurses (to hire) who are used to dealing with kids.'

(Nakanishi and Tomioka 2004:136)

(44) Malay

- a. Mereka pelajar-pelajar cemerlang.  
they student.PL excellent  
'They are excellent students.'
- b. Hospital itu sedang mencari jururawat-jururawat yang biasa menjaga  
hospital that PROG look.for nurse.PL REL used.to take.care.of  
kanak-kanak.  
child  
√look.for > nurse(s)  
'That hospital is looking for nurses (to hire) who are used to dealing with children.'

---

The unnaturalness is thought to arise due to the general restriction against indefinite subjects in Mandarin.

Though the judgments are not always clear, the presence of modifiers improves the acceptability in Mandarin as well.<sup>13</sup> Compare (23) and (25) above and (45) below.

- (45) a. Tamen shi {??congming/ ?gang lai riben} de xuesheng-men.  
 they be smart just come Japan MOD student-PL  
 ‘They are {smart students/students who just came to Japan}.’
- b. Na jia yiyuan zhengzai zhao shuxi zhaogu xiaohai de hushi-men.  
 that CLF hospital currently look.for familiar take.care child MOD nurse-PL  
 √look.for > nurse(s)  
 ‘That hospital is looking for nurses (to hire) who are used to dealing with children.’

The disappearance of definite-like behaviors of plurals with *-tati* in the presence of modifiers has been also reported by Nakanishi and Tomioka (2004). They attempt to understand the phenomenon as a result of modifiers minimizing exceptions encoded in the semantics of *-tati* (in their analysis, but not mine). Such a line of analysis is not applicable to other languages including Malay and Mandarin, as plurals of common nouns in these languages are not associative unlike *-tati* plurals in Japanese, that is, exceptions are not allowed at all.

The phenomenon above can be generalized as follows in terms of referentiality:

- (46) Semantic licensing by modification: referentiality  
 Modifiers are able to make otherwise referential expressions less referential.

The following two facts justify the appeal to referentiality rather than plurality or other notions related to plurality in the generalization in (46). First, while unmodified predicate nominals in Persian cannot co-occur with the referential indefinite marker *-i* (cf. (34)), modified ones must occur with *-i* (Ghomeshi 2003).

- (47) Unmodified
- a. Ânhâ dânešju-and.  
 they student-be.3PL  
 ‘They are students.’

<sup>13</sup>I speculate that the unclear judgments are related to the modification construction with *de*, because modifiers in Japanese and Malay do not contain anything comparable to *de*.

- b. \*Ânhâ dâneşju-hâ-i-and.  
they student-PL-INDF-be.3PL
- (48) Modified
- a. \*Ânhâ dâneşju-ye xub-and.  
they student-MOD good-be.3PL
- b. Ânhâ dâneşju-hâ-ye xub-i-and.  
they student-PL-MOD good-INDF-be.3PL  
'They are good students.'

This contrast is orthogonal to plurality, as shown by the examples in (49)–(50), which do not contain the plural marker *-hâ*.

- (49) Unmodified
- a. Bižan dâneşju-e.  
Bijan student-be.3SG  
'Bijan is a student.'
- b. \*Bižan dâneşju-i-e.  
Bijan student-INDF-be.3SG

(Ghomeshi 2003:61)

- (50) Modified
- a. \*Bižan dâneşju-ye xub-e.  
Bijan student-MOD good-be.3SG
- b. Bižan dâneşju-ye xub-i-e.  
Bijan student-MOD good-INDF-be.3SG  
'Bijan is a good student.'

(Ghomeshi 2003:61)

According to Ghomeshi (2003:61), modification requires an indefinite marker also in names of professions and occupations in French, Spanish, Romanian and European Portuguese, and in bare predicate nominals in Dutch and Hawaiian English Creole.

Secondly, similar effects are also observed with the distal demonstratives *that* and *those* in English. Being indexical, demonstratives are unscoped and virtually take the widest scope. However, when a demonstrative phrase contains a restrictive postnominal modifier, it can serve as a predicate nominal as in (51) and take narrow scope under an intensional operator as in

(52) (Wolter 2007; Kaneko 2007).<sup>14</sup> The sentences in (52) do not have a sensible interpretation unless the demonstrative phrase scopes under the modal.

(51) They are those acts [which keep one's reputation bright without reference to a specific previous indebtedness to another person].

(R. Benedict, *The Chrysanthemum and the Sword*, cited in Kaneko (2007))

(52) a. Those employees [who were the least productive] might have been different people.

b. That person [at the top of the list] could have been someone else.

(Wolter 2007:612, 617)

Non-restrictive relative clauses do not make an opaque reading of a demonstrative phrase possible.

(53) #That presidential candidate, who won the election, might have been someone else.

(Wolter 2007:617)

The distribution of the indefinite marker *-i* in Persian confirms the parallelism between referential determiners and demonstratives: demonstrative phrases can have a narrow scope interpretation where *-i* occurs unexpectedly. According to Ghomeshi (2003:65), *-i* occurs in a noun phrase modified by a restrictive relative clause, but not if the relative clause is non-restrictive, as in (54).

(54) a. Ahmad-i ke diruz âmad, injâ-st.  
 Ahmad-INDF that yesterday come.PST.3SG here-3SG  
 'The Ahmad who came yesterday is here.' (as opposed to the one who came today)

---

<sup>14</sup>With prenominal modifiers, demonstrative phrases remain scopally inert, as shown in (i) (Wolter 2007).

(i) #Those [least productive] employees might have been different people.

(Wolter 2007:612)

- b. Ahmad, ke diruz âmad, injâ-st.  
 Ahmad that yesterday come.PST.3SG here-3SG  
 ‘Ahmad, who came yesterday, is here.’

(Thackston 1983:82)

Given the generalization in (46) and my analysis that plural markers are licensed by a referential determiner (or quantifiers), the fact that plurals can serve as predicate nominals and take narrow scope with respect to intensional verbs when accompanied by modifiers does not affect the semantics of plural markers. It is not the case that modifiers license plural markers directly. Instead, the licensing of plural markers takes place indirectly, by way of determiners. That is, modifiers license an opaque reading for referential determiners, which in turn license plural markers syntactically. The relevant determiner is overt in Persian (i.e. *-i*) but covert in languages such as Japanese, Malay and Mandarin.

Two questions arise at this point. First, what is the mechanism underlying behind the generalization in (46)? I will discuss it at length in section 5.4.2 below. Second, the generalization in (46) seems to contradict what some authors have noted about the effect of modifiers on referentiality. For instance, Fodor and Sag (1982:361) state that “[a]ny relative clause modifying an indefinite adds to its descriptive content and thus tends to favor a referential understanding”. Givón (2001:449) also points out a positive correlation between the amount of modification and the strength of referentiality, as illustrated by the sentences in (55).

- (55) a. Did you see *a man* there?  
 b. Did you see *a tall man* there?  
 c. Did you see *a man wearing a blue shirt* there?  
 d. Did you see *a man there wearing a blue shirt and sitting on a red barrel and twirling a silver baton in his left hand*?

(Givón 2001:449)

I argue that this tendency does not contradict my generalization. This is because as Fodor and Sag (1982) state, the tendency is pragmatic in nature, whilst my generalization is governed by a purely semantic mechanism, as I will show shortly. Hence, the tendency does not override the semantic generalization in (46). It is only concerned with indefinites occurring in contexts in which the effect of (46) is not observed, e.g. in sentences with no other scope-taking element

such as those in (55).

### 5.4.2 Analysis

The phenomena that have been reported in the literature under the label of ‘licensing by modification’ to date can be classified into two types, i.e. syntactic and semantic. The syntactic type makes otherwise ungrammatical structures grammatical. Modifiers license the definite article *the* in English (Kayne 1994), the bare classifier construction in Thai (Jenks 2011) and *de* nominals in French (Mathieu 2012), for example. The semantic type makes otherwise impossible interpretations possible. The facts presented above and generalized in (46) are of this type. Putting aside the type of determiners involved (indefinite or demonstrative), these facts can be regarded as different manifestations of the same phenomenon and hence should be given a single analysis.

The analysis that I propose is a minimally modified version of the semantic analysis of demonstrative phrases proposed by Wolter (2007) but rejected in favor of a pragmatic analysis. The basic idea of Wolter’s analysis is that a modifier provides a spatio-temporal/situation variable independent of that of the head noun (Dayal 2004a),<sup>15</sup> and this situation variable mediates between the situation of a higher predicate and that of the demonstrative phrase, whose situation variable is otherwise completely free. She assumes that a demonstrative determiner bears a numerical index and it saturates the situation argument position of the nominal complement. This is captured by the lexical entry in (56).

- (56)  $\llbracket \text{that}_n \rrbracket : \lambda P_{\langle s, et \rangle} : P(s_n)$  is a singleton set and  $s_n$  is free.  
If defined, denotes  $\iota x.P(x)(s_n)$

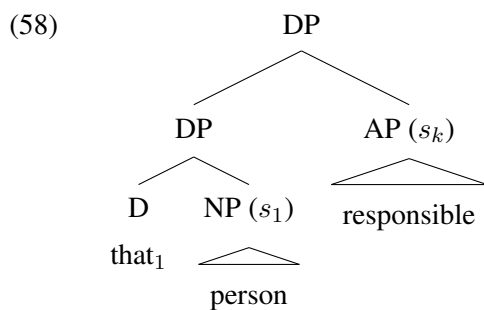
(Wolter 2007:620)

The situation variable of a prenominal modifier is dependent on the index of the demonstrative as in (57a) whereas that of a postnominal modifier is not as in (57b).

- (57) a.  $\llbracket \text{that}_1 \text{responsible person} \rrbracket = \iota x.\text{person}(x)(s_1) \wedge \text{responsible}(x)(s_1)$   
b.  $\llbracket \text{that}_1 \text{person responsible} \rrbracket = \iota x.\text{person}(x)(s_1) \wedge \text{responsible}(x)(s_k)$

<sup>15</sup>Dayal’s (2004a) original proposal is more restrictive. She contends that only postnominal modifiers (in English) introduce a new situation variable, based on Sadler and Arnold’s (1994) claim that postnominal modifiers are phrasal and have more structure than prenominal modifiers. My analysis is compatible with this more restricted view.

The situation variable of a postnominal modifier is not saturated by the index of the demonstrative because postnominal modifiers are attached to DP and hence fall outside of the c-command domain of the demonstrative, as shown in (58).



As  $s_1$  is free, with its value determined by the context, the entire noun phrase takes the widest scope in the case of (57a). By contrast,  $s_k$  in (57b), with a postnominal modifier, needs to be bound by some operator in the sentence. In other words, it interacts with operators above it. This explains why demonstrative phrases can have an opaque interpretation when they contain a postnominal modifier. In (59), the demonstrative phrase *that person responsible* can take narrow scope under *believes*. That is, the identity of the person at issue depends on John's belief world. This interpretation is impossible if the adjective *responsible* precedes the noun as in *that responsible person*.

(59) John believes that that person responsible left.<sup>16</sup>

(Wolter 2007:622)

In sentence (59), the situation variable introduced by *responsible* is bound by the modal operator associated with *believes*, as in (60).

(60)  $\lambda s[\forall s_2 \in \text{Dox}_j(s, s_2)[\text{left}(\iota x.\text{person}(x)(s_1) \wedge \text{responsible}(x)(s_2))(s_2)]]$

Consequently, the entire noun phrase refers to a unique person in some contextually salient situation ( $s_1$ ) who is responsible in John's belief world ( $s_2$ ). Demonstrative phrases with modifiers can be predicates as in (51) and other kind of generic sentences if the operator binding the situation variable introduced by the modifier is a generic operator instead of a modal operator.

<sup>16</sup>Not all speakers judge this sentence as grammatical.

Extending the analysis above to the cases involving the referential indefinite marker *-i* in Persian and plural markers licensed by the covert counterpart of *-i* in Japanese, Malay and Mandarin is straightforward. These cases differ from demonstrative phrases in English in the determiner involved. One has only to substitute the indefinite counterpart of the demonstrative determiner in (56), i.e. a choice function with a numerical index. The relevant determiner has the denotation in (61).

$$(61) \quad \llbracket -i_n / \emptyset_{\text{indefinite } n} \rrbracket = \lambda P_{\langle s, et \rangle} : s_n \text{ is free. CH}(P)(s_n)$$

The Persian and Japanese noun phrases meaning ‘good students’ in (62) and (63) have the structure in (64a), and the whole sentence receives the interpretation in (64b). I analyze the premodifier positioning of the constituent ‘student-*hâ*’ in Persian as a result of a movement of this phrase presumably to Spec,DP (specificity-driven movement). The modification marker known as ‘*ezafe*’, *ye* is thought to be a reflex of this movement.

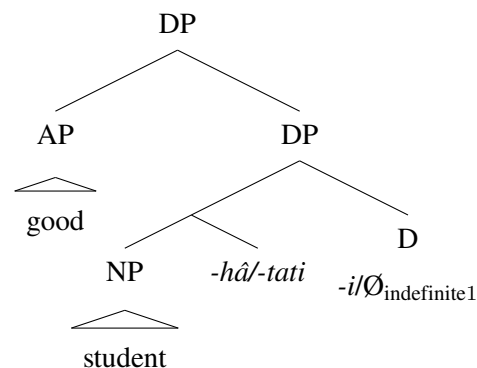
(62) Persian

Ânhâ dânešju-hâ-ye xub-i-and.  
 they student-PL-MOD good-INDF-be.3PL  
 ‘They are good students.’

(63) Japanese

Karera-wa yoi gakusei-tati da.  
 they-TOP good student-PL be  
 ‘They are good students.’

(64) a.



b.  $\lambda s[\text{GEN}_{s_2}[\text{3PL}(x) \wedge \text{C}(x, s_2)]] [x = \text{CH}(\text{students}(s_1) \wedge \text{good}(s_2))](s_2)]$

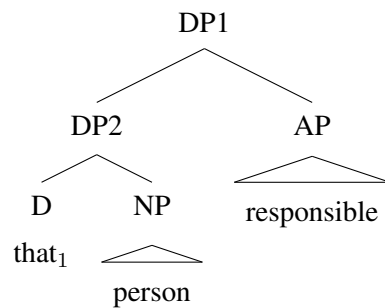
‘Every situation of the appropriate type containing the relevant third person

plural individual is a situation in which that individual is certain students in  $s_1$  picked out by a choice function who are good.’ ( $\approx$  ‘They are those students who are good.’)

In fact, plurals with modifiers, though they can serve as predicative nominals, sometimes leave a sense of referentiality. That is why the generalization about the semantic licensing by modification concerning referentiality in (46) above states that modifiers are able to make otherwise referential expressions “less referential,” but not “non-referential.” The interpretation shown in (64b) captures this fact. Making use of a choice function, it basically provides a referential interpretation. However, the choice function cannot make a completely independent choice due to  $s_2$  of *good*. It is for this reason that plurals with modifiers are neither totally referential nor totally non-referential.

What remains to be worked out is how to derive the meaning in (57b) compositionally from the structure in (58), reproduced below with a few minor notational changes.

(65)



- a.  $\llbracket \text{DP2} \rrbracket = \iota x.\text{person}(x)(s_1)$  (presupposition:  $s_1$  is free)
- b.  $\llbracket \text{AP} \rrbracket = \lambda s \lambda y.\text{responsible}(y)(s_j)$
- c.  $\text{Bach-Cooper}'(\llbracket \text{DP2} \rrbracket) = \lambda R_{(s,et)} \lambda s[\iota x.\text{person}(x)(s_1) \wedge R(x)(s_k)]$
- d.  $\llbracket \text{DP1} \rrbracket = \lambda s[\iota x.\text{person}(x)(s_1) \wedge \text{responsible}(x)(s_k)] (\approx (57b))$

Let the situation arguments of the NP *person* and the AP *responsible*,  $s_i$  and  $s_j$  respectively.  $s_i$  is saturated by the numerical index of the demonstrative to give (65a). The semantic type of DP2 is thus  $e$ . The crucial point at the derivation is when the modifier AP is attached to DP2. Some sort of type-shifting is required to enable the relevant composition. I propose a variant of Bach and Cooper’s (1978) semantics for high-adjoined relative clauses, *Bach-Cooper’*. Bach and Cooper propose a procedure that adds a property argument that is saturated by an adjoined relative

clause, as in (66b). This operation enables to obtain the meaning of NP2 in (66) compositionally.

- (66) [NP2 [NP1 every man ] [S' who loves Mary ]]
- a.  $\llbracket \text{NP1} \rrbracket = \lambda P. \forall x [\text{man}(x) \rightarrow P(x)]$
  - b.  $\text{Bach-Cooper}(\llbracket \text{NP1} \rrbracket) = \lambda R \lambda P. \forall x [[\text{man}(x) \wedge R(x)] \rightarrow P(x)]$
  - c.  $\llbracket \text{S}' \rrbracket = \lambda z. \text{love}(z, m)$
  - d.  $\llbracket \text{NP2} \rrbracket = \lambda P. \forall x [[\text{man}(x) \wedge \text{love}(x, m)] \rightarrow P(x)]$

Bach-Cooper in (66b) shifts a generalized quantifier of type  $\langle et, t \rangle$  to a function type of  $\langle et, \langle et, t \rangle \rangle$ . Bach-Cooper' in (65c) also adds a property argument that is saturated by an adjoined modifier, and is a type-shifter, specifically one from  $\langle s, e \rangle$  (or  $e$  if the situation variable is free as in (65a)) to  $\langle \langle s, et \rangle, \langle s, e \rangle \rangle$ . Bach-Cooper' differs from Bach-Cooper in that the former can handle situation variables.<sup>17</sup>

### 5.4.3 Licensing by contrast

Nakanishi and Tomioka (2004) point out that a clear sense of contrast facilitates an otherwise unavailable non-referential interpretation for plurals with *-tati* in Japanese. For example, while a generic interpretation is very difficult for (67a), it is readily available for (67b).

- (67) a. Itariazin-tati-wa yooki-da.  
 Italian-PL-TOP cheerful-COP  
 ???'Italians are cheerful.'
- b. Kodomo-tati-wa itumo otona-tati-no mane-o suru.  
 child-PL-TOP always adult-PL-GEN imitation-ACC do  
 √'Children always imitate adults.'

(Nakanishi and Tomioka 2004:136)

Like the case of licensing by modification, their exception-based analysis only works for Japanese, but not for other languages with similar phenomena. The effect of contrast is also observed in other languages. Contrast licenses an otherwise unavailable non-referential interpretation.

<sup>17</sup>Wolter (2007) proposes the extensional version of Bach-Cooper', i.e. a type-shifter from  $e$  to  $\langle et, e \rangle$ . The problem of her type-shifter is that it is indifferent to situation variables. As a consequence, the situation variable of the head noun is imposed on the modifier, and the important difference between pre- and postnominal modifiers as shown in (57), which she does acknowledge, is lost.

- (68) Malay
- a. Penyanyi-penyanyi ceria.<sup>18</sup>  
 singer.PL cheerful  
 ??‘Singers are cheerful.’
- b. Budak-budak selalu meniru orang dewasa.  
 kid.PL always imitate person adult  
 √‘Children always imitate adults.’
- (69) Mandarin
- a. Yidaliren-men hen kailang.  
 Italian-PL very cheerful  
 \*‘Italians are cheerful.’
- b. Haizi-men zongshi mofang daren-men.  
 child-PL always imitate adult-PL  
 √‘Children always imitate adults.’

My analysis for this licensing by contrast phenomenon is essentially the same as that for licensing by modification above. Like modifiers, contrast introduces a new situation variable that is independent of the situation variable of the head noun. This new situation variable interacts with operators within the sentence, giving rise to characteristics of non-referential noun phrases.

The question, of course, is how contrast introduces a situation variable. Contrast is a kind of focus. Hence, a contrastively used noun phrase involves the semantics of focus. I adopt Rooth’s (1992) theory of focus interpretation, which is accepted as the standard analysis of focus in formal semantics. Rooth proposes that focus introduces a free variable that can be anaphoric to

<sup>18</sup>This sentence describes a particular situation (‘The singers are being cheerful now’). Importantly, the sentence does not contain the overt copula *adalah*. Although the relevant interpretive contrast between bare noun phrases and plurals, i.e. only the latter is compatible with generics, is observed in sentences without *adalah*, it is not observed with sentences with *adalah*.

- (i) a. √Penyanyi/ ??Penyanyi-penyanyi ceria.  
 singer singer.PL cheerful
- b. √Penyanyi/ √Penyanyi-penyanyi adalah ceria.  
 singer singer.PL COP cheerful  
 ‘Singers are cheerful.’

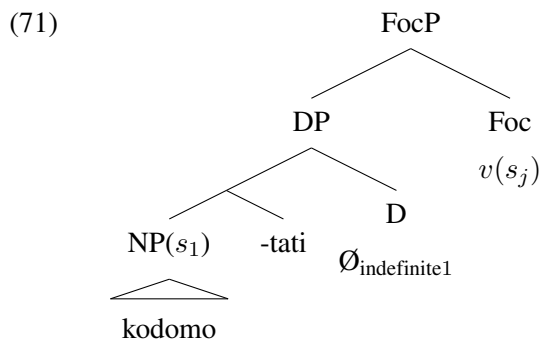
It is unclear why the presence of *adalah* makes an otherwise non-generic sentence generic. I thus leave it for future research.

a variety of pragmatic and semantic objects. The interpretation of this variable is pragmatically constrained by the set consisting of alternatives to the expression in focus. For example, in (70), where *American* is contrasted with *Canadian*, the focus operator adjoined to *American farmer* introduces a new free variable. The antecedent of this new variable is selected from the alternatives to *American farmer*, of which *Canadian farmer* is one.

(70) An [American]<sub>F</sub> farmer was talking to a [Canadian]<sub>F</sub> farmer . . .

(Rooth 1992:80)

I claim that the free variable introduced by focus comes with a situation variable independent of the noun phrase that is brought into focus. I assume that a focus phrase (FocP) is projected in syntax when a noun phrase receives focus. The focus head, Foc, triggers focus semantics and phonology. The noun phrase *kodomo-tati* [child-PL] ‘children’ in (67b) thus has the following structure:



Notice that the situation variable under Foc is not saturated by the indefinite determiner unlike that of the NP. Hence, it can be bound by a generic operator in generic sentences such as (67b). Consequently, the plural form *kodomo-tati* behaves like a non-referential noun phrase.

Finally, it is interesting to note that Chierchia (1998b) also points out the relevance of focus in discussing licensing of bare arguments in Italian. In his analysis, bare arguments have a null determiner. Hence, licensing of bare arguments amounts to licensing of the null determiner. Bare arguments are possible in a lexically governed position, but not elsewhere. The object *patete* ‘potatoes’ in (72a) is governed by a lexical verb (i.e. *mangiato* ‘to eat’) whereas the subject *studenti* ‘students’ in (72b–c) is not.

- (72) a. Leo ha mangiato patate.  
 Leo PST eat potatoes  
 'Leo ate potatoes.'
- b. \*Studenti hanno telefonato.  
 students have telephoned
- c. \*Hanno telefonato studenti.  
 have telephoned students

(Chierchia 1998b:383–384)

However, there are some contexts in which bare arguments are acceptable. These contexts include focus positions (73a), contrastive settings (73b) and modification constructions (73c).

- (73) a. POLLO io voglio, non pesce.  
 chicken I want not fish  
 'I want CHICKEN, not fish.'
- b. I butteri maremmani domani un cavallo in 2 o 3 sedute. Invece,  
 the cowboys Maremma tame a horse in 2 or 3 sessions instead  
*butteri di qui lo fanno in una seduta.*  
 cowboys from here it do in one session  
 'The cowboys from Maremma tame a horse in 2 or 3 sessions. Conversely, the  
 cowboys from here do it in one session.'
- c. Hanno telefonato *studenti* che volevano sapere la data dell'esame.  
 have telephoned students that want know the date of.the.exam  
 'The students who want to find out the date of the exam have called.'

(Chierchia 1998b:384–385, 388)

Chierchia claims that these cases all involve focus structure, i.e. FocP, and the Foc head licenses the null determiner. In fact, he also considers an alternative analysis proposed by Delfitto and Schrotten (1991) in a footnote (footnote 32, p. 386): "the availability of the null determiner is linked to the presence of 'audible' plural features on the noun." However, he doubts the viability of such an alternative, as it leads one to posit abstract features in common with plurals for mass nouns.

The Italian facts surprisingly resemble the facts concerning plurals in classifier languages. Both involve a null determiner. The environments in which unexpected licensing takes place are also similar. It is thus possible that the two cases are in fact related and receives the same account. If that is the case, the alternative analysis in line of Delfitto and Schrotten (1991),

specific details being put aside, is actually quite plausible. This is especially so because in the next chapter, I will propose a crosslinguistic typology of number systems where some mass nouns in languages like English and Italian indeed share a number feature (i.e. [-Sg]) with plurals.

## 5.5 Summary

In this chapter, I have pointed out the semantic parallelism between classifiers and plural markers in classifier languages. The semantics of both classifiers and plural markers consists of two parts, i.e. the asserted meaning concerning number and conventional implicatures concerning the characteristics of the nouns combining with classifiers/plural markers. In the next chapter, I will present more facts pointing to the parallelism between the two, and propose a number system of classifier languages based on the findings so far.

In discussing the relation between plurals in classifier languages and definiteness, I have claimed that plurals in classifier languages involve a null determiner that is semantically related to referentiality and licenses a plural marker. The present study thus lends support to the view that bare arguments in classifier languages project DPs, even though the D head is normally not audible (e.g. Cheng and Sybesma 1999, 2005; Simpson 2005; Watanabe 2006). Furthermore, it also endorses the universal connection between the syntactic category of D and the meaning of referentiality (Longobardi 1994, 2005).

## Chapter 6

# Classifier languages and the typology of number marking

### 6.1 Introduction

In the last two chapters, I have discussed the semantics of classifiers, numerals and plural markers in classifier languages. We have seen that a striking parallelism exists between the semantics of classifiers and plural markers. This chapter first summarizes the number system of classifier languages emerging from the discussions so far (section 6.2). I claim that classifiers and plural markers should be treated as genuine number markers, singular and plural markers respectively, on a par with those of non-classifier languages. That is, classifier languages do not lack genuine number marking, contrary to popular belief. The chapter then proposes a typology of number marking that accommodates classifier languages naturally in relation to other types of number systems (section 6.3). I argue that there are four basic number categories made available by UG through the combinations of two binary number features, and that classifier languages make the most fine-grained distinction among them. Non-classifier languages morphologically conflate two or more basic number categories into one. Whether a language has classifiers thus depends on how the language realizes the four basic number categories morphologically, but not on what nouns or numerals denote (cf. Krifka 1995; Chierchia 1998b; Wilhelm 2008). The proposed typology allows apparently optional uses of classifiers and plural markers in classifier languages, which is sometimes raised as support for not regarding classifiers and plural markers in classifier languages as genuine number marking. Finally, section 6.4 discusses an implication of

the proposed typology for number and definiteness marking patterns in kind terms across languages. The present study offers an account for the puzzle of bare “singular” kind terms in languages such as Brazilian Portuguese.

## 6.2 The number system of classifier languages

### 6.2.1 Classifiers and plural markers in classifier languages are true number morphology

So far, we have seen that the meanings of classifiers and plural markers in classifier languages are structured essentially in the same way, consisting of the asserted meaning concerning number and conventional implicatures concerning the characteristics of the nouns combining with them. Specifically, I have proposed the following semantics for them, where CLASS is a shorthand for a conjunction of properties of the nouns such as ‘animate’ and ‘human’.

- (1) a.  $\llbracket \text{CL} \rrbracket = \lambda P \lambda x [P(x) \wedge \neg \exists y \in P[y < x]] \blacklozenge \lambda P [P \subseteq \text{CLASS}]$   
 b.  $\llbracket \text{PL} \rrbracket = \lambda P \lambda x [P(x) \wedge \exists y \in P[y < x]] \blacklozenge \lambda P [P \subseteq \text{CLASS}]$

Classifiers and plural markers only differ in the assertion portion (i.e. the material before  $\blacklozenge$ ) concerning number. They restrict the domain to singularities and pluralities respectively. In terms of assertions, they are no more than ordinary number markers and do not differ from more established number markers in other languages such as the suffix *-s* in English. The only difference is that they contain conventional implicatures (i.e. the material after  $\blacklozenge$ ). In this sense, classifiers and plural markers in classifier languages are sophisticated type of singular and plural markers. The content of conventional implicatures varies from item to item. In general, the conventional implicatures of plural markers are not as rich as those of classifiers. Some plural markers in classifier languages have no conventional implicature, in which case plural markers in classifier languages indeed behave like those of non-classifier languages and apply to any count nouns. For example, while Malay employs different classifiers for different referents, it has only one device for plural marking, i.e. reduplication, as shown in (2).

- |     |    |                           |    |                         |
|-----|----|---------------------------|----|-------------------------|
| (2) | a. | se-orang lelaki ‘a man’   | b. | lelaki-lelaki ‘men’     |
|     |    | se-ekor kucing ‘a cat’    |    | kucing-kucing ‘cats’    |
|     |    | se-helai baju ‘a clothes’ |    | baju-baju ‘clothes’     |
|     |    | se-buah harapan ‘a hope’  |    | harapan-harapan ‘hopes’ |

One might think that classifiers cannot be a singular marker because they co-occur with numerals greater than one. It is true that a phrase consisting of a numeral greater than one and ‘CL NP’ (e.g. *tiga orang lelaki* [three CLF man] ‘three men’) refers to pluralities. However, the plural reference of the phrase is due to the numeral in this case. As seen in chapter 3, the semantic composition of a numeral and ‘CL NP’ involves multiplication: ‘three CLF man’ =  $3 \times 1$  man.

Regarding plural markers, I have already refuted the counterarguments that they encode individuality, associativity or definiteness instead of plurality in chapters 3 and 5. It was pointed out that not all plural forms express the relevant meanings. It is thus implausible to think that these meanings are semantically encoded meanings of plural markers. The meaning of individuality arises from pragmatic inference based on the presence of two possible forms to refer to multiple referents. Some plural markers have a homophonous associative marker; the two markers have different distributions and can co-occur simultaneously. Plurals are associated with referentiality, but not definiteness; this association exists not because plural markers encode referentiality in addition to plurality, but because they are syntactically licensed by determiners encoding referentiality.

One may wonder why only plural markers, but not classifiers, require syntactic licensing by determiners. Such an asymmetry would weaken my claim that classifiers are a sophisticated kind of singular number marker. However, classifiers seem to be licensed by determiners in many, though not all, cases. Previous studies have pointed out a close connection between classifiers and referentiality/specificity. Studying the use of classifiers in a Classical Malay text, *Hikayat Abdullah*, Hopper (1986:314) states that noun phrases with classifiers (‘Num CL NP’ as opposed to ‘Num NP’) are “typically new but referential, and refer to individual participants in the discourse.” Bisang (1999:152–155) reports a similar role played by classifiers in discourse for Hmong and Weining Miao: a new participant is introduced into the discourse by ‘one CL NP’ (specific indefinite) and then referred to by ‘CL NP’ (bare classifier constructions; definite)

and even by ‘NP’ (definite).<sup>1</sup> According to Jenks (2011:207–211), noun modification constructions with classifiers (‘NP CL modifier’) are referential unless they contain a true indefinite quantifier such as *baaŋ* ‘some’. They can occur as complements of the equative copula *khuu*, but cannot occur as complements of the predicative copula *pen*, as in (3).<sup>2</sup> An overt NP does not appear in (3). They cannot occur in generics either, unlike noun modification constructions without a classifier, as shown in (4).

- (3) a. Nik khuu [khon thii chán rák].  
 Nick COP.EQ CLF REL I love  
 ‘Nick is the person that I love.’  
 b. \*Nik pen [khon thii chán rák].  
 Nick COP.PRED CLF REL I love

(Ruangjaroon 2005:105)

- (4) a. \*Taam-thamadaa [thúrian lúuk sùk-sùk] wǎan mâak.  
 generally durian CLF ripe sweet very  
 For: ‘Generally, ripe durians are very sweet.’  
 b. Nát chōp [thúrian sùk].  
 Nat like durian ripe  
 ‘Nat likes ripe durians.’

(Jenks 2011:91, 207)

All these facts suggest that classifiers are subject to the same syntactic licensing as plural markers, i.e. by means of agreement with referential determiners. Moreover, Jenks’s qualification about quantifiers is in accordance with the fact that plural markers are licensed also by quantifiers (cf. chapter 5). In fact, classifiers are also licensed by quantifiers. In this case, a noun

<sup>1</sup>This fact invalidates Cheng and Sybesma’s (1999) hypothesis that languages either have definite bare classifier constructions or definite bare NPs, but not both. Simpson et al. (2011) present further evidence against the hypothesized complementarity from Bangla, Cantonese, Hmong and Vietnamese.

<sup>2</sup>Licensing by modification discussed in chapter 5 (modifiers are able to make otherwise referential expressions less referential) does not apply to this Thai construction. This is due to the syntactic structure of the construction. Unlike the cases discussed in chapter 5, the modifier is within the c-command domain of the determiner in this construction, like pronominal modifiers in English. Adopting Kayne’s (1994) analysis of relative clauses, Jenks (2011) posits the structure in (ib) for the noun phrase in (ia).

- (i) a. thúrian lúuk thii mǎn  
 durian CLF REL smelly  
 ‘the durian that is smelly’  
 b. [DP D [CP [ClFP [NP thúrian ] lúuk ] [C’ thii mǎn ]]] (ClFP: classifier phrase)

phrase receives a non-specific interpretation. The ‘Num CL NP’ pattern is thus typically ambiguous between referential/specific and non-specific interpretations in a position where both interpretations are generally allowed, as in (5) and (6). The (a) sentence can be continued by either (b) or (c).

(5) Japanese

- a. San/Suu nin-no gakusei-ga siken-ni oti-ta.  
three/a.few CLF-LINK student-NOM exam-DAT fail-PST  
‘Three/A few students failed in the exam.’
- b. Yuko to Aiko to Ryoko-da.  
Yuko and Aiko and Ryoko-COP  
‘They are Yuko, Aiko and Ryoko.’
- c. Sono uti-no hito-ri-ga watasi-ja-nai to ii-kedo.  
that among-GEN one-CLF-NOM me-COP-NEG if good-though  
‘I hope I’m not one of them.’

(6) Malay

- a. Tiga/Beberapa orang pelajar gagal dalam peperiksaan itu.  
three/some CLF student fail in exam that  
‘Three/Some students failed in the exam.’
- b. Mereka ialah Abu, Hasan dan Rasyid.  
they are Abu Hasan and Rasyid  
‘They are Abu, Hasan and Rasyid.’
- c. Harap-harap aku bukan salah se-orang daripada-nya.  
hopefully I not among one-CLF from-3  
‘Hopefully I’m not one of them.’

Another instance of the parallelism between classifiers and plural markers is the default definite interpretation of bare classifier constructions (‘CL NP’), which are free from the influence of other elements such as numerals and quantifiers. We have seen in chapter 5 that plurals in classifier languages are interpreted as definite by default, allowing indefinite interpretations in certain contexts. The same holds true for bare classifier constructions. It has been noted that bare classifier constructions are interpreted as definite, at least when they occur in the subject position, as in (7) (Cheng and Sybesma 1999, 2005; Simpson 2005).

- (7) Cantonese (Cheng and Sybesma 1999:511)

Jek gáu gāmyaht dahkbiht tēngwah.  
 CLF dog today special obedient  
 ‘The dog is specially obedient today.’

- (8) Bangla (Simpson et al. 2011:170)

kEmera-Ta khub dami.  
 camera-CLF very expensive  
 ‘The camera was/is expensive.’

- (9) Hmong

Lub koob thajj duab kim kim heev.  
 CLF camera expensive expensive very  
 ‘The camera was/is expensive.’

However, this does not mean that indefinite interpretations are totally unavailable in languages with the bare classifier construction, though Bangla seems to be an exception. I will return to Bangla below. Bisang (1999:146, 152, 157) states that bare classifier constructions in Vietnamese and Cantonese are ambiguous between definite and indefinite interpretations whilst the numeral ‘one’ is obligatory in indefinite contexts in Hmong. Remember that plurals can have otherwise unavailable non-referential interpretations when accompanied by modifiers. The same semantic licensing seems to be at work in the case of bare classifier constructions in Hmong. All three consultants of mine accepted the bare classifier construction as well as the canonical ‘one CL NP’ for the indefinite interpretation for sentence (10) below. One of them commented on the bare classifier option that one could say it because there was a “modifier.”

- (10) Kuv toob kas {ib tug<sup>3</sup> tub kawm ntawv/ tus tub kawm ntawv} pab kuv tu lub  
 I need one CLF student CLF student help me clean CLF  
 chav kawm ntawv.  
 classroom  
 ‘I need a student to help me clean the classroom.’

The obligatory definite interpretation of bare classifier constructions in Bangla is thought to be related to their special word order. While the NP follows the classifier in numeral modification

<sup>3</sup>The classifier *tus* (s-tone) is realized as *tug* (g-tone) after a b-tone word like *ib* ‘one’. Currently, this tone sandhi rule is dropping out of the language in both Asia (Ratliff 1987) and Minnesota (p.c. Vincent Carveth).

constructions, it precedes the classifier in bare classifier constructions, as shown in (11).

- (11) a. Ek-Ta kEmera  
 one-CLF camera  
 ‘a camera’  
 b. kEmera-Ta  
 camera-CLF  
 ‘the camera’.

Previous studies on Bangla nominals such as Bhattacharya (1999) and Chacón (2011) analyze bare classifier constructions as involving a specificity/definiteness-driven leftward NP movement, assuming a head-initial syntax, as schematically summarized in (12a). However, given that the language is unquestionably head-final in the verbal domain (but see Simpson and Bhattacharya 2003),<sup>4</sup> one could also adopt the head-final DP structure proposed for Japanese by Watanabe (2006), as shown in (12b). The NP also moves leftward in this case, triggered by Case, but not D.

- (12) a. [DP NP<sub>i</sub> [ D [ (numeral) [ CL t<sub>i</sub> ] ] ] ] (head-initial analysis)  
 b. [DP [CaseP NP<sub>i</sub> [ [ (numeral) [ t<sub>i</sub> CL ] ] Case ] ] D ] (head-final analysis)

Two possible accounts are conceivable for the definite interpretation. First, the classifier moves to D (via Case) or agrees with D in situ, and virtually functions as a definite article (or has become one already through grammaticalization). The second possibility is that the language in fact has an indefinite marker on a par with *-i* in Persian, which either occurs with classifiers only rarely or has been analyzed differently in the literature; bare classifier constructions are interpreted as definite in the absence of this indefinite marker.<sup>5</sup>

<sup>4</sup>I thank Dustin Chacón for bringing Simpson and Bhattacharya’s work to my attention.

<sup>5</sup>In passing, under the head-final analysis along the line of Watanabe (2006), approximate noun phrases like (ia) discussed by Chacón (2011) have the structure in (ib). According to Watanabe (2010), vague quantity expressions are merged in a position higher than precise numerals, specifically Spec,QP above CaseP. The internal structure of the vague quantity expression *goTa du=Ek* ‘two or so’ is given in (ic). The classifier-initial word order results from a remnant movement of XP. Chacón derives the same word order from a head-initial structure by a successive head movement of the classifier to the numeral head and then to D. His analysis hinges on the following three assumptions that my analysis does not make: (i) head movement is achieved by left-adjunction; (ii) numerals head their own projection; and (iii) no numerals including complex ones (e.g. 150) are phrasal.

- (i) a. goTa du=Ek kham  
 CLF two=or.so envelope  
 ‘two envelopes or so’

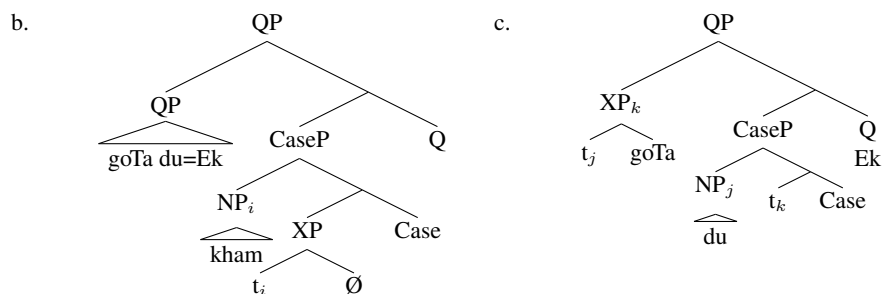
The last shared feature of classifiers and plural markers in classifier languages is their optionality. While noun phrases with these markers are singular and plural (save the influence of other materials such as numerals), not all noun phrase that can refer to singular and plural referents occur with these markers. For some, this kind of optionality may be an insurmountable obstacle keeping them from regarding any number-related markers in classifier language as genuine number morphology on a par with those in non-classifier languages such as *-s* in English. However, I will show below that this kind of optionality is in fact expected in the number system I am proposing, if classifiers and plural markers in classifier languages are true number markers.

To summarize, the denotations of classifiers and plural markers in classifier languages (cf. (1)) are nothing but those of singular and plural markers. Moreover, I have pointed out a few commonalities that justify analyzing them as belonging to the same grammatical category of number. Classifier languages thus have true number markers, i.e. classifiers as singular markers and plural markers, which means that they distinguish at least two number categories morphologically.

### 6.2.2 General number as the third number

The conclusion that classifier languages distinguish between the singular and plural may sound surprising, as they have been often described as lacking true number morphology. However, it is not the end of the story. Classifier languages are famous for their extensive use of bare NPs. As seen in chapter 3, bare NPs in classifier languages are associated with the general number, which expresses number-neutral properties. I argue that the general is the third number category on a par with the singular and plural. In other words, classifier languages morphologically distinguish not only the singular and the plural but also the general.

Concepts such as general number and number-neutrality/transnumerality have been widely



known to researchers of nominal syntax and semantics. However, the general number has been kept separate from more established number categories such as the singular and plural. For example, Corbett (2000:9–10) characterizes the general number as “non-number” and states that “it is outside the number system.” His decision to exclude the general from the number system is in fact only definitional and is not grounded by empirical facts. Corbett starts the section on the general as follows:

In English we are usually forced to choose between singular and plural when we use a noun. However, there are languages for which number is less dominant, languages in which the meaning of the noun can be expressed without reference to number. We shall call this ‘general number’, by which we mean that it is outside the number system. (Corbett 2000:9–10)

It may be true that the meaning of bare NPs in classifier languages *can* be described without using the terms ‘singular’ and ‘plural’. However, that does not necessarily mean that it *should* be described in such a way and that the linguistic mechanism involved does not refer to number. Corbett’s decision is also influenced by his view on the English number system, which he regards as lacking the general. This classification is inadequate. As seen in chapter 3, English does have forms associated with the general number, i.e. morphologically plural forms as in *How many cats do you have?* Therefore, nothing prohibits us from considering the general as a genuine number category insofar as doing so captures empirical facts better than not doing so.

Now that some languages mark two number categories overtly and use bare NPs without the relevant markers, it is a natural move for any sensible linguist to render the unmarked NPs an equal status as a number category. Putting aside the issues of articles and mass nouns, the absence of the suffix *-s* in NPs in English (e.g. *(a) book*) is considered to represent the singular number as opposed to the plural. No one would contend that it means a lack of number and the plural is the only number category available in English.

It should be clear now that the number system of classifier languages emerging from the discussions so far is quite contrary to the prevalent belief. Classifier languages do not lack number morphology; they in fact distinguish three number categories morphologically, i.e. singular, plural and general. I consider these three number categories as basic number categories in natural languages, as opposed to less common number categories such as the dual and paucal. Table 6.1 gives examples of each basic number category in Malay and Japanese. The numeral

‘one’ in the singular forms is irrelevant for number distinction. It is added here because these languages do not have the bare classifier construction ‘CL NP’.

Table 6.1: The realizations of basic number categories in Malay and Japanese.

	Malay	Japanese
Singular:	classifiers <i>se-ekor harimau</i> [one-CLF tiger]	classifiers <i>tora ip-piki</i> [tiger one-CLF]
Plural:	reduplication <i>harimau-harimau</i> [tiger.PL]	<i>-tati, -ra, -gata, -domo</i> <i>tora-tati</i> [tiger-PL]
General:	∅ <i>harimau</i> [tiger]	∅ <i>tora</i> [tiger]

### 6.2.3 Optionality

Classifier languages are most particular about number distinction as they have morphological devices to distinguish all three basic number categories. At the same time, however, they deceptively look indifferent to number distinction as well, due to the frequent use of general number forms, i.e. bare NPs. The latter characteristic of classifier languages brings us back to the issue of optionality of plural markers in classifier languages.

I claim that the observed optionality is merely an ostensible one and is not a genuine one. In the absence of numerals greater than one, the plural marker is obligatory to refer exclusively to pluralities. The use of general NPs would also include unwanted singularities. Nomoto (2013) calls this type of optionality ‘ostensible optionality’ and distinguishes it from ‘genuine optionality’, where the denotations of the forms with and without the grammatical marker in question are identical. In ostensible optionality, the denotation of the marked form is a proper subset of that of the unmarked form. Given this distinction, marking plurality in classifier languages is in fact obligatory, as in well-known number-languages such as English, and the observed optionality is only ostensible. Plural markers in languages like English appear to be obligatory because they encode not only the plural but also the general number. Thus, the “plural” marker is present whenever pluralities are referred to, regardless of whether singularities are also included or not. That is to say, plural markers that we know well is not as pure as plural markers in classifier

languages.

The ostensible optionality of plural marking in classifier languages also follows from analyses that treat bare common nouns as denoting kinds and make the semantics of plural markers only yield pluralities (e.g. Yang 1998; Kurafuji 2004). The property extracted from kinds is always number-neutral, and it thus subsumes pluralities and singularities (cf. chapter 3). The difference between such an analysis and mine lies in whether the kind reference and number-neutrality of bare NPs in classifier languages are a semantic choice or a morphological one. In the former approach, bare NPs in classifier languages denote kinds because of some *semantic* principle, say Chierchia's (1998b) Nominal Mapping Parameter; the number-neutrality of bare NPs follows from the kind reference. By contrast, under the analysis I propose here, the number-neutrality of bare NPs is a result of a *morphological* choice, i.e. marking the three basic number categories distinctly, and the kind reference of bare NPs follows from the general, associated with number-neutral properties, being the morphologically unmarked category in the three-way number marking system.

As the general subsumes not only the plural but also the singular, singular marking by classifiers is also expected to be optional ostensibly. This prediction is borne out to different degrees in different languages. The absence of classifiers is allowed in basically all classifier constructions in optional classifier languages such as Malay whereas it is only allowed in certain restricted nouns/constructions in obligatory classifier languages such as Japanese (cf. chapter 2). The question to ask then is: why is classifier use generally obligatory in obligatory classifier languages?

As I have shown in chapter 2, semantic accounts based on the "classifiers for counting" thesis do not really answer this question. There is no difference between obligatory, optional and non-classifier languages, either in the denotations of nouns and numerals or in the process whereby numerals combine with nouns. Therefore, I suggest that the crosslinguistic difference in question is morphosyntactic rather than semantic.

I argue that obligatory and optional classifier languages differ in the extent to which nouns can serve as classifier substitutes. Simpson (2005) maintains that it is generally possible in optional classifier languages for a noun to fill the syntactic position for classifiers. According to this analysis, the Malay expressions with and without a classifier in (13) have the structures shown in (14). I mark the syntactic position occupied by classifiers with CL, without committing myself to the issue of what it really is. Notice that the classifier position is occupied by the

noun *buku* in (14b).

- (13) a. dua buah buku  
two CLF book  
'two books'
- b. dua buku  
two book  
'two books'
- (14) a. dua [CL buah ] [N buku ]  
b. dua [CL [N buku ]]

Simpson (2005) implements this idea by means of head movement. Specifically, a noun moves to the classifier position when a classifier is not selected from the lexicon. Such an analysis is justified by the so-called repeater construction, where the classifier position is occupied by the noun itself, as in (15).

- (15) a. Thai  
hôn sãm hôn  
room three room/CLF  
'three rooms'
- b. Burmese  
cún tã cún  
island one island/CLF  
'one island'

(Simpson 2005:832)

Under the copy theory of movement (Chomsky 1995), which decomposes movement into Copy and Merge, these examples can be understood as a result of pronouncing both the original and its copy ( $\alpha_{\text{original}}\beta\alpha_{\text{copy}}$ ). In the more traditional type of movement, the original becomes phonologically inert and only the copy is pronounced ( $\beta\alpha_{\text{copy}}$ ). Simpson compares classifiers to expletives such as *there* and *it* in English. When they are not selected from the lexicon, the syntactic position that is supposed to be filled by them is filled by a noun phrase. Compare (14) and the sentences in (16) below.

- (16) a. [EXP It ] is likely for [DP John ] to be promoted next year.  
 b. [EXP [DP John ]] is likely to be promoted next year.

To use the analogy of expletives, optional classifier languages are comparable to languages in which the expletive position (Spec,TP in the case of (16)) can be filled by either expletives or raised DPs, whilst obligatory classifier languages are comparable to languages in which the expletive position must be filled by expletives in normal cases. Non-classifier languages can be compared to languages that lack overt expletives; they have null classifiers, as Sharvy (1978), Muromatsu (1998) and Borer (2005) claim. Interesting though Simpson's head movement analysis is, the same range of data can be also accounted for by hypothesizing that a noun is simply merged in the classifier position without involving any movement. No empirical reason seems to exist for choosing one analysis over the other.

In obligatory classifier languages, substituting nouns for classifiers is not a general phenomenon, but is limited to certain classes of nouns, i.e. (i) some animate nouns as in (17) and (ii) many abstract nouns, including words meaning 'kind', as in (18).<sup>6</sup>

- (17) Korean (Kang 1994:8)
- a. [N haksayng] han [CL myeng]  
 student one CLF  
 'a student'
- b. han [CL [N haksayng]]  
 one student  
 'a student'

---

<sup>6</sup>These are the first two of the four contexts where classifier use becomes optional in obligatory classifier languages pointed out in chapter 2 (section 2.3.2), the other two contexts being (iii) when the numeral expresses large or non-specific numbers and (iv) in the 'NP(-case) Num CL' pattern (in Japanese). The analysis of the latter two cases should be different from that given to the first two. The factor governing classifier optionality for (iii) is obviously not nouns but numerals. One possibility is that large or non-specific numbers are ambiguous between numerals and vague quantity expressions such as *ooku* 'many', which only require the linker *no*, but not a classifier, when modifying a noun (see Watanabe (2010) for a syntactic account of why vague quantity expressions in Japanese do not co-occur with classifiers). Regarding (iv), the noun is not occupying the classifier slot in the absence of a classifier, given that classifiers never precede numerals in Japanese. Hence, classifier optionality here is due to some special character of the relevant word order, which is arguably shared by many other languages, as the 'NP Num (CL)' sequence is commonly used in lists even in languages that normally do not allow this word order (Greenberg 1975; Simpson 2005).

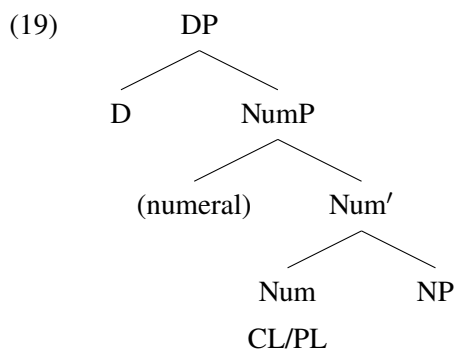
- (18) Japanese<sup>7</sup>
- a. mit- [CL tu]-no [N gengo/shurui]  
 three CLF-LINK language/kind  
 ‘three languages/kinds’
- b. san [CL [N gengo/shurui]]  
 three language/kind  
 ‘three languages/kinds’

The word order in (17b) and the loss of the linker *no* in (18b) support the analysis whereby the noun occupies the classifier position. To summarize, obligatory and optional classifier languages differ in that nouns serving as classifier substitutes is a fully productive phenomenon in optional classifier languages whereas it is only allowed in limited cases in obligatory classifier languages.

## 6.2.4 The syntax of DPs in classifier languages

### 6.2.4.1 The DP structure

So far, I did not name the node consisting of an NP and a classifier/plural marker intentionally, as it did not affect the argument. Based on the discussion above, we can now identify this node as Num(ber)P. As classifiers and plural markers are singular and plural number morphology, they occupy the same functional head, i.e. Num. Noun phrases thus have the following structure:



The structure above is the same as the structure proposed by Watanabe (2006), except for the idea that plural markers in classifier languages occupy the same head as classifiers. Watanabe

<sup>7</sup>The numerals meaning ‘three’ in (a) and (b) are the native Japanese and Sino-Japanese forms respectively. See Table 4.1 in chapter 4 for the list of native Japanese and Sino-Japanese numerals.

posits two other functional heads between D and Num, i.e. Q and Case:

(20) [DP D [QP Q [CaseP Case [NumP Num NP ]]]]

I omitted QP and CaseP in the tree above for simplicity. The structure in (19) differs from the structures proposed by many others (e.g. Li 1999; Cheng and Sybesma 1999; Borer 2005; Simpson 2005; Huang et al. 2009) in that classifiers and numerals are in the same projection and do not project two distinct projections, e.g. Cl(assifier)P and Num(ber/eral)P.

An argument that initially appears to support positing an independent functional head for numerals is that numerals do not occur adjacent to classifiers in some (head-initial) languages (Simpson 2005). The Benue-Congo language Ejagham is one such language discussed by Simpson.

(21) a-mege ‘ i-cokud a-bae  
 NC-CLF GEN NC-orange.seed NC-two  
 ‘two orange seeds’ (NC: noun class marker)

(Watters 1981)

If numerals project an independent projection above ClP, which I refer to NumP, this word order can be easily obtained by a leftward movement of ClP (the case marker is put aside): [DP [ClP CL NP]<sub>i</sub> [NumP numeral t<sub>i</sub>]]. However, the same word order can be also derived from Watanabe’s structure in (20). The derivation proceeds as shown in (22). First, the NP ‘orange seed’ moves to Spec,CaseP for Case to give (22b). Next, the classifier in Num moves up to D via Case presumably due to referentiality. Movement of classifiers to D is also posited by Simpson (2005) to capture definiteness.

(22) a. [DP [CaseP GEN [NumP two [Num’ CL [NP orange.seed ]]]]]  
 b. [DP [CaseP orange.seed<sub>i</sub> [Case’ GEN [NumP two [Num’ CL t<sub>i</sub> ]]]]]  
 c. [DP CL<sub>j</sub>+GEN<sub>k</sub> [CaseP orange.seed<sub>i</sub> [Case’ t<sub>k</sub> [NumP two [Num’ t<sub>j</sub> t<sub>i</sub> ]]]]]

For similar facts in Nung and Thai (both Tai languages) as in (23) and (24), I assume that the numeral ‘one’ is not merged in the numeral position but in the determiner position, as the word order is peculiar to ‘one’ and is not observed in numerals in general. Simpson (2005) analyzes this word order by a head movement of ‘one’ to D. However, it is not clear why this movement

does not occur with other numerals. ‘One’ in the determiner position does not co-occur the numeral ‘one’ as in (24c) arguably for the same reason why *a* does not co-occur with *one* in English (*\*a one child*).

(23) Nung (Saul and Wilson 1980:27, 29)

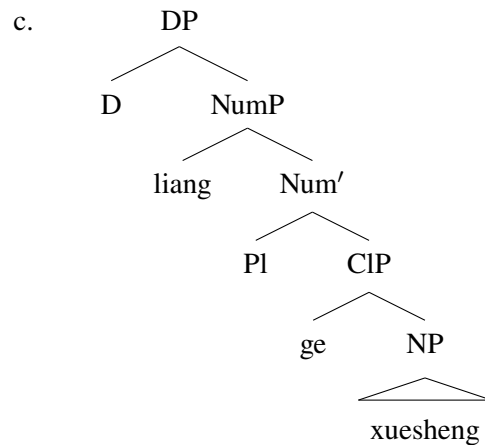
- a. áhn táhng nưhng  
CLF chair one  
‘one/a chair.’
- b. slám áhn vệt  
three CLF spoon  
‘three spoons’

(24) Thai (Simpson 2005:827–828)

- a. dèk khon nàŋ  
child CLF one  
‘one/a child’
- b. dèk sǎam khon  
child three CLF  
‘three children’
- c. \*dèk sǎam/nàŋ khon nàŋ  
child three/one CLF one

Li (1999) proposes a structure in which classifiers head a projection distinct from NumP, as in (25c). She accounts for the incompatibility between classifiers and the plural marker *-men* in Mandarin as in (25b) based on this structure. She claims that they cannot co-occur because the plural feature in Num cannot be realized on N because of the intervening classifier head.

- (25) a. xuesheng-men  
student-PL  
‘(the) students’
- b. \*san-ge xuesheng-men  
three-CLF student-PL



This argument is valid for Mandarin, but not for many other classifier languages in which classifiers and plural markers can co-occur, freely or in a restricted fashion (cf. chapter 4). The Japanese equivalent of (25b) is grammatical beyond any doubt.

- (26) san nin-no gakusei-tati  
 three CLF-LINK student-PL  
 'three students'

Neither the proposed structure in (19), where classifiers and plural markers occupy the same position, nor Li's structure in (25c), where classifiers and plural markers/features occupy two distinct heads, can handle the observed crosslinguistic variation as it stands. Hence, the Mandarin data above does not support one analysis over the other.

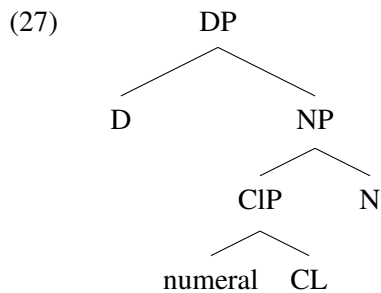
I propose the following morphological solution for the crosslinguistic variations in the morphological realization of the number head in (19). In the feature system that I will propose below, the number head hosts a pair of number-related features, i.e. [ $\pm$ Sg] and [ $\pm$ PI], and morphology assigns particular lexical items to different combinations of these features. It is common for a head to host more than one feature, whereas multiple lexical items occupying the same head normally leads to ungrammaticality. A classifier and a plural marker co-occur when the feature combination is '[+Sg], [+PI]'. The features [+Sg] and [+PI] are normally realized by separate morphemes, i.e. by a classifier and a plural marker respectively. In this case, there is a one-to-one correspondence between features and morphemes. However, one syntactic head ends up with hosting two morphemes. Some languages avoid this mismatch by realizing the two features by a single word. Such words are known as 'plural classifiers' and include *dī* in

Cantonese and *cov* in Hmong, for example. In this case, the one-to-one correspondence between features and morphemes is sacrificed to attain a one-to-one correspondence between heads and morphemes. Mandarin takes a radical strategy to resolve a potential loss of correspondence. Its morphology gives up realizing the ‘[+Sg], [+Pl]’ feature combination and makes the structure ungrammatical. These three different ways of handling the feature-head mismatch can be summarized as in Table 6.2. One can think of REALIZE FEATURES, 1 MORPHEME FOR 1 FEATURE and 1 MORPHEME FOR 1 HEAD in the top row like Optimality Theoretic constraints. “(√)” in the bottom row means that the relevant constraint is vacuously satisfied.

Table 6.2: Crosslinguistic variations in ways of handling the feature-head mismatch.

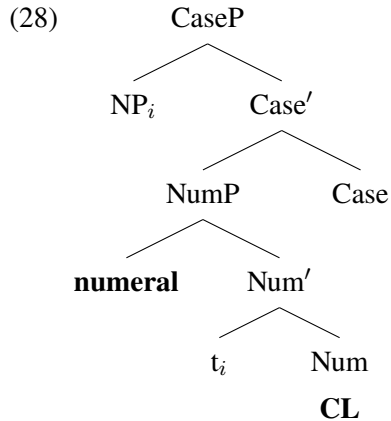
Language type	REALIZE FEATURES	1 MORPHEME FOR 1 FEATURE	1 MORPHEME FOR 1 HEAD
Japanese	√	√	*
Cantonese	√	*	√
Mandarin	*	(√)	(√)

Another structure that is commonly assumed in the literature is one in which classifiers forms a constituent with numerals, as in (27).



I have already pointed out a semantic problem of this structure in chapter 4: it neglects the truth-conditional meaning of classifiers, i.e. restriction to singularities. There must be a point in the structure-building process at which a classifier and an NP form a constituent to the exclusion of a numeral, because number morphology is concerned with nouns, but not with numerals. See Simpson (2005) for other arguments against a structure like (27). I admit that speakers of classifier languages, including myself, have the intuition that classifiers are more tightly linked to numerals than to NPs. I understand that this intuition reflects either the phonological dependency between numerals and classifiers or/and a later stage of syntactic derivation, at

which point there is a phrase containing a numeral and a classifier alone, but not an NP. For instance, Watanabe (2006) claims that NPs obligatorily move from the initially merged position to the specifier of a higher projection (CaseP) in Japanese. After this movement, the numeral and classifier can move around independently of the NP, as shown in (28).



This explains speakers' intuition that the classifier is closely tied to the numeral. Speakers' intuition and some surface phenomena implying the 'numeral CL' constituency thus do not necessarily mean that the structure in (27) is correct. The speakers' intuition and relevant surface phenomena are also compatible with (19), which is most plausible, given the semantics of classifiers proposed in this study.

#### 6.2.4.2 D-Num agreement

In section 6.2.1 above, I claimed that both classifiers and plural markers are licensed by referential determiners or quantifiers by means of agreement. What this means is that the presence of classifiers and plural markers entail the presence of a DP projection even if it is not apparent on the surface, unlike languages with (in)definite articles. The determiners under D are covert in most classifier languages. However, some languages do have determiners with phonological content. We have already seen in chapter 5 that Persian has an overt referential indefinite determiner, i.e. *-i*. Plurals with *-hâ* in Persian are indefinite when *-i* occurs and definite otherwise. According to Jiang and Hu (2012), the Tibeto-Burman language Yi has a definite determiner, i.e. *su*. Noun phrases with a classifier in Yi are definite when *su* occurs and indefinite otherwise,

as in (29).<sup>8</sup>

- (29) a. mu<sup>33</sup> sɔ<sup>33</sup> ma<sup>33</sup> su<sup>44</sup> li<sup>44</sup>ndo<sup>33</sup> o<sup>44</sup>.  
horse three CLF DEF lose PART  
‘The three horses got lost.’
- b. mu<sup>33</sup> sɔ<sup>33</sup> ma<sup>33</sup> li<sup>44</sup>ndo<sup>33</sup> o<sup>44</sup>.  
horse three CLF lose PART  
‘Three horses got lost.’

(Jiang and Hu 2012)

Classifiers and plural markers take different forms depending on definiteness in Wenzhou Chinese (Cheng and Sybesma 2005) and Weining Ahmao (Miao) (Bisang 1999; Gerner and Bisang 2010). Neither language has overt (in)definite determiners. The Wenzhou classifiers in (30) differ in their tones (indicated by superscript).

(30) Wenzhou (Cheng and Sybesma 2005:266)

- a. ŋ<sup>4</sup> ɕi<sup>3</sup> ma<sup>4</sup> paŋ<sup>7</sup> si<sup>1</sup>.  
I want buy CLF.DEF book  
‘I want to buy the book.’
- b. ŋ<sup>4</sup> ɕi<sup>3</sup> ma<sup>4</sup> paŋ<sup>3</sup> si<sup>1</sup>.  
I want buy CLF.INDF book  
‘I want to buy a book.’

(31) Weining Ahmao (Gerner and Bisang 2010:591, 594)

- a. a<sup>33</sup>lau<sup>33</sup> mfi<sup>a35</sup> dzfi<sup>a35</sup> pi<sup>31</sup>dfiau<sup>31</sup> li<sup>31</sup> ŋi<sup>55</sup>: ....  
old.people have CLF.DEF story manner this  
‘The ancestors have a saying (= little story) which goes like this: ...’<sup>9</sup>
- b. a<sup>33</sup>lau<sup>33</sup> mfi<sup>a35</sup> dza<sup>35</sup> pi<sup>31</sup>dfiau<sup>31</sup>.  
old.people have CLF.INDF story  
‘The ancestors have a saying (= little story).’
- c. pi<sup>55</sup> ti<sup>55</sup> ŋie<sup>55</sup> hi<sup>44</sup> zau<sup>55</sup>.  
our PL.DEF tooth not good  
‘Our teeth are not in a good state.’

<sup>8</sup>Bare NPs without number markers can be either definite or indefinite in both Persian and Yi.

<sup>9</sup>This is the translation given by Gerner and Bisang (2010). An alternative translation that reflects the definiteness of the noun phrase is ‘The ancestors have *this* story: ...’.

- d. pi<sup>55</sup> di<sup>34</sup>      nie<sup>55</sup> hi<sup>44</sup> zau<sup>55</sup>.  
 our PL.INDF tooth not good  
 ‘Some of our teeth are not in a good state.’

Given the presence of determiners that can co-occur with classifiers in Persian and Yi, it is not adequate to analyze these classifiers and plural markers as determiners *per se*. A more plausible analysis is that they inflect agreeing with null determiners.

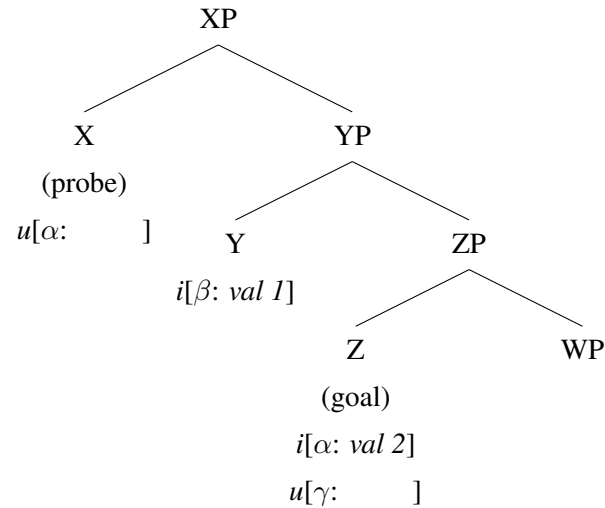
Let us now look at the details of the agreement between D and Num. D hosts determiners whereas Num hosts classifiers and plural markers. I adopt the theory of Agree proposed by Chomsky (2000, 2001). Agree occurs between two features. Two kinds of features are distinguished, i.e. those which are interpreted at LF and those not; both can, but do not have to, be reflected in the phonology. I will mark interpretable and uninterpretable features with prefixes *i* and *u* respectively. Features consist of an attribute and a value: [ATT: *val*]. Interpretable features have a value whereas uninterpretable ones do not.<sup>10</sup> Uninterpretable features remaining at LF make a syntactic derivation crash, and must be deleted in the course of derivation. Chomsky suggests that deletion is carried out by Spell-Out/Transfer. Agree is an operation that matches an uninterpretable feature with an interpretable feature with the same attribute and copies the value of the latter to the former. In the Agree operation, an uninterpretable/unvalued feature functions as a ‘probe’ and searches its c-command domain for the closest interpretable/valued feature (‘goal’). One of the important properties of Chomsky’s Agree is the requirement that both probe and goal be active by having an uninterpretable feature. This property captures interdependency between two features, which explains licensing of one category by another, including the licensing of Num (classifiers and plural markers) by D.

(32) shows the workings of Agree.

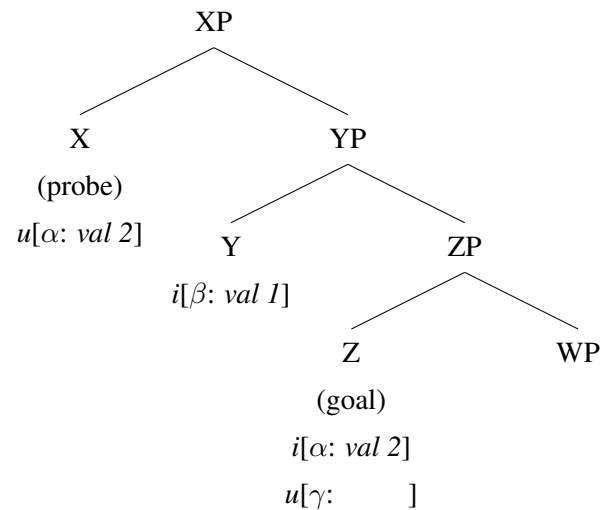
---

<sup>10</sup>Pesetsky and Torrego (2007) reject this biconditional relationship between interpretability and valuation. The analysis of D-Num agreement proposed below does not hinge on whether the biconditional relationship is assumed.

(32) a. Before Agree



b. After Agree

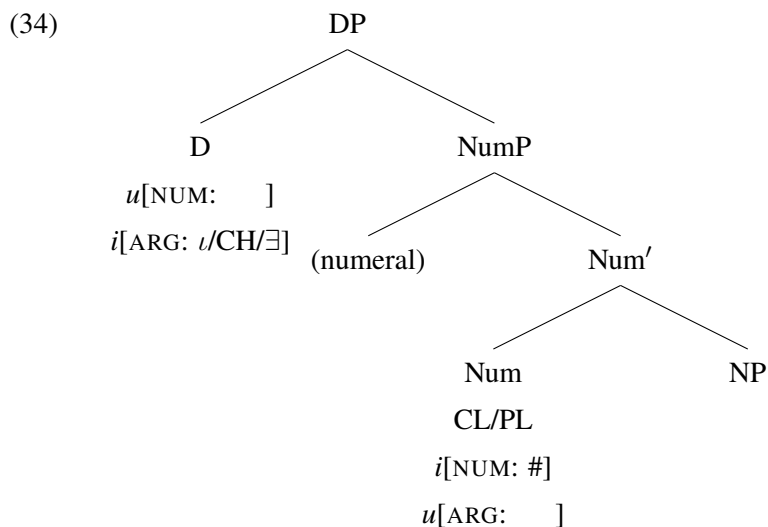


The uninterpretable  $\alpha$  feature functions as a probe and searches its c-command domain for a feature with the same attribute. Upon detecting the closest goal, i.e. the  $\alpha$  feature in Z, it is valued by the goal. Z is active, as it also bears an uninterpretable feature, i.e. the  $\gamma$  feature. To take the noun-verb agreement (e.g. *She plays the violin*) for example, the relevant probe and goal are T and D respectively. T has uninterpretable  $\phi$ -features whereas D has interpretable  $\phi$ -features, as shown in (33a). In addition, D has an uninterpretable Case feature. The presence of this uninterpretable feature makes D active. T has an ability to assign structural nominative Case. While Chomsky does not implement Case assignment by means of Case feature matching, Radford

(2009:404) suggests that T has an uninterpretable Case feature with the value ‘Nom(inative)’, differing from Chomsky in allowing for uninterpretable features with values like Pesetsky and Torrego (2007). I adopt Radford’s implementation here. This decision is largely for the sake of exposition and is orthogonal to my analysis of D-Num agreement presented below. The uninterpretable  $\phi$ -features of T trigger Agree. They are assigned the values ‘3SgF’ by their interpretable counterpart in D, and concurrently the uninterpretable Case feature of D is assigned the value ‘Nom’ by (the Case feature of) T, as shown in (33b).

- (33)
- |    |  |  |                     |                                       |      |  |
|----|--|--|---------------------|---------------------------------------|------|--|
|    | T  |  | [ <sub>vP</sub> [DP | D                                     | ...] |  |
| a. | Before Agree: $u$ [CASE: Nom], $u$ [ $\phi$ : ]    |  |                     | $u$ [CASE: ], $i$ [ $\phi$ : 3SgF]    |      |  |
| b. | After Agree: $u$ [CASE: Nom], $u$ [ $\phi$ : 3SgF] |  |                     | $u$ [CASE: Nom], $i$ [ $\phi$ : 3SgF] |      |  |

Let us turn to the D-Num agreement responsible for the licensing of classifiers and plural markers. (34) shows the relevant part of a DP containing them.



I hypothesize that D has an ARGument feature. This feature is concerned with the argumenthood of DPs. The possible values of the feature thus include  $\iota$  (definite, referential), CH (indefinite, referential/specific) and  $\exists$  (indefinite, non-referential/non-specific), which are interpreted as respective property-to-argument type-shifting operators. I assume that property-denoting noun phrases, including incorporated noun phrases, do not project to the DP level. Given that determiners inflect for number in many languages, D should also have a NUMBER feature. However, it

is not D but Num where the number information is encoded and interpreted. Hence, the number feature in D is uninterpretable. Num is thought to have an uninterpretable argument feature, whose presence makes Num active. Argumenthood is thus comparable to Case in the verbal domain. Whether it is implemented by a feature of the probe or other means, it is necessary for number (= a  $\phi$ -feature) agreement to take place. Positing it explains not only the licensing of classifiers and plural markers by referential determiners ( $\iota$ , CH) and quantifiers ( $\exists$ ) but also inflecting classifiers and plural markers in Yi and Weining Ahmao discussed above. The value of the number feature in Num is just specified as # in the tree above. I will elaborate on it in the next section.

The agreement proceeds as follows. The uninterpretable number feature of D triggers Agree, which matches the number features of D and Num. Consequently, it is assigned a value by the interpretable number feature of Num, and the uninterpretable argument feature of Num is assigned a value. If Chomsky's view on Case is extended to argumenthood, the latter valuation takes place "as a kind of 'bonus,'" to borrow Pesetsky and Torrego's (2007) words. This is because in Chomsky's system, the valuation of the uninterpretable Case feature of D in (33) is no more than a reflex of the  $\phi$ -feature agreement and is not a result of a distinct Case feature agreement. One way of making sense of such "bonus" valuation is to think that the path established by the valuation process involved in Agree can be reused by other features hosted by the same heads for valuation. The relevant valuation path is 'T-D' for  $\phi$ -features in (33) and 'D-Num' for the number feature in (34). Additional valuation can occur in any direction because unlike the complete Agree operation, the valuation component alone does not require the probing of the syntactic tree, and hence directionality and locality are not important.<sup>11</sup> The valued uninterpretable features are deleted later when the structure is spelled out.

The proposed analysis suggests that truly bare argument noun phrases, i.e. argument noun phrases that do not have either overt number markers, numerals or determiners as in *hon* 'book' in Japanese, also project NumP and DP. Otherwise they cannot function as arguments. This is because Chomsky's Agree necessarily involves two different features. If a noun phrase is an

<sup>11</sup>In this analysis, the "bonus" downward valuation appears as if a second Agree had taken place in the opposite direction. Thus, it is also possible that the "bonus" valuation in fact involves a second application of Agree. This alternative analysis calls for some modification to Chomsky's theory of Agree, as the first and second instances of Agree, probing as well as valuation, appear to operate in opposite directions. There is a debate about the directionality of probing and valuation in Agree (e.g. Baker 2008; Zeijlstra 2012; Wurmbrand 2012; Diercks et al. 2012; Carstens 2013). Discussing it is beyond the scope of this study.

argument, it must contain D, which has an interpretable argument feature, the source of the argument interpretation. Crucially, D also contains an uninterpretable number feature. Projecting only DP without NumP is impossible, because the uninterpretable number feature of D must be valued by agreeing with its interpretable counterpart in Num. This agreement is impossible unless Num is active by having an uninterpretable feature, which I identified as an argument feature. Number and argument features are interdependent in this manner.

The number feature involved in truly bare argument noun phrases has a value associated with the general number. Therefore, it is quintessential for any theory of number features to articulate the feature value(s) associated with the general number, which I do in the next section.

## 6.3 The typology of number marking and number features

### 6.3.1 The typology of number marking

Classifier languages distinguish three number categories, i.e. singular (SG), plural (PL) and general (GN). I regard these three number categories as universally available basic number categories. The singular and the plural refer to ‘one’ and ‘more than one’ respectively. The general number subsumes both the singular and plural, i.e. it does not distinguish between the two or number-neutral. I will decompose these three number categories into two number features in section 6.3.2 below. However, in this section, I will continue to use the labels ‘singular’, ‘plural’ and ‘general’ as if they were unanalyzable primitive concepts, to avoid unnecessary complications that could arise due to the use of features.

Languages vary in how many basic number categories are morphologically distinguished and how the distinction is made (cf. Corbett 2000). (35) shows the four possible language types based on these criteria.

- (35) The typology of basic number systems
- a. Three-way distinction
    - SG : GN : PL
  - b. Two-way distinction
    - (i) SG/GN : GN/PL
    - (ii) SG/GN : PL
    - (iii) SG : GN/PL

- c. No distinction  
SG/GN/PL

Assuming that all languages have at least one morphologically unmarked number category and that the unmarked number category is the singular when the general is morphologically conflated with another category (Corbett 2000:17),<sup>12</sup> only languages of the ‘SG : GN : PL’ type can have phonologically overt singular markers, which are normally identified as classifiers.<sup>13</sup> Classifier languages thus make the finest possible distinction of the basic number categories, with the three-way distinction ‘SG : GN : PL’ number system, where the unmarked category is GN. Not all languages with the ‘SG : GN : PL’ basic number system are known as classifier languages. The singular morphology is analyzed as a true singular marker in the Cushitic language Bayso, for example.<sup>14</sup>

(36) Bayso (Dick Hayward, p.c. cited in Corbett 2000:11)

- a. lúban-titi  
lion-SG  
‘a lion’
- b. lúban  
lion.GN  
‘lion’ (it could be one, or more than that)
- c. lúban-jool  
lion-PL  
‘(a lot of) lions’

<sup>12</sup>This assumption can be formulated as a markedness constraint of the Optimality Theory, as in (i) or (ii) in terms of the feature system to be presented shortly, where ‘X/∅’ means ‘category/feature X is morphologically unmarked’.

(i) \*PL/∅ >> \*SG/∅ >> \*GN/∅

(ii) \*[+Pl]/∅ >> \*[+Sg]/∅

<sup>13</sup>Similar conclusions have been reached by Greenberg (1972) and Rullmann and You (2006). Greenberg links the appearance of classifiers to number systems with a three-way distinction found in some Arabic varieties and Slavic languages, where the transnumeral category of the collective contrasts with the singulative, which in turn divides into the singular and the plural (and the dual in Arabic varieties). Rullmann and You hypothesize that languages with classifiers form a proper subset of the class of languages in which morphologically unmarked nouns are associated with the general number.

<sup>14</sup>The language also has the paucal number, which I omit here.

Corbett (2000:13) states that languages of the ‘SG : GN : PL’ type are not widespread. However, this is only because he does not regard classifiers as an expression of number. According to his description, the ‘SG : GN : PL’ system is not stable or restricted to certain nouns in non-classifier languages, coexisting with other number systems.

In other language types, the singular must be morphologically unmarked. Brazilian Portuguese and Singlish (Colloquial Singapore English) have the ‘SG/GN : GN/PL’ number system, where some GN nominals are conflated with SG and morphologically unmarked while others are conflated with PL and bear plural morphology (cf. Schmitt and Munn 2002; Gil 2003; Kim et al. 2009).

(37) Brazilian Portuguese

- a. Eles têm filho? — Sim, eles têm {um filho/ dois filho-s}.  
 they have child yes they have one child two child-PL  
 ‘Do they have children? — Yes, they have {one child/two children}.’
- b. Eles têm filho-s? — Sim, eles têm {um filho/ dois filho-s}.  
 they have child-PL yes they have one child two child-PL  
 ‘Do they have children? — Yes, they have {one child/two children}.’

(38) Singlish

- a. You got cat? — Yes, I got {one/two}.<sup>15</sup>  
 ‘Do you have cats? — Yes, I have {one/two}.’
- b. You got cats? — Yes, I got {one/two}.  
 ‘Do you have cats? — Yes, I have {one/two}.’

The judgments shown in (37) is my consultant’s. Müller and Oliveira (2004) present sentences that are similar to (37) but judged differently, as in (39).

- (39) a. Ela tem filho? — Sim, ela tem {um filho/ dois filho-s}.  
 she has child yes she has one child/ two child-PL  
 ‘Does she have children? — Yes, she has {one child/two children}.’

<sup>15</sup>The unmarked noun *child* is not felicitous in this context. I consider this as a lexical idiosyncrasy related to the word’s irregular inflection or high frequency of use.

- b. Ela tem filho-s? — Sim, ela tem {\*um filho/ dois filho-s}.  
 she has child-PL yes she has one child/ two child-PL  
 ‘Does she have children? — Yes, she has {one child/two children}.’

(Müller and Oliveira 2004:24)

Unlike the dialect of my consultant, morphologically plural forms refer exclusively to pluralities in their dialect. The latter dialect should then have a slightly different number system, i.e. the ‘SG/GN : PL’ system, where the general is conflated with the singular, but not with the plural.<sup>16</sup>

English belongs to the ‘SG : GN/PL’ type, though Corbett (2000:19) classifies it into a category of languages that lack the category of the general (‘SG : PL’ type).<sup>17</sup> The so-called “plural” nouns in English with the suffix *-s* can actually entail the general number (e.g. *How many cats do you have?*), with the plural meaning obtained pragmatically. This does not necessarily mean that English lacks the category of the plural. Farkas and de Swart (2010) propose a generalization that morphologically plural forms are interpreted as plural in upward entailing contexts as in (40a) whereas they are interpreted as number-neutral in downward entailing contexts and questions as in (40b–c).<sup>18</sup>

- (40) a. Mary saw horses. (\*a horse,  $\sqrt{\text{more than one horse}}$ )  
 b. Mary has never seen horses. ( $\sqrt{\text{a horse}}$ ,  $\sqrt{\text{more than one horse}}$ )  
 c. Has Mary ever seen horses? ( $\sqrt{\text{a horse}}$ ,  $\sqrt{\text{more than one horse}}$ )

(adapted from Farkas and de Swart 2010:2–3)

It is thus plausible to think that the categories of the plural and general both exist in English, but the morphology of English does not differentiate them. As Farkas and de Swart (2010) claim, the suffix *-s* is polysemous. Finally, languages that do not have number marking on nouns such

<sup>16</sup>It must be noted, however, that this conclusion is incompatible with another fact reported by Müller and Oliveira (2004), namely that bare plurals denote kinds and are scopeless. As seen in chapter 3, plurals in classifier languages, whose exclusive association with the plural is unquestionable, do not show this property.

<sup>17</sup>The ‘SG : PL’ type does not exist in my system, as the general is hypothesized to be a universally available number category.

<sup>18</sup>They explain this generalization based on the Strong Meaning Hypothesis (Dalrymple et al. 1998), which states that when an expression is assigned multiple interpretations ordered by entailment, the strongest interpretation compatible with the context is chosen. In upward entailing contexts, a plural interpretation is stronger than a number-neutral interpretation. In the case of (40a), Mary seeing two or more horses entails her seeing one or more horses, but not the other way round. In downward entailing context, the reverse holds. A number-neutral interpretation entails a plural interpretation. Mary not seeing one or more horses entails her not seeing two or more horses in (40b). Questions resemble downward entailing contexts in many respects (e.g. negative polarity item licensing).

as Dëne Sùliné (Wilhelm 2008) belong to the ‘SG/GN/PL’ type.

The proposed typology of basic number systems predicts that number marking (on nouns) may look optional in languages with the ‘SG : GN : PL’ and ‘SG/GN : (GN/PL)’ systems. We have already seen that classifiers (i.e. singular) and plural markers in classifier languages are ostensibly optional, to different degrees in different languages. As for ‘SG/GN : (GN/PL)’ languages, variability of plural marking has been noted in the literature (e.g. Brazilian Portuguese by Thomas (1969), Singlish by Ho and Platt (1993)). In both types of languages, the general number (conflated with the singular) is morphologically unmarked. As the general subsumes both the singular and plural, plural number marking appears to be optional in these languages.

Integrating classifiers and general number into the basic number system, my theory offers a novel way of understanding the difference between classifier and non-classifier languages. The difference originates from morphology, specifically a number morphology difference, which in turn leads to certain syntactic and semantics differences. Many previous studies have sought the difference in the semantics of nouns or numerals (cf. Krifka 1995; Chierchia 1998b; Wilhelm 2008). As I have pointed out in chapters 2 and 3, these studies are based on an assumption that I think is problematic, i.e. classifiers enable counting through individuation. In this study, I have kept myself from this assumption and investigated classifier languages. The conclusion thus reached is nothing sensational. Crosslinguistic syntactic variations are due to morphological differences, but not to semantic parameters such as Chierchia’s (1998b) Nominal Mapping Parameter. Classifier languages have a three-way distinction of basic number categories, where classifiers encode the singular. Non-classifier languages, conversely, morphologically conflate the general with one or both of the other categories. Given the unmarked status of the singular, classifiers cannot exist in these languages. Li (1999) proposes a syntactic parameter that distinguishes classifier languages from non-classifier languages, i.e. the former, but not the latter, project a classifier phrase (CIP). Analyses of a similar vein has been proposed by other authors (e.g. Sato 2008; Butler 2011). I think that a syntactic parameter analysis of crosslinguistic variations should be avoided if a morphological account is possible that can handle the facts at issue.

### 6.3.2 Number features

As I suggested at the outset of the last section, the three basic number categories, i.e. singular, plural and general, are not primitive concepts but only labels. They can be decomposed into

Table 6.3: Four basic number categories and their feature compositions.

Number	Features	Morphology	Denotation
SG	[+Sg], [-PI]	singular	singularities alone
PL	[-Sg], [+PI]	plural	pluralities alone
GN <sup>+</sup>	[+Sg], [+PI]	singular + plural	singularities + pluralities (= number-neutral)
GN <sup>-</sup>	[-Sg], [-PI]	none	non-singularities + non-pluralities (= pluralities + singularities = number-neutral)

two binary features, i.e. [ $\pm$ Sg] and [ $\pm$ PI]. These number features are associated with the Num head in the DP structure presented in section 6.2.4.1 above. As four possible combinations are logically available for them, there are in fact four basic number categories, rather than just three. Specifically, two kinds of general can be distinguished, which I refer to as GN<sup>+</sup> and GN<sup>-</sup>. The four number categories and their feature compositions are summarized in Table 6.3. As can be seen in the table, the two number features have both morphological and semantic reflections.

### 6.3.2.1 Morphology

With regard to morphology, typically the positive value indicates the presence of the relevant marking while the negative one its absence. There are four possible number marking patterns as shown in (41).

- (41) a. SG: [+Sg], [-PI]  $\leftrightarrow$  singular marking (no plural marking)  
 b. PL: [-Sg], [+PI]  $\leftrightarrow$  plural marking (no singular marking)  
 c. GN<sup>+</sup>: [+Sg], [+PI]  $\leftrightarrow$  singular and plural marking  
 d. GN<sup>-</sup>: [-Sg], [-PI]  $\leftrightarrow$  no marking

Although GN<sup>+</sup> and GN<sup>-</sup> have the same denotation, they differ morphologically. GN<sup>+</sup> involves singular as well as plural marking whereas GN<sup>-</sup> involves neither. Crosslinguistic variations in number marking are understood as variations concerning how the two features are realized, which can be summarized as in Table 6.4. ‘—’ indicates that the relevant category is missing in syntax. The singular marker for ‘SG : GN : PL’ languages is represented by CL (classifier) because it is analyzed as such in many languages.

Table 6.4: Crosslinguistic variations in number feature realizations.

	SG	GN <sup>-</sup>	GN <sup>+</sup>	PL
Number system	[+Sg], [-Pl]	[-Sg], [-Pl]	[+Sg], [+Pl]	[-Sg], [+Pl]
SG : GN : PL	CL	bare	CL + PL	PL
SG/GN : GN/PL	∅	bare	∅ + PL	PL
SG/GN : PL	∅	bare	—	PL
SG : GN/PL	∅	—	∅ + PL	PL
SG/GN/PL	∅	bare/—	—/∅ + ∅	∅

Many classifier languages distinguish all four number categories morphologically, as exemplified by the Japanese phrases in (42). The number morphology is indicated by *italics*.

- (42) a. SG: [+Sg], [-Pl]  
 (san) *nin-no*    *gakusei*  
 three CLF-LINK student  
 ‘(three) students’
- b. PL: [-Sg], [+Pl]  
*gakusei-tati*  
 student-PL  
 ‘students’
- c. GN<sup>+</sup>: [+Sg], [+Pl]  
 (san) *nin-no*    *gakusei-tati*  
 three CLF-LINK student-PL  
 ‘(three) students’
- d. GN<sup>-</sup>: [-Sg], [-Pl]  
*gakusei*  
 student  
 ‘student(s)’

The presence of languages in which classifiers and plural markers co-occur justify hypothesizing the feature combination ‘[+Sg], [+Pl]’. These languages realize the two features as separate morphemes, specifically as a classifier and a plural marker respectively. Another way of realizing them is to realize them as a single morpheme. Such morphemes are known as ‘plural classifiers’. Plural classifiers are found in languages with bare classifier constructions such as

Cantonese, Hmong and Bangla. Examples in Cantonese are given in (43).

- (43) a. SG: [+Sg], [−PI]  
*go* hohksāang  
 CLF student  
 ‘(the) student’
- b. PL: [−Sg], [+PI]  
*dī* hohksāang  
 PL student  
 ‘(the) students’
- c. GN<sup>+</sup>: [+Sg], [+PI]  
*dī* hohksāang  
 CLF.PL student  
 ‘(the) students’
- d. GN<sup>−</sup>: [−Sg], [−PI]  
 hohksāang  
 student  
 ‘student(s)’

My analysis of the plural classifier as being ambiguous between PL and GN<sup>+</sup> here is tentative. It is also possible that Cantonese lacks a realization for ‘[−Sg], [+PI]’ (i.e. the realization of [+PI] is conditioned by [+Sg]) and plural interpretations arise pragmatically, but not semantically. At any rate, there is evidence that noun phrases with *dī* denote number-neutral properties. According to Au-Yeung (2007), besides bare NPs, which are categorized into GN<sup>−</sup>, noun phrases with *dī* can also refer to kinds, as in (44). In my analysis of kinds presented in chapter 3, the kind-deriving operator <sup>□</sup> can only apply to number-neutral properties.

- (44) (Dī) sāigwāa jauh faai jyuhtjúng la.  
 CLF.PL watermelon FOC soon extinct PART  
 ‘Watermelons will become extinct soon.’

(Au-Yeung 2007:2)

Moreover, I expect that the negative sentence in (45) is true not only when there exist two or more doctors who are as bad as the interlocutor but also when at least one such doctor exists.

- (45) Saigaai seuhng móuh (dī) yīsāng hóuchíh néih gam sēui.  
 world up not.exist CLF.PL doctor like you so bad  
 ‘In the world, there are no such doctors so bad as you.’

(Au-Yeung 2007:10)

*Dī* plurals in Cantonese thus pattern with plurals in English in that a number-neutral interpretation is normally unavailable/unnoticed, but becomes available/visible in certain contexts such as kind-referring expressions and negatives (cf. (40)).

In non-classifier languages, the singular marker is phonologically null. They have a null classifier, as it were (cf. Sharvy 1978). As a consequence, no audible difference exists between SG and GN<sup>-</sup>, and between PL and GN<sup>+</sup>. At the descriptive level, these situations amounts to saying that bare singular/plural nouns can be number-neutral. In languages with the ‘SG/GN : GN/PL’ such as Brazilian Portuguese and Singlish, all four number categories are available in syntax like classifier languages. As SG and GN<sup>-</sup> have the same form, a morphologically singular unmarked noun phrase can refer to multiple referents. Moreover, it triggers singular, but not plural, verbal agreement even when receiving such an interpretation as in (46)–(47), because neither SG nor GN<sup>-</sup> contains [+PI].

- (46) Brazilian Portuguese (Schmitt and Munn 1999:342)

- a. Chegou criança.  
 arrived.3SG child  
 ‘A child/children arrived.’
- b. Chegaram criança-s.  
 arrived.3PL child-PL  
 ‘Children arrived.’

- (47) Turkish (Kornfilt 1996:118–119)<sup>19</sup>

- a. İki öğrenci ben-imi-le gör-üş-mek iste-di(\*-ler).  
 two student I-GEN-INST see-RECP-INF want-PST-3PL  
 ‘Two students wanted to meet with me.’
- b. Bazı öğrenci-ler ben-imi-le gör-üş-mek iste-di-ler.  
 some student-PL I-GEN-INST see-RECP-INF want-PST-3PL  
 ‘Some students wanted to meet with me.’

<sup>19</sup>I follow Kornfilt (1997) and assume that Turkish is not a classifier language, though it has a few classifier-like elements (e.g. *tane* ‘item’).

In other non-classifier languages, either  $GN^-$  or  $GN^+$  is missing in syntax. By “missing in syntax” I mean that the associated number features are relegated to the lexicon (cf. Acquaviva 2004), and hence do not trigger count syntax. I claim that relegated, or lexicalized, number features are dedicated to the so-called ‘fake mass nouns’ found in non-classifier languages, such as *furniture* and *footwear* in English. Syntactically, fake mass nouns pattern with true mass nouns such as *water*. They do not take plural markers, trigger singular agreement, and so on, because no number feature exists in syntax. However, they are conceptually count, that is, they have clearly perceivable atoms (i.e. naturally atomic), because they have number features built in themselves. While these features are not accessible to syntax, they are interpreted (as number-neutral properties) by semantics. This captures Chierchia’s (1998a, 1998b) insight that fake mass nouns denote the entire lattice.

One might wonder if it is necessary to hypothesize multiple number categories for ‘SG/GN/PL’ languages, which appear to lack not only morphological number distinction but also number categories. However, examples in Dëne Sųliné in (48) suggest the possibility that distinct number categories actually exist and affect the verb forms, possibly by means of number agreement. The object noun phrases in (48a) and (48b) are identical in form (i.e. *tth’áy*) but interpreted differently, either as singular or plural, depending on the number inherent to the verb stem.

- (48) a. Tth’áy thıł-tşı sı.  
dish PRF.1SG.SBJ-make.SG.OBJ.PRF assert  
‘I made a (one) dish.’
- b. Tth’áy ghı-ghą sı.  
dish PRF.1SG.SBJ-make.PL.OBJ.PRF assert  
‘I made several dishes.’

(Wilhelm 2008:45)

### 6.3.2.2 Semantics

As for the semantic reflection, the values of the two features determine the denotation of the ‘Num NP’ constituent, which is subject to further modification by numerals, if any (cf. chapter 4, section 4.3). The two features are defined as in (49). I assume that NPs (before undergoing number specification by Num) universally denote the entire lattice, i.e. the property  $P$  in (49) is number-neutral. Under this assumption, [+Sg] and [+Pl] concern atomicity and divisibility respectively. They restrict the entire lattice to singularities and pluralities respectively.

- (49) a.  $\llbracket [+Sg] \rrbracket$  is defined iff  $\text{MIN}(P) \neq \emptyset$ , where  
 $\text{MIN}(P) = \{x \in P : \neg \exists y \in P[y < x]\}$  (Minimality)  
 When defined,  $\llbracket [+Sg] \rrbracket = \lambda P \lambda x. x \in \text{MIN}(P)$
- b.  $\llbracket [+Pl] \rrbracket = \lambda P : \underline{\text{MIN}(P) \neq \emptyset} . \lambda x. x \in \text{DIV}(P)$ , where  
 $\text{DIV}(P) = \{x \in P : \exists y \in P[y < x]\}$  (Divisibility)
- c. Feature Negation (Harbour 2007:68)  
 $\llbracket -F \rrbracket = \neg \llbracket +F \rrbracket$   
 (i)  $\llbracket [-Sg] \rrbracket = \lambda P \lambda x. x \notin \text{MIN}(P) (= \llbracket [+Pl] \rrbracket)$ ; defined iff  $\text{MIN}(P) \neq \emptyset$   
 (ii)  $\llbracket [-Pl] \rrbracket = \lambda P : \text{MIN}(P) \neq \emptyset . \lambda x. x \notin \text{DIV}(P) (= \llbracket [+Sg] \rrbracket)$

The underlined portion in (49b) is a presupposition to the effect that the lattice denoted by an NP is atomistic. Count nouns satisfy the presupposition with no problem. ‘Num NP’ refers to multiple objects, a “plural reading” in the normal sense. The presupposition is not satisfied by mass nouns (, which do not include fake mass nouns such as *furniture* and *footwear*, as they have inherent number specification in my theory).<sup>20</sup> In this case, an implicit container/unit is introduced as in *two ice creams* to accommodate the atomicity presupposition. Pelletier (1979) refers to this function of plural morphology as ‘universal packager’. Plural mass nouns are infelicitous if such accommodation is impossible. The atomicity presupposition is either weak or absent in some languages. In such languages, plural markers occur not only with count but also with mass nouns.<sup>21</sup> Such mass plurals express abundance (‘lots of X’), as nothing in the semantics of [+Pl] imposes the lower bound at which further division becomes impossible.<sup>22</sup> Mass plurals expressing abundance have been reported in Persian (Sharifian and Lotfi 2003; Hamedani 2011), Greek (Tsoulas 2006; Alexiadou 2011) and Innu-aimun (Gillon 2010). According to Myint Soe’s (1999) description, Burmese has such plurals too.

<sup>20</sup>Remember that nouns are systematically ambiguous between object and subkind readings (cf. chapter 3). Since all nouns are count in their subkind reading, nouns that are mass in the object reading satisfy the atomicity presupposition in the subkind reading.

<sup>21</sup>As compatibility with plural markers is just one of the diagnostics for count-mass distinction, being compatible with plural markers does not automatically mean that nouns are count. Other diagnostics include quantifiers, numerals and classifiers.

<sup>22</sup>Harbour (2009) proposes a similar analysis. One crucial difference between my analysis and his is that the same feature, i.e. [+Pl], applies to both count and mass nouns in my analysis whereas Harbour’s analysis invokes two different features, i.e. [–singular] for count nouns and [+augmented] for mass nouns. One fact that my analysis captures better than Harbour’s is that the same plural morphology is employed for both count and mass plurals. In this connection, the phenomenon of mass plurals discussed here should not be confused with the phenomenon referred to as ‘greater plural’ by Corbett (2000), where a single count noun has two plural forms, one expressing ordinary plurality and the other expressing abundance.

- (50) Persian (Hamedani 2011:7)  
 Vaqti dâšt keyk dorost mi-kard, šîr(-hâ)-ro rixt, ru-ye zamin.  
 when had cake make DUR-do.PST.3SG milk-PL-OM spill.PST.3SG on-MOD floor  
 ‘When he/she was making cake, he/she spilt the milk on the floor.’  
 cf. pesar-hâ [boy-PL], ketâb-hâ [book-PL]
- (51) Burmese (Myint Soe 1999:60)  
 Hsì twèi hpei’ thwá pì.  
 oil PL spill go PUNC  
 ‘The oil has spilt.’  
 cf. hké=dàn twèi [pencil PL], sà=ou’ twèi [book PL]
- (52) Innu-aimun (Gillon 2010:17)
- |    |                                   |                            |
|----|-----------------------------------|----------------------------|
| a. | âshûkan [bridge]                  | âshûkan-a [bridge-PL]      |
|    | pâushtik <sup>u</sup> [waterfall] | pâushtiku-a [waterfall-PL] |
|    | metuâkan [toy]                    | metuâkan-a [toy-PL]        |
| b. | nekâu [sand]                      | nekâu-a [sand-PL]          |
|    | neneu [breath]                    | neneu-a [breath-PL]        |
|    | nîpîsh [tea]                      | nîpîsh-a [tea-PL]          |

These facts are important, because they show that  $[\pm\text{Pl}]$  has a motivation different from  $[\pm\text{Sg}]$ . Hence, even though the two features are effectively redundant in most languages, as seen in (49c) ( $[-\text{Sg}] = [+Pl]$ ,  $[-Pl] = [+Sg]$ ), they are both necessary as they are conceptually distinct.

A combination of two features is interpreted as the union of them as in (53).

- (53) a.  $[[+\text{Sg}], [-\text{Pl}]] = [[+\text{Sg}]] \cup [[+\text{Sg}]] = [[+\text{Sg}]]$   
 $= \lambda P \lambda x. x \in \text{MIN}(P)$  (singular)
- b.  $[[-\text{Sg}], [+Pl]] = [[+Pl]] \cup [[+Pl]] = [[+Pl]]$   
 $= \lambda P : \text{MIN}(P) \neq \emptyset. \lambda x. x \in \text{DIV}(P)$  (plural)
- c.  $[[+\text{Sg}], [+Pl]] = [[+\text{Sg}]] \cup [[+Pl]]$   
 $= \lambda P : \text{MIN}(P) \neq \emptyset. \lambda x. x \in \text{MIN}(P) \cup \text{DIV}(P)$  (number-neutral)
- d.  $[[-\text{Sg}], [-Pl]] = [[+Pl], [+Sg]] = (53c)$

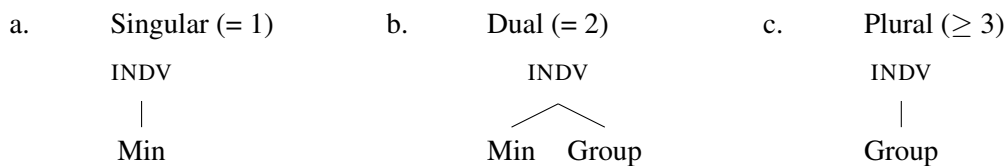
Notice that the denotations of ‘ $[\text{+Sg}], [\text{+Pl}]$ ’ (53c) and ‘ $[-\text{Sg}], [-\text{Pl}]$ ’ (53d) are identical.

### 6.3.2.3 A comparison with previous theories of number features

Many other recent studies on number also attempt to capture grammatical and semantic numbers by two UG number features (e.g. Noyer 1992; Harley and Ritter 2002; McGinnis 2005; Harbour 2007, 2011; Nevins 2011). My theory of number features differs from those proposed by these studies (“previous studies” hereafter) in that it includes the general number as one of the basic number categories besides the singular and plural. Previous studies, on the other hand, do not pay much attention to the general. In what follows, I point out three problems of the previous studies that do not arise in my theory.

First, previous studies cannot handle the general number well. The two number features employed in these studies consist of one concerning atomicity/minimality and the other indicating whether an individual is made up of more than one smallest element. They are Minimal and Group in Harley and Ritter (2002) and [ $\pm$ singular] and [ $\pm$ augmented] in Harbour (2007, 2011).<sup>23</sup> The combinations of these features represent three number categories, i.e. singular, dual and plural, as in (54)–(55). INDV and Min in (54) stand for INDIVIDUATION and Minimal respectively. The INDIVIDUATION node represents features independent of the discourse, i.e. features related to number and gender, but not person (see Figure 6.1 below). [ $+$ augmented][ $+$ singular] is undefined in Harbour’s system.

(54) Harley and Ritter (2002)



<sup>23</sup>Harley and Ritter (2002) do not provide explicit definitions for Minimal and Group. They write “[s]ingular entities are represented by the Minimal node,” whereas “Group identifies multiple entities, and its presence represents plural number” (p. 490). They also associate Minimal to the intuition of “smallest possible” (p. 492). Harbour (2007, 2011), on the other hand, defines the features explicitly. In essence, his [ $\pm$ singular] and [ $\pm$ augmented] do not differ from my [ $\pm$ Sg] and [ $\pm$ PI] (see for (49) definitions) in that they are concerned with atomicity and divisibility respectively. (i) shows the definitions of the two features in Harbour (2011), where “ $\lambda z : \phi$ ” means that  $\phi$  is a presupposition on  $z$ .

(i) a. [ $+$ singular] =  $\lambda x[\text{atom}(x)]$   
 b. [ $+$ augmented] =  $\lambda P.\lambda x : P(x).\exists y[P(y) \wedge y \sqsubset x]$

- (55) Harbour (2007)
- a. Singular (= 1)  
[−augmented]( [+singular])
  - b. Dual (= 2)  
[−augmented]( [−singular])
  - c. Plural ( $\geq 3$ )  
[+augmented]( [−singular])

In Harley and Ritter’s system, languages without the dual lack (54b) and omit Min in (54a), as it is the default value of the INDIVIDUATION node. Without (54b), (54c) is associated with the plural in the normal sense ( $\geq 2$ ). Conversely, in Harbour’s system, such languages conflate the dual with the plural ( $\geq 3$ ). It is not clear how the general is expressed in either system.

McGinnis (2005) analyzes the general as a bare INDIVIDUATION node, as in (56a). Likewise, one could argue that the general lacks the number features altogether, with no number-related functional projection (e.g. NumP) projected.

- (56)
- |            |             |           |                        |
|------------|-------------|-----------|------------------------|
| a. General | b. Singular | c. Dual   | d. Plural ( $\geq 3$ ) |
| INDV       | INDV        | INDV      | INDV                   |
|            |             | / \       |                        |
|            | Min         | Min Group | Group                  |

Although this analysis is able to distinguish the general from other number categories semantically, it cannot account for the crosslinguistic variations in the morphological realizations of the general number discussed above and summarized in Table 6.4. Indeed, general (more specifically  $GN^-$ ) noun phrases are morphologically unmarked as (56a) predicts, in all language types other than the ‘SG : GN/PL’ type (e.g. English). However, the general number (specifically  $GN^+$ ) is also associated with morphologically plural forms. For ‘SG : GN/PL’ type languages, one could argue that General (56a) is conflated with Dual (56c) and Plural (56d). However, a conflation analysis does not work for ‘SG : GN : PL’ (e.g. Cantonese, Japanese) and ‘SG/GN : GN/PL’ (e.g. Singlish, Brazilian Portuguese). This is because in such an analysis one would have to split (56a) into two kinds: one is conflated with Singular (56b) and morphologically unmarked whilst the other is conflated with Dual (56c) and Plural (56d) and triggers plural

morphology. In short, the analysis above loses the tight connection between morphology and semantics that Harley and Ritter's feature geometry theory intends to capture. By contrast, the feature system that I propose captures the connection successfully.

The second problem pertains to the way in which the meaning of a combination of two number features is calculated. I explain this point by taking the dual for example, as it involves two features in both Harley and Ritter (2002) and Harbour (2007). The meaning of the dual is calculated by restricting the set of non-singular individuals to those consisting of the minimal number (= 2) of elements. While this is unproblematic in Harbour's system, it is unnatural in Harley and Ritter's system, where number features belong to the same feature geometry as person features shown in Figure 6.1. The features under the PARTICIPANT node are the person features.

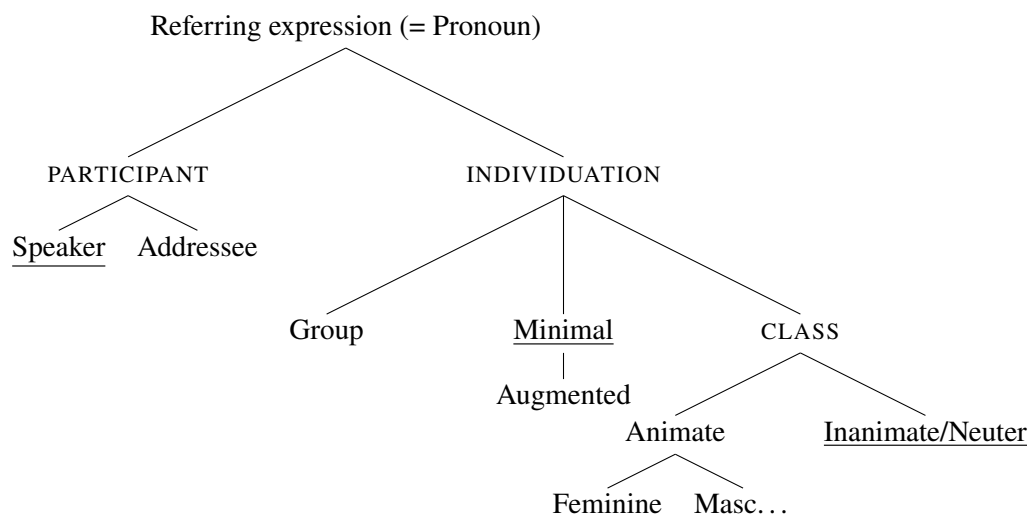
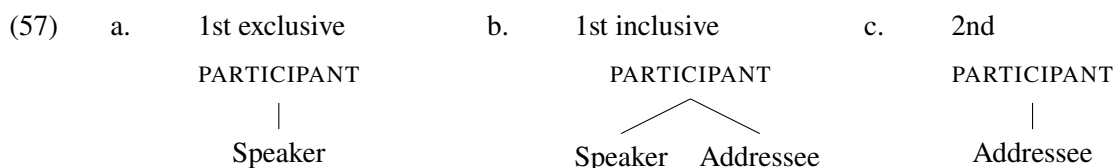


Figure 6.1: Harley and Ritter's (2002) feature geometry. The underlined features are underspecified (i.e. default) values.

(57) shows three kinds of person representations corresponding to the number representations in (54).



While a combination of two number features is interpreted by restriction, a combination of person features is interpreted by union: [Speaker]  $\cup$  [Addressee] = 1st inclusive (inclusive ‘we’). Interpreting the two features by restriction fails to refer anything: [Speaker]([Addressee]) =  $\emptyset$ , [Addressee]([Speaker]) =  $\emptyset$ . Number and person features are thus interpreted differently. In my theory, the basic number categories, i.e. singular, plural and general, are interpreted by union.

I analyze the dual as a subcategory of the plural. That is, the dual is not a basic number category, but a derived number category. Its meaning is derived by restricting the denotation of the plural ( $\geq 2$ ) to the subset whose members are not divisible as in (58).

$$(58) \quad \text{Dual: } [-\text{Pl}]([\underline{-\text{Sg}}, \underline{+\text{Pl}}])$$

The trial and above are derived similarly, following the general formula in (59a).

$$(59) \quad \begin{array}{ll} \text{a. } n\text{-al:} & [-\text{Pl}]([\underline{+\text{Pl}}]_1(\cdots([\underline{+\text{Pl}}]_{n-2}([\underline{-\text{Sg}}, \underline{+\text{Pl}}])))]) \\ \text{b. } \text{Trial:} & [-\text{Pl}]([\underline{+\text{Pl}}]([\underline{-\text{Sg}}, \underline{+\text{Pl}}])) \\ \text{c. } \text{Quadral:} & [-\text{Pl}]([\underline{+\text{Pl}}]([\underline{+\text{Pl}}]([\underline{-\text{Sg}}, \underline{+\text{Pl}}]))) \end{array}$$

The more complex the derivation, the less common the relevant number becomes. The dual is not as common as the three basic numbers; the trial is yet less common; and the existence of the quadral is doubtful (Corbett 2000:26–30). Furthermore, Greenberg’s (1963) universal 34, i.e. no language has the dual/trial unless it has the plural/dual, follows from the formula. The arguments of  $[-\text{Pl}]$  in the dual (58), trial (59b) and quadral (59c) (underlined) represent “plural” ( $\geq 2$ ), “plural–dual” ( $\geq 3$ ) and “plural–dual–trial” ( $\geq 4$ ) respectively.

Lastly, the singular-dual-plural ( $\geq 3$ ) distinction considered basic in previous studies is not as common as number systems without the dual. Only 21 of the 91 genetically distinct languages surveyed by Harley and Ritter (2002) have the dual. This is unexpected especially in Harbour’s (2007) system, where the dual has the same status as the singular and plural ( $\geq 3$ ). In contrast, the general number, which is a basic number category in my theory, seems to exist in most (arguably all) languages, though its existence is often obscured by morphological conflation with other number categories.

## 6.4 Implications for the theory of number and definiteness marking in kind terms

As seen in chapter 3, kind reference is closely related to nominal number. I claimed there that Chierchia's (1998b) kind-deriving operator  $\cap$  is defined only for number-neutral properties. The typology of number marking proposed in the last section thus has implications for crosslinguistic variations in the forms of kind terms. It predicts that languages with different number marking systems allow different forms to denote kinds.

To recapitulate, there are two kinds of kind terms, i.e. kind emulations and true kinds. Kind emulations are the kind derived from number-neutral properties in the [object] subdomain by  $\cap$ . They are also known as “plural” kind terms because morphologically plural forms are used in English, as in (60a). As pointed out in chapter 3, they are in fact general forms. True kinds are the kind that denotes in the [kind] subdomain. They are singular [kind] individuals, and are known as “singular” kind terms because singular forms are used in English, as in (60b).

- (60) a. Dinosaurs became extinct.  
 b. The dinosaur became extinct.

Dayal (2004b) claims that a true kind term denotes in the [kind] subdomain that only contains the relevant [kind] individual, but not its subkinds. This explains the singular morphology of true kinds and the occurrence of the definite article *the* in (60b), which encodes maximality/uniqueness by  $\iota$ . Therefore, two operators are involved in kind term formation, i.e.  $\cap$  and  $\iota$ .

The forms of kind terms vary across languages not only in their number morphology but also in the presence/absence of a definite article. Dayal (2004b) argues that this is because the lexicalization patterns of  $\cap$  and  $\iota$  vary crosslinguistically. She claims that there are four possible patterns as shown in Table 6.5a. ‘D’ in the table means that the relevant operator is lexicalized by definite articles. ‘(D)’ means that the use of definite articles is optional. ‘ $\emptyset$ ’ means that the relevant operator applies freely, that is, its application is not signaled overtly.

Crosslinguistic variations are limited to the four patterns in Table 6.5a because Dayal's theory makes the following two assumptions. The first assumption is the Blocking Principle, which states that a type-shifting operation, including  $\iota$  and  $\cap$ , cannot apply covertly if it is associated with an overt morpheme (Chierchia 1998b). The definite article in kind terms becomes optional

Table 6.5: Possible and impossible lexicalization patterns of  $\cap$  and  $\iota$  (Dayal 2004b).

(a) Possible lexicalization patterns.

	$\iota$	$\cap$	True kind (“singular”)	Kind emulation (“plural”)	Languages
a.	D	D	D NP	D NP <sub>PL</sub>	Italian, French
b.	D	(D)	D NP	(D) NP <sub>PL</sub>	German
c.	D	$\emptyset$	D NP	NP <sub>PL</sub>	English
d.	$\emptyset$	$\emptyset$	NP	NP <sub>PL</sub>	Hindi, Russian

(b) Impossible lexicalization patterns.

	$\iota$	$\cap$	True kind (“singular”)	Kind emulation (“plural”)
a.	(D)	D	(D) NP	D NP
b.	(D)	(D)	(D) NP	(D) NP
c.	$\emptyset$	D	NP	D NP

when the relevant operator is exempt from the Blocking Principle. Secondly, it is assumed that  $\iota$ , but not  $\cap$ , is the basic function of definite articles. Thus, if the Blocking Principle is relevant to only one of the operators that are lexicalized by definite articles, it is  $\iota$  rather than  $\cap$  that adheres to the principle. Hence, while German-type languages, where the definite article is optional for plural kind terms but obligatory for singular kind terms, are expected to exist, the reverse of the German pattern, where the definite article is optional for singular kind terms but obligatory for plural kind terms, should not exist ((a) in Table 6.5b). Furthermore, in order for the Blocking Principle to be meaningful at all, it should not be possible that the definite article is optional in both singular and plural kind terms ((b) in Table 6.5b). Since the basic function of definite articles is  $\iota$ , languages in which only  $\cap$  is lexicalized by definite articles should be ruled out too ((c) in Table 6.5b).

Examples illustrating each type in Table 6.5a are given below. Hindi does not have overt definite articles.

(61) Italian (Dayal 2004b:438)

- a.  $\Pi$ /\* $\emptyset$  cane abbaia.  
the dog barks  
‘The dog barks.’

- b. I/\*Ø cani abbaiano.  
the dog.PL bark  
'Dogs bark.'
- (62) German (Dayal 2004b:441–442)
- a. Der/\*Ø Pandabär ist vom Aussterben bedroht.  
the panda is of extinction threatened  
'The panda is facing extinction.'
- b. Die/Ø Pandabären sind vom Aussterben bedroht.  
the panda.PL are of extinction threatened  
'Pandas are facing extinction.'
- (63) English
- a. The/\*Ø dinosaur became extinct.
- b. \*The/Ø dinosaurs became extinct.
- (64) Hindi (Dayal 2004b:402)
- a. Ø kutta aam janvar hai.  
dog common animal is  
'The dog is a common animal.'
- b. Ø kutte yehaaN aam haiN.  
dog.PL here common are  
'Dogs are common here.'

The typology of number marking proposed in this study enriches Dayal's theory of number and definiteness marking in kind terms. Dayal assumes that the basic number system consists of two categories, i.e. the singular and plural, which is categorized into the 'SG : GN/PL' type in my typology. She does not take the general into account. Since the number marking system of natural languages is much richer, more variations in number and definiteness marking should exist across languages than Dayal's system allows. Such variations should be found in languages whose basic number marking system is other than 'SG : GN/PL'. In the rest of this section, I show that this is indeed the case in Brazilian Portuguese and Singlish, which belong to the 'SG/GN : GN/PL' type in my typology.

As Dayal (2004b) herself admits, Brazilian Portuguese does not fit well with her theory. Since the language encodes  $\iota$  with definite articles, Dayal's system predicts that the definite article is obligatory for singular kind terms, though it could be either optional or absent for

plural kind terms. More generally, bare singular kind terms are not allowed in languages with definite articles. However, Brazilian Portuguese has been claimed to allow bare singular kind terms, as in (65).

- (65) %*(O)* panda logo estará extinto.  
       the panda soon will.be extinct  
       ‘Pandas will soon become extinct.’

What complicates the matter is that the empirical status of bare singular kind terms in the language remains unclear as indicated by the ‘%’ sign (Schmitt and Munn 1999, 2002; Müller 2002; Dobrovie-Sorin and Pires de Oliveira 2008; Ionin et al. 2011). This empirical unclarity makes us wonder whether our theory of kind terms should include bare singular kind terms as an option available for languages with definite articles. Fortunately, unlike Brazilian Portuguese, the empirical status of bare singular kind terms is clear in Singlish, which also encodes  $\iota$  with a definite article like Brazilian Portuguese. In Singlish, bare singular kind terms are not only acceptable but preferred to their counterparts with the definite article, as in (66).

- (66) (The) dinosaur extinct already.  
       ‘The dinosaur became extinct.’

Therefore, it can be said that bare singular kind terms are an available option for languages with definite articles and thus need to be accounted for by any adequate theory of kind terms.

In the traditional conception of number with the singular-plural dichotomy, the unmarked nominal forms in these languages are considered to be “singular.” However, in my number system with two kinds of generals, they are ambiguous between SG and  $\text{GN}^-$ , as Brazilian Portuguese and Singlish have a ‘SG/GN : GN/PL’ type number marking system. The fact that the definite article is optional in “plural” kind terms in these languages suggests that the two kind-generating operators are lexicalized in the same way as in German in Table 6.5a, i.e. the definite article is obligatory for  $\iota$  and optional for  $\cap$ . Table 6.6 shows the number and definiteness marking patterns in languages with the ‘SG/GN : GN/PL’ basic number system.

Brazilian Portuguese and Singlish follow the (b) pattern. Kind terms with plural morphology are derived not from plural forms but from  $\text{GN}^+$  forms by means of  $\cap$ . The definite article is optional here. Kind terms that have been traditionally regarded as “singular” are derived from either SG or  $\text{GN}^-$  forms. The definite article is optional for those “singular” kind terms which

Table 6.6: Number and definiteness marking in ‘SG/GN : GN/PL’ languages.

	$\iota$	$\cap$	True kind (“singular”)	Kind emulation (“plural”)	
			SG	GN <sup>-</sup>	GN <sup>+</sup>
a.	D	D	D NP	D NP	D NP <sub>PL</sub>
b.	D	(D)	D NP	(D) NP	(D) NP <sub>PL</sub>
c.	D	∅	D NP	NP	NP <sub>PL</sub>
d.	∅	∅	NP	NP	NP <sub>PL</sub>

are derived from GN<sup>-</sup> forms, because they are derived by  $\cap$ . On the other hand, the definite article is obligatory for those “singular” kind terms which are derived from singular forms, because they are derived by  $\iota$ . When the two types are seen as a whole, the definite article of “singular” kind terms appears to be optional. Thus, Dayal’s (2004b) theory can account for Brazilian Portuguese and Singlish facts if my theory of number marking is adopted.

Incidentally, Hungarian also ceases to be a problem for Dayal’s theory. The language is thought to follow the (a) pattern in Table 6.6. When unmarked nominals are allowed, they are interpreted as number-neutral as shown in the English translation of the sentence in (67). This indicates that Hungarian conflates SG and GN morphologically.<sup>24</sup>

- (67) Mari verset olvasott ma délután.  
 Mari poem.ACC read.PST today afternoon  
 ‘Mary read a poem/poems/poetry this afternoon.’

(Farkas and de Swart 2003:101)

The definite article *a* is obligatory for both “singular” and “plural” kind terms, as in (68).

- (68) a. \*(A) medve okos.  
 the bear intelligent  
 ‘The bear is intelligent.’  
 b. \*(A) medvék okosak.  
 the bear.PL intelligent.PL  
 ‘Bears are intelligent.’

(Farkas and de Swart 2003:122)

Although bare “singular” kind terms are an available option, there is a difference in their

<sup>24</sup>This analysis runs counter to some authors’ view that the language is a classifier language (Gil 2005; Csirmaz and Dékány 2010). If such a view is correct, then Hungarian has a mixed number marking system.

acceptability between Brazilian Portuguese and Singlish. While they are doubtlessly acceptable in Singlish, their empirical status is not so clear in Brazilian Portuguese. The unclear empirical status of bare “singular” kind terms in Brazilian Portuguese has been sometimes ascribed to the speakers’ awareness of the formal variety of the language, in which bare “singular” kind terms are ungrammatical (Ionin et al. 2011). However, I believe that the influence of the knowledge of the formal variety is trivial, if any, because not only Brazilian Portuguese but also Singlish speakers are conscious of the fact that the definite article cannot be omitted in the formal/standard variety of the respective languages. Therefore, the reason must to be sought elsewhere.

I suggest that the difference between the two languages arises because when recovering the number features of NPs (NumPs to be more precise), speakers can associate ‘D NP’ forms with  $GN^-$  NPs in Singlish, but not in Brazilian Portuguese. In other words, the difference is related to processing. In Brazilian Portuguese, the processing mechanism recovers from all “singular” kind terms singular NPs with features ‘[+Sg], [−Pl]’, for which the definite article is obligatory. However, the grammar of the language derives “singular” kind terms also from  $GN^-$  NPs, in which case the definite article is optional. Hence, a contradiction arises between what the grammar allows and what is possible in the recovered system. This is why Brazilian Portuguese speakers are uncertain about whether the definite article of bare “singular” kind terms can be omitted.

Let us now see in more detail how the number features of NPs are recovered from the surface nominal form ‘D NP’. To begin with, it is necessary to understand how determiners and NPs interact. The key data is given in (69). The data shows that the number-neutrality of  $GN^-$  NPs disappears when they are modified by a determiner. When  $GN^-$  NPs are used without a determiner, they are number-neutral (69a). However, when they are accompanied by a determiner as in (69b), the number-neutrality disappears and only singular reference is possible. In order to refer to plural entities, plural forms as in (69c) must be used.

(69)	BRAZILIAN PORTUGUESE	SINGLISH	
a.	cachorro	dog	‘one or more than one dog’
b.	o/este cachorro	the/dis dog	‘the/this dog’, *‘the/these dogs’
c.	os/estes cachorros	the/dis dogs	‘the/these dogs’, *‘the/this dog’
	cf. *o/este cachorros		

The disappearance of number-neutrality is not surprising in Brazilian Portuguese because the determiners in (69b) are in the singular forms. However, in Singlish, the forms of the determiners are invariant regardless of number (Gil 2003). *The* and *dis* ‘this, these’ can be used regardless of the number of the NPs to which they are attached. Thus, if *dog* is number-neutral, it is expected that *the/dis dog* is also number-neutral.<sup>25</sup>

What the data in (69) suggests is that in these languages, overt determiners are lexically specified for number and the number system used for this purpose is the singular-plural dichotomy, lacking the general number category. In terms of features, they only have the  $[\pm\text{Pl}]$  feature, assuming that the *-s* in Brazilian Portuguese plural determiners in (69c) realizes a positive value of a feature ( $[\text{+Pl}]$ ) rather than a negative one ( $[\text{-Sg}]$ ). The relevant feature is  $[\pm\text{Pl}]$  also because some languages have distinct determiners for the dual, which involves  $[\text{-Pl}]$  (cf. (58)).  $\text{GN}^-$  NPs lose their number-neutrality when combined with determiners because the  $[\pm\text{Sg}]$  feature either gets deleted or becomes invisible, as shown in (70a). This process is thought to happen in order for the concord between the determiner and the Num head to take place successfully. The number-neutrality is retained when no determiner with its own number specification is present in the structure, as in (70b-i). In this case, the number feature is assigned a value through agreement between D and Num (cf. section 6.2.4.2), as in (70b-ii).

- (70) a. *the dog* (singular)  
           [DP the<sub>u</sub>[NUM:[-Pl]] [NumP Num<sub>i</sub>[NUM:[-Sg],[ -Pl]] [NP dog]]]  
       b. *dog* (number-neutral)  
           (i) [DP D<sub>u</sub>[NUM:                  ] [NumP Num<sub>i</sub>[NUM:[-Sg],[ -Pl]] [NP dog]]]  
           (ii) [DP D<sub>u</sub>[NUM:[-Sg],[ -Pl]] [NumP Num<sub>i</sub>[NUM:[-Sg],[ -Pl]] [NP dog]]]

Since overt determiners only have the  $[\pm\text{Pl}]$  feature, they do not determine the value of the  $[\pm\text{Sg}]$  feature of NumP/NPs. Having said that, they imply the value of the  $[\pm\text{Sg}]$  feature. The following two implicational relationships should hold.

- (71) a.  $[\text{-Pl}] \Rightarrow [\text{+Sg}]$   
       b.  $[\text{+Pl}] \Rightarrow [\text{-Sg}]$

<sup>25</sup>Similar facts are also reported in Turkish (Bliss 2004) and Western Armenian (Bale to appear). In Japanese and Malay, a plural interpretation is strongly preferred for bare (i.e.  $\text{GN}^-$ ) NPs modified by (singular) demonstratives, though examples showing that such forms may receive a number-neutral interpretation are found in naturally occurring texts (cf. Downing 1996:205).

I conjecture that (71a) is invoked in the recovery process in Brazilian Portuguese (at least for some speakers), but not in Singlish, probably because the nominal concord system is more complex in the former. The inference when recovering the number features of NumP/NPs proceeds as shown in (72).

- (72)
1. ‘D NP’ (surface form)
  2. NP is unmarked.  $\Rightarrow$  ‘[+Sg], [-Pl]’ (SG) or ‘[-Sg], [-Pl]’ ( $GN^-$ )
  3. D is the singular form.  $\Rightarrow$  [-Pl]
  4. (i) SINGLISH  
Both SG and  $GN^-$  are compatible.  
 $\Rightarrow$  D can be optional. = Grammar (Table 6.6 (b))
  - (ii) BRAZILIAN PORTUGUESE  
Given the inference rule (71a) ( $[-Pl] \Rightarrow [+Sg]$ ), only SG is compatible.  
 $\Rightarrow$  D is obligatory.  $\neq$  Grammar (Table 6.6 (b))

As can be seen in (72–4), in Brazilian Portuguese, kind terms of the form ‘D NP’ are not associated with  $GN^-$  NPs in the system recovered by the speakers although they should be according to the grammar in their minds. It is such a mismatch that is thought to degrade the acceptability of bare “singular” kind terms in Brazilian Portuguese. By contrast, no such mismatch occurs in Singlish. Bare “singular” kind terms are not only acceptable, but even preferred to the form with the definite article.

## 6.5 Summary

This chapter has revealed that classifier languages do not lack number morphology, but on the contrary, they make the finest basic number distinction available in the typology of number marking. Specifically, they have distinct forms for the singular, plural and general, where the singular is realized by classifiers. Recognizing the general number as one of the basic number categories is a crucial aspect of my typology of number marking systems. Languages vary in how they realize the basic number categories morphologically. The difference between classifier and non-classifier languages thus stems from morphology, but not syntax or semantics; the syntactic and semantic differences are the consequences of the morphological difference.

I have argued that the three basic number categories of ‘singular’, ‘plural’ and ‘general’ can

be decomposed into two binary number features, i.e. [ $\pm$ Sg] and [ $\pm$ Pl]. This decomposition allows us to distinguish between two types of generals, which I refer to as  $GN^+$  and  $GN^-$ . I have demonstrated that the presence of two kinds of generals account for the co-occurrence of classifiers and plural markers and bare “singular” kind terms in Brazilian Portuguese and Singlish.

## Chapter 7

# Conclusion

As I conclude, I briefly summarize the chief proposals advanced in this study (section 7.1) and discuss their implications for the number system in general, both in the nominal and verbal domains (section 7.2). I also make short remarks on the ramifications of this study for two related areas in cognitive science, i.e. language and thought, and first language acquisition (section 7.3).

### 7.1 Summary of the dissertation

In essence, this study is an attempt to convince the linguistics and philosophy society of the flaw of two widely held assumptions about classifier languages and demonstrate what new perspectives abandoning them will offer to us on the natural language number system. The two assumptions are:

**Misbelief 1** Numeral classifiers enable counting by individuating an unindividuated mass.

**Misbelief 2** Classifier languages lack genuine number morphology.

The first assumption is partly the basis of the second one, as uncountable nouns are incompatible with number morphology in languages such as English (e.g. *\*waters*, *\*furnitures*).

I invalidated the first assumption in chapter 2. Classifier use in numeral modification constructions is optional in some classifier languages such as Malay and Persian. Moreover, optional classifier use is also observed in the so-called obligatory classifier languages such as Japanese and Vietnamese. Therefore, common nouns are countable without classifiers in all

classifier languages, though the optionality is limited to certain nouns/constructions in obligatory classifier languages. This has led us to the following conclusion:

**Proposal 1** No difference exists between classifier and non-classifier languages regarding the semantics of either nouns or numerals. Common nouns universally denote properties, contra Chierchia's (1998b) Nominal Mapping Parameter Hypothesis.

Of course, such a conclusion suggests we rethink the function of classifiers. With direct numeral modification as in *three books* possible in all languages, though with different degrees of generality, classifiers appear redundant except for their noun classification function. In chapter 4, I analyzed classifiers as follows:

**Proposal 2** Classifiers are a sophisticated kind of singular number marker in that their meanings consist of two components:

1. Assertion: Restriction of the domain to singularities.
2. Conventional implicature: Specification of the characteristics of the noun.

I put forward a generalization that classifiers prevent reference to subkinds in chapter 2 and attributed it to their conventionally implicated meanings in chapter 3.

In chapter 5, I pointed out that the semantics of plural markers in classifier languages is structured in the same way as classifiers, consisting of asserted and conventionally implicated meanings. I also showed that classifiers and plural markers in classifier languages are subject to the same licensing condition, which involves either referential determiners or quantifiers (chapters 4 and 5). Hence, they obviously belong to the same grammatical category, and I concluded that:

**Proposal 3** Plural markers in classifier languages are genuine plural number markers.

I have explained away major potential counterarguments to this claim, which include claims that plural markers do not count as genuine number markers because they encode meanings other than plurality such as individuality and associativity (chapter 3), or because their use is not obligatory (chapter 6).

In fact, optionality is evidence for regarding classifiers and plural markers in classifier languages as genuine singular and plural number markers, as it is a natural consequence of the following claim, which I made in chapter 6:

Table 7.1: Basic number categories in natural languages.

Number	[±Sg]	[±Pl]	Morphology	Semantics
SG	+	–	singular	singularities alone
PL	–	+	plural	pluralities alone
GN <sup>+</sup>	+	+	singular + plural	number-neutral (singularities + pluralities)
GN <sup>–</sup>	–	–	none	number-neutral (non-singularities + non-pluralities)

**Proposal 4** General number, associated with number-neutral properties, is one of the universally available basic number categories, along with the singular and plural.

In chapter 6, I argued that number categories such as ‘singular’, ‘plural’ and ‘general’ are labels for combinations of two number features, i.e. [±Sg] and [±Pl], which are concerned with atomicity/minimality and divisibility respectively. These features have both morphological and semantic reflections, as summarized in Table 7.1. Notice that there are two kinds of general numbers, i.e. GN<sup>+</sup> and GN<sup>–</sup>, which are morphologically distinct but semantically identical. Indeed, languages such as Brazilian Portuguese and Singlish have two number-neutral forms, one with a plural marker and the other without.

I have proposed a typology of number marking systems as in Table 7.2, whereby languages are classified according to how many basic number categories are morphologically distinguished. This typology is similar in spirit to that proposed by Corbett (2000), but different from his in that it takes into account classifiers and two kinds of general numbers. According to my typology, Misbelief 2 above (classifier languages lack genuine number morphology) is far from the reality:

**Proposal 5** Classifier languages make the most fine-grained basic number distinction.

Moreover, my typology shows that the difference between classifier and non-classifier languages reduces not to semantics (Krifka 1995; Chierchia 1998b; Wilhelm 2008) or syntax (Li 1999), but to a morphological difference, specifically one concerning number. I have demonstrated that the proposed basic number system and typology make it possible to account for a hitherto unaccounted fact about number and definiteness marking in kind terms in Brazilian Portuguese and other languages of the ‘SG/GN : GN/PL’ type.

Table 7.2: The typology of number marking systems.

Number system	SG	GN <sup>-</sup>	GN <sup>+</sup>	PL	Languages
<i>Three-way distinction</i>					
SG : GN : PL	CL	bare	CL + PL	PL	Cantonese, Japanese, Korean, Malay, Persian, Yucatec Maya
<i>Two-way distinction</i>					
SG/GN : GN/PL	∅	bare	∅ + PL	PL	Brazilian Portuguese, Hungarian, Singlish
SG/GN : PL	∅	bare	—	PL	Brazilian Portuguese (Müller-Oliveira variety)
SG : GN/PL	∅	—	∅ + PL	PL	English, French, German, Hindi, Italian, Russian
<i>No distinction</i>					
SG/GN/PL	∅	bare/—	—/∅ + ∅	∅	Dëne Sųfiné

## 7.2 D as the interface between two number systems

The number system proposed in this study is much richer than commonly assumed. There are four basic number categories expressed by combinations of two number features. However, it must be noted that such a complex number system is not attested in verbal agreement. It is not attested in pronouns either. The number system employed in these areas is normally a simple singular-plural dichotomy. It is not unusual to find languages with no number distinction at all in these areas (i.e. no verbal agreement, the same pronominal form for both singular and plural referents). Languages with a three-way distinction ‘singular-dual-plural ( $\geq 3$ )’ are found less commonly. Crucially, extensive typological studies such as Corbett (2000, 2006), while acknowledging the existence of the general number for common nouns, report no instance of general agreement or general pronouns that contrast with singular or plural ones. Corbett (2000:126) points out that in languages with two number systems “the system with general number is almost always the second system”, where “the second system” means the system employed for nouns belonging to low categories on the following animacy hierarchy:

- (1) The Animacy Hierarchy (Corbett 2000:56)
- 1st person pronouns > 2nd person pronouns > 3rd person pronouns > kin > human > animate > inanimate

The qualification “almost always” is due to pronouns in Asheninca (an Arawakan language of the central Peruvian highlands), which show the same plural-general opposition that its common nouns do (Corbett 2000:15). However, according to Reed and Payne’s (1986) description, this statement only applies to free forms, which are marked compared to the obligatory affixal forms bound to the verbs. Pronominal affixes do not seem to show the plural-general opposition. The data available in their paper suggests that pronominal affixes in the language make person and gender distinctions, but no number distinction. Hence, the absence of the general number category in pronouns is also observed in this language.

These facts may be taken as a counterargument to my analysis of nominal number system. However, I would like to suggest a different possibility. That is, different number systems are at work in the nominal and verbal domains. From a conceptual point of view, such a situation is quite plausible. While the number information of nouns has semantic contribution, that of verbs is mostly for syntactic purposes. D, as the outermost layer of the nominal domain, interfaces the two number systems.

Although this idea is still immature and requires further investigation, I know of at least two facts that seem to support it. The first fact has been discussed in relation to bare “singular” kind terms in chapter 6: the presence of overt determiners nullifies the number-neutrality of otherwise number-neutral noun phrases. The relevant data is repeated below. The demonstrative *dis* in Singlish is neutral to number distinction, encompassing both *this* and *these* in English. *The* is also neutral with respect to number. Hence, the whole noun phrase in (2b) is expected to inherit the number-neutrality of (2a), which is surprisingly not the case.

(2)	BRAZILIAN PORTUGUESE	SINGLISH	
a.	cachorro	dog	‘one or more than one dog’
b.	o/este cachorro	the/dis dog	‘the/this dog’, *‘the/these dogs’
c.	os/estes cachorros	the/dis dogs	‘the/these dogs’, *‘the/this dog’
	cf. *o/este cachorros		

I interpreted this phenomenon as indicating that overt determiners are lexically specified for number and the number system used for this purpose is the singular-plural dichotomy (in these languages), lacking the general number category; they only have the  $[\pm\text{Pl}]$  feature. I speculate that overt determiners have this property because D, located at the edge of the nominal

domain, has a function of facilitating the communication between the verbal and nominal domains, which work with different number systems.<sup>1</sup> The verbal domain can handle  $[\pm\text{PI}]$ , but not  $[\pm\text{Sg}]$ . This is because while dual agreement distinct from singular and plural agreement exists, general agreement distinct from singular and plural agreement is unattested. The dual is distinguished from the plural by an additional  $[-\text{PI}]$  feature in the former (cf. chapter 6, section 6.3.2): Dual ( $[-\text{PI}][-\text{Sg}, +\text{PI}]$ ) vs. PL ( $[-\text{Sg}, +\text{PI}]$ ). The two general number categories are distinguished from the singular and plural by the  $[\pm\text{Sg}]$  feature:  $\text{GN}^-$  ( $[-\text{Sg}], [-\text{PI}]$ ) vs. SG ( $[+\text{Sg}], [-\text{PI}]$ ) and  $\text{GN}^+$  ( $[+\text{Sg}], +\text{PI}]$ ) vs. PL ( $[-\text{Sg}], +\text{PI}]$ ). The verbal domain is insensitive to the latter distinction, which involves  $[\pm\text{Sg}]$ . Overt determiners only have the  $[\pm\text{PI}]$  feature and lack the  $[\pm\text{Sg}]$  feature to conform to the number system of the verbal domain. If no overt determiner occurs, no facilitation is provided, and both features of the nominal domain, i.e.  $[\pm\text{Sg}]$  and  $[\pm\text{PI}]$ , are brought up to the border to the verbal domain. Figure 7.1 depicts how determiners with inherent number specification link the verbal and nominal domains with different number systems.

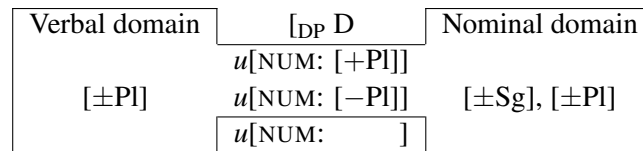


Figure 7.1: Overt determiners and two different number systems.

If this speculation is on the right track, there should be a positive correlation between the presence of verbal agreement and that of overt determiners, more specifically articles. Moreover, it also predicts that classifier languages are less likely to have verbal agreement than non-classifier languages. This is because the number system of classifier languages are the least compatible with that of the verbal domain. Verbal agreement deprives languages of two number distinctions made possible by the  $[\pm\text{Sg}]$  feature, i.e. SG ( $[+\text{Sg}], [-\text{PI}]$ ; CL NP) vs.  $\text{GN}^-$  ( $[-\text{Sg}], [-\text{PI}]$ ; NP) and PL ( $[-\text{Sg}], +\text{PI}]$ ; NP PL) vs.  $\text{GN}^+$  ( $[+\text{Sg}], +\text{PI}]$ ; CL NP PL). A language cannot have a classifier without these distinctions, at least the former, if not both. Both of these predictions seem to be true, though the second of them is also predicted by Misbelief 2

<sup>1</sup>The statement also applies to covert determiners that have overt counterparts such as the covert indefinite article in bare plurals in English (e.g. *A lion is a big cat*/ $\emptyset$  *Lions are big cats*). I put aside such covert determiners for simplicity here.

above (classifier languages lack genuine number morphology).

The second fact supporting the idea that different number systems are at work in the nominal and verbal domains is “optional agreement” in Yucatec Maya (a classifier language) reported by Butler (2011:52–54). She describes that “[p]lural agreement between nominal and verbal elements is not obligatory”, as shown in (3). A3 in the gloss is a cross-reference marker.

- (3) a. Táan u k'aay le x-ch'úupal-o'.  
 PROG A3 sing DEF F-girl-PART  
 'The girl is singing./'The girls are singing.'
- b. Táan u k'aay le x-ch'úupal-o'ob-o'.  
 PROG A3 sing DEF F-girl-PL-PART  
 'The girls are singing.'
- c. Táan u k'aay-o'ob le x-ch'úupal-o'ob-o'.  
 PROG A3 sing-PL DEF F-girl-PL-PART  
 'The girls are singing.'
- d. Táan u k'aay-o'ob le x-ch'úupal-o'.  
 PROG A3 sing-PL DEF F-girl-PART  
 'The girls are singing.'

(Butler 2011:52–53)

Butler analyzes this fact as indicating that the plural marker in Yucatec Maya is adjoined to DP and outside of agreement morphology, extending Wiltscko's (2008) analysis of plural marking in Halkomelem Salish. However, the fact can be also interpreted as indicating the independence of the number system in the verbal domain from that in the nominal domain. The translations for (3a) suggest that the definite marker in this language does not nullify the number-neutrality of unmarked NPs, unlike the two languages discussed above.<sup>2</sup> Hence, both [ $\pm$ Sg] and [ $\pm$ Pl] features are accessible to the verbal domain. If my speculation is correct, normal noun-verb agreement is impossible in such a case because of the presence of [ $\pm$ Sg]. If so, no noun-verb number agreement is actually taking place in the examples above. This line of analysis also explains the apparently unconstrained distribution of the plural marker on both the noun and verb.

A detailed study of the relation between the number systems in the nominal and verbal domains is left for future research.

<sup>2</sup>The plural interpretation of the subject in (3d) is due to the plural marking on the verb.

## 7.3 Ramifications for related areas

### 7.3.1 Language and thought

The dominant view about classifier and non-classifier languages among linguists and philosophers has been that fundamental differences exist between the two language types in the meanings of nouns (e.g. Quine 1968; Krifka 1995; Chierchia 1998b). This view (Misbelief 1) has generated a number of psychological studies that compare classifier and non-classifier languages to test whether the alleged grammatical difference (i.e. all common nouns, including conceptually count ones, are grammatically uncountable and mass-like/substance-referring in classifier languages, but not in non-classifier languages) leads to non-linguistic differences in how speakers perceive the world (e.g. Lucy 1992; Imai and Gentner 1997; Barner et al. 2009; Li et al. 2009). The general conclusion of these studies except Lucy (1992) is that no drastic difference exists between speakers of classifier and non-classifier languages in their perception of objects (conceptually count), though conclusions differ in details depending on the experimental method adopted (similarity judgment or quantity judgment). These studies thus deny a strong form of Whorfian linguistic relativity hypothesis.

The present study has rejected the assumption upon which all these studies rely and claimed that no difference exists between the two language types regarding the semantics of either nouns or numerals (Proposal 1). If this claim is correct, the results of the above-mentioned studies receive different interpretations. They cannot be about whether speakers of one language type perceive objects as substance. Instead, they are about the likelihood in which individuated objects are perceived as a whole (cf. Nisbett 2003). As shown in Table 7.2, the crucial difference between classifier and non-classifier languages is that the linguistic forms associated with number-neutral properties (i.e.  $GN^+$ ,  $GN^-$ ) are distinct from those associated with singular and plural properties (i.e. SG, PL).

### 7.3.2 First language acquisition

The number marking typology in Table 7.2 suggests that children's language acquisition starts from the top of the table and proceeds downwards. Assuming that language acquisition processes involve search for expressions for innately available categories, children's first hypothesis about the number system is that their target languages have the classifier language type number

system (i.e. 'SG : GN : PL'), where the correspondence between the four basic number categories and linguistic forms is the most transparent, with none of them conflated with another category. At this stage, bare nouns are extensively used as in adults' classifier languages. The acquisition of numerals in this type of language may take more time compared to other types, for it proceeds in tandem with the acquisition of classifiers and plural markers, which occur less frequently in adults' speech than in the latter languages due to the availability of bare noun forms (GN<sup>-</sup>). The only reliable evidence for this language type is the presence of classifiers. If classifiers are not found, children will next hypothesize that their languages are of the Brazilian Portuguese types (i.e. 'SG/GN : GN/PL', 'SG/GN : PL'). The evidence for these types is the use of bare nouns for referring to number-neutral properties. If children fail to find a plural marker, which should be available in these language types, they conclude that their target languages are of the Dëne Sùliné type (i.e. 'SG/GN/PL'). Alternatively, if they learn that bare noun forms, which they have associated with number-neutral properties, in fact only refer to singularities, they conclude that their target languages are of the English type (i.e. 'SG : GN/PL').

In the scenario above, number marking acquisition does not involve either semantic or syntactic parameter setting; instead, children learn the form and meaning of number-related lexical items. Furthermore, it is not a process of creating a new category, but one of reorganizing the preexisting categories. In this respect, number marking acquisition resembles phoneme acquisition (e.g. Werker and Tees 1984; Kuhl 1994). It goes without saying that the suggested scenario needs to be tested by empirical studies.

## References

- Acquaviva, Paolo. 2004. Plural mass nouns and the compositionality of number. *Verbum* 26:387–401.
- Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- Alexiadou, Artemis. 2011. Plural mass nouns and the morpho-syntax of number. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 33–41. Somerville: Cascadilla Proceedings Project.
- Alwi, Hasan, Soenjono Dardjowidjojo, Hans Lapoliwa, and Anton M. Moeliono. 1998. *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka, edisi ketiga edition.
- Asmah Haji Omar. 1972. Numeral classifiers in Malay and Iban. *Anthropological Linguistics* 14:87–96.
- Au-Yeung, Ben. 2007. Cantonese classifier *di* and genericity. In *Studies in Cantonese Linguistics 2*, ed. Joanna Ut-Seong Sio and Sze-Wing Tang, 1–16. Hong Kong: Linguistic Society of Hong Kong.
- Bach, Emmon, and Robin Cooper. 1978. The NP-S analysis of relative clauses and compositional semantics. *Linguistics and Philosophy* 2:145–150.
- Baker, Mark C. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.
- Bale, Alan. to appear. Number, competition and syntactic complexity. In *Proceedings of the 41st Annual Meeting of the North East Linguistic Society (NELS 41)*.
- Bale, Alan C., and David Barner. 2012. Semantic triggers, linguistic variation and the mass-count distinction. In *Count and Mass Across Languages*, ed. Diane Massam, 238–260. Oxford: Oxford University Press.

- Bale, Alan, Michaël Gagnon, and Hrayr Khanjian. 2011. Cross-linguistic representations of numerals and number marking. In *Proceedings of Semantics and Linguistic Theory (SALT) XX*, ed. Nan Li and David Lutz, 582–598.
- Bale, Alan, and Hrayr Khanjian. 2009. Classifiers and number marking. In *Proceedings of Semantics and Linguistic Theory (SALT) XVIII*, ed. Tova Friedman and Satoshi Ito, 73–89.
- Barner, David, Shunji Inagaki, and Peggy Li. 2009. Language, thought, and real nouns. *Cognition* 111:329–344.
- Bhattacharya, Tanmoy. 1999. The Structure of the Bangla DP. Doctoral Dissertation, University College London.
- Bisang, Walter. 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In *Numeral Types and Changes Worldwide*, ed. Jadranka Gvozdanović, 113–185. Berlin: Mouton de Gruyter.
- Bliss, Heather. 2004. The semantics of the bare noun in Turkish. *Calgary Papers in Linguistics* 25:1–65.
- Borer, Hagit. 2005. *Structuring Sense*. Oxford: Oxford University Press.
- Butler, Lindsay Kay. 2011. The Morphosyntax and Processing of Number Marking in Yucatec Maya. Doctoral Dissertation, University of Arizona.
- Carlson, Gregory N. 1977. Reference to Kinds in English. Doctoral Dissertation, University of Massachusetts, Amherst. Published from Garland Publishing (New York) in 1980.
- Carlson, Greg. 2009. Generics and concepts. In *Kinds, Things, and Stuff: Mass Terms and Generics*, ed. Francis Jeffrey Pelletier, New Directions in Cognitive Science Series, 16–35. Oxford: Oxford University Press.
- Carstens, Vicki. 2013. Delayed Valuation: A reanalysis of “upwards” complementizer agreement and the mechanics of Case. Ms. University of Missouri. URL <http://ling.auf.net/lingbuzz/001432>.
- Chacón, Dustin Alfonso. 2011. Head movement in the Bangla DP. *Journal of South Asian Linguistics* 4:3–25.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 1998. Yi-wan tang, yi-ge tang: Classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies, New Series* 28:385–412.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.

- Cheng, Lisa Lai-Shen, and Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 259–292. Oxford: Oxford University Press.
- Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of “semantic parameter”. In *Events and Grammar*, ed. Susan Rothstein, 53–104. Kluwer Academic Publishers.
- Chierchia, Gennaro. 1998b. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174:99–149.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chung, Sandra. 2000. On reference to kinds in Indonesian. *Natural Language Semantics* 8:157–171.
- Conklin, Nancy Faires. 1981. The Semantics and Syntax of Numeral Classification in Tai and Austronesian. Doctoral Dissertation, University of Michigan, Ann Arbor.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Croft, William. 1994. Semantic universals in classifier systems. *Word* 45:145–171.
- Csirmaz, Aniko, and Éva Dékány. 2010. Hungarian classifiers. Paper presented at Word Classes: Nature, Typology, Computational Representation.
- Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo, and Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21:159–210.
- Dalrymple, Mary, and Suriel Mofu. 2012. Plural semantics, reduplication, and numeral modification in Indonesian. *Journal of Semantics* 29:229–260.

- Daniel, Michael, and Edith Moravcsik. 2011. The associative plural. In *The World Atlas of Language Structures Online*, ed. Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library.
- Dayal, Vaneeta. 2004a. Licensing by modification. *Ilha do Desterro Florianópolis* 47:217–238.
- Dayal, Vaneeta. 2004b. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27:393–450.
- Delfitto, D., and J. Schroten. 1991. Bare plurals and the number affix in DP. *Probus* 3:155–186.
- Diercks, Michael, Michael Putnam, and Marjo van Koppen. 2012. Upward- vs. downward-probing Agree: Defining a single feature valuation relation. Paper presented at the 30th West Coast Conference on Formal Linguistics (WCCFL).
- Dobrovie-Sorin, Carmen, and Roberta Pires de Oliveira. 2008. Reference to kinds in Brazilian Portuguese: Definite singulars vs. bare singulars. In *Proceedings of SuB12*, ed. Atle Grønn, 107–121. Oslo: ILOS.
- Dölling, Johannes. 1995. Ontological domains, semantic sorts and systematic ambiguity. *International Journal of Human-Computer Studies* 43:785–807.
- Downing, Pamela. 1996. *Numeral Classifier System: The Case of Japanese*. Amsterdam: John Benjamins.
- Dryer, Matthew S., and Martin Haspelmath, ed. 2011. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, 2011 edition. URL <http://wals.info/>.
- Farkas, Donka, and Henriëtte de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford, CA: CSLI Publications.
- Farkas, Donka F., and Henriëtte E. de Swart. 2010. The semantics and pragmatics of plurals. *Semantics and Pragmatics* 3:1–54.
- Fodor, Janet Dean, and Ivan A. Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5:355–398.
- Gebhardt, Lewis. 2008. Classifiers, plural, and definiteness in Persian. In *Aspects of Iranian Linguistics*, ed. Simin Karimi, Vida Samiian, and Donald Stilo, 33–55. Newcastle upon Tyne: Cambridge Scholars Press.
- Gebhardt, Lewis. 2009. Numeral Classifiers and the Structure of DP. Doctoral Dissertation, Northwestern University.

- Gerner, Matthias, and Walter Bisang. 2010. Classifier declinations in an isolating language: On a rarity in Weining Ahmao. *Language and Linguistics* 11:579–623.
- Ghaniabadi, Saeed. 2010. The Empty Noun Construction in Persian. Doctoral Dissertation, University of Manitoba.
- Ghomeshi, Jila. 2003. Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57:47–74.
- Gil, David. 2003. English goes Asian: Number and (in)definiteness in the Singlish noun phrase. In *Noun Phrase Structure in the Languages of Europe*, ed. Frans Plank, 467–514. Berlin: Mouton de Gruyter.
- Gil, David. 2005. Numeral classifiers. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew Dryer, David Gil, and Bernard Comrie, 226–229. Oxford: Oxford University Press.
- Gillon, Carrie. 2010. The mass/count distinction in Innu-aimun: Implications for the meaning of plurality. In *Papers for WSCLA 15: The Fifteenth Workshop on Structure and Constituency in Languages of the Americas*, ed. Beth Rogers and Anita Szakay, volume 29 of *The University of British Columbia Working Papers in Linguistics*, 12–29.
- Givón, Talmy. 2001. *Syntax: An Introduction*, volume I. Amsterdam: John Benjamins.
- Görgülü, Emrah. 2011. Plural marking in Turkish: Additive or associative? *Working Papers of the Linguistics Circle of the University of Victoria* 21:70–80.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, ed. Joseph H. Greenberg, 73–113. Cambridge, MA: MIT Press.
- Greenberg, Joseph. 1972. Numeral classifiers and substantial number: Problems in the genesis of a linguistic type. In *Working Papers in Language Universals*, volume 9, 1–39. Stanford, CA: Stanford University.
- Greenberg, Joseph. 1975. Dynamic aspects of word order in the numeral classifier. In *Word Order and Word Order Change*, ed. Charles N. Li, 27–46. Austin, Texas: University of Texas Press.
- Grice, H. Paul. 1975. Logic and conversation. In *Syntax and Semantics: Speech Acts*, ed. Peter Cole and Jerry L. Morgan, volume 3, 41–58. New York: Academic Press.

- Hamedani, Ladan. 2011. The Function of Number in Persian. Doctoral Dissertation, University of Ottawa.
- Harbour, Daniel. 2007. *Morphosemantic Number: From Kiowa Nouns Classes to UG Number Features*. Dordrecht: Springer.
- Harbour, Daniel. 2009. Mass, non-singularity, and augmentation. In *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT*, ed. Claire Halpert, Jeremy Hartman, and David Hill, 349–360. Cambridge, MA: MITWPL.
- Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42:561–594.
- Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78:482–526.
- Hla Pe. 1965. A re-examination of Burmese classifiers. *Lingua* 15:164–186.
- Ho, Mian Lian, and John Platt. 1993. *Dynamics of a Contact Continuum: Singapore English*. Oxford: Clarendon Press.
- Hopper, Paul. 1986. Some discourse functions of classifiers in Malay. In *Noun Classes and Categorization*, ed. Colette G. Craig, 309–325. Amsterdam: John Benjamins.
- Hosoi, Hironobu. 2005. Japanese *-tachi* plurals. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, volume 31, 157–168.
- Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Iljic, Robert. 1994. Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32:91–116.
- Imai, Mutsumi, and Dedre Gentner. 1997. A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62:169–200.
- Ionin, Tania, and Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23:315–360.
- Ionin, Tania, Silvina Montrul, and H elade Santos. 2011. The expression of genericity in English and Brazilian Portuguese: An experimental investigation. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 115–123. Somerville, MA: Cascadilla Proceedings Project.

- Iwasaki, Shoichi, and Preeya Ingkaphirom. 2005. *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.
- Jenks, Peter Spiros Eric. 2011. The Hidden Structure of Thai Noun Phrases. Doctoral Dissertation, Harvard University.
- Jespersen, Otto. 1927. *A Modern English Grammar*. Heidelberg: Carl Winter's Universitätsbuchhandlung.
- Jiang, Li Julie, and Suhua Hu. 2012. An overt determiner and complementizer in a classifier language. In *Proceedings of GLOW in Asia VIII 2010: Universals and Variation*, ed. Ming-Le Gao. Beijing: Beijing Language and Culture University.
- Kaneko, Makoto. 2007. Indexical plural marker TATI in Japanese. Paper presented at Workshop on Nominal and Verbal Plurality.
- Kang, Beom-Mo. 1994. Plurality and other semantic aspects of common nouns in Korean. *Journal of East Asian Linguistics* 3:1–24.
- Kawasaki, Noriko. 1989. *Jibun-tati* and non-coreferential anaphora. In *University of Massachusetts Occasional Papers: Papers on Quantification*, 115–146.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Keller, Frank. 2000. Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality. Doctoral Dissertation, University of Edinburgh.
- Keller, Frank. 2006. Linear Optimality Theory as a model of gradience in grammar. In *Gradience in Grammar: Generative Perspectives*, ed. Gisbert Fanselow, Caroline Féry, Ralph Vogel, and Matthias Schlesewsky, 270–287. Oxford: Oxford University Press.
- Kester, Ellen-Petra, and Cristina Schmitt. 2007. Papiamentu and Brazilian Portuguese: A comparative study of bare nominals. In *Noun Phrases in Creole Languages*, ed. Marlyse Baptista and Jacqueline Guéron, 107–143. Amsterdam: John Benjamins.
- Khazriyati, Salehuddin, and Heather Winskel. 2009. An investigation into Malay numeral classifier acquisition through an elicited production task. *First Language* 29:289–311.
- Kim, Chonghyuck, Qizhong Chang, and Leslie Lee. 2009. Number marking in Colloquial Singapore English. *Journal of Cognitive Science* 10:149–172.
- Kishimoto, Hideki. 2000. Locational verbs, agreement, and object shift in Japanese. *The Linguistic Review* 17:53–109.

- Kornfilt, Jaklin. 1996. Naked partitive phrases in Turkish. In *Partitives: Studies on the Syntax and Semantics of Partitive and Related Constructions*, ed. Jacob Hoeksema, 107–142. Berlin: Mouton de Gruyter.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Kratzer, Angelika. 1998. Scope or pseudo-scope? Are there wide-scope indefinites? In *Events in Grammar*, ed. Susan Rothstein, 163–196. Dordrecht: Kluwer.
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and Contextual Expressions*, ed. Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, 75–115. Dordrecht: Foris.
- Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In *The Generic Book*, ed. Gregory N. Carlson and Francis Jeffrey Pelletier, 398–411. Chicago: Chicago University Press.
- Krifka, Manfred, Francis Jeffrey Pelletier, Gregory N. Carlson, Alice ter Meulen, Godehard Link, and Gennaro Chierchia. 1995. Genericity: An introduction. In *The Generic Book*, ed. Gregory N. Carlson and Francis Jeffrey Pelletier, 1–124. Chicago: University of Chicago Press.
- Kuhl, Patricia K. 1994. Learning and representation in speech and language. *Current Opinion in Neurobiology* 4:812–822.
- Kurafuji, Takeo. 2004. Plural morphemes, definiteness, and the notion of semantic parameter. *Language and Linguistics* 5:211–242.
- Kuroda, S.-Y. 1992. *Japanese Syntax and Semantics: Collected Papers*. Dordrecht: Kluwer.
- Kwon, Song Nim, and Anne Zribi-Hertz. 2004. Number from a syntactic perspective: Why plural marking looks ‘truer’ in French than in Korean. In *Empirical Issues in Formal Syntax and Semantics* 5, ed. Olivier Bonami and Patricia Cabredo Hofherr, 133–157.
- Lan, Haifan. 2010. Possible interpretations for suffix *men* in Mandarin Chinese. Master’s thesis, Utrecht University.
- Landman, Fred. 1989. Groups I. *Linguistics and Philosophy* 12:559–605.
- Landman, Fred. 2004. *Indefinites and the Type of Sets*. Oxford: Blackwell.
- Lee, Iksop, and S. Robert Ramsey. 2000. *The Korean Language*. Albany, NY: State University of New York Press.

- Li, Peggy, Yarrow Dunham, and Susan Carey. 2009. Of substance: The nature of language effects on entity construal. *Cognitive Psychology* 58:487–524.
- Li, Yen-Hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8:75–99.
- Liaw, Yock Fang. 1999. *Malay Grammar Made Easy: A Comprehensive Guide*. Singapore: Times Books International.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, Use and Interpretation of Language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 302–323. Berlin: de Gruyter.
- Link, Godehard. 1984. Hydras. on the logic of relative clause constructions with multiple heads. In *Varieties of Formal Semantics: Proceedings of the Fourth Amsterdam Colloquium, September 1982*, ed. Fred Landman and Frank Veltman, 245–258. Dordrecht: Foris.
- Löbel, Elisabeth. 2000. Classifiers versus genders and noun classes: A case study in Vietnamese. In *Gender in Grammar and Cognition*, ed. Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, and Mirja Saari, 259–320. Berlin: Mouton de Gruyter.
- Longobardi, Giuseppe. 1994. Reference and proper names. *Linguistic Inquiry* 25:609–666.
- Longobardi, Giuseppe. 2005. Toward a unified grammar of reference. *Zeitschrift für Sprachwissenschaft* 24:5–44.
- Lucy, John A. 1992. *Grammatical Categories and Cognition*. Cambridge: Cambridge University Press.
- Mathieu, Éric. 2012. Licensing by modification: The case of French *de* nominals. *Journal of Linguistics* 48:389–426.
- Matsumoto, Yo. 1993. Japanese numeral classifiers: A study of semantic categories and lexical organization. *Linguistics* 31:667–713.
- Matthews, Stephen, and Virginia Yip. 2011. *Cantonese: A Comprehensive Grammar*. London: Routledge, 2nd edition.
- McCawley, James. 1968. Review of *Current Trends in Linguistics 3. Language* 44:556–593.
- McCready, Eric. 2009. Conventional implicature and classifiers. *Journal of Cognitive Science* 10:195–208.
- McCready, Eric. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3:1–57.

- McCready, Eric. 2012. Classification without assertion. In *Proceedings of the Fifth Formal Approaches to Japanese Linguistics Conference (FAJL5)*, ed. Matthew A. Tucker, Anie Thompson, Oliver Northrup, and Ryan Bennett, volume 64 of *MIT Working Papers in Linguistics*, 141–154.
- McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language* 81:699–718.
- Mintz, Malcolm W. 2002. *An Indonesian & Malay Grammar for Students*. Perth: Indonesian / Malay Texts and Resources.
- Mizuguchi, Shinobu. 2004. *Individuation in Numeral Classifier Languages: A Case of Japanese Classifiers and Plurals*. Tokyo: Shohakusha.
- Moltmann, Friederike. 2004a. Properties and kinds of tropes: New linguistic facts and old philosophical insights. *Mind* 113:1–41.
- Moltmann, Friederike. 2004b. Two kinds of universals and two kinds of collections. *Linguistics and Philosophy* 27:793–776.
- Müller, Ana. 2002. The semantics of generic quantification in Brazilian Portuguese. *Probus* 14:279–298.
- Müller, Ana, and Fátima Oliveira. 2004. Bare nominals and number in Brazilian and European Portuguese. *Journal of Portuguese Linguistics* 3:1–30.
- Muromatsu, Keiko. 1998. On the Syntax of Classifiers. Doctoral Dissertation, University of Maryland, College Park.
- Murphy, Gregory L. 2002. *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Myint Soe. 1999. A Grammar of Burmese. Doctoral Dissertation, University of Oregon.
- Nakanishi, Kimiko. 2008. The syntax and semantics of floating numeral quantifiers. In *The Oxford Handbook of Japanese Linguistics*, ed. Shigeru Miyagawa and Mamoru Saito, 287–319. Oxford: Oxford University Press.
- Nakanishi, Kimiko, and Elizabeth Ritter. 2008. Plurality in languages without a count-mass distinction. Paper presented at Mass/Count Workshop.
- Nakanishi, Kimiko, and Satoshi Tomioka. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13:113–140.
- Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42:413–444.

- Nguyen, Đinh Hoa. 1957. Classifiers in Vietnamese. *Word* 13:124–152.
- Nguyễn, Đình-Hoà. 1997. *Vietnamese: Tiếng Việt Không Sơn*. Amsterdam: John Benjamins.
- Nisbett, Richard. 2003. *The Geography of Thought: How Asians and Westerners Think Differently... and Why*. New York: Free Press.
- Nomoto, Hiroki. 2009. On indefinite subjects of pivot verbs in Malay. *International Journal of Corpus Linguistics* 14:221–254.
- Nomoto, Hiroki. 2010. Reference to subkinds and the role of classifiers. Paper presented at 2010 Annual Meeting of the Linguistic Society of America (LSA).
- Nomoto, Hiroki. 2013. On the optionality of grammatical markers: A case study of voice marking in Malay/Indonesian. In *Voice Variation in Austronesian Languages of Indonesia*, ed. Alexander Adelaar, volume 54 of *NUSA*, 121–143. Jakarta and Tokyo: Universitas Katolik Indonesia Atma Jaya and Tokyo University of Foreign Studies.
- Noyer, Rolf. 1992. Features, Positions and Affixes in Autonomous Morphological Structure. Doctoral Dissertation, MIT.
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, ed. Jeroen Groenendijk and Martin Stokhof, 115–143. Dordrecht: Foris.
- Pelletier, Francis Jeffrey. 1979. Non-singular reference. In *Mass Terms: Some Philosophical Problems*, ed. Francis Jeffrey Pelletier, 1–14. Dordrecht: Kluwer.
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation, In Honor of Joseph E. Emonds*, ed. Wendy K. Wilkins Simin Karimi, Vida Samiian, 262–294. Amsterdam: John Benjamins.
- Piriyawiboon, Nattaya. 2010. Classifiers and Determiner-less Languages: The Case of Thai. Doctoral Dissertation, University of Toronto.
- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
- Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

- Quine, W. V. 1968. Ontological relativity. *The Journal of Philosophy* 65:185–212.
- Radford, Andrew. 2009. *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press.
- Ratliff, Martha. 1987. Tone sandhi compounding in White Hmong. *Linguistics of the Tibeto-Burman Area* 10:71–105.
- Reed, Judy, and David L. Payne. 1986. Asheninca (Campa) pronominals. In *Pronominal Systems*, ed. Ursula Wiesemann, 323–331. Tübingen: Narr.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Rothstein, Susan. 2007. Counting and the mass-count distinction. Ms. Bar-Ilan University.
- Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27:343–397.
- Ruangjaroon, Sugunya. 2005. The Syntax of Wh-Expressions as Variables in Thai. Doctoral Dissertation, University of British Columbia.
- Rullmann, Hotze, and Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In *Where Semantics Meets Pragmatics*, ed. Klaus von Stechow and Ken Turner, 175–196. Amsterdam: Elsevier.
- Sadler, Louisa, and Douglas J. Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30:187–226.
- Sato, Yosuke. 2008. Minimalist Interfaces: Selected Issues in Indonesian and Javanese. Doctoral Dissertation, University of Arizona.
- Sauerland, Uli. 2003. A new semantics for number. In *Proceedings from Semantics and Linguistic Theory XIII*, ed. Rob Young and Yuping Zhou, 258–275. Ithaca, NY: Cornell University.
- Saul, Janice E., and Nancy F. Wilson. 1980. *Nung Grammar*, volume 62 of *Summer Institute of Linguistics Publications in Linguistics*. Dallas: Summer Institute of Linguistics.
- Schmitt, Cristina, and Alan Munn. 1999. Against the nominal mapping parameter: Bare nouns in Brazilian Portuguese. In *Proceedings of NELS 29*, ed. Tamanji Pius, Masako Hirotsu, and Nancy Hall, 339–354.
- Schmitt, Cristina, and Alan Munn. 2002. The syntax and semantics of bare arguments in Brazilian Portuguese. *Linguistic Variation Yearbook* 2:185–216.
- Sharifian, Farzad, and Ahmad R. Lotfi. 2003. “Rices” and “waters”: The mass-count distinction in modern Persian. *Anthropological Linguistics* 45:226–244.

- Sharvy, Richard. 1978. Maybe English has no count nouns: Notes on Chinese semantics. *Studies in Language* 2:345–365.
- Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 806–838. Oxford: Oxford University Press.
- Simpson, Andrew, and Tanmoy Bhattacharya. 2003. Obligatory overt wh-movement in a wh-in-situ language. *Linguistic Inquiry* 34:127–142.
- Simpson, Andrew, Hooi Ling Soh, and Hiroki Nomoto. 2011. Bare classifiers and definiteness: A cross-linguistic investigation. *Studies in Language* 35:168–193.
- Sneddon, James Neil Sneddon, Alexander K. Adelaar, Dwi N. Djenar, and Michael Ewing. 2010. *Indonesian: A Comprehensive Grammar*. London: Routledge, 2nd edition.
- Soh, Hooi Ling, and Hiroki Nomoto. 2009. Progressive aspect, the verbal prefix *meN-*, and the stative sentences in Malay. *Oceanic Linguistics* 48:148–171.
- Soh, Hooi Ling, and Hiroki Nomoto. 2011. The Malay verbal prefix *meN-* and the unergative/unaccusative distinction. *Journal of East Asian Linguistics* 20:77–106.
- Soh, Hooi Ling, and Hiroki Nomoto. to appear. Degree achievements, telicity and the verbal prefix *meN-* in Malay. *Journal of Linguistics*.
- Tang, Chih-Chen Jane. 1990. Chinese Phrase Structure and the Extended X'-Theory. Doctoral Dissertation, Cornell University.
- Thackston, Wheeler M. 1983. *An Introduction to Persian*. Cambridge, MA: Department of Near Eastern Languages and Civilizations, Harvard University.
- Thomas, Earl W. 1969. *The Syntax of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.
- Tsoulas, George. 2006. Plurality of mass nouns and the grammar of number. Paper presented at the 29th GLOW Colloquium.
- Watanabe, Akira. 2006. Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language and Linguistic Theory* 24:241–306.
- Watanabe, Akira. 2010. Vague quantity, numerals, and natural numbers. *Syntax* 13:37–77.
- Watanabe, Akira. 2012. Measure words as nouns: A perspective from silent years. *Studia Linguistica* 66:181–205.

- Watters, John R. 1981. A phonology and morphology of Ejagham—with notes on dialect variation. Doctoral Dissertation, UCLA.
- Werker, Janet F., and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49–63.
- Wilhelm, Andrea. 2008. Bare nouns and number in Dëne Sųliné. *Natural Language Semantics* 16:39–68.
- Wilkinson, Karina. 1995. The semantics of the common noun *kind*. In *The Generic Book*, ed. Gregory N. Carlson and Francis Jeffrey Pelletier, 383–397. Chicago, IL: The University of Chicago Press.
- Wiltsecko, Martina. 2008. The syntax of non-infective plural marking. *Natural Language and Linguistic Theory* 26:639–694.
- Windfuhr, Gernot L. 1979. *Persian Grammar: History and State of Its Study*. The Hague: Mouton Publishers.
- Wolter, Lynsey. 2007. Situation variables and licensing by modification in opaque demonstratives. In *Proceedings of Sinn und Bedeutung 11*, ed. Estela Puig-Waldmüller, 612–625. Barcelona.
- Wurmbrand, Susi. 2012. The syntax of valuation in auxiliary-participle constructions. In *Coyote Papers: Proceedings of the Poster Session of the West Coast Conference on Formal Linguistics (April, 2012)*, ed. Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, and Alex Trueman, volume 20, 154–162. Tucson, AZ: University of Arizona Linguistics Circle.
- Yang, Rong. 1998. Chinese bare nouns as kind-denoting terms. *RuLing Papers* 1:247–288.
- Yi, Byeong-Uk. 2010. Numeral classifiers and the mass/count distinction. Ms. University of Toronto.
- Yorifuji, Atsushi. 1976. Men nitsuite. *Toukyou Gaikokugo Daigaku Ronshuu* 26:73–88.
- Zamparelli, Roberto. 1998. A theory of kinds, partitives and of/z possessives. In *Possessors, Predicates and Movement in the Determiner Phrase*, ed. Artemis Alexiadou and Chris Wilder, 259–304. Amsterdam: John Benjamins.
- Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistics Review* 29:491–539.

- Zubin, David A., and Mitsuaki Shimojo. 1993. How “general” are general classifiers? with special reference to *ko* and *tsu* in Japanese. In *Proceeding of the 19th Annual Meeting of the Berkeley Linguistic Society*, ed. Joshua S. Guenter, Barbara A. Kaiser, and Cheryl C. Zoll, 490–502. Berkeley, CA: BLS.
- Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32:353–407.