

UNIVERSITY OF CALGARY

The Semantics of Number in Malay Noun Phrases

by

Jana Cécile Carson

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF MASTER OF ARTS

DEPARTMENT OF LINGUISTICS

CALGARY, ALBERTA

AUGUST, 2000

© Jana Cécile Carson 2000



National Library  
of Canada

Acquisitions and  
Bibliographic Services

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque nationale  
du Canada

Acquisitions et  
services bibliographiques

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-55138-5

Canada

## Abstract

Many authors describe Malay as a language which treats all nouns as mass nouns. Languages like English make a distinction between count nouns such as *cow* and mass nouns such as *water*. In Malay all nouns are treated alike. All Malay nouns require classifiers when a numeral is present, and can be reduplicated to mark plurality. In Malay the root noun may be used to refer to one or more than one entity, and so is neutral with respect to number. This is different from English count nouns, where the root noun is typically used to refer to only one entity. In this thesis, Link's (1983) semantic theory of plurality and Krifka's (1989, 1991, 1995) analysis of classifiers are extended to Malay. It is argued that in Malay the root noun denotes a set of entities which includes atomic entities as well as sums, or groups of entities. The reduplicated noun denotes a set which includes only sums of entities but not atoms. Sortal classifiers denote functions which map entities onto the number of atoms they contain. A compositional analysis of Malay noun phrases is developed, and the pragmatic use of root versus reduplicated nouns is investigated.

## Acknowledgements

It is truly a pleasure to thank all the people who have encouraged and assisted me through the writing of this thesis. First, I wish to thank my supervisor, Dr. Hotze Rullmann, for his enthusiasm and encouragement. He read all the many drafts that I produced, took the time to discuss ideas, and was a wonderful person to learn about semantics from. I feel lucky to have had the chance to work with him. As well, I would like to thank Dr. Elizabeth Ritter and Dr. Amanda Pounder for serving on my examination committee. I thank Betsy for her enlightening discussions about the structure of DPs, and for her thoughtful questions at my defense. I thank Amanda for her careful editing of my thesis, through which it has been greatly improved. I would also like to thank several members of the Linguistics Department who I have had the opportunity to learn from the last few years. Thanks to Dr. John Archibald, who I worked with over the past two summers, for giving me the opportunity to learn more about child language acquisition, and to Dr. Michael Dobrovolsky, who first told me that I could do this stuff. For inspiring me to look at the semantics of reduplication, I owe a big thank you to Dr. Suzanne Urbanczyk. As well, Linda Toth deserves an award for her humor and perceptiveness. I would also like to thank the Linguistics Department and the Faculty of Graduate studies for providing scholarships and financial support which eased the financial burden of being a student.

Over the last year, I had the opportunity to work with native speakers of Malay, without whom this thesis would not exist. I thank Safarah Salim, Mohd Nawi Wahid, and Su Lin Lee for their friendly cooperation and insights into Malay. They cheerfully translated, pondered and explained their language, and I am grateful to them for their sense of humor when I (often) mispronounced words or constructed ridiculous sentences. *Terima kasih*. I thank Catherine Goode for providing judgments on some of the Chinese data.

My fellow grad students in the Linguistics Department have provided an immense amount of support, encouragement and sympathy. A special thanks goes to Susan Atkey, who suffered through sympathetic departmentals and defenses with me, and who was a great person to talk to about the insanity of being a grad student. I am also grateful to

Susan Armstrong, Val Baggaley, Leslie Blair, Olga Karpacheva, Melanie Strickland (Blair), and Andrea Wilhelm for their support and perspective. I have had a wonderful time!

Finally, I wish to thank some people in my personal life who have provided support through their confidence in me. First, for discovering with me how much fun linguistics is, I thank my classmates from Simon Fraser University's Linguistics Field School in Fiji, especially Catherine Etmanski, Robyn Gathercole, Kimberley Gatrell, Chris Seguin, and Jillian Snider. I thank also Dr. Ross Saunders for allowing us that incredible opportunity, and for guiding us through it. Here in Calgary, Janelle Douglas, Melissa Grainger and Katherine Ruelle reminded me how much fun life outside of school is. I thank them for movie marathons, and for being such great friends. I also deeply appreciate the support of Allen and Bettilou Carson, and I thank them for providing a restful retreat in Fernie when Calgary became too busy.

My family has been a constant source of encouragement ever since I can remember, and especially over the last two years. I thank my parents, Greg and Pat Belland, for believing in me and for offering sound advice whenever I asked for it. I cannot possibly repay them for all they have provided me over the years, but I thank them now for everything. To my sister, Jill, who took me on long walks to 'air me out', I say thank you for your humor and strength. And last, but definitely not least, I thank my husband, JAC, for showing up at the syntax final in his green velvet suit. Things have never been the same since.

*For Jeremy Allen Carson*

## TABLE OF CONTENTS

Approval page .....	ii
Abstract .....	iii
Acknowledgements .....	iv
Dedication .....	vi
Table of Contents .....	vii
Chapter One: Introduction .....	1
1.1 Introduction to the Problem.....	1
1.2 Object Language and Methodology of Study.....	2
1.3 Classifiers in Malay.....	3
1.4 Reduplication.....	6
1.4.1 Reduplication in Malay.....	6
1.4.2 Reduplication and the Root Noun.....	8
1.4.3 A Note on the Historical Development of Reduplication.....	9
1.5 Number and the Count / Mass Distinction.....	11
1.6 Background Information on Malay Grammar.....	15
1.6.1 Lexical Categories in Malay.....	15
1.6.2 Basic Word Order and Clause Structure.....	16
1.6.3 Verbs .....	18
1.6.3.1 Tense / Aspect Markers.....	18
1.6.3.2 Verbal Morphology.....	19
1.7 Summary.....	20
1.8 Organization of the Thesis.....	21
Chapter Two: The Internal Structure of Malay DPs.....	23
2.1 The Structure of Malay DPs .....	23
2.2 Reduplication.....	24
2.3 Classifiers.....	25
2.4 Numerals.....	28
2.5 Determiners.....	30
2.5.1 Definite Determiners.....	30
2.5.2 Indefinites.....	32
2.6 Other Aspects of Malay DPs.....	35
2.6.1 Attributive Nouns.....	35
2.6.2 Attributive Adjectives and Relative Clauses.....	37
2.7 Alternative Structures for DP in Classifier Languages.....	40
2.8 Summary.....	42
Chapter Three: Distribution of Malay Root versus Reduplicated Nouns.....	43
3.1 Malay Nouns are Neutral for Number.....	43
3.2 Vagueness or Ambiguity?.....	44
3.3 Indefinite Noun Phrases.....	47
3.3.1 Grice's View of Pragmatics.....	47

3.3.2 Problems with Indefinites.....	51
3.4 Definite Noun Phrases.....	53
3.4.1 Definiteness and <i>Itu</i> and <i>Ini</i> .....	53
3.4.2 Definite Determiners and Restrictions on Root Nouns.....	54
3.5 The Nature of Field Work.....	58
3.6 Summary.....	58
 Chapter Four: Formal Semantics of Malay Nouns and Reduplication.....	59
4.1 Theories of Plural Nouns in English.....	59
4.2 Link's (1983) Theory of Plurals in English.....	60
4.3 The Semantics of Malay Nouns.....	62
4.4 Translation of Definites in English.....	67
4.5 Translation of Definites in Malay.....	68
4.6 Link's Classification of Predicates.....	71
4.7 Link's Classification of Predicates in Malay.....	74
4.7.1 Distributive Predicates.....	74
4.7.2 Collective Predicates.....	75
4.7.3 Mixed Predicates.....	76
4.8 Compositional Analysis of Malay Sentences.....	76
4.9 The Problem with Definite Noun Phrases.....	81
4.10 Summary.....	83
 Chapter Five: Classifiers and the Mass / Count Distinction.....	84
5.1 Characteristics of Mass Nouns.....	84
5.2 Count versus Mass Nouns?.....	86
5.3 Formal Properties of Classifiers.....	89
5.3.1 Link (1983, 1991).....	89
5.3.2 Krifka (1989, 1991, 1995).....	92
5.3.3 Kang's (1994) Analysis of Korean Nouns.....	95
5.3.4 Chierchia (1998) and Classifiers.....	97
5.4 Classifiers in Malay.....	98
5.5 'All kinds of' Nouns.....	106
5.6 Summary.....	107
 Chapter Six: Directions for Further Research and Conclusion.....	108
6.1 Conclusions.....	108
6.2 Extensions.....	109
6.3 Typological Issues.....	111
6.3.1 Gil's (1987) Typology of NP Structure.....	111
6.3.2 Chierchia's (1998) Typology of NP Structure.....	116
6.3.3 Further Typological Issues.....	117
6.4 Implications for Child Language Acquisition.....	117
 References.....	119

## Chapter One – Introduction

### 1.1 Introduction to the Problem

Malay noun phrases<sup>1</sup> pose an interesting problem to formal theories for number. They are different from English noun phrases in two significant ways: 1) Malay is a classifier language, which means that classifiers are obligatory when the noun phrase contains a numeral, and 2) Malay does not always require plural marking when more than one object is being referred to. These two differences lead us to pose the following question: what is the semantics of number in Malay? This is an important question because the semantics and pragmatics of plural marking in classifier languages has not been explored in depth, although Malay is not the only classifier language that can mark plurality. Also, Malay is interesting in that it does not have classifiers across the board. Classifiers only occur in cases where a numeral is in the noun phrase. Most formal theories of classifier languages have focused on Chinese, which requires classifiers both in construction with demonstratives and with numerals. Malay therefore provides a different and interesting perspective on classifier languages. This thesis presents Malay as a language in which noun roots are neutral for number and explores the formal semantics of number in Malay noun phrases. Unreduplicated Malay nouns are found to denote a set which contains individual entities as well as groups of entities. The plural in Malay is marked by reduplication. Reduplicated nouns denote a set of groups of entities. Classifiers make the entities denoted by the noun “countable” by measuring the number of individual entities in a group.<sup>2</sup>

In this introductory chapter, I first situate Malay genetically and areally. Next, information is given about the nature of the field work conducted. Then, I support the above observations by examining Malay as a classifier language and the role of reduplication as a number marker in Malay. Finally, in order to affirm that Malay is a

---

<sup>1</sup> I will use the term ‘noun phrase’ to refer to the maximal projection of a nominal constituent, regardless of its syntactic category. The terms NP and DP are reserved for specific maximal projections.

<sup>2</sup> Here the terms “individual” and “group” are used in a non-technical sense. I introduce formal terminology in Chapter 4.

language in which noun roots are neutral for number, and all nouns are treated as “mass” nouns, I examine the count / mass distinction as well as the question whether Malay makes this distinction syntactically.

## **1.2 Object Language and Methodology of the Study**

Malay is a Western Austronesian language (Asher 1994). It is the national language of Indonesia, Malaysia, Borneo, and Singapore, and is spoken by approximately twenty million people (Bright 1991). Although each country has its own standard variety, the different varieties of Malay are mutually intelligible and are collectively known as “Bahasa”, which literally means ‘language’. Modern Malay evolved from Old Malay, the trade language used in ports and coastal areas throughout Southeast Asia in the 1<sup>st</sup> century CE or earlier. For purposes of this thesis, Indonesian will refer to the variety of Malay spoken in Indonesia, and Malay will refer to the variety of Malay spoken in Malaysia. Although not genetically related, Malay also shares some morpho-syntactic features with languages in the East and Southeast Asian area including Korean, Chinese, Vietnamese, Thai and Japanese, especially in the area of number marking and classifiers.

The data for this thesis were collected over an eight-month period between October of 1999 and June of 2000.<sup>3</sup> Three native speakers of Malay with fluency in English were consulted in a series of interviews held at the University of Calgary. One male speaker, age 22, and two female speakers, age 22 and 26, participated. All participants speak the Kuala Lumpur dialect of Malay, and were exchange students from Malaysia at the University of Calgary. The interviews were conducted mainly in English, with later interviews conducted partly in Malay. Consultants were given several tasks. In beginning interviews, consultants were simply asked to translate sentences. In later interviews, consultants were asked to tell stories, or to give judgments on Malay sentences. For this last task, the Malay sentences were of two types: isolated sentences without a context, or

---

<sup>3</sup> All data, unless otherwise cited, are from my own field notes.

sentences with the context given in either English or Malay. In sentences where the context was given, subjects were asked whether the Malay sentence was an acceptable sentence, and whether it would be true or false in the situation described.

### 1.3 Classifiers in Malay

According to observation 1) above, Malay is a classifier language. Unlike English, Malay requires the presence of classifiers when objects named by a noun are being counted. One can say that, classifiers render the noun “countable” (Allan 1977) and thereby function as number markers. Obligatory classifiers occur in constructions with numerals and usually give information about the size, shape, and animacy of the entity in question. For example, the classifier *biji* can only be used with small, round objects, whereas *ekor*, which literally means ‘tail’, is only used with animals:

- |     |  |      |             |                    |             |                     |             |
|-----|--|------|-------------|--------------------|-------------|---------------------|-------------|
| (1) | Budak-budak                            | itu  | bermain     | dengan             | <b>tiga</b> | <b>biji / *ekor</b> | <b>bola</b> |
|     | Child-pl                               | the  | play        | with               | three       | cl.                 | ball        |
|     | ‘The children played with three balls’ |      |             |                    |             |                     |             |
| (2) | Saya                                   | mahu | <b>lima</b> | <b>ekor/ *biji</b> | <b>kuda</b> |                     |             |
|     | I                                      | want | five        | cl.                | horse       |                     |             |
|     | ‘I want five horses’                   |      |             |                    |             |                     |             |

Lyons (1977) distinguishes two types of classifiers: sortal classifiers, which pick out permanent characteristics of the noun, and mensural<sup>4</sup> classifiers, which pick out a temporary characteristic of the noun. For example, in Malay, the classifier *helai* picks out the properties of being thin and occurring in layers, and can only occur with objects having these properties (such as cloth, leaves or hair). Mensural classifiers, on the other hand, represent units of measure. A typical mensural classifier is *pound*. Mensural classifiers can be applied to all different kinds of nouns. For example, in Malay the

---

<sup>4</sup> The term “measural” classifiers is also used.

mensural classifier *kilo* 'kilogram' may be applied to sand or apples. Count nouns are often associated with sortal classifiers; mensural classifiers are more often associated with mass nouns (this is investigated further in Chapter Five). Examples of each type of classifier in English and Malay are given below:

- (3) Sortal
- a) two **heads** of cattle
  - b) two **slices** of bread
  - c) Dua           **helai**   sarong / \*kuda  
Two           cl.    sarong / \*horse  
      'Two sarongs' / \*'two horses'
- (4) Mensural
- a) two **piles** of cotton
  - b) two **pounds** of sand
  - c) Dua           **timbangan**   pasir / epal  
Two           pile            sand / apple  
      'Two piles of sand / apples'

According to T'sou (1976), all languages exhibit both categories, but the range of each category may vary. Malay, Chinese and other Southeast Asian languages are good examples of languages which have a rich system of sortal classifiers. This correlates with the observation that many Southeast Asian languages have nouns which are number-neutral, and must therefore be rendered "countable" by means of a classifier. Nonetheless, even Southeast Asian languages differ in the range and distribution of classifiers. Malay has classifiers only in noun phrases which contain numerals. This is different from the more extensively studied language Chinese, which has classifiers both in noun phrases which contain numerals and in noun phrases which contain demonstratives. This is illustrated below:

## Malay

(5) Dua        **helai**   sarong

Two        cl.    sarong

‘Two sarongs’

(6) Sarong    itu

Sarong    that

‘That sarong’

## Chinese

(Catherine Goode p.c.)

(7) Liang     zhi     quianbi

Two        cl.     pencil

‘Two pencils’

(8) \*Zhe      quianbi

This        pencil

‘This pencil’

(9) Zhe        zhi     quianbi

This        cl.     pencil

‘This pencil’

Another interesting aspect of Malay is that many classifiers may also be used as regular nouns. For example, *buah* is the classifier for roundish, medium-sized objects, but it may also occur in a noun phrase as the noun meaning ‘fruit’. In situations such as (10), *buah* is not a classifier, but a regular noun. An argument in favor of this analysis involves the possibility of reduplication. Reduplication is explored in depth in section 1.4. When used as a classifier, *buah* may not be reduplicated, but when used as a noun, *buah* may be reduplicated. This is illustrated below in (11)-(13).

(10) **Buah**    itu     masam

Fruit    that    sour

‘That fruit is sour’

- (11) **Buah-buahan** itu masam  
Fruit the sour  
'The fruits are sour' – the different varieties of fruit are sour
- (12) Budak-budak itu telah menerima **empat buah hadiah**  
Child-pl the alreadyreceive four cl. prize  
'The children received four prizes'
- (13) \*Budak-budakitu telah menerima **empat buah-buah hadiah**  
Child-pl the alreadyreceive four cl. prize  
'The children received four prizes'

## 1.4 Reduplication

### 1.4.1 Reduplication in Malay

In addition to classifiers, Malay also marks number by means of reduplication. Reduplication is a process where a word is copied in whole or in part (Moravcsik 1978, Marantz 1982, Abbi 1985). The resulting word differs in form from the source word. Reduplication in Malay is a productive process. Although reduplication in Malay applies to many syntactic categories, I will be concentrating on reduplication in nouns.

There has been some disagreement over the semantic effect of reduplication in nouns in traditional descriptions of Malay and Indonesian. Macdonald (1976) describes the reduplicated form in Indonesian as generally meaning "the same as that of the unreduplicated form but with the added connotation of variety, randomness, or repetition" (p. 33). Haji (1974a) describes reduplication in nouns as indicating the plural. Lewis (1969) describes reduplicated nouns as indicating "indefiniteness". These descriptions all capture the idea that while reduplication in nouns conveys the meaning of plurality, it can also convey the meaning 'all kinds of'. Although the meaning of 'all kinds of' is not strictly plural, I will argue that the "regular" plural meaning evolves from the meaning of 'all kinds of'. This evolution is explored in section 1.4.3. I therefore

conclude that the primary function of reduplication in Malay nouns is indeed to indicate plurality, whether it is simply ‘more than one X’ or ‘all kinds of X’.

**Table 1 – Effects of Nominal Reduplication in Malay**

Category	Semantic Effect	Source	Target
1) Noun	Plural	buku ‘book’	buku-buku ‘books’
2) Noun	Kind	epal ‘apple’	epal-epal ‘different kinds of apple’

We have now seen the two primary number markers used in Malay: reduplication and classifiers. In the next section, the interaction between classifiers and reduplication is explored. Classifiers and reduplication occur in complementary distribution, and are ungrammatical when combined. In example (14), the plural is indicated by the reduplication of the noun *buku*. In example (15), a specific number of books is being referred to by means of a numeral and the classifier *buah* is used. Example (16) is ungrammatical because we have a classifier co-occurring with reduplication.

- (14) **Buku-buku** ini berat  
 Book-pl those heavy  
 ‘Those books are heavy’
- (15) Dia ada **tiga buah buku**  
 She has three cl. book  
 ‘She has three books’
- (16) \*Dia ada **tiga buah buku-buku**  
 She has three cl. book  
 ‘She has three books’

### 1.4.2 The Root Noun is Number-Neutral

Although reduplication can be used to form the plural, this is not strictly obligatory. The root form is often used in place of the reduplicated form. Example (17) may indicate one or more entities. This is different from English, which requires plural marking whenever the noun refers to more than one object.<sup>5</sup> This is shown in (18).

(17) Bill    melihat        **kuda**  
        Bill    see            horse  
        ‘Bill saw a horse / horses’

(18) \*Bill saw **horse**

Because the root noun can be used to refer to more than one entity as well as a single entity, we can say that the root is neutral for number. This has been observed by many authors for a wide variety of languages (Boas 1911, MacDonald 1976, Gil 1987, Mithun 1988, Link 1991, Kang 1994, Chierchia 1998, Ojeda 1998, Gil 1999). Interestingly, Sanches (1973) proposes an implicational universal whereby if a language has classifiers, it will not have obligatory plural marking. Malay provides support for this assertion. The use of number-neutral versus plural nouns in Malay has not yet been explored. In Chapter Three, I explore this use and the factors involved in choosing the neutral versus plural form.

We have seen that one fundamental difference between English and Malay noun phrases is in the way they encode number. I propose that, formally, Malay nouns are unlike English nouns in terms of what the noun root denotes. In English, noun roots usually denote sets of atomic entities. I propose that Malay noun roots denote sets of entities which may include both atomic and polyatomic entities, which can be thought of as “groups” of atomic entities. The claim that Malay noun roots include both atomic and

---

<sup>5</sup> There are exceptions to this claim. For example, in an English sentence such as: *Bill saw every horse*, the noun *horse* is singular although in some sense more than one horse is being referred to.

polyatomic entities is motivated by the observation that Malay noun roots may be used to refer to more than one entity, as seen in (17). Intuitively, we can describe Malay nouns as referring to a collectivity; the classifier makes the noun countable (Allan 1977, Chan 1993). In Malay the root noun is not marked for number, and so it can be used in situations where there is more than one entity as well as in situations where there is only one entity. The classifier serves to “count” the number of individual entities. Classifiers are combined with numerals in order to pick out a specific number of entities. This is different from English count nouns, which do not require classifiers when combined with a numeral, but which require plural marking. In English, mass nouns do require classifiers in order to be counted. For example, *water* is a mass noun. In order to count a quantity of water, one must use a classifier, such as *drop* (e.g. *one drop of water*). English has a small number of nouns that function like typical Malay nouns. For example, the noun *furniture* is a mass noun although it refers to a set of discrete entities. Example (19) is ungrammatical because we have not made this set countable by means of a classifier. Example (20) is grammatical because we have made the noun countable via the classifier *piece (of)*.

- (19) \*I bought a furniture  
 (20) I bought a **piece of** furniture

### 1.4.3 A Note on the Historical Development of Reduplication

The central claim of this thesis is that in Malay the root of the noun is neutral for number, and can therefore be used when speaking of one or many entities. Why, then, does Malay use reduplication to indicate ‘more than one’? There is evidence that the meaning of plurality in reduplication evolves from a meaning other than simply ‘more than one’. Mithun (1988, p. 218) explores the evolution of number marking and remarks that “the most common form of number marking over multiple lexical categories is reduplication”. She examines Franz Boas’ (1911) grammar of Nass Tsimshian (of the Penutian language family) and notes that although he does list reduplication as marking plurality,

reduplication does not mark plurality in the way English plural marking simply indicates ‘more than one X’. She notes that “reduplicated nouns function as distributives, emphasizing temporal, locative, or conceptual distribution: the separateness of the entities they identify” (p. 220). She also quotes Sapir (1930) who notes that reduplicated nouns in Southern Paiute are “not plural, though sometimes, particularly in the case of animate nouns, practically equivalent to such” (p. 257). This description of reduplicated nouns matches the interpretation of Malay where the reduplicated noun may indicate ‘all kinds of X’. It is likely that the use of reduplication to mark the simple plural evolved from this meaning. Boas (1911) noted that younger generations learning Tsimshian used reduplication to mark the simple plural, losing the “distributive” meaning. He also noted that many North American languages do not distinguish count and mass nouns the way languages like English do. “The idea of plurality is not clearly developed. Reduplication of a noun expresses rather the occurrence of an object here and there, or of different kinds of a particular object, than plurality. It is therefore a distributive rather than a true plural. It seems that this form is gradually assuming a purely plural significance.” (Boas 1911, p. 219). Sapir (1921) also noted that languages may differ in whether plurality is marked by derivational or inflectional morphology. He cites Nootka as an example of a language which expresses plurality derivationally, with the meaning of ‘several X’s here and there’. English, on the other hand, expresses plurality through inflections, with the meaning of ‘more than one X’.

Often, reduplication does not remain a productive means of marking plurality. Suzanne Urbanczyk (p.c.) has observed that in Salish, reduplication is often lost as an inflectional marker. Reduplication in Malay also has a tendency to take on meanings other than plural. This is known as lexicalization. For instance:

- |      |       |         |             |                    |
|------|-------|---------|-------------|--------------------|
| (20) | Kuda  | ‘horse’ | kuda-kuda   | ‘sawhorse’         |
| (21) | Semut | ‘ant’   | semut-semut | ‘pins and needles’ |

This lexicalization suggests that reduplication, when lost as a “distributive”, or derivational marker, does not tend to remain productive. It is possible that, as reduplication loses its “distributive” meaning the reduplicant becomes an inflectional affix. However, because of the processing demands involved in reduplication (every morpheme has a different reduplicant), once it has lost its derivational meaning, it does not remain productive as an inflectional marker and tends to evolve toward derivational meanings expressed in lexicalized forms. Malay appears to be in an intermediate stage of this process. Reduplication is mainly used to mark simple plurality, but is optional in many cases and is becoming lexicalized in others. We will see more of this when we examine the interaction of Malay reduplication with definiteness and grammatical function (i.e. subject, object, etc.) in Chapter Three.

### **1.5 Number and the Count / Mass Distinction**

I have proposed that root nouns in Malay may refer to one or more entities, and are considered neutral for number. Another way to describe this property is to say that all nouns in Malay behave as mass nouns (Gil 1999). In this section, I examine the syntactic properties of mass nouns, and see how they apply to Malay. We will see that there is no syntactic distinction between count and mass nouns in Malay.

Syntactically, several properties of count and mass nouns have been identified. These may or may not match up with the actual characteristics of the entity being referred to by the noun. Entities may be one or many, and may be made up of discrete natural units (such as chairs), or not (water). The extent to which languages grammaticalize these differences, however, varies considerably. English is a language which grammaticalizes singular (one entity), plural (more than one entity) and mass (entity without natural units) expressions. A detailed examination of the relation between form and meaning in English is beyond the scope of this thesis. Even in English, however, the relation between form and meaning is not one-to-one. This can be seen from examples like *furniture* or *change*

(in the sense of ‘coins’). These nouns behave syntactically like mass nouns, although they denote discrete natural units.

Chierchia (1998) reviews the main criteria used to distinguish count and mass nouns syntactically in English. Firstly, count nouns generally allow plural morphology, while mass nouns do not. For example, one may say *shoes* or *cups*, but not *furnitures* or *bloods*. Secondly, mass nouns may not combine directly with numerals, where count nouns may. For example, one may say *two shoes* or *two cups* but not *two water(s)* or *two blood(s)*. As a third characteristic, Chierchia points out that mass nouns require classifiers in order to be able to combine with numerals, for example, *two glasses of water* or *two pools of blood*. The fourth characteristic noted by Chierchia is that determiners are sensitive to the count / mass distinction. Some determiners occur only with count nouns, some only with mass nouns, some with mass and plural nouns, and some are unrestricted. He presents a number-based catalogue of English determiners, reproduced in Table 2:

**Table 2 – Determiners and the Count / Mass Distinction**

Count Nouns		Mass Nouns	Plurals and Mass Nouns	Unrestricted
Singular Determiners	Plural Determiners			
every, each, a	several, few, a few, many, both	little, much	a lot of, plenty of, more, most	the, some, any, no

Count nouns are thus distinguished syntactically from mass nouns in English. There are also languages where all nouns behave as “mass” nouns syntactically. I argue that Malay is such a language. In Malay, all nouns behave the same with respect to the syntax. As

illustrated below, all nouns can be reduplicated, and all require classifiers in order to combine with numerals.<sup>6</sup>

- (22) Sup-sup        itu    panas  
 Soup-pl        the    hot  
 ‘The varieties of soup are hot’
- (23) Tiga    mankuk        sup    panas  
 Three bowl        soup    hot  
 ‘Three bowls of soup are hot’
- (24) \*Tiga sup    panas  
 Three soup    hot  
 ‘Three soups are hot’
- (25) Kuda-kuda    itu    panas  
 Horse-pl        the    hot  
 ‘The horses are hot’
- (26) \*Tiga kuda    panas  
 Three horse    hot  
 ‘Three horses are hot’
- (27) Tiga ekor    kuda    panas  
 Three cl.        horse    hot  
 ‘Three horses are hot’

Although there is no syntactic mass / count distinction in Malay, there may be semantic differences between nouns like *kuda* ‘horse’ and those like *sup* ‘soup’. In Chapter Five, possible semantic criteria for a count / mass distinction in Malay are explored.

The implications for the proposal that Malay nouns are number-neutral are explored in detail in Chapter Three. The presence of number-neutral nouns seems to be characteristic

---

<sup>6</sup> Chierchia’s other criteria have not been applied to Malay. The criteria examined above are sufficient to demonstrate my point.

of East and Southeast Asian languages such as Tagalog, Japanese, Thai and Vietnamese. Thompson (1965) characterizes Vietnamese nouns as behaving like English mass nouns. Quine (1969) does the same for Japanese, Stein (1981) for Thai, Kang (1994) for Korean, Link (1991) for Chinese and Gil (1999) for Tagalog.

For example:

(28) Tagalog (from Gil 1999)

Mansana      ang    kinain ni      boy  
 Apple            TOP    eat    pers    boy  
 ‘Boy(s) ate apple/ an apple/ apples’

(29) Vietnamese (from Gil 1999)

Su      ?n      tao  
 Sue    eat      apple  
 ‘Sue ate apple / an apple / apples’

In the above examples, the sentence is true whether the person denoted by the subject is eating apple-matter, one apple, or many apples. These languages also characteristically have nominal classifiers which serve to individuate entities denoted by the noun, or allow the noun to refer to atomic entities. However, many of these languages also have plural markers and determiners which restrict the use of the root noun. These properties of number-neutral languages have not been explored in detail, and have not all been explained in a formal theory of semantics. This examination of Malay nouns and noun phrases then provides an important step toward the development of a formal semantics of languages which have number-neutral nouns.

## 1.6 Background Information on Malay Grammar

### 1.6.1 Lexical Categories in Malay

Now that we have seen some basic facts about number in Malay, let us examine more general properties of the language. The properties of Malay reviewed here will become relevant in our discussion of Malay sentences in later chapters. We have seen that nouns may refer to one or more entities, but how are nouns identified morphosyntactically? Lexical bases in Malay may belong to more than one word class. Words are identified as nouns and verbs by either their morphology or their position in the sentence. The former possibility is illustrated in examples (30)-(31) with the root *kumpul*. In (30), the suffix *-an* is being used to form a noun; in (31) the prefix *ber-* is being used to form a verb. These are the typical functions of these affixes.<sup>7</sup>

- (30) Setiap orang berjumpa dengan sepuluh **kumpul-an**  
 Every person meeting with ten group  
 ‘Every person met with ten groups’
- (31) Kanak-kanak itu **ber-kumpul** di koridor  
 Kid-pl the gather in corridor  
 ‘The kids gathered in the corridor’

The role of derivational morphology in identifying lexical categories is also illustrated in (32) and (33), where (32) is an adjective, which becomes a noun through the addition of the suffix *-an* in (33).

- (32) kotor ‘dirty’  
 (33) kotor-an ‘trash’

---

<sup>7</sup> For a more detailed examination of Malay affixes and their functions, see MacDonald (1976), Hassan (1974), Darajowidjodjo (1978) and Haji (1974), among others.

In colloquial speech, many words are stripped of their affixes, and consequently the identification of the lexical category of words in Malay is largely dependent on syntactic context. In fact, Gil (2000a) has gone so far as to say that the Riau dialect of Malay does not distinguish lexical categories, with meaning being entirely derived from the context of the utterance. I will not be adopting this view, but will note that unaffixed words in Malay are much more flexible in their category status than in a language such as English.

### 1.6.2 Basic Word Order and Clause Structure

Given that syntactic context is important in establishing the lexical category of words, let us examine the basic word order of Malay. Malay is a Subject Verb Object (SVO) language (MacDonald 1976). More generally, a Malay sentence normally consists of a subject followed by a predicate. Example (34) demonstrates SVO order, with a verbal predicate. Example (35) shows an adjectival predicate, and example (36) shows a nominal predicate. The subject is in italics and the predicate is in bold. Examples (34)-(36) also demonstrate the absence of a copular verb, such as ‘be’, in Malay.<sup>8</sup>

- (34) *Budak-budak itu*      **memetik**      **rambutan**  
 Child-pl      the      pick      rambutan  
 ‘The children pick rambutan’
- (35) *Budak-budak lelaki itu*      **sangat tinggi**  
 Child-pl      male      the      very      tall  
 ‘The boys are very tall’
- (36) *Yang ini*                      **mangga,**      *yang itu*      **rambutan**  
 Rel.      these                      mango              rel.      those      rambutan  
 ‘These ones are mangoes, those ones are rambutans’

There are also cases, however, where the subject is not the first element in the sentence. For pragmatic reasons, other constituents may occur at the beginning of the sentence.

<sup>8</sup> There are instances where translations of sentences involve the copular *be*. However, Malay is traditionally considered to be a language without a copula.

This is known as topicalization, and is illustrated below in examples (37) and (38). The suffix *-nya* serves to nominalize the predicate (Macdonald 1976).

(37) **Berat-nya** buku-buku ini  
 Heavy-nom book-pl these  
 ‘These books are heavy’

(38) **Tinggi-nya** lelaki tadi  
 Tall-nom boy just now  
 ‘That boy who just passed by is tall’

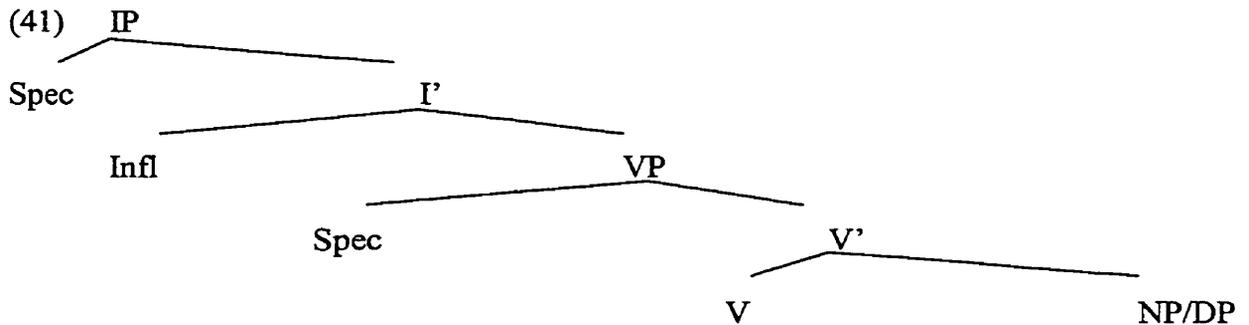
There is also evidence of NP-topicalization, as seen by the passive sentence in (39). The corresponding active sentence is shown below it in example (40).

(39) Anjing itu Ali/ saya/ ku- pukul  
 Dog the Ali I 1sg clitic hit  
 ‘The dog was hit by Ali/ me’

(40) Ali / saya pukul anjing itu  
 Ali / I hit dog the  
 ‘I / Ali hit the dog’

The status of subjects and topics in Austronesian has been subject to considerable debate for some time. For a detailed discussion of the status of subjects and topics in Malay, the reader is referred to works by Verhaar (1988), Hopper (1988), Uhrbach (1988), among others. Guilfoyle, Hung and Travis (1992) examine the word order of four Austronesian languages, including Malay, and found evidence for a VP-internal subject position (Spec of VP) in addition to the traditional subject position Spec of IP. According to Guilfoyle et al., all subjects are generated in Spec of VP but they may move to Spec of IP in order to check Case. They argue that subject-sensitive properties may be split between Spec of VP and Spec of IP, where Spec of VP is associated with prototypical agents (subjects) and Spec of IP is associated with topics in Austronesian. They make this argument based on

properties of subjects cross-linguistically, and the observation that tests for subjects such as wh-extraction and quantifier raising seem to target topic positions whereas reflexivization and Equi-NP deletion may target either the topic or the agent of a sentence. They use this structure to explain three types of passives in Austronesian languages, where the Agent of the sentence remains in Spec of VP and the Topic moves to Spec of IP. I will be assuming the following structure proposed by Guilfoyle et al. for Malay:



### 1.6.3 Verbs

#### 1.6.3.1 Tense / Aspect Markers

Verbs in Malay are not marked for person or number agreement. Verb phrases in Malay may however contain tense or aspect markers, which occur as separate words such as:

telah	‘already’
sudah	‘already’
tadi	‘just now’
sedang	‘progressive’
akan	‘will’

(Taken from Soh 1994, p. 61)

This is not intended to be an exhaustive list of tense and aspect markers. It is also possible that these are adverbs rather than aspect markers. I will refer to them as aspect markers, but do not commit myself to either analysis. See examples below for illustration.

- (42) Dia **telah** menghabiskan semua air itu  
 He already finish all water the  
 ‘He drank all the water’
- (43) Pisang-pisang itu **sudah** masak  
 Banana-pl the already ripe  
 ‘The bananas are ripe’
- (44) Budak-budak itu **sedang** mengelilingi se-ekor anjing  
 Child-pl the prog. surround one+cl. dog  
 ‘The children are surrounding a dog’
- (45) Semua orang **akan** gembira kejayaan pasukan itu  
 All person will happy win team the  
 ‘Everyone will be happy for the team that won’
- (46) **Tadi**, budak itu bermain di taman  
 Just now child the play at park  
 ‘The child played at the park just now’

### 1.6.3.2 Verbal Morphology

In addition to the tense / aspect markers discussed above, Malay also has a complex system of verbal morphology. Common verbal markers include: *men-* which typically attaches to roots to signal an active interpretation (see example (47)), *di-* occurs with the word *oleh* ‘by’ in passive constructions (see example (48)). Other verb-forming affixes are: *per-*, *-i*, *-kan*, *ter-*, and *ke- -an* (shown in (47)- (52)).<sup>9</sup>

<sup>9</sup> For a complete description of the function of verb-forming affixes, the reader is referred to Haji (1974a), Darajowidjodjo (1978), and Hassan (1974).

- (47) Budak-budak itu **men-geliling-ianjing** itu  
 Child-pl the surround dog the  
 ‘The kids are surrounding the dog’
- (48) Dua tandan pisang habis **dimakan** oleh nam ekor monyet  
 Two cl. banana finish eaten by six cl. monkey  
 ‘Two big bunches of bananas were eaten by six monkeys’
- (49) Didalam **per-jalan,** mereka telah berjumpa  
 In travel they already meet  
 beberapa kumpulan pengembara  
 a few group traveler.  
 ‘While traveling, they met a few groups of travellers’
- (50) Dia **mem-bahagi-kan** guli-guli-nya kepada tiga  
 He divide marble-pl-poss into three  
 biji se-kumpulan  
 cl. one+group  
 ‘He divided his marbles into three per group’
- (51) Tasik-tasik itu telah **ter-cemai**  
 Lake-pl the already polluted  
 ‘The lakes are polluted’
- (52) **Ke-banyak-an** budak-budak lelaki itu sangat tinggi  
 Many child-pl male the very tall  
 ‘Many of the boys are very tall’

### 1.7 Summary

In this chapter we have seen that Malay differs from English in terms of its number-marking system. I motivated this claim by observing that in Malay root nouns can be used to refer to one or more entities, while English typically requires plural marking in order to refer to more than one entity. I also observed that in Malay reduplication can be used to indicate ‘more than one’, and that Malay nouns require classifiers in order to

become countable. The use of reduplication to mark plurality poses an interesting problem to formal theories of number because the root noun can be used to refer to more than one object or entity. An investigation of the use of reduplicated versus root nouns is therefore an important step in understanding languages with number-neutral nouns as well as plural markers. The requirement of numerals with classifiers in Malay is interesting in that most formal analyses of classifier languages have focused on Chinese, which does not have this requirement. We will see in Chapter Four that this difference affects both the syntactic and semantic representation of Malay nouns in that the classifier is an operator which requires two arguments rather than one. It is therefore important to explore a classifier language such as Malay, which provides a different perspective on classifier languages and, as we will see in Chapter Six, bears on current typological proposals in allowing us to tease apart two parameters which were previously considered a single parameter. I motivated Malay as a language with number-neutral nouns by showing that Malay treats all nouns syntactically as mass nouns, while English makes a grammatical distinction between count and mass nouns. Finally, I examined some basic aspects of Malay grammar. I found that Malay is an SVO language which, although it has a complex system of verbal morphology, often relies on syntactic context to identify grammatical functions.

### **1.8 Organization of the Thesis**

In Chapter Two, the internal syntactic structure of Malay noun phrases is examined. I demonstrate that classifiers and reduplicants function as heads of a functional projection Num, because they mark number and Num is the locus of number marking. I also argue that demonstratives in Malay can function without deictic force as definiteness markers, and are heads of D. Establishing the internal syntactic structure of Malay noun phrases will enable us to give a compositional analysis of the semantics of Malay noun phrases in Chapters Four and Five.

In Chapter Three the use of reduplicated versus non-reduplicated nouns is considered. First, it is argued that the root noun is semantically vague, and not ambiguous for number. In other words, the root noun has one reading which can be true of one entity as well as true of more than one entity. I argue in Chapter Three that the choice between reduplicated and root nouns is affected by pragmatic factors. The observations about the use of reduplicated versus non-reduplicated nouns are formalized in Chapter Four.

In Chapter Four, a formal semantic analysis is given of Malay noun phrases involving bare nouns and reduplicated nouns, indefinite as well as definite. In order to accomplish this formal analysis, Link's (1983) analysis of plurals is adopted. The semantics of Malay nouns is compared with English, and a compositional analysis of Malay noun phrases is given. Chapter Four also examines the Link-style classification of predicates into mixed, distributive and collective. The compositional analysis of Malay noun phrases is then extended to simple sentences.

Chapter Five extends the analysis developed in Chapter Four to include classifiers and "mass" nouns. The distinction between count and mass nouns is revisited, and it is argued that Malay does not distinguish, either syntactically or semantically, between mass and count nouns. Three analyses of classifiers are reviewed, and a compositional semantics of phrases containing mensural and sortal classifiers is given, following the analysis of Krifka (1989, 1991, 1995).

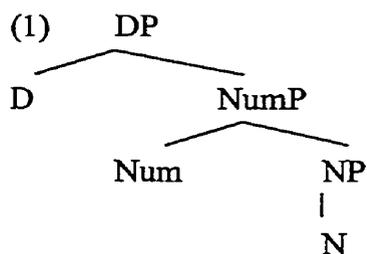
Chapter Six proposes extensions of the analysis developed in this thesis, and problems for further research. Two typologies of noun phrases proposed by Gil (1987) and Chierchia (1999) are reviewed, and it is argued that Malay provides an interesting case to compare with languages like English and Chinese. Problems for further research include: a closer look at child language acquisition of classifier languages, and a closer look at the distribution of quantifiers in Malay.

## Chapter Two – The Internal Structure of Malay DPs

### 2.1 The Structure of Malay DPs

Now that we have seen the basic structure of Malay sentences as well as some basic data on reduplication, classifiers and number, let us examine the internal syntactic structure of Malay noun phrases in more detail. In this chapter, I propose that classifiers and reduplicants occur as heads of Number Phrase (NumP), while demonstratives occur as heads of DP. I also propose that demonstratives can be used without deictic force to mark definiteness. Bare indefinite noun phrases are found to have only NP projections, unlike English bare indefinites which require NumP projections. I will use the syntax proposed in this chapter to develop a compositional analysis of the semantics of Malay noun phrases in Chapters Four and Five.

Following Abney (1987), Ritter (1991) and others, I assume that noun phrases are in fact DPs. DPs are the maximal projection of a functional element, D, which marks definiteness. DPs also contain a second functional projection between DP and NP. This intermediate projection houses information about number, and is therefore named NumP (Number Phrase) (Ritter 1991). The hierarchical structure of a DP is given in (1) below (linear order irrelevant):



Given this structure, we can examine the internal syntax of Malay DPs, which have the following noteworthy characteristics: Malay nouns are neutral for number, reduplication and classifiers are used to mark number in Malay, and Malay does not have a “definite”

determiner, according to traditional descriptions of the language (Macdonald 1976). I extend the structure above to account for the Malay facts by taking a closer look at each of the characteristics of Malay, examining their syntactic behavior, then relating this behavior to the structure proposed. Let us first examine the position of reduplication.

## 2.2 Reduplication

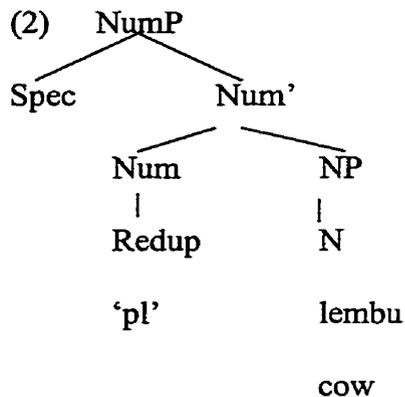
I analyze reduplication in Malay as an abstract inflectional affix that is spelled out as a copy of the root.<sup>10</sup> Recall that Sapir (1921) found that across languages the morphological expression of number can range from derivational to inflectional. I believe that as a number marker in Malay, reduplication is an inflectional affix. Several criteria are commonly used to distinguish inflectional from derivational affixes (Matthews 1991). Derivational affixes normally change the category membership of a word, for instance, adding the affix *-er* to the verb *work* creates the noun *worker*. Inflectional affixes indicate a grammatical subclass of a word; for example, in English number is normally considered an inflectional category. Inflections are also very productive, and apply with few exceptions; for example, in English, the plural *-s* can apply to almost all count nouns. Inflections are also often obligatory, for example, English nouns must agree in person and number features with verbs. In Malay, reduplication does not change the category membership of the word, for example, *budak-budak* ‘children’ is a noun, as is the unreduplicated form *budak* ‘child’. Reduplication is also very productive in that it may apply to almost any noun to indicate the plural in Malay (Hassan 1974). Although we will see in Chapter Three that reduplication is not obligatory, I believe that the above criteria provide sufficient evidence for proposing that, as a number marker, reduplication in Malay is inflectional.

Following recent developments in generative syntax (e.g. Chomsky 1986, Pollock 1989, Ritter 1991 and many others), I assume that inflectional morphemes project their own

---

<sup>10</sup> In other languages reduplication can be derivational rather than inflectional (Kiyomi 1995). Derivational reduplication would require a different analysis than the one presented here.

heads and are affixed to the root via head movement. Because reduplication marks number, I propose that it occurs in the functional projection that houses information about number, namely NumP. A reduplicated noun phrase such as *lembu-lembu* ‘cows’ therefore has the following syntactic structure:



### 2.3 Classifiers

Classifiers also mark number in the sense that they allow the noun to become countable. We saw in Chapter One that Malay has a rich system of sortal classifiers. To demonstrate further, some of the most common classifiers are: *batang*, used with nouns denoting long, large objects; *helai*, used with nouns denoting flat, thin objects; *orang*, used with nouns denoting humans; and *utas*, used with nouns denoting thread-like things. Examples (3)-(6) illustrate these classifiers.

(3) **Tiga**      **batang**      **pokok** berwarna      hijau  
 Three      cl.      tree      color      green  
 ‘Three trees are green’

(4) Dia      sedang      membasuh      **tiga**      **helai**      baju  
 He      prog.      wash      three      cl.      shirt  
 ‘He’s washing three shirts’

(5) **Dua orang budak** sedang bermain

Two cl. child prog. play

'Two children are playing'

(6) Dia membeli **dua utas rantai**

He buy two cl. chain

'He bought two chains'

In the above examples, note that the classifier and the noun always occur together. Also notice from the above examples that classifiers are always used with numerals. In fact, nothing may interrupt the numeral-classifier-noun sequences. In cases where adjectives occur, they must occur after the noun, as in (8).

(7) Tiga ekor ikan

Three cl. fish

'Three fish'

(8) Tiga ekor ikan merah

Three cl. fish red

'Three red fish'

(9) \*Tiga ekor merah ikan

Three cl. red fish

'Three red fish'

(10) \*Tiga merah ekor ikan

Three red cl. fish

'Three red fish'

Numerals are also obligatory with mensural classifiers, introduced in Chapter One. These classifiers "measure out" (T'sou 1976) a quantity of the substance denoted by the noun. Examples (11)-(13) further illustrate this.

- (11) Dia            telah            minum            **dua liter air**  
 He            already            drink            two liter water  
 ‘He drank two liters of water’
- (12) Ada            **se-timbun pasir** putih di sana  
 There is            one pile            sand white over there  
 ‘There is a pile of white sand over there’
- (13) Terdapat            enam ekor            monyet            pada **se-tandan pisang**  
 There is            six cl.            monkey            to one bunch banana  
 ‘There are six monkeys to each big bunch of bananas’

I propose that classifiers occur as heads of NumP. There are several reasons to believe this is the case. Firstly, I note that numeral-classifier sequences are in complementary distribution with reduplication. We therefore have reason to suggest that either the numeral or the classifier is the head of Num. I argue that the classifier is the head of Num for the following reasons. The first reason is syntactic: in many classifier languages, classifiers may occur without numerals but numerals may not occur without classifiers. For instance, in Chinese, classifiers occur without numerals in demonstrative constructions, but numerals cannot occur without classifiers, shown below:

- (14) Zhe zhi quianbi  
 This cl. pencil  
 ‘This pencil’
- (15) \*Liang quianbi  
 Two pencil  
 ‘Two pencils’

(Catherine Goode, p.c.)

The possibility of classifiers without numerals but not vice versa leads me to posit that classifiers, and not numerals, are heads of NumP. Another reason to believe that classifiers are heads of NumP is semantic. We will see in Chapter Five that semantically,

the classifier is a functor that makes the noun countable. As a functor, the classifier takes the numeral and the noun as its arguments. In this way, the classifier is the head of the phrase, and the numeral and noun are in non-head positions. It therefore follows that the classifier should be the head of NumP.

## 2.4 Numerals

Numerals always occur with classifiers, and pick out a specific quantity of the object denoted by the noun. Numerals do not occur with reduplication.<sup>11</sup> In Malay, as we have seen, no material may intervene between the numeral and the classifier.

(16) \*Tiga merah ekor ikan  
 Three red cl. fish

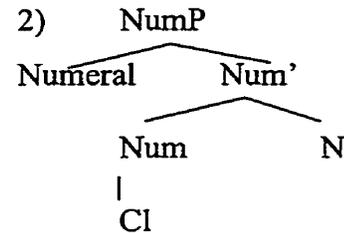
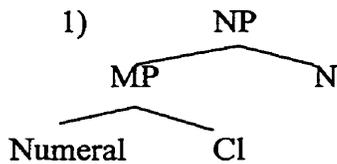
Turning to the position of the numeral, if the numeral is specifying a precise number of the entity being referred to by the noun, then semantically it is acting as a specifier. We therefore have reason to place it in Spec of NumP, rather than in the position of the head of the phrase. This order in Malay of numeral-classifier is also consistent with the proposed universal that specifiers precede heads (Kayne 1994). It therefore makes sense to propose that classifiers occur as heads of NumP, while numerals occur in Spec of NumP.

I have argued that the numeral occurs in Spec of NumP. In terms of how the numeral, classifier and noun are combined in the syntax, however, there are theoretically two possibilities. In the first possibility, assumed by Krifka (1995) as well as Link (1991), the Numeral first combines with the Classifier to produce a Measure Phrase. The Measure Phrase then combines with the Noun to produce the Noun Phrase. In the second

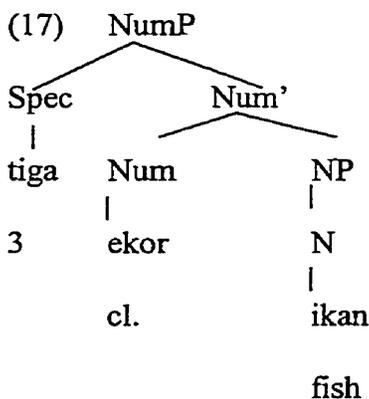
---

<sup>11</sup> A similar phenomenon is also attested for Hungarian, where numerals and plural markers occur in complementary distribution (Ritter 1991).

possibility, the Classifier first combines with the noun to produce Num'; the Number phrase then combines with the Numeral to produce a Number phrase.



It is not clear how to distinguish empirically between these two analyses, as nothing may intervene between the numeral, classifier and noun. However, assuming the second structure, we are able to maintain a parallel between the syntax commonly assumed for DPs ( see Abney, 1987 and Ritter 1991, among others), and the syntax of Malay DPs. I therefore assume that a noun phrase consisting of a numeral, classifier and noun such as *tiga ekor ikan* 'three fish' has the following syntactic structure:



Now that I have established the syntactic structure for phrases which contain numerals and classifiers, I note that the order numeral-classifier-noun is a very common word order for languages which have classifiers. Often, too, the classifier and the numeral must occur adjacent to one another (Rijkhoff 1990, Link 1991, Krifka 1995, Li 1999).

The present formalization of Malay noun phrases makes the important prediction that reduplication and classifiers are mutually exclusive, for the following reason. Both the abstract morpheme, Redup, and the classifier are competing for the same position.

Therefore, since only one item may fill the position, either the Redup morpheme or the classifier may occur as the head of Num, but not both.

To recapitulate, indicators of number occur as the heads of NumP, while numerals, which indicate specific quantities, occur in Spec of NumP. Heads of the Number Phrase in Malay include classifiers and the reduplicant, which is an abstract morpheme that is spelled out as a copy of the noun and is an inflectional affix.<sup>12</sup>

## 2.5 Determiners

### 2.5.1 Definite Determiners

Having established the content and structure of the Number Phrase in Malay, which is the locus of number specification, I will now examine the content and structure of the Determiner Phrase. I will propose that the demonstratives *itu* and *ini*, which always occur at the rightmost edge of the noun phrase, are the heads of DP.

The most common determiners are *itu*, translated as ‘that’, and *ini*, translated as ‘this’. These determiners are traditionally described as demonstratives (Haji 1974b), but also mark definiteness and are frequently translated as ‘the’. Examples of *itu* and *ini* functioning as deictics are shown in (18) and (19). An example of *itu* being used as a marker of definiteness is shown in (20). In cases such as (20), the determiner is used to mark definiteness only, and does not have deictic force. This property of determiners is explored in more detail in Chapter Three.

---

<sup>12</sup> Recent work by Travis (2000) and Ghomeshi (2000) suggests that reduplication is in some cases syntactic, rather than affixal. I do not explore this possibility here.

- (18) **Meja itu** senget  
 Table that crooked  
 ‘That table is arranged crookedly’
- (19) **Beras ini** berkualiti  
 Rice this good quality  
 ‘This rice is good-quality’
- (20) **Budak-budak itu / ini** sentiasa memetik rambutan  
 Child-pl the always pick rambutan  
 ‘The children always pick rambutans’

The head of the determiner phrase, D, is considered the locus of definiteness (Abney 1987). I assume that *itu* and *ini* are determiners, and heads of the Determiner Phrases. This makes sense, as demonstratives are usually considered Ds cross-linguistically. Baggaley (1998), found that *itu* and *ini* occur in complementary distribution with pronouns, which are also considered Ds.<sup>13</sup>

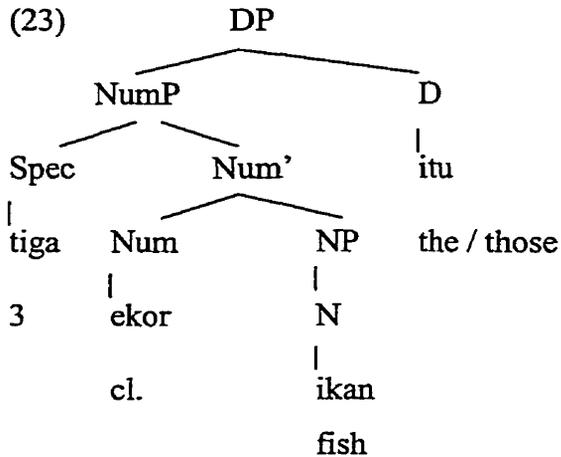
- (21) \*saya itu  
 I that  
 ‘That me’
- (22) \*kamu itu  
 You that  
 ‘That you’

(Taken from Baggaley 1998, p. 101)

We see from (18) – (20) above that *itu* and *ini* always occur at the right-most edge of the noun phrase. The head of the DP, then, is on the on the right. A DP such as *tiga ekor ikan itu* ‘the three fish’ has the following structure:

---

<sup>13</sup> Guilfoyle, Hung and Travis (1992) treat Malay demonstratives as nominal modifiers. I will not pursue this possibility here.

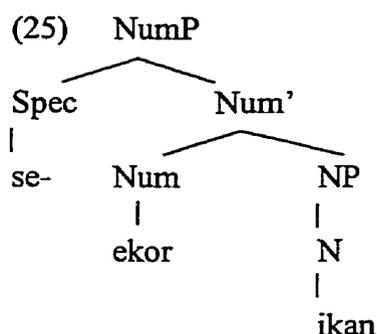


### 2.5.2 Indefinites

We have seen that definite noun phrases are syntactically represented as DPs. We now examine the representation of indefinite noun phrases. In cases which require the indefinite article *a* in English, Malay may use *se-* + classifier, which is historically derived from the numeral *satu* ‘one’ (Guilfoyle et al. 1992).

- (24) Terdapat      **se-buah**      pokok      yang      renek      dikawasan      itu  
 There is      one+cl.      tree      rel.      small area      that  
 ‘There is a small tree in that area’

In (24), the (in)definiteness of the noun is marked through the clitic *se-* ‘one’ which is attached to the classifier. We also notice that in cases such as these, classifiers are obligatory. *Se-*, like numerals, occurs in Spec, NumP; NumP is required in these cases. In cases where definiteness is not marked, I follow Li (1999) and Déchaine and Wiltschko (2000) in assuming that no DP is projected. The syntactic structure of a phrase such as *se-ekor ikan* ‘a fish’ is shown below:



Indefinite phrases also occur without classifiers, however. Examples of bare indefinite noun phrases are shown in (26) and (27).

(26) **Kuda** sedang makan  
 Horse prog. eat  
 ‘Horse(s) is/are eating’

(27) Saya makan **epal**<sup>14</sup>  
 I eat apple  
 ‘I ate an apple / apples’

In the examples above, (in)definiteness is not marked, and the question arises of whether DP and NumP are projected. I argued in Chapter One that root nouns are neutral for number, and we see in the above examples that no overt marker of definiteness is present. I therefore propose that neither NumP nor DP is projected in bare indefinite noun phrases. Further evidence for this proposal can be found by examining the differences between English and Malay. Recall from Chapter One that English makes a syntactic distinction between mass nouns and count nouns, and that Malay treats all nouns as mass-like. English also has agreement for number, where Malay does not. Now let us examine

<sup>14</sup> Examples with bare indefinite nouns such as example (28) have an alternative analysis. In the alternative, the noun is incorporated into the verb, and nouns such as *epal* lose their status as separate arguments. This possibility is not likely, however, as sentences with bare indefinite objects may be passivized. For example, the passive of (28) is: *Epal dimakan oleh saya*. ‘Apple(s) were eaten by me’. The possibility of passivization demonstrates that the object maintains its status as a separate argument, and has not been incorporated (Mithun 1984).

the distribution of bare noun phrases in English and Malay. In English, bare count nouns without number agreement are ungrammatical, whereas in Malay, they are not. This is illustrated below:

- (28) \*Horse is eating  
 (29) **Kuda** sedang makan  
       **Horse** prog. eat  
       ‘Horse(s) is/are eating’

The above examples illustrate a crucial difference in the syntactic representation of English and Malay. While English requires NumP for its count nouns, Malay does not require NumP for any of its nouns, as they are all number-neutral. Note that in English, bare indefinite mass nouns are grammatical. This is shown below:

- (30) Mud is messy

The structure for a bare indefinite noun phrase in Malay and mass nouns in English is therefore simply the NP; no additional nodes are projected. I also assume that if definiteness is not overtly marked, no DP is projected (Li 1999).

- (31) NP  
       |  
       N  
       |  
       ikan

We have now found an important difference between Malay and English. While bare indefinite nouns in Malay require only an NP projection, bare indefinite nouns in English may be of two types. Bare count nouns require number marking and therefore require a NumP projection, and mass nouns do not exhibit number agreement and therefore require only the NP projection. Thus, Malay does not make a syntactic distinction between mass

nouns and count nouns. In Chapter Five, I explore whether a semantic distinction can be made.

## 2.6 Other Aspects of Malay DPs

### 2.6.1 Attributive Nouns

We have seen that classifiers function as heads of NumP when they occur in constructions with numerals. Classifiers may also function as regular nouns, however. This is shown in (32):

- (32) **Buah** itu masam  
 Fruit the sour  
 ‘The fruit is sour’
- (33) Budak-budak itu memetik **buah rambutan**  
 Child-pl the pick fruit rambutan  
 ‘The children are picking rambutan fruit’

In (33) we see that a noun like *buah* ‘fruit’ can also be used as a head noun followed by another, attributive, noun. In such cases where the “classifier” is being used as a regular noun, no numeral is required. How can one distinguish between true nouns and nouns which have undergone zero derivation into classifiers?<sup>15</sup> I will argue that the crucial determining factors are the possibility of reduplication, and the presence of a numeral. Let us first begin by examining regular nouns which are being modified by other nouns, and the restrictions on reduplication. Here we have sentences of the following type:

---

<sup>15</sup> In terms of the direction of the derivation, there are two possibilities: either the noun undergoes zero derivation into a classifier, or the classifier undergoes zero derivation into a noun. As semantic bleaching often occurs in lexical items which are grammaticalized, I assume that the direction of derivation is from noun to classifier. Nothing hinges on this assumption, however.

- (34) **Budak-budak**      **lelaki** itu      sedang      bermain  
 Child-pl              male the      prog.      play  
 ‘The boys are playing’
- (35) **Budak-budak**      **perempuan** itu      sedang      bermain  
 Child-pl              female the      prog.      play  
 ‘The girls are playing’
- (36) **Perempuan-perempuan**      itu      sedang memasak  
 Woman-pl                      the      prog.      cook  
 ‘The women are cooking’
- (37) **Budak-budak**      itu      telah      memakan      semua anggur-anggur  
 Child-pl              the      already      eat      all      grape-pl  
 diatas meja itu  
 on      table the  
 ‘The children ate all the grapes on the table’

In (34) and (35) we have *budak*, which means ‘child’, combining with *lelaki*, meaning ‘male person’, or *perempuan*, meaning ‘female person’. If the word for *male* or *female* is reduplicated alone, it refers to the adult, as in example (36). *Budak-budak* may be used only when referring to a group of children of mixed gender (as in example (37)), and *budak-budak lelaki / perempuan* is the most common form when referring to ‘boys’ or ‘girls’ (as in examples (34)-(35)). If the head noun is present, then only the head noun may be reduplicated; the attributive noun may not. This is illustrated in (38).

- (38) \***Budak**      **perempuan-perempuan**      itu      sedang      bermain  
 Child              female                      the      prog.      play  
 ‘The girls are playing’

This shows that there is only one functional projection that marks number in the noun phrase, and that it targets the head noun. The fact that head nouns may be reduplicated, however, provides us with a good test for whether or not a lexical item is functioning as a

classifier. We have asserted that some nouns can appear to function as classifiers, but that classifiers must always occur with numerals. When functioning as common nouns they can undergo reduplication. When functioning as classifiers, they cannot. This is shown below:

- (39) Budak-budak itu memetik **buah rambutan**  
 Child -pl the pick fruit rambutan  
 ‘The children are picking rambutan fruit’
- (40) Budak-budak itu memetik **buah-buahan**<sup>16</sup> **rambutan**  
 Child -pl the pick fruit-pl rambutan  
 ‘The children are picking different varieties of rambutan fruit’
- (41) Budak-budak itu memetik **tiga buah rambutan**  
 Child -pl the pick three fruit / cl. rambutan  
 ‘The children are picking three rambutan fruits’
- (42) \*Budak-budak itu memetik **tiga buah-buah(an) rambutan**  
 Child -pl the pick three cl. rambutan  
 ‘The children are picking three rambutan fruit’

We have seen that although many nouns appear to function as classifiers, when functioning as nouns, they behave differently. Common nouns may be reduplicated while classifiers may not, and classifiers require numerals while bare nouns do not.

## 2.6.2 Attributive Adjectives and Relative Clauses

Attributive adjectives are also possible in Malay. Attributive adjectives occur after the head noun. When they are reduplicated they do not have the interpretation ‘more than one X’. In (43) an example is given of an attributive adjective, and in (44), we see that the reduplicated adjective has the interpretation ‘reddish’.

<sup>16</sup> The suffix *-an* is usually described as a nominal marker (MacDonald 1976). *-an* is obligatory in (41); however, it is not clear to me why this is the case.

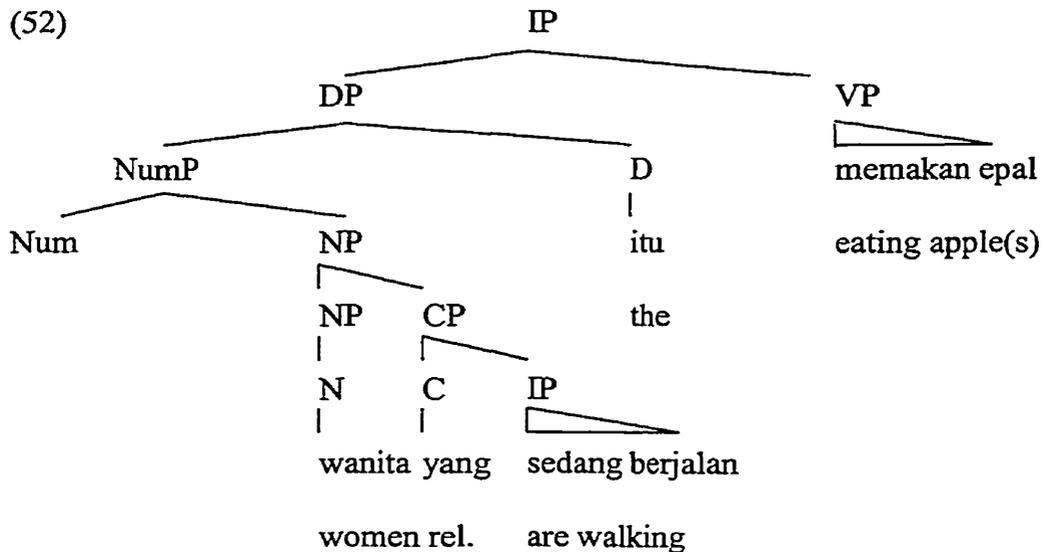
- (43) Tiga ekor ikan merah  
 Three cl. fish red  
 'Three red fish'
- (44) Tiga ekor ikan merah-merah  
 Three cl. fish reddish  
 'Three reddish fish'

Relative clauses in Malay follow the head noun and are formed with the relativizer *yang*. There are several different types of relative clauses in Malay. Examples (45) and (46) show relative clauses containing full VPs, and (47) shows a relative clause consisting only of a predicative adjective. (48) and (49) contain relative clauses without a head noun (or with an empty head noun). Examples (50) and (51) show that classifiers and determiners may occur in simple as well as complex noun phrases. (50) shows a simple noun phrase which combines classifiers and determiners; (51) shows a complex noun phrase containing a relative clause as well as classifiers and determiners.

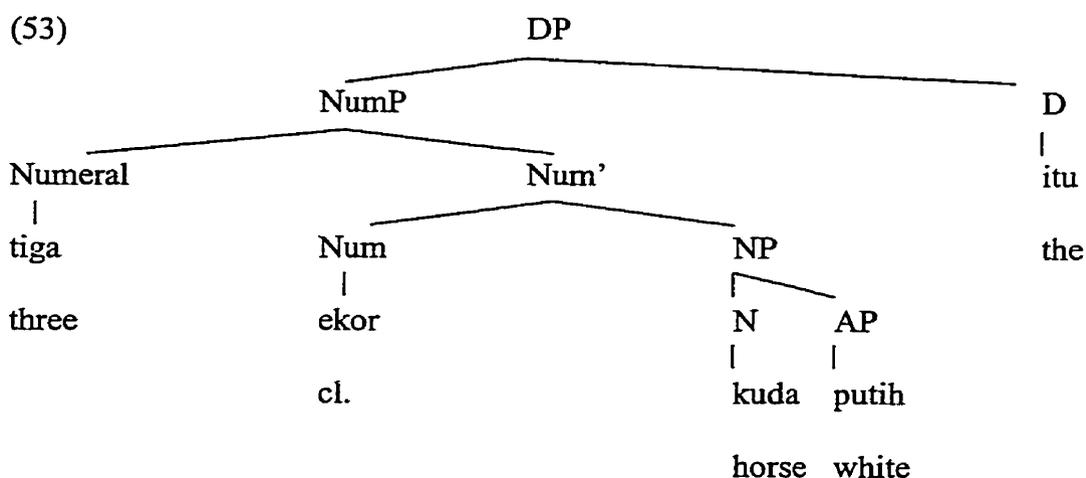
- (45) Wanita yang sedang berjalan itu memakan epal  
 Woman rel. prog. walk the eat apple  
 'The women who are walking are eating apples'
- (46) Lelaki yang berjalan itu  
 Boy rel. walk the  
 'The boy who is walking'
- (47) Buku yang berat itu  
 Book rel. heavy the  
 'The book that is heavy'
- (48) Yang cantik itu sedang berjalan  
 Rel. beautiful the prog. walk  
 'The beautiful ones are walking'

- (49) **Yang sedang berjalan itu** cantik  
 Rel. prog. walk the beautiful  
 'The one who is walking is beautiful'
- (50) **Tiga ekor ikan itu** merah  
 Three cl. fish the red  
 'The three fish are red'
- (51) **Tiga ekor ikan yang merah itu** berat  
 Three cl. fish rel. red the heavy  
 'The three fish that are red are heavy'

Relative clauses provide further evidence that D appears at the right edge of the phrase in Malay, as relative clauses must always appear before the determiner. Syntactically, *yang* can be considered a relativizer which occurs as the head of the CP, and *itu* maintains its status at the end of the noun phrase as a determiner. Although it is not clear where the relative clause is attached in the syntax, one possibility is that CP is an adjunct to NP. The following tree gives the structure for the phrase *wanita yang sedang itu memakan epal* 'the women that are walking are eating apples'.



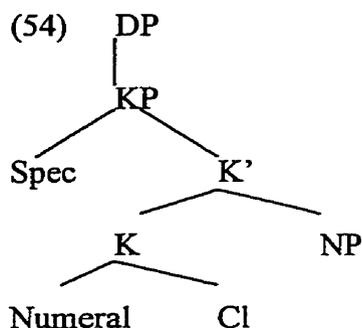
I also assume that attributive adjectives are adjoined to NP, as seen in the following phrase for the sentence *tiga ekor kuda putih itu* ‘those three white horses’:



## 2.7 Alternative Structures for DP in Classifier Languages

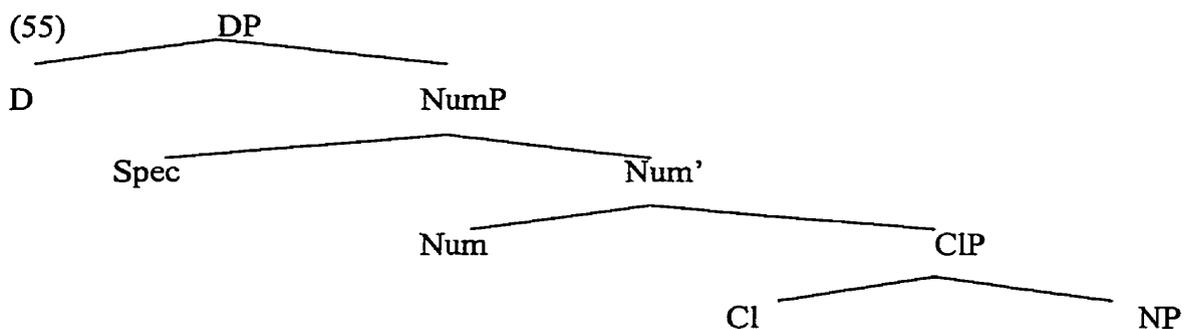
We have seen how the structure proposed for Malay DPs accounts for the Malay data. However, alternative proposals for DP in classifier languages have been made. In this section, I will examine the alternative proposals and show how the one proposed for Malay accounts most straightforwardly for the Malay facts. Although the structure that I have posited for Malay noun phrases is compatible with the data so far, there are alternative proposals which have been posited for other classifier languages such as Chinese, Burmese, Korean and Japanese (see Tang 1990, Gao 1994, Li 1999). The references cited all postulate the existence of at least one intermediate projection between DP and NP, though the nature and number of these projections has been subject to debate. I will review two of the most widely used analyses.

In 1990, Tang proposed the existence of a Classifier Phrase (KP) as the intermediate projection between DP and NP. To a large extent, this proposal makes sense for classifier languages which make use of numeral-classifier constructions rather than singular and plural inflections. Tang proposes the following structure:



In this structure, the numeral and the classifier are under the same lexical head, K. Tang justifies this on the basis that the numeral and classifier always act as a single unit. There is also a universal tendency for the classifier and numeral to be grouped together (Rijkhoff 1990). One disadvantage this analysis faces is that, according to recent syntactic theory, all lexical and functional heads project their own nodes. The proposed structure for Malay avoids this problem by placing the numeral in Spec of Num.

Li (1999) also proposes an analysis of Chinese noun phrases with an intermediate projection between DP and NP. In fact, Li's analysis calls for both NumP and CIP. In Li's analysis, we have the following structure.



This structure has been successful in describing certain Chinese word-order facts. Li (1999) points out that this structure allows us to account for word order facts involving the plural morpheme *-men*. *-men* is analyzed as a plural marker which may only appear on pronouns or proper names, illustrated below:

- (56) \*sange            xuesheng-*men*  
          three-cl.        student-*men*  
          ‘Three students’
- (57) ta-*men* liangge            ren  
          they-*men*        two-cl.            person  
          ‘They two’

According to Li, *-men* is generated in Num, but cannot be realized on common nouns because of the intervening CIP projection. Therefore, *-men* must raise to D, where it may be realized on pronouns and proper nouns base-generated in D. The facts in Malay are slightly different, however, and do not require such an intermediate projection. In fact, positing a Classifier Phrase leads us to make incorrect predictions about possible word orders in Malay. In Malay the plural morpheme is realized as reduplication, which we assume is base-generated in Num. Were we to assume that there was an intermediate Classifier Phrase, the prediction would be that reduplication and classifiers could co-occur. This is clearly not the case, as we saw in Chapter 1, example (16) .

## 2.8 Summary

In this chapter, I examined the internal DP syntax of Malay. I proposed that classifiers and reduplicants are of the category Num, and that *itu* ‘that’ and *ini* ‘this’ are of the category D. I also argued that, unlike English count nouns, all bare nouns in Malay do not have NumP projections. I supported this with the observation that Malay does not require plural marking in nouns, where English does. Lastly, I considered alternate proposals of DPs for languages with classifiers and argued that the structure proposed for Malay best fits the Malay facts. In Chapter Three, I look more closely at the use of root versus reduplicated nouns with the goal of showing that Malay root nouns are indeed neutral for number.

## Chapter Three – The Use of Malay Root versus Reduplicated Nouns

### 3.1 Malay Nouns are Neutral for Number

One of the central claims of this thesis is that Malay nouns are neutral for number. Many have noted this property of nouns in various languages including Malay (Boas 1911, MacDonald 1976, Gil 1987, 1999, Mithun 1988, Link 1991, Krifka 1991, Kang 1994, Chierchia 1998, Ojeda 1998). The possibility of number-neutral nouns has consequences for the representation of nouns in the grammar. Languages like English make a distinction between count nouns such as *cow* (one cow) and typical mass nouns such as *water* (\*one water), whereas languages like Malay essentially treat all nouns as mass nouns. In this chapter, I show that Malay nouns are indeed neutral for number because the unreduplicated noun can be used to refer to one or many entities. Specifically, I will compare Malay with English and show that, in places where English requires the plural form of the noun, Malay allows the root form. However, this issue is obscured because there are pragmatic factors restricting the use of reduplicated versus non-reduplicated nouns. A part of this chapter will be devoted to factoring out these external constraints and distinguishing them from the truth-conditional semantics of reduplication.

Concentrating now on number-neutral nouns, the following sentences may be true if one is referring to one entity as well as if one is referring to more than one entity.

- (1) **Kuda** sedang makan  
**Horse** prog. eat  
 ‘Horses are / a horse is eating’
- (2) **Epal** sangat mahal  
**Apple** very expensive  
 ‘Apples are / An apple is very expensive’

- (3) Mereka sedang menyembelih **lembu** untuk majliskeramaian  
 They prog. slaughter **cow** for feast  
 malam ini  
 night this  
 ‘They are slaughtering cows / a cow for the feast tonight’
- (4) Saya makan **epal**  
 I eat **apple**  
 ‘I eat apples / an apple’

Examples (1)-(4) demonstrate that in both subject and object position, Malay root nouns appear in places where non-singular English nouns would require plural marking. Sentences (1)-(4) may also be used to refer to only one entity (“singular”) in Malay. Hence, the root is considered neutral for number, because it can be used regardless of how many entities one is referring to. This does not mean, however, that Malay does not have a grammaticalized way of expressing number. For example, in (5) reduplication marks ‘more than one’, and in (6) the numeral *tiga* ‘three’ and the classifier *ekor* specify exactly how many entities are being referred to.

- (5) **Anjing-anjing** itu tidur pade waktu tengahari  
 Dog-pl the sleep at time afternoon  
 ‘Those dogs sleep in the afternoon’
- (6) **Tiga ekor** semut dijumpai didapur  
 Three cl. ant found in kitchen  
 ‘Three ants were found in the kitchen’

### 3.2 Vagueness or Ambiguity?

I have described Malay as a language in which noun roots are neutral for number. In this section, I argue that the non-reduplicated noun is vague, and not ambiguous, for number. A lexical item is ambiguous if there are two different meanings associated with it. A

lexical item is vague if it has only one meaning which permits various possibilities (see, for instance, Cruse 1986<sup>17</sup>). In Malay nouns, both possibilities (“singular” and “plural”) are possible under a single reading. Let us demonstrate the idea intuitively. In example (7), imagine a situation in which friend A is watching your things, and you say to friend B about A:

- (7) Kalau **penjat**, dia akan berlari  
 If thief, he will run  
 ‘If there is a **thief** / **thieves**, he will run’

In this case, it does not matter how many thieves come – if any do, friend A will run. Similarly, if you are a day-care worker and you’re leaving for your break, you might utter sentence (8). Here also, it does not matter how many children begin to cry, if any do, you are to be called.

- (8) Kalau **budak** menangis, panggil saya  
 If child cry, call me  
 ‘If a **child** cries / **children** cry, call me’

We see the same thing with predicative noun phrases in (9) and (10), where the non-reduplicated noun *kucing* ‘cat’ may refer to more than one entity, as well as exactly one entity.

- (9) **Callie dan Tisa** kucing  
 Callie and Tisa cat  
 ‘Callie and Tisa are cats’

- (10) **Callie** kucing  
 Callie cat  
 ‘Callie is a cat’

---

<sup>17</sup> Cruse uses the term ‘general’ rather than ‘vague’.

What exactly is the difference between vagueness and ambiguity and how can we distinguish between the two? Cruse (1986) outlines several tests to distinguish between ambiguous and vague meanings. To take an English example, let us examine the lexical items *bank* and *kick*. The noun *bank* is ambiguous, because it has two different meanings: one being ‘a financial institution’, the other being ‘edge of a river’. The word *kick*, however, is vague, because it has only one meaning ‘to strike with the foot’. The word *kick* does not tell us how the kicking was executed i.e. with the right foot or the left foot, because *kick* is vague in this respect. We can see that *bank* is ambiguous and *kick* is vague by applying Cruse’s VP deletion test. This test states that if a deleted VP contains an ambiguous element, then two readings are possible but the deleted element must have the same interpretation as the overt antecedent. If, on the other hand, a deleted VP contains a vague element, then the deleted VP one reading is possible which permits various interpretations. In English, this means that if we take the word *bank* in a conjoined sentence with VP deletion in the second conjunct, only one meaning of *bank* is permissible. In sentence (11), John and Bill must both see either the edge of a river or a financial institution. It is not possible to give this sentence an interpretation in which John sees a financial institution and Bill sees the edge of a river. The word *kick*, on the other hand, is vague. In sentence (12), it doesn’t matter if John kicked the ball with his right foot and Bill kicked the ball with his left foot.

(11) John saw the bank and Bill did too

(12) John kicked the ball and Bill did too

We can apply the VP-deletion test to our Malay data to demonstrate that non-reduplicated forms are vague, rather than ambiguous.

(13) John melihat        **kuda** dan     Bill juga  
        John see            **horse** and     Bill also  
        ‘John saw a horse / horses and Bill did too’

- a) John saw a horse and Bill saw a horse or horses
- b) John saw a horses and Bill saw a horse or horses

(13) is true in a situation in which John saw one horse and Bill saw more than one horse (or vice versa). This is evidence that the root noun is vague, or neutral for number, and not ambiguous.

Now that we have seen that the root noun is indeed neutral for number, the existence of number markers raises the question: exactly what factors govern the use of root versus reduplicated forms? For example, reduplication marks plural forms, and root nouns can be used to indicate one or more than one entity. What, then, is the use of root nouns and reduplicated nouns? I believe that two main factors contribute to the use of root versus reduplicated nouns.<sup>18</sup> These factors are: (in)definiteness of the noun phrase, and the pragmatic context of the linguistic utterance. I will examine both these factors in the following discussion. I examine first indefinite and then definite noun phrases, paying particular attention to the influence of definiteness and pragmatic context.

### 3.3 Indefinite Noun Phrases

#### 3.3.1 Grice's View of Pragmatics

Now, let us examine bare indefinite noun phrases. As seen below, there are differences in the meaning of reduplicated versus non-reduplicated nouns. For example, sentences (14) and (16) are true when referring to one or more than one entity (one or more than one horse was eating / seen), whereas sentences (15) and (17) are only true when referring to more than one entity (more than one horse was eating / seen).

---

<sup>18</sup> There is one other peripheral factor that may contribute to the distribution of root versus reduplicated nouns, and that is lexicalized forms. Lexicalized forms were introduced in section 1.4.3. In some nouns, the reduplicated form has become lexicalized and means something other than 'more than one'. For example *kuda-kuda*, in addition to meaning 'more than one horse', also has the lexicalized meaning 'sawhorse'. In cases such as this, the two forms are in competition and the plural interpretation can be lost. Examples such as these are relatively rare, however.

- (14) **Kuda** sedang makan  
 Horse prog. eat  
 ‘A horse / horses is/are eating’
- (15) **Kuda-kuda** sedang makan  
 Horse-pl prog. eat  
 ‘Horses are eating’
- (16) Saya terlihat **kuda**  
 I see horse  
 ‘I saw a horse / horses’
- (17) Saya terlihat **kuda-kuda**  
 I see horse-pl  
 ‘I saw horses’

In cases such as (14)-(17), how can we account for speakers’ choice between the non-reduplicated form and the reduplicated form? If only one entity is being referred to, only the root form is possible because the sentences are false if we use the reduplicated form as in (15) and (17). If more than one entity is being referred to, however, then the choice of which form to use is not forced, but rather the speaker has to make a choice between the root and reduplicated forms on the basis of pragmatic factors. If it is important for the speaker to emphasize the fact that more than one entity is being referred to, then he or she will choose the reduplicated form. If it doesn’t matter how many entities are being referred to, or the speaker does not know how many entities are being referred to, the root form is used.

One approach to the choice of the non-reduplicated versus the reduplicated form presents itself in Grice’s view of pragmatics. Grice (1975) was interested in the relationship between logic and conversation. He proposed a general principle which participants in a conversation are expected to observe, generally known as the cooperative principle. The co-operation principle is:

“Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.”

Grice (1975, p. 46)

Subsumed under this principle are conversational maxims, which the hearer in the conversation expects the speaker to adhere to. When faced with an apparent violation of the principle, the hearer can figure out the intent of the speaker by calculating conversational implicatures.

Grice’s maxims are quoted below:

Quantity:

1. Make your contribution as informative as is required.
2. Do not make your contribution more informative than is required.

Quality:

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Relevance: Be relevant

Manner:

1. Avoid obscurity of expression.
2. Avoid ambiguity. (Be clear)
3. Be brief.
4. Be orderly.

(Taken from Schiffrin 1994)

Conversational implicatures are not part of the truth-conditional meaning of the sentence, but rather arise from the conversation. They may be cancelled or suspended by explicitly denying them. For example, in English in the context of conjoined clauses the word *and* carries the implicature that the event reported in the first conjunct occurred before the event reported in the second conjunct, as in (18).

(18) Susan took a shower and went for a run

Example (18) carries the implicature that Susan took a shower before she went for a run. This implicature follows from the assumption that the speaker adheres to the Maxim of Manner (submaxim “be orderly”). The implicature can be suspended, however, by simply adding: *but not in that order*.

In the case of Malay indefinites, the sentence containing the reduplicated form is more informative, because it has a narrower meaning than the one with the non-reduplicated form – it is true in fewer situations. However, if the speaker cannot (or does not wish to) be specific, then the non-reduplicated form captures this intended vagueness well. As an example, let us examine the sentence in (19), which may be the response to the question what Bill saw.

(19) Bill    melihat        **kuda**  
        Bill    see               horse  
        ‘Bill saw **a horse / horses**’

Two cases can be distinguished here. In the first case, the speaker cannot be specific about how many horses Bill saw (for example, the speaker doesn’t know). According to Grice’s Maxim of Quantity, which states that one must “make [one’s] contribution as informative as is required”, the hearer assumes that if the speaker was in a position to be more informative, she would have used the reduplicated form. A second case is possible in which the hearer concludes that only one horse was seen. In other words, the non-reduplicated form receives a ‘singular’ reading. Because the reduplicated form is more specific, or denotes a subset of the set that the root noun denotes, the hearer assumes that if the speaker meant to convey that Bill saw more than one horse, he or she would have used the reduplicated form, following the Maxim of Quantity “make [one’s] contribution as informative as is required”. The hearer therefore calculates the implicature that Bill saw one horse.

In terms of articulatory effort, all other factors being equal, longer forms are more cumbersome (Maxim of Manner: “be brief”), and so if the unreduplicated form is being used and the discourse context clearly implies more than one horse (e.g. the participants in the discourse are standing in a field of horses), then the hearer will assume that the speaker means more than one horse, but is simply using the shorter form.

### 3.3.2 Problems with Indefinites

One important complication arises from the above discussion of indefinite Malay noun phrases. I have asserted that Malay indefinites are different from English in that they are number-neutral and can be used to refer to more than one entity. However, in English, sentence (20) is also true if more than one horse was eating. For example, in a situation where one farmer asks the other (21), a reasonable reply could be that in (22). Sentences such as these, then, are true if either one or more than one object is being referred to.

- (20) A horse was eating  
 (21) Was a horse eating here?  
 (22) Yes, in fact, many horses were eating here

What is the difference, then, between Malay and English? One difference manifests itself in the use of anaphoric pronouns. In English, only a singular pronoun may refer back to a singular noun. In Malay, either a singular or a plural pronoun may refer back to the unreduplicated noun. This is illustrated below.

- (23) a1. [A horse]<sub>i</sub> was eating. \*They<sub>i</sub> were white.  
       b1. [A horse]<sub>i</sub> was eating. It<sub>i</sub> was white.
- (24) a2. Kuda<sub>i</sub>        telah        makan. Ia<sub>i</sub>        berwarna        putih.  
       Horse        already        eat        It        color        white.  
       ‘Horse(s) were/are eating. It is white.’

- b2. Kuda<sub>i</sub> telah makan. Mereka<sub>i</sub> berwarna putih.  
 Horse already eat They color white.  
 ‘Horse(s) were/are eating. They are white.’

In these Malay sentences, *kuda* introduces a discourse referent whose number is unspecified. Consequently, both singular and plural pronouns are permitted as anaphors which pick up this discourse referent. In English, the situation is different. *A horse* introduces a singular discourse referent, and so may only be the antecedent of a singular pronoun. Both *horses* and *kuda-kuda* introduce plural discourse referents. As a result, we expect that they will only allow plural pronouns as anaphors. This is indeed the case, as seen below.

- (25) a3. [Some horses]<sub>i</sub> were eating. They<sub>i</sub> were white.  
 b3. [Some horses]<sub>i</sub> were eating. \*It<sub>i</sub> was white.
- (26) a4. [Kuda-kuda]<sub>i</sub> telah makan. Mereka<sub>i</sub> berwarna putih.  
 Horse already eat. They color white.  
 ‘Horses were eating. They were white.’  
 b4. [Kuda-kuda]<sub>i</sub> telah makan. \*Ia<sub>i</sub> berwarna putih.  
 Horse-pl already eat. \*It color white.  
 ‘Horses were eating. \*It was white.’

Given the evidence above, then, we can say that English pronouns have stricter restrictions on discourse referents than Malay nouns do, explaining the fact that English nouns must agree in number with their antecedents. Malay exhibits number-neutral nouns which accommodate a broader range of anaphors.

### 3.4 Definite Noun Phrases

#### 3.4.1 Definiteness and *Itu* and *Ini*

Although it has been established that Malay nouns are neutral for number, in definite noun phrases the data are less straightforward. In many definite noun phrases, the use of the reduplicated form to mark the plural is highly preferred and non-reduplicated forms are normally interpreted as singular.<sup>19</sup> Such cases arise when the DP is headed by the determiners *itu* or *ini*. Traditionally, *itu* and *ini* are described as demonstratives which modify the noun phrase (Hassan 1974). I will show, however, that they can also be used as definite determiners, without deictic force, and that they can restrict the use of unreduplicated nouns.

Definiteness and indefiniteness are illustrated well through a conversation or story, because new discourse referents are introduced by means of indefinite NPs, whereas old discourse referents are referred to by definite NPs (Prince 1981, Heim 1982).

(27)

- |                         |                 |             |               |            |            |            |          |            |
|-------------------------|-----------------|-------------|---------------|------------|------------|------------|----------|------------|
| 1. <b>Seorang</b>       | <b>wanita</b>   | pulang      | dari          | kerja      | dan        | terkerjut  | apabila  | <b>dia</b> |
| One+cl.                 | woman           | came home   | from          | work       | and        | shocked    | when she |            |
| 2. <b>terjumpa</b>      | <b>beberapa</b> | <b>ekor</b> | <b>kuda</b>   | dihalaman  | belakang   | rumahnya.  |          |            |
| find                    | some            | cl.         | horse         | in yard    | back       | house-her. |          |            |
| 3. <b>Kuda-kuda</b>     | <b>itu</b>      | sedang      | memakan       | pokok      | bunganya   | dan        | juga     |            |
| Horse-pl                | the             | prog,       | eat           | tree       | flower-her | and        | also     |            |
| 4. <b>sayur-sayuran</b> | yang            | ditanamnya. | <b>Wanita</b> | <b>itu</b> | berasa     | sangat     | marat.   |            |
| vegetable-plrel.        | plant-her.      | Woman       | the           | felt       | very       | angry.     |          |            |

“A woman came home from work and was shocked when she found some horses in her backyard. The horses were eating her flowers and the vegetables that she had planted. The woman was very angry.”

<sup>19</sup> Gil (1999) found this effect in Singaporean Malay; Kang (1994) found this effect in Korean.

We notice here that the story begins by introducing a character by means of the indefinite noun phrase *seorang wanita* ‘a woman’. After the woman has been introduced, however, she can be referred to by a definite noun phrase *wanita itu* ‘the woman’, as seen in line 4. The same holds for the horses, which are introduced with an indefinite *beberapa ekor kuda* ‘some horses’, and subsequently referred to with a definite noun phrase *kuda-kuda itu* ‘the horses’. We see in this story that *itu* is being used to mark definites, and is translated as ‘the’ rather than ‘this’. Expressions such as *seorang* ‘one’ and *beberapa* ‘a few’ are normally used to mark indefinites.<sup>20</sup> In cases where “plural” discourse referents have been introduced, the reduplicated form of the noun must be used. For example, to begin the second sentence, one may not use *kuda itu* ‘the horse’ instead of *kuda-kuda itu* ‘the horses’ to refer to more than one horse.

We have seen, then, that *itu* may mark definiteness in certain noun phrases, and that this can restrict the use of reduplicated versus non-reduplicated forms. This effect of *itu* is not entirely regular, as we will see in the next section.

### 3.4.2 Definite Determiners and Restrictions on Root Nouns

Now that we have established *itu* and *ini* as markers of definiteness, let us examine their effect on the distribution of root versus reduplicated nouns more closely. Sentences such as (17)–(20), repeated here as (28)–(31) (below), pose a problem at first glance for my assertion that the root noun is neutral for number. In isolation, (28) and (30) are only true when referring to exactly one entity. In demonstrative or definite constructions, speakers clearly prefer a noun phrase referring to more than one entity to be marked with reduplication as in (29) and (31), and an unreduplicated noun to be interpreted as singular. Example (28) shows the ‘only one’ interpretation of the non-reduplicated noun when followed by *itu*. If speakers want to indicate that there was more than one horse,

---

<sup>20</sup> It is not clear what lexical category items such as *beberapa* ‘some’ belong to. They do not follow the N, as *itu* and *ini* do, and therefore it is not clear whether they are Ds. I leave this to further research. See also Chapter Six.

they prefer to use reduplication as shown in (29). Similarly, example (30) would normally be interpreted as referring to one horse, and example (31) would normally be interpreted as referring to many horses.

- (28) **Kuda itu** sedang makan  
 Horse the prog. eat  
 ‘The horse is eating’
- (29) **Kuda-kuda itu** sedang makan  
 Horse-pl the prog. eat  
 ‘The horses are eating’
- (30) Saya terlihat **kuda itu**  
 I see horse the  
 ‘I saw the horse’
- (31) Saya terlihat **kuda-kuda itu**  
 I see horse-pl the  
 ‘I saw the horses’

There are instances, however, where the unreduplicated form + *itu* may be used to refer to more than one entity. For example, one may imagine a situation where two women are looking at a man with a group of horses, and one asks the other: *what’s he doing?* The other may reply using sentence (32). Similarly, one may imagine a situation in which a restaurant worker must clean the floor and the tables and the boss asks the employee what she or he has cleaned. In this case, the answer may be (33). There are also cases where the noun is not overtly realized in the sentence and *itu* may be interpreted as plural. In (34), there is a headless relative clause which refers to a plural subject, meaning something like ‘the ones who...’. If you are standing outside after school and a group of pretty girls walk by, you may utter sentence (34).

- (32) Dia membeli **kuda itu**  
 He buy **horse the**  
 'He's buying the horses'
- (33) Saya membersihkan **meja itu**  
 I clean table the  
 'I've cleaned the tables'
- (34) **Yang cantik itu** berjalan  
 Rel. pretty the walk  
 'The pretty ones are walking'

In each of these examples, it is clear from the extra-linguistic context that the speaker is referring to more than one entity.

We have seen that in definite noun phrases, reduplicated forms are highly preferred, although not obligatory, for marking the plural form of the noun. Non-reduplicated forms are preferred for marking the singular, although they may also be used to refer to more than one entity. What factors can be used to explain this? It is clearly not a purely syntactic phenomenon, as is the case in English. Otherwise, we would expect reduplicated forms to be obligatory whenever more than one entity is being referred to.

I will suggest that *itu* and *ini* are in transition in Malay. Historically, *itu* was unmarked for number, and could therefore be used with the non-reduplicated form to refer to more than one entity, as the non-reduplicated form is neutral for number. I examined some texts from Lewis (1969), and found that at least until recently, *itu* was available for use with the unreduplicated form for plural interpretations.

- (35) Batang keladi itu.... (Lewis 1969, p. 240)  
 Stalk keladi the  
 'The keladi stalks...'

- (36) Maka sepatong itu... (Lewis 1969, p. 238)  
 And dragonfly the  
 ‘And the dragonflies...’

Although in these texts<sup>21</sup> N + *itu* constructions could be used for plural, there is pragmatic pressure to use the reduplicated form when referring to more than one entity because the reduplicated form is stronger and more specific (Maxim of Quantity). This greater informativeness of the reduplicated form creates a conversational implicature that the non-reduplicated form is only used to refer to a single entity. It is often the case that conversational implicatures are grammaticalized, and this process has clearly begun in Malay. When I say that conversational implicatures are grammaticalized, I mean that what was once simply an implicature has become part of the truth conditions associated with a word or phrase. In other words, it is “the fixing of discourse strategies in syntactic and morphological structure” (Traugott and Heine 1991). For example, in English, the word *since* has grammaticalized a causal meaning, which was once only an implicature. Historically, *since* had only a temporal meaning associated with it, and was used to indicate that one event followed the other. However, often when one event follows the other, the first event causes the second. In this way, *since* came to take on a causal interpretation, which has since been grammaticalized. For instance, the sentence (37) would be generally interpreted as meaning something like (38).

- (37) Since Pat started working full time, the house is a mess  
 (38) Because Pat started working full time, the house is a mess

As a tentative hypothesis, I propose that in Malay the implicature that a non-reduplicated noun refers to a single entity is in the process of being grammaticalized in the meaning of the definite determiner *itu*. Thus, in the more “innovative” variety of the language, there are two homophonous *itus*, a plural *itu* which requires the reduplicated form and refers to

---

<sup>21</sup> These texts were collected during the early 1960s or late 1950s (according to Lewis’ 1964 foreword), and are therefore at least a generation before my consultants.

more than one entity, and a singular *itu* which requires the non-reduplicated form and refers to only one entity. The nature of the two *itus* is further explored in section 4.9.

### 3.5 The Nature of Fieldwork

One difficulty with the data on definites is that without spontaneous natural data, it may be impossible to discover what kinds of situations require the reduplicated form to indicate ‘more than one’, and what kinds of situations allow the non-reduplicated form to indicate this. Because of the nature of the interviews that I conducted, the sentences that I received tended to be slightly more formal than what one would normally find in relaxed conversation. It is possible, therefore, that because my consultants were using more formal language, the ratio of reduplicated versus non-reduplicated forms to indicate ‘more than one’ is not the same as one would find in a corpus that contained only naturalistic data. The tendency to use the more precise reduplicated form may be a significant factor in my data. It is also possible that, due to their significant exposure to English, my consultants were more likely to use the reduplicated form only for the plural, parallel to English *-s*. This would be a case of L2 affecting L1 (Kaufman & Aronoff 1991, Archibald 1996, Blair 2000), which is all the more likely since the interviews were conducted predominantly in English.

### 3.6 Summary

In sum, in this chapter we have seen that Malay root nouns are in fact neutral for number. I motivated this claim with explicit examples in which Malay may use the root noun where a plural would normally be required in English. I have also examined cases where the use of unreduplicated nouns is restricted, and have explained these according to pragmatic principles and definiteness. In the next chapter, we will see how some of these findings can be expressed in a formal theory of semantics. My formal analysis of Malay is based on a certain idealization of the data, and on the whole does not take into account the pragmatic factors discussed in this chapter.

## Chapter Four – Formal Semantics of Malay Nouns and Reduplication

### 4.1 Theories of Plural Nouns in English

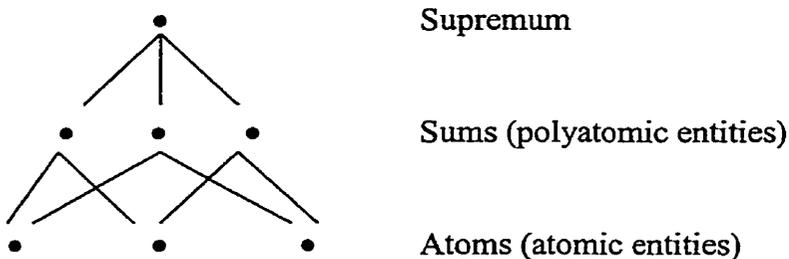
In this chapter, I represent Malay noun phrases in a formal theory of semantics. Link's (1983) theory of plurals and mass nouns in English is extended to the Malay data, with some modifications. I argue that Link's \* operator is not necessary for Malay, and that there are cases in which his use of the  $\sigma$  operator must be modified to account for the Malay facts. A compositional analysis of Malay noun phrases and sentences is developed, and Link's classification of predicates is discussed for Malay.

In order to develop a formal semantics of Malay nouns, we must first review the current approaches to the semantics of plurality in English. There are two general theoretical approaches to plurals in the current literature. One is the approach set forth by Link (1983), which views singulars and plurals both as 'individuals', or entities. This kind of approach is called a domain-oriented approach (Landman 1991). The other is the approach advocated by Landman (1989), which views singulars and plurals as subsets of the set of entities. This kind of approach is called a set-theoretic approach. It is often possible to give a set-theoretic account of a domain-oriented approach and vice versa. Set-theoretic approaches, however, have the disadvantage of not adequately representing mass nouns, as sets presuppose atomicity, and mass nouns are typically thought of as not necessarily atomic (but see Chierchia 1998 for an alternative proposal). I therefore concentrate chiefly on domain-oriented approaches. Also, although both accounts are equally 'valid' for plurals, different authors have focused on different questions, and this has led to several accounts of various aspects of the problem. These issues have not been resolved, and I do not propose to do so here. I do hope, however, that looking at a language such as Malay will shed light on which approaches may be more useful and what a unified account of plurals must achieve in a language like Malay.

## 4.2 Link's (1983) Theory of Plurals in English

Link's theory of plurals in English was developed based on his observation that there is often a parallel in the behavior of plurals and mass nouns. In Link's theory, the domain of individuals (or entities) includes "singular" individuals (atomic entities) such as *John* and "plural" individuals (polyatomic entities) which consist of more than one atom, such as *John and Bill* (together) or *the boys*. While Link uses the terms "singular", "plural" and "individual", I will be using the terms "atomic" and "polyatomic" when referring to semantic objects (entities), reserving the terms "singular" and "plural" for the morphological form of nouns. I also use the term "entity" rather than "individual". This is done in order to maintain consistent terminology for both Malay and English, and to avoid confusion over the terms "singular", "plural" and "individual". Atomic entities are entities with only themselves as parts. Polyatomic entities are made up of more than one atom. Atomic entities can be added together to create polyatomic entities. Atomic and polyatomic entities are related by a part-whole relation, and the maximal element formed from the sum of all elements in a set is called the supremum. Link's domain of entities can be represented by a structure called a semilattice.

### The Semilattice



Link uses the join operation,  $+$ , to form sums of entities, or polyatomic entities. The sum of  $j$  and  $b$  is  $j+b$ , which is itself an entity. Also important in Link's theory is the fact that entities are considered to be in a part-of relation, symbolized by  $\leq_i$ . For example,  $j \leq_i j+b$ . The part-of relation is a partially-ordered relation between entities. It has the formal properties of being reflexive, transitive, and antisymmetric (Landman 1991). A relation is

reflexive iff every element is a part of itself;  $\forall x (x \leq_i x)$ . A relation is transitive iff  $x$  is part of  $y$ , and  $y$  is part of  $z$ , then  $x$  is part of  $z$  ( $\forall x, y, z ((x \leq_i y \wedge y \leq_i z) \rightarrow x \leq_i z)$ ). A relation is antisymmetric, iff  $x$  is a part of  $y$ , and  $y$  is part of  $x$ , then  $x$  is identical to  $y$  ( $\forall x, y ((x \leq_i y \wedge y \leq_i x) \rightarrow x=y)$ ). These three properties of reflexivity, transitivity, and antisymmetry are properties of a semilattice.

Another important aspect of Link's theory of plurals involves his use of the \* operator. \* is an operation that applies to a predicate  $P$  and forms a predicate  $*P$  which denotes the set of all sums that can be formed out of the entities in the denotation of  $P$ . For example, suppose that in a certain model  $[\text{boy}]$ <sup>22</sup> is a set of entities,  $\{j, b, s\}$ , say John, Bill and Sam. Then  $[\text{*boy}]$  is the set of all sums of elements of  $[\text{boy}]$   $\{j, b, s, j+b, j+s, b+s, j+b+s\}$ . According to Link, the denotation of the plural noun *boys* is the set of all the sums in  $[\text{*boy}]$ , minus the set of atomic entities  $At$ . This can be formalized as:  $[\text{*boy}] - At$ , where we take out all atomic boys from the set of boys, leaving us with the set  $\{j+b, j+s, b+s, j+b+s\}$ . This is summarized in Table 1.

**Table 1 - English Translations and Extensions**

Noun	Translation	Sample Extension
boy	$[\text{boy}]$	$\{j, b, s\}$
boys	$[\text{*boy}] - At$	$\{j+b, j+s, b+s, j+b+s\}$
	$[\text{*boy}]$	$\{j, b, s, j+b, b+s, j+b+s\}$

Finally, Link distinguishes three types of predicates: distributive predicates, collective predicates, and mixed predicates. These three types of predicates each have different restrictions on the types of entities they take into their denotations. Link notes that plurals can occur with collective predicates, such as *gather*, *be numerous*, and *be a happy couple* but singulars cannot. For example:

<sup>22</sup> I will be using single square brackets [ ] for denotation. Also, in the notation, I will normally omit reference to a specific model: i.e. I will write [...] rather than [...] <sup>M</sup>.

- (3) a. John and Mary are a happy couple

This sentence could be translated as:  $\text{HappyCouple}(j+m)$ , where  $\text{HappyCouple}$  is a predicate that applies to the sum of John and Mary. However, if John and Mary are a happy couple, this does not imply that John is a happy couple and Mary is a happy couple.

- b. \*John is a happy couple

In this way, collective predicates only apply to polyatomic entities, but not to their atomic parts. In Link's account, a predicate is collective if it contains only polyatomic entities in its denotation. The different types of predicates proposed by Link are explored in detail in Section 4.6.

### 4.3 The Semantics of Malay Nouns

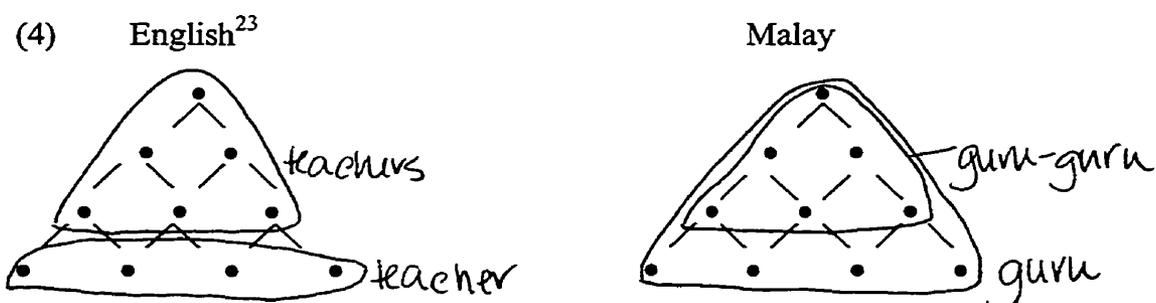
In this section I look at the formal semantics of English nouns and compare their representation with that of Malay nouns. I will argue that in Malay the root noun denotes a set which includes both atomic entities and polyatomic entities (sums). Kang (1994) analyses Korean along the same lines, but in less detail. I also argue that reduplication takes out all the atoms from the denotation of the noun. This semantics of Malay nouns is significantly different from the semantics of English nouns. In English, the root noun denotes a set which includes only atomic entities, and plural nouns denote sets which include only polyatomic entities. For instance, in English,  $[\text{teacher}]$  is the set of all atomic elements that have the property of being a teacher.  $[\text{teachers}]$  is the set of all the non-atomic sums of atomic elements that have the property of being a teacher. It includes only polyatomic elements: To give a concrete example, consider a model where the teachers are John, Bill and Sam. For future reference, I will be calling this Model 1.

### Model 1 – English Nouns

Teacher: [Teacher] = {j, b, s}

Teachers: [\*Teacher] – At = {j+b, b+s, j+s, j+b+s}

Malay differs from English in the following significant respect. In Malay, the denotation of the root noun *guru* ‘teacher(s)’ is neutral for number. It includes **both** atomic entities (such as in English), and polyatomic entities. The denotation of the reduplicated form, *guru-guru* ‘teachers’, includes only polyatomic entities. This is the set [guru] – At. For convenience I introduce an operator, #, which takes the atoms out of the denotation of the noun. Thus [#guru] = [guru] – At. Formally, the # operator can be defined as follows: [#P] = { x | x ∈ [P] ∧ x ∉ At}. This operator captures the semantic effect of reduplication.

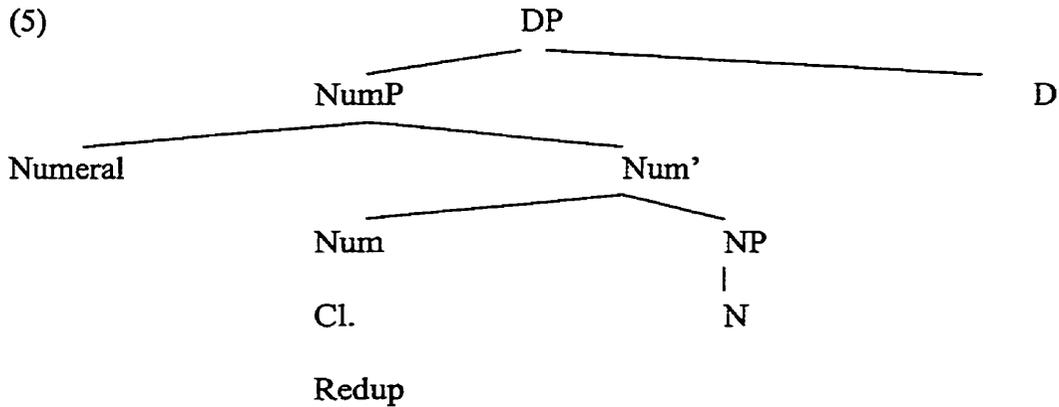


One important similarity between English nouns and Malay nouns is the following: [\*teacher] in English picks out the same set as Malay [guru]. Notice, however, that in Malay, *guru* is an actual word that may be used in natural language whereas in English, there is no simple word or phrase that would denote [\*teacher].

Given the semantic analysis of Malay nouns developed above, I now illustrate how the meanings of noun phrases and sentences are built up from the meanings of the individual lexical items. The way words are combined into phrases and sentences is captured in the

<sup>23</sup> These lattices are simplified versions of complete atomic lattices.

syntactic structure; we therefore recall from Chapter Two that the syntactic structure of Malay DPs is the following:



The syntactic structure of the DP shows us how the words in the phrase are combined. Because the meanings of lexical items are combined into the meanings of phrases, the way in which words are combined into phrases determines the meaning of the phrase. Therefore, given the proposed structure for Malay DPs, we can derive the meanings of Malay expressions by building them up step by step from the bottom of the tree to the top.

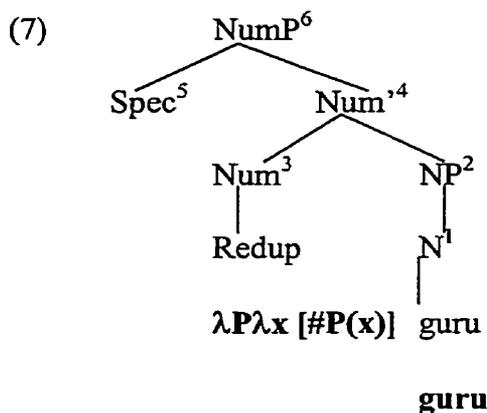
Let us first look at the representation of a simple Malay noun such as *guru* 'teacher'. In the following trees the superscripts on the syntactic labels show the semantic composition, once the individual words have been attached to the bottom of the tree. The semantics of each of the nodes is listed below the tree, showing the derivation of the phrase from the bottom of the tree to the top.



- 1) guru
- 2) guru

As we can see from node 2, the semantics of an NP containing only the noun *guru* ‘teacher’ is *guru*. The semantics of the N *guru* is also *guru*. The reason the semantics of the NP and the N are the same is the following: if nodes are non-branching, the semantic representation is simply projected up the tree. Since we have only the noun *guru* in this phrase, the representation of the noun is projected all the way up to NP, and the semantics of the NP is the same as the semantics of N. The NP *guru* thus denotes a set which includes both atoms and sums in its denotation. Consequently, in our sample model the denotation of the NP is simply  $[guru] = \{j, b, s, j+b, j+s, b+s, j+b+s\}$ . Notice in (6) that the last maximal projection is NP. I assume, as in Chapter Two, that if definiteness and number are not marked, no additional nodes are projected. However, nothing hinges on this assumption, as translations are projected up the tree when the nodes are empty.

We can also give a compositional analysis of noun phrases which consist only of reduplicated nouns such as *guru-guru*. Recall that in reduplicative structures I assume an abstract morpheme Redup that is spelled out phonologically as a copy of the root. Recall also that the # operator captures the semantic effect of reduplication by taking the atoms out of the denotation of the noun.



- 1) guru

- 2) *guru*  
 3)  $\lambda P \lambda x [\#P(x)]$   
 4)  $\lambda P \lambda x [\#P(x)](\text{guru}) \equiv \lambda x [\#\text{guru}(x)]$   
 6)  $\lambda x [\#\text{guru}(x)]$

The translation of *guru-guru* ‘teachers’ is  $\lambda x [\#\text{guru}(x)]$ . Notice that in example (7) the translation of *guru-guru* involves lambda abstraction. Lambda abstraction is a way of defining complex properties in terms of properties already given. It allows us to express sets of objects with complex properties as one-place predicates. An expression of the form  $\lambda x [\phi]$  is a one-place predicate which denotes the set of all objects  $x$  of which  $\phi$  is true. Since  $\lambda x [\phi]$  is a one-place predicate, we can apply it to terms to obtain well-formed formulas. For example, let us take the property of being an unmarried adult male (a bachelor). This property can be defined as follows:

$$(8) \lambda x [\neg \text{married}(x) \wedge \text{male}(x) \wedge \text{adult}(x)]$$

(8) denotes the set of all objects  $x$  such that  $x$  is not married,  $x$  is male and  $x$  is an adult. We can apply the above expression to terms, such as John ( $j$ ). This is shown below.

$$(9) \lambda x [\neg \text{married}(x) \wedge \text{male}(x) \wedge \text{adult}(x)](j)$$

(9) says that the property denoted by (8) applies to John. To further reduce this expression, we drop the  $\lambda x$  at the beginning, substitute  $j$ 's for all the  $x$ 's, and drop the  $j$  at the end. This is known as lambda-conversion and gives us the following expression:

$$(10) \quad [\neg \text{married}(j) \wedge \text{male}(j) \wedge \text{adult}(j)]$$

Two-place predicates can also be formed using lambda abstraction. For example,  $\lambda x \lambda y [\text{love}(x,y)]$  expresses the relation of loving, which is a two-place predicate denoting the set of pairs  $\langle x, y \rangle$  such that  $x$  loves  $y$ .

With this explanation of lambda abstraction in mind, we can discuss example (7). Essentially, in example (7) the denotations of Redup (node 3) and *guru* (node 1) are amalgamated at node 4 to give us the denotation of *guru-guru*. At node 1), *guru* denotes a set which includes both sums and atoms, just as in example (6). In example (7), however, Num is filled with Redup, which acts as an operator and takes the atoms out of the denotation of the noun. Redup is represented semantically as:  $\lambda P \lambda x [\#P(x)]$ . The semantic representation of Redup applies to *guru* at node 4, giving us  $\lambda P \lambda x [\#P(x)](\text{guru})$ , which can be further reduced to  $\lambda x [\#guru(x)]$  by lambda-conversion. The semantic representation of the whole noun phrase is then the set of sums in the denotation of *guru*, translated in reduced form as  $\lambda x [\#guru(x)]$  at node 6.  $\lambda x [\#guru(x)]$  is the set of all objects  $x$  that are in the denotation of  $\#guru$ . Thus,  $[guru-guru] = \{j+b, j+s, b+s, j+b+s\}$ . We note that the  $*$  operator is not necessary here, as the noun already denotes a set that includes sums. Instead, I use the  $\#$  operator to take the atoms out of the denotation of the noun.

#### 4.4 Translation of Definites in English

We now come to the task of translating definite DPs in Malay. Let us first examine the representation of definite phrases in English.  $\sigma x.*Px$  is Link's translation of the definite phrase 'The Ps'.  $\sigma$  is a term-creating operation that picks out the supremum of a set. A formal definition of  $\sigma x.Px$  can be given as follows:  $[\sigma x.Px] = \text{Sup}([P])$  if  $\text{Sup}([P]) \in [P]$ , otherwise undefined.  $\sigma x.Px$  is only defined if the denotation of 'P' forms a semilattice, and  $\sigma x.Px$  is the maximal element (supremum). In the case of definite plurals such as *the teachers*,  $*\text{teacher}$  denotes the semilattice generated by the atoms in *teacher*. Therefore, *the teachers* denotes the maximal element of the set denoted by  $*\text{teacher}$  (Landman 1989, p. 563). In the case of singulars, *teacher* denotes a set of atoms, which is only a semilattice if it contains only one element. Hence, *the teacher* is only defined if there is only one teacher. For example, let us take a model (Model 2) where the denotation of *man* is  $\{j, b, s\}$  and the denotation of *judge* is  $\{j\}$ . In this case, we have the semilattice generated by  $*\text{man}$  whose supremum is  $j+b+s$ . Here the denotation of the phrase *the men*, is defined because  $j+b+s$  is the supremum of the semilattice denoted by  $*\text{man}$ . The

denotation of the phrase *the judge* is also defined, because the denotation of *judge* consists of only one element, and so the supremum of the semilattice formed by *judge* is  $j$ . The denotations of the phrases *the judges* or *the man*, however are not defined, because there is only one judge, and more than one man. We can see this formally by looking at the sets denoted by *judges* or *man*, and applying the definition of  $\sigma x.Px$ . In the case of *the judges*, we end up with the empty set because  $[*judge] - At$  is the empty set.  $\sigma x.man(x)$  is undefined, because the sum of the elements of the set  $[man]$ , i.e.  $j+b+s$ , is not an element of  $[man]$  which only contains atoms. This is summarized in Table 2.

### Model 2

$$[*judge] = [judge] = \{j\}$$

$$[judges] = [*judge] - At = \emptyset$$

$$[man] = \{j, b, s\}$$

$$\sigma x.man(x) = \text{undefined because in Model 2 } \text{Sup}(\{j, b, s\}) = j+b+s \text{ and } j+b+s \notin \{j, b, s\}.$$

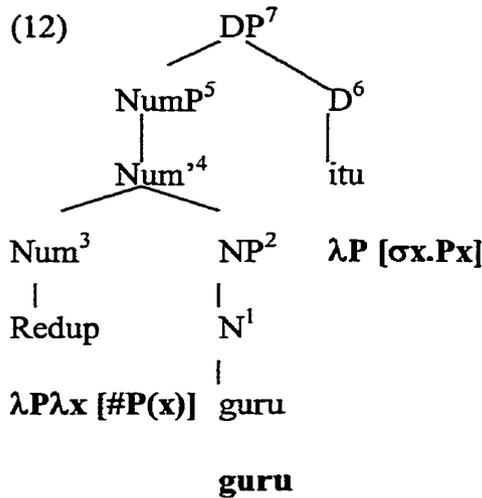
**Table 2 English Definites**

Root	Translation	Extension
the judge	$[\sigma x.Judge(x)]$	J
the men	$[\sigma x.*man(x)]$	$j+b+s$
the man	$[\sigma x.man(x)]$	Undefined because more than one man
the judges	$[\sigma x.*judge(x)]$	Undefined because $[judges]$ is empty

### 4.5 Translation of Definites in Malay

Link's translation of definites in English can be extended to Malay. I will first give the representation for the phrase *guru itu*. Example (11) refers to either one or more teachers.





- 1) guru
- 2) guru
- 3)  $\lambda P \lambda x$  [#P(x)]
- 4)  $\lambda P \lambda x$  [#P(x)](guru)  $\equiv \lambda x$  [#guru(x)]
- 5)  $\lambda x$  [#guru(x)]
- 6)  $\lambda P$  [σx.Px]
- 7)  $\lambda P$  [σx.Px]( $\lambda y$  [#guru(y)])  $\equiv \sigma x.(\lambda y$  [#guru(y)](x))  $\equiv \sigma x.$  [#guru(x)]

We see in the derivation of example (12) that the translation of the phrase *guru-guru itu* is  $\sigma x.(\#guru(x))$ . This denotes the supremum of the set denoted by *guru-guru*. In Model 1, the denotation of  $\sigma x.(\#guru(x))$  is j+b+s. Nodes 1 and 2 are simply the semantic reading of the noun *guru*. Next the # operator (node 3) is applied (node 4) in order to take the atoms out of the set that *guru* denotes. This is projected up to node 5. We then take the supremum of this set in node 7 through the application of the  $\sigma$  operator, which is introduced in node 6 to get the translation of the whole phrase:  $\sigma x.(\#guru(x))$ .

To sum up, the table below gives the translations of phrases involving Malay noun *guru* and their extensions in Model 1.

**Table 3 – Malay Translations and Extensions**

Malay	Translation	Extension
<b>Guru</b>	[guru]	{j, b, s, j+b, j+s, b+s, j+b+s}
<b>Guru-guru</b>	$\lambda x$ [#guru(x)]	{j+b, j+s, b+s, j+b+s}
<b>Guru itu</b>	$\sigma x$ .guru(x)	j+b+s
<b>Guru-guru itu</b>	$\sigma x$ .(#guru(x))	j+b+s <sup>24</sup>

#### 4.6 Link's Classification of Predicates

In order to give a complete compositional analysis, we must show how the meanings of entire sentences can be built up from the meanings of their lexical items. Therefore, I will now extend the analysis developed so far to simple sentences.<sup>25</sup> In order to accomplish this, we take a closer look at the different predicates that Link proposes for English. We will see that Malay also has the three types of predicates Link describes. In this section, translations for sentences with Link-style predicates in English are given; in the next section, Link's proposal is extended to Malay and a compositional analysis of Malay sentences is given.<sup>26</sup> Link recognizes distributive, collective and mixed predicates.

A distributive predicate denotes a set of atomic entities. Thus, the denotation of *blond* is a set of atomic entities: for instance {j, b, s}. A sentence such as *John is blond* then has the following translation:

(13) John is blond

**Translation:** Blond(John)

<sup>24</sup> The difference between *guru itu* and *guru-guru itu* is discussed in section 4.9.

<sup>25</sup> I abstract away from tense in the translations of sentences.

<sup>26</sup> My translations are in the spirit of Link, but are not always identical to the ones he gives. Link gives his translations within a Montague-style grammar.

The “plural” of a predicate like *blond* is formed by using the \* operator, thus [*\*blond*] has both atoms and sums in its extension. For example, if John, Bill and Sam are blond, then [*\*blond*] is the following set: {j, b, s, j+b, j+s, b+s, j+b+s}. We translate a predicate which applies to a subject referring to a polyatomic entity with the \* operator because the denotation of the “starred” predicate contains sums, where that of the “un-starred” predicate does not.<sup>27</sup>

(14) John and Bill are blond

**Translation:** *\*Blond*(j+b)

Collective predicates, on the other hand, denote sets of polyatomic entities. For example, let us take a situation where John, Bill, Sam gather in one place, and Mary, Angie, and Kristy gather in another place. The denotation of *gather* is in this case: {jbs, mak}<sup>28</sup>, which does not contain atomic entities. A sentence such as (15) predicates the property of gathering of the polyatomic entity jbs.

(15) John, Bill and Sam gathered

**Translation:** *Gather*(jbs)

Although Link does not use the \* operator with non-distributive predicates, technically, the \* operator can be applied to any predicate to yield its sums (Schwarzschild 1996). The \* operator can then apply here to create sums of sums, and so [*\*gather*] is: {jbs, mak, jbsmak} In this case, the \* operator simply allows us to form a predicate that is true of the polyatomic entity jbsmak together. For example, let us take a sentence such as (16). (16) will be true in cases where John, Bill and Sam gathered in one place and Mary, Angie and Kristy gathered in another place as well as in cases where they all gathered together in one place. In (16) we have a polyatomic entity with the property of gathering. The

---

<sup>27</sup> A possible motivation for Link’s use of the \* predicate is that it reflects number agreement of the predicate with the subject in English.

<sup>28</sup> The term j+b+s represents a polyatomic entity which may also be written *jbs*. Henceforth, I will use the notation *jbs* for convenience.

sentence will be true if everyone gathered in one place, or if sub-groups gather separately. For further discussion of such cases, see Schwarzschild (1996).

(16) John, Bill, and Sam, and Mary, Angie and Kristy gathered

**Translation:** \*Gather(jbsmak)

Mixed predicates denote sets of entities that are not necessarily atomic. For example, let us take a situation in which John wrote songs by himself, Bill, Sam and Angie wrote songs together, and Mary and Kristy wrote songs together. In this case, [WroteSongs] is a set that may contain both polyatomic and atomic entities: {j, bsa, mk}. Thus, the following sentence could be translated as follows:

(17) Bill, Sam and Angie wrote songs

**Translation:** WroteSongs(bsa)

The plural of a mixed predicate includes both atomic and polyatomic entities, and so [\*WroteSongs] is a set containing all the sums that can be formed from these atomic and non-atomic entities. Using the example given above, this is {j, bsa, mk, jbsa, jmk, jbsamk}. (18) is an example of a sentence whose translation would be \*WroteSongs(jbsamk). In this case, it doesn't matter whether all the individuals wrote songs together or whether John wrote songs by himself – the sentence will be true in either situation.

(18) John, Bill, Sam, Mary, Kristy and Angie wrote songs.

**Translation:** \*WroteSongs(jbsamk)

To take another example, consider the sentence *Rodgers and Hammerstein wrote songs*. This sentence will be true if Rodgers and Hammerstein both wrote songs individually, or if they wrote songs together. Both interpretations are available (Schwarzschild 1996).

## 4.7 Link's Classification Predicates in Malay

### 4.7.1 Distributive predicates

Now that we have seen the three types of predicates posited for English, let us examine the three different types of predicates and their translations in Malay. I will show that Malay has all three types of predicates by Link, and that the kinds of entities different predicates take into their denotation can influence which form of the noun (root vs. reduplicated) is permitted to occur. A complete compositional analysis of Malay sentences is given in Section 4.8.

Distributive predicates must apply to atomic elements. If a group being referred to, then the denotation of the predicate must also include the atoms making up the group. Therefore, in a sentence such as (19), we have the interpretation that either one or more than one cow is being slaughtered, but they are being slaughtered individually. Because of the nature of distributive predicates, if  $a+b$  is being slaughtered, then  $a$  and  $b$  are individually being slaughtered. If there is a group that is being slaughtered, then each of the individuals must also be slaughtered. Example (19) is also true if only one cow is slaughtered. In these translations,  $x$  is a variable that ranges over atomic and polyatomic entities.

(19) Mereka            menyembelih lembu  
       They             slaughter        cow  
       'They are slaughtering cow(s)'

**Translation:**  $\exists x (\text{lembu}(x) \wedge \text{menyembelih}(x))$ <sup>29</sup>

Example (20) only has the interpretation that more than one cow is being slaughtered.

---

<sup>29</sup> Translation given only of relevant part of the sentence.

- (20) Mereka            menyembelih lembu-lembu  
 They                slaughter        cow-pl  
 ‘They are slaughtering cows’  
**Translation:**  $\exists x (\#lembu(x) \wedge menyembelih(x))$

It is important to notice here that the \* operator is not used in the translations of Malay sentences, as predicates may already denote sets which may include sums.<sup>30</sup>

#### 4.7.2 Collective predicates

Collective predicates apply only to polyatomic elements. A sentence such as (21) can therefore only be interpreted as meaning that more than one child gathered in the corridor. This is because of the nature of the predicate *berkumpul* ‘gather’, which only takes sums into its extension.

- (21) Budak berkumpul    di        koridor  
 Child gather            in        corridor  
 ‘Child(ren) gathered in the corridor’  
**Translation:**  $\exists x (\text{budak}(x) \wedge \text{berkumpul}(x))$

Example (22) has the interpretation that children gathered. The only interpretation available is that more than one child gathered. This is because of reduplication, which refers to sums.

- (22) Kanak-kanak berkumpul    di        koridor  
 Kid-pl            gather            in        corridor  
 ‘Kids gathered in the corridor’  
**Translation:**  $\exists x (\#\text{kanak}(x) \wedge \text{berkumpul}(x))$

---

<sup>30</sup> I assume that, because of the loose distinction between lexical categories, both nouns and verbs can denote sets which include sums.

### 4.7.3 Mixed Predicates

Recall that mixed predicates may include both atomic and polyatomic elements in their denotation. Therefore, example (23) is true iff one or more books has the property of being heavy. In other words, there is an entity (atomic or polyatomic) that is heavy. This entity may be heavy because each of the individual atoms in it are heavy, or because they are heavy as a group. In the case of one book, that book must be heavy. I assume here that predicates like *berat* also include both atomic and polyatomic elements in their denotations. The translation of example (23) gives us the set of objects which are both books and are heavy.

- (23) Buku berat  
 Book heavy  
 ‘Book(s) is/are heavy’  
**Translation:**  $\exists x (\text{buku}(x) \wedge \text{berat}(x))$

In a phrase where the noun has been reduplicated, we only have the interpretation that more than one book has the property of being heavy. In other words, there is a polyatomic entity that is heavy. This could be either a) because all the individual books are heavy, or b) because all the books together are heavy. This is shown in (24).

- (24) Buku-buku berat  
 Book-pl heavy  
 ‘Books are heavy’  
**Translation:**  $\exists x (\# \text{buku}(x) \wedge \text{berat}(x))$

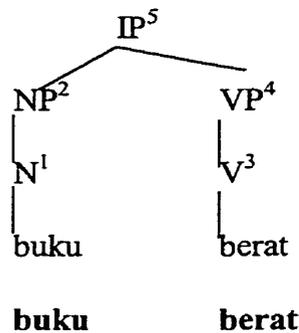
## 4.8 Compositional Analysis of Malay Sentences

We have now looked at the semantic representation of Malay nouns as well as predicates. We have seen that predicates in Malay fit into Link’s classification of predicates, so now

let us see how the meanings of sentences in Malay can be built up compositionally from lexical items. We will first examine sentences with indefinite noun phrases, such as *buku berat* ‘book(s) is / are heavy’.

- (25) *Buku berat*  
 Book heavy  
 ‘Book(s) is/are heavy’

**Translation:**  $\exists x (buku(x) \wedge berat(x))$



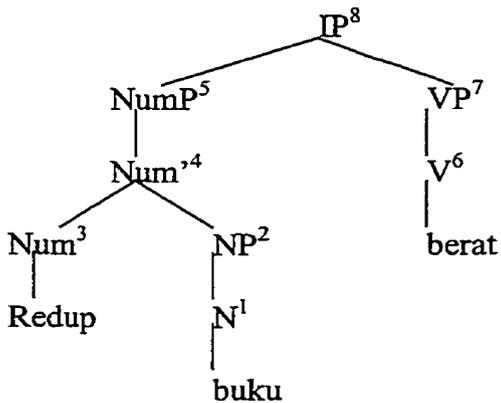
- 1) buku
- 2) buku
- 3) berat
- 4) berat
- 5)  $\exists x(buku(x) \wedge berat(x))$

The translation of example (25) amalgamates the semantics of *buku* and *berat* to give us the set of objects which are both books and are heavy. This could be either because the books individually are heavy, or because they are heavy as a group. The translation of the sentence includes the existential quantifier  $\exists$ . In using the existential quantifier  $\exists$ , I adopt the standard assumption that a rule of existential closure applies at the sentence level (Heim 1982), giving us the translation:  $\exists x(buku(x) \wedge berat(x))$ .

Next, we examine a sentence where the noun has been reduplicated. In the case of (26), we only have the interpretation that more than one book has the property of being heavy. In other words, there is a polyatomic entity that is heavy. This could be either a) because all the individual books are heavy, or b) because all the books together are heavy. The translation of the sentence reflects this, as the # operator takes out all the atoms from the denotation of the noun *buku* at node 4), which then combines with *berat* at node 5) to give us the set of polyatomic entities that are heavy. Again, existential closure applies at sentence level.

(26) Buku-buku berat  
 Book-pl heavy  
 ‘Books are heavy’

**Translation:**  $\exists x (\#buku(x) \wedge berat(x))$



- 1) buku
- 2) buku
- 3)  $\lambda P \lambda x [\#P(x)]$
- 4)  $\lambda P \lambda x [\#P(x)](buku) \equiv \lambda x [\#buku(x)]$
- 5)  $\lambda x [\#buku(x)]$
- 6) berat
- 7) berat
- 8)  $\exists x (\#buku(x) \wedge berat(x))$

We can also derive translations of intransitive sentences with definite subjects.

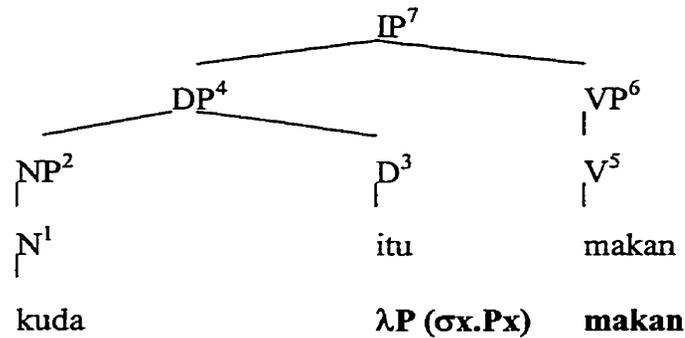
Sentence (27) predicates of one or more horses that they are eating. Recall that  $\sigma$  is a term-creating operator that picks out the supremum of the sub-lattice denoted by the noun. There is one interpretation which is vague as to the number of horses. Recall that if there is only one element in the set,  $\sigma x.Px$  picks out that element. If there is more than one element in the set,  $\sigma x.Px$  picks out the supremum. In sentence (27), the property of eating is being predicated of a unique entity (either atomic or polyatomic); therefore the translation is  $\text{Makan } (\sigma x.kuda(x))$ .

(27) **Kuda itu** sedang makan

Horse the prog. eat

‘The horse(s) is/are eating’

**Translation:**  $\text{Makan } (\sigma x.kuda(x))$



**kuda**

1) kuda

2) kuda

3)  $\lambda P (\sigma x.Px)$

4)  $\lambda P[\sigma x.Px] (\text{kuda}) \equiv \sigma x.[\text{kuda}(x)]$

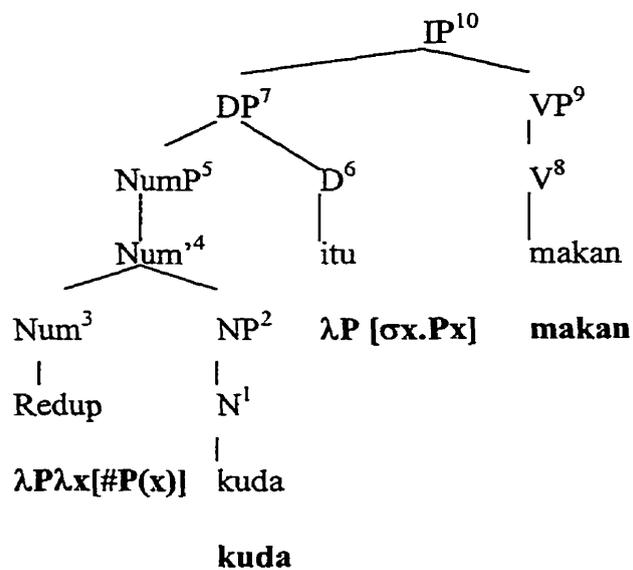
5) makan

6) makan

7) makan  $[\sigma x.kuda(x)]$

In example (28), more than one horse has the property of eating. Because of the presence of reduplication, we make use of the # operator at node 4 to take out the atoms from the denotation of the noun. The  $\sigma$  operator picks out the supremum of the set denoted by the noun at node 7. This translation yields one reading of the sentence that is true in either situation. In other words, the property of eating is being predicated of a unique polyatomic entity.

- (28) **Kuda-kuda itu sedang makan**  
 Horse the prog. eating  
 ‘The horses are eating’  
**Translation:** makan ( $\sigma x.\#kuda(x)$ )



- 1) kuda
- 2) kuda
- 3)  $\lambda\text{P}\lambda x(\#P(x))$
- 4)  $\lambda\text{P}\lambda x [\#P(x)](\text{kuda}) \equiv \lambda x [\#kuda(x)]$
- 5)  $\lambda x [\#kuda(x)]$
- 6)  $\lambda\text{P} [\sigma x.\text{Px}]$
- 7)  $\lambda\text{P} [\sigma x.\text{Px}] (\lambda x [\#kuda(x)]) \equiv \sigma x.\lambda y[\#kuda(y)](x) \equiv \sigma x.(\#kuda(x))$

8) makan

9) makan

10) makan( $\sigma x.\#kuda(x)$ )

#### 4.9 The Problem with Definite Noun Phrases

So far, our translation of definite noun phrases has assumed that the unreduplicated form can be used to refer to a polyatomic entity; in these sentences, the translation of *itu* does not pose a problem to our analysis. However, we saw in Chapter Three that the use of definite noun phrases with unreduplicated nouns is in some instances different from of indefinite noun phrases. We found that nouns used in combination with *itu* can be interpreted as obligatorily singular so that the only the reduplicated form can indicate ‘more than one’.<sup>31</sup> In these cases, *guru itu* picks out a single atom from a singleton set and *guru-guru itu* picks out a polyatomic entity, the supremum of a non-singleton set. This does pose a problem for our translation of *itu*. In these cases we cannot use  $\sigma x.Px$  to force a singular reading because the denotation of *guru* always includes sums. In order to solve this problem, I propose that in the “innovative” variety of Malay, there are two homophonous *itus*: a singular *itu<sub>2</sub>* that takes the non-reduplicated form of the noun, and a plural *itu<sub>1</sub>* that takes the reduplicated form.<sup>32</sup> *Itu<sub>2</sub>* has grammaticalized the implicature discussed in section 3.4.2. *Itu<sub>2</sub>* picks out the element of a singleton set. *Guru itu<sub>2</sub>* is only defined if *guru* denotes a singleton set e.g.  $[guru] = \{j\}$ . I will use, the symbol  $\rho$  to capture the meaning of *itu<sub>2</sub>*.  $\rho x.Px$  can be defined as follows: if  $[P]$  is a singleton set  $\{j\}$ , then  $[\rho x.Px] = j$ . Otherwise,  $\rho x.Px$  is undefined.<sup>33</sup>  $\sigma x.Px$  is used to pick out the supremum of a non-singleton set, in the case of a reduplicated noun + *itu*. The phenomenon of a determiner requiring the noun to be either singular or plural is not unusual cross-linguistically, as number can be reflected in the determiner due to syntactic

<sup>31</sup> Kang (1994) notes that this is also a problem in Korean definite noun phrases.

<sup>32</sup> Gil (1999) also found this effect in Singlish. However, one must note that neither determiners nor nouns are normally marked for definiteness or number; in this case, when combined, they produce an expression which is marked for both definiteness and number.

<sup>33</sup> This is the classical definition of the iota operator (Chierchia & McConnell-Ginet 1996).

agreement. For example, in English *this* occurs only with singular nouns and *these* occurs only with plural nouns.

**Table 4 – Malay Innovative Definite Translations and Extensions**

Root	Translation	Extension
<b>guru itu<sub>2</sub></b>	$[\rho x.guru(x)]$	$j$ – if $j$ is the only element in set
<b>guru-guru itu<sub>1</sub></b>	$[\sigma x.\#guru(x)]$	$j+b+s$ (supremum)

Given the definitions above for the two respective *itus*, we can translate “innovative” definite phrases.

In sentence (29) with *itu<sub>2</sub>* and a non-reduplicated noun, one book has the property of being heavy.

(29) Buku itu<sub>2</sub> berat  
 Book the heavy  
 'The book is heavy'

**Translation:** berat( $\rho x.buku(x)$ )

Sentence (30) with *itu<sub>1</sub>* and a reduplicated noun is vague as to whether all the books as a group have the property of being heavy (collective reading), or each individual book has the property of being heavy.

(30) Buku-buku itu<sub>1</sub> berat  
 Book-pl the heavy  
 'The books are heavy'

**Translation:** berat ( $\sigma x.\#buku(x)$ )

#### 4.10 Summary

In this chapter I demonstrated that Link's theory of plurals in English can be successfully extended to plurals in Malay. I argued that Malay nouns denote sets which include both atomic and polyatomic entities. I also proposed that reduplication takes the atoms out of the denotation of the noun. I used the # operator to capture the semantic effect of reduplication. We also saw that Link's translation of definites in English can also account for some of the Malay data. The formal representation of Malay nouns included a compositional analysis for both noun phrases and sentences. Link's classification of predicates involving distributives, collectives and mixed predicates was extended to Malay, and an examination of these predicates allowed us to develop a compositional analysis of Malay sentences. We will see in Chapter Five that this analysis can be extended to account for classifiers in Malay, following the analyses of Krifka (1989) and Chierchia (1998).

## Chapter Five – Classifiers and the Mass / Count Distinction

### 5.1 Characteristics of Mass Nouns

In order to accomplish a formal semantic analysis of classifiers, we must take a closer look at whether there is a distinction between count and mass nouns in Malay. If there is a distinction between count nouns and mass nouns, our analysis of classifiers will need to reflect this fact. I noted in Chapter One that in terms of syntactic criteria all bare nouns in Malay behave like mass nouns. All bare nouns denote entities that may include more than one atom, and require classifiers in order to be countable. In this chapter I explore the question whether there is a semantic distinction between “mass” and “count” nouns in Malay. I find that Malay does not make a semantic distinction between mass and count nouns, but that the presence of natural units in the entity being referred to can affect the choice of classifier. I then extend the formal analysis of Malay nouns developed in Chapter Four to include classifiers, following the approach of Krifka (1989, 1991, 1995).

Before examining the semantic properties of nouns in Malay, let us review characteristics of mass nouns proposed for languages in general. Syntactically, they often share the following two properties. Mass nouns do not take plural morphology. Nor do mass nouns combine directly with numerals. These properties are illustrated below.

- (1) \*I bought waters
- (2) \*I bought two water

In addition to syntactic criteria, a count / mass distinction can be made in terms of semantic criteria. Bosveld-de Smet (1997) reviews the literature on mass nouns (Quine 1960, Parsons 1979, Bunt 1979, Lonning 1987a), in which two key semantic properties of mass nouns are proposed. The first property is known as the homogeneous reference property. To illustrate, let us take *water* as our example of a typical mass noun. The

homogeneous reference property is fulfilled whenever any sum of parts that are water is water, and any part of a substance that is water is water.

Another characteristic thought to be shared by mass nouns is known as the minimal-parts property. The denotation of a mass noun is typically thought of as lacking minimal parts. In other words, mass nouns have some part-whole structure without singling out any particular part as atomic. This works well for substances like water and flour, which are typically not thought of as consisting of discrete, individuated parts. Generally, substances like gases, powders and liquids do not have perceptually salient individual parts, and so we think of them as lacking minimal parts. In this way mass nouns are like plurals in the sense that they denote a set of entities<sup>34</sup> that are related in a part-of structure. The difference between plurals and mass nouns is that in a mass noun denotation, there are not necessarily any atoms, and in a plural noun denotation there are.

However, there are cases where the minimal parts hypothesis is problematic. Substances like water and flour are referred to by mass nouns and do not have minimal-parts, but other objects denoted by mass nouns such as *furniture* or *change* (in the sense of ‘coins’) do.<sup>35</sup> Furniture and change are typically thought of as consisting of discrete, individual parts. For example, change is made up of individual coins, and we can say that these individual coins act as minimal-parts. Another potential problem with the minimal parts hypothesis is the fact that although they do not have perceptually salient individual parts, many mass substances may still be divisible into discrete individuated singularities at some level. For example, we can refer to a drop of water, a blade of grass, or a grain of sand. I will follow Chierchia (1998) in assuming that mass nouns do not **presuppose** the existence minimal parts, whereas count nouns do.

---

<sup>34</sup> Here I use the term “entity” so as to include “portion of matter”.

<sup>35</sup> In this way, the homogeneous reference hypothesis faces the same problem as the minimal parts hypothesis (Gillon 1991). If mass nouns have the property of distributivity, then they should not have minimal parts. Different authors adopt different solutions to this problem. Some authors follow a weak version of the homogeneous reference hypothesis in which only the property of cumulativity applies. I follow Chierchia (1998) in assuming that mass nouns do not presuppose minimal parts.

## 5.2 Count versus Mass Nouns?

Given the above assumptions about mass nouns, we may ask ourselves the following question: can one truly make a distinction between mass and count nouns in Malay? English has a grammaticalized distinction between mass nouns and count nouns; Malay does not. If we take the noun *water* as an example of a typical mass noun, and *chair* as an example of a typical count noun, we see this difference in the syntax very clearly. *Chair* can be combined directly with numerals, and pluralized, whereas *water* may not. *Water* requires a classifier, such as *liter* in order to combine with a numeral.

- (3) I bought one chair
- (4) I bought chairs
- (5) \*I drank one water
- (6) \*I drank waters
- (7) I drank two liters of water

As seen in Chapter One, all Malay nouns require classifiers when combined with numerals, and all nouns can be reduplicated. There is therefore no syntactic difference between count nouns and mass nouns in Malay.

- (11) Bill    membeli      dua    buah    kerusi  
        Bill    buy                two    cl.     chair  
        ‘Bill bought two chairs’
- (12) \*Bill   membeli      dua    kerusi  
        Bill    buy                two    chair  
        ‘Bill bought two chairs’
- (13) Bill    minum      dua    liter    air  
        Bill    drink            two    liter    water  
        ‘Bill drank two liters of water’

- (14) \*Bill minum            dua    air  
       Bill    drink            two    water  
       ‘Bill drank two waters’

Although there is clearly no syntactic distinction between count and mass nouns in Malay, it is possible that a semantic distinction can be made. Given that the homogeneous reference property and the minimal parts property both seem to target the nature of the entity being denoted by the noun, we can propose that a possible semantic criterion that distinguishes mass nouns and count nouns is what I term the atomicity of the entity denoted by the noun. Some languages may be sensitive to whether the entity denoted by the noun consists of natural units. Korean is such a language. Kang (1994) asserts that Korean is a language in which the noun is number-neutral. Nevertheless, he distinguishes between what he calls “count” and “mass” nouns. I suggest that the distinction that Kang makes in terms of “count” and “mass” is actually a distinction in terms of atomicity. Kang (1994) shows that the plural marker *-tul* can be attached only to “count” nouns such as *sakwa* ‘apple’. For instance *-tul* may not apply to a noun such as *mwul* ‘water’. He also shows that some suffix particles denoting distributivity can only be attached to “count” nouns and that some quantifiers are sensitive to the “count” / “mass” distinction. I propose that these affixes and quantifiers are picking up on the atomicity of the entity denoted by the noun; if the entity denoted by the noun does not contain atoms, then it is impossible to combine them with these inflections / lexical items. In this way, Korean grammaticalizes the semantic property of atomicity.

Given that a language like Korean is sensitive to the atomicity of the entity being referred to, let us examine English and Malay. English distinguishes between entities which consist of atoms, or natural units, such as chairs, and entities which do not have natural units, such as water. One might say there is a continuum between those entities which consist of natural units and those which do not. For example, chairs clearly have natural units and water typically does not, but entities such as sand and grass have natural units (grains, blades) while still being thought of as mass-like because the natural units are so

small as to be imperceptible in many ordinary situations. English ignores this gradation, however. For example, *sand* is a noun which behaves syntactically as a mass noun like *water* even though sand consists of atoms (the individual grains) and water does not (unless one considers the individual molecules). Also, the noun *furniture* behaves as a mass noun even though it contains obvious natural units (the individual pieces of furniture).

Malay also distinguishes between entities which consist of atoms, such as chairs, and entities which do not, such as water. In Malay, this distinction is reflected in the choice of classifier. Entities which consist of natural units combine with sortal classifiers, and entities which do not have natural units combine with mensural classifiers. There are also mixed cases where the entity has natural units, and so can combine with sortal classifiers, and yet can also be thought of as substance-like and therefore can also combine with mensural classifiers. An example of an entity which has natural units while remaining substance-like is sand. Sand has natural units, grains, but is also thought of as a substance that may be measured out in pounds. This distinction in Malay between entities which have natural units and those that do not is not grammaticalized, however. We will see in later sections that all classifiers have specific selectional restrictions on the properties of the objects they are classifying; atomicity is only one such restriction. I therefore propose that Malay grammar makes neither a syntactic nor a semantic distinction between count and mass nouns.<sup>36</sup> As I will show, the formal semantics of all nouns in Malay is the same. They all denote sets which include sums.

Now that we have seen that Malay does not distinguish between count and mass nouns, let us take a closer look at the nature of classifiers. I suggested above that the choice of sortal versus mensural classifier is a selection restriction based on the presence or

---

<sup>36</sup> Here the term ‘semantics’ has two senses: 1) ‘semantics’ as lexical semantics which refers to the meaning of the word in the lexicon and its selection restrictions and 2) ‘semantics’ as formal semantics which refers to meaning in terms of the denotation of a word in a specific model. When I say that Malay does not make a semantic distinction between count and mass nouns, I use the latter sense of “semantic”. Of course, the word for ‘sand’ has a different lexical meaning than the word for ‘chair’, even in Malay.

absence of atoms in the entity denoted by the noun. Any mismatch between classifiers and nouns is semantically anomalous because one has chosen a classifier which picks out the “wrong” properties of the object. For example, using a classifier that applies to flat, thin objects with a noun that denotes an entity which is large and bulky will be semantically anomalous, because there is a mismatch in the properties of the noun and the classifier. This is exemplified below:

- (15) Saya pergi ke kedai untuk membeli lima helai sarong  
 I go to shop for buy five cl. sarong  
 ‘I’m going to the shop to buy five sarongs’
- (16) \*Saya pergi ke kedai untuk membeli lima buah sarong  
 I go to shop for buy five cl. sarong  
 ‘I’m going to the shop to buy five sarongs’

We have seen that different classifiers simply target certain properties of the entity denoted by the noun. If the properties targeted by the classifier do not match the properties of the entity denoted by the noun, we have a mismatch and an anomalous expression. So, although Malay does distinguish the atomicity of objects by means of choice of classifier, there is no need to distinguish grammatically between mass and count nouns in Malay. I examine the formal properties of classifiers further in the next section.

### 5.3 Formal Properties of Classifiers

#### 5.3.1 Link (1983, 1991)

We saw in section 5.2 that all nouns are treated similarly in the grammar. We also saw that classifiers are sensitive to properties such as atomicity, species, shape and size of the entity being referred to. We now examine ways to formalize the semantic effect of classifiers. Before looking in detail at the different views of classifiers, I first introduce

some formal notions designed to formally capture the difference between nouns which denote entities which have natural units and those which do not. Link's view of mass nouns and classifiers is therefore introduced.

According to Link (1983), count terms denote atomic lattices and mass terms denote lattices that are **not necessarily** atomic. A mass term denotes a set of portions of matter closed under sum-formation. Instead of being related via an individual-part-of relation  $\leq_i$ , portions of the mass domain are connected by a material-part-of relation  $\leq_m$ . Here, we do not need to make a commitment as to the presence or absence of minimal parts (sand may have minimal parts; water may not). As Malay classifiers are sensitive to whether the entity being referred to is atomic or not, Link's division in terms of individual-part-of relations and material-part-of relations is useful in reflecting this sensitivity.

Link (1991) extends his view of the count / mass domain to classifiers. His proposal takes into account the fact that classifiers apply only to nouns denoting entities or substances with particular characteristics. According to Link, classifiers denote measure functions. A measure function is a (partial) function from entities onto real numbers. In other words, measure functions apply to (complex) entities in order to obtain the quantity of the entity; formally if  $cl.$  is a classifier,  $x$  is an entity and  $n$  is a natural number, we can say that  $cl(x) = n$ . To take a concrete example, the mensural classifier *liter*, applied to a portion of matter, gives us the exact number of liters in that portion of matter. *Liter* applied to an entity which consists of three liters of water, yields the value 3. Measure functions preserve certain structures in the object domain. One structure preserved in the object domain is the part-of relation. If  $x$  is part of  $y$  in the object domain, the measure function maintains that order, and  $cl(x)$  will be less than  $cl(y)$ . So if  $x \leq_m y$ , then  $cl(x) \leq cl(y)$ . Further, the measure function is considered additive, that is  $cl(x) + cl(y) = cl(x+y)$  where  $x$  and  $y$  are non-overlapping entities. To take an example, if I add two liters of water to three liters of water, I have five liters of water. Link also states that classifiers are lexically specified for the types of entities with which they occur. For example, in order

to be used with the classifier *ekor*, the object being referred to must have the property of being an animal.

We can extend Link's (1991) formalization of mensural classifiers to sortal classifiers.<sup>37</sup> To take a concrete example, let us look at the classifier *ekor* in Malay. To take a concrete example, the sortal classifier *ekor*, applied to an entity, gives us the exact number of animals in that entity. *Ekor* applied to an entity which consists of three animals yields the value 3. *Ekor* is only defined for those entities which are animals. Mensural classifiers also require particular properties. For example, the classifier *liter* requires the entity to which it applies to be a non-solid substance; otherwise it is undefined. The domain of the function *liter* is the set of portions of matter with the property of being non-solid. Link's account of classifiers elegantly expresses the observation that different types of classifiers (sortal vs. mensural) apply to different domains, and that classifiers require certain properties of the entities denoted by the nouns with which they are combining.

In the next section, I review three views of classifiers which build on the general idea of classifiers presented above. I review these three views of classifiers in order to show that Krifka's theory of classifiers can be made to account most straightforwardly for the Malay facts. Krifka (1989, 1991, 1995), states that all classifiers are measure functions. Recall that a measure function is a function from objects onto numbers which preserves certain structures of the object domain, such as the property of addition, and the ordering relation between the parts of the domain (i.e. the part-of structure). The second view advanced by Kang (1994), states that classifiers allow mass nouns to shift from the mass to the count domain. The third, advanced by Chierchia (1998), distinguishes between sortal and mensural classifiers. In Chierchia's theory, sortal classifiers are functions from entities onto sets of atoms, where mensural classifiers are functions onto real numbers. I adopt Krifka's view of classifiers for my analysis of Malay.

---

<sup>37</sup> Link (1991) does not give this formalization of sortal classifiers. The analysis proposed above is compatible with Link's analysis, and is more completely worked out than the analysis Link presents in his 1991 paper.

### 5.3.2 Krifka (1989, 1991, 1995)

Krifka (1989, 1991, 1995) extends and modifies Link’s analysis of classifiers. Krifka uses the term “measure phrase” for constructions such as *five liters* which consist of a numeral (*five*) and a measure word (*liters*). Measure words like *liter*, are functions which apply to entities in order to yield the number of liters the entity consists of.

Krifka postulates that while languages like Chinese have overt classifiers to measure out objects and substances, in English only mass nouns occur with overt classifiers, whereas count nouns have a (non-overt) classifier built into the head noun which measures out objects.

In Chinese, the classifier is overt, and the phrase *san zhi xiong* ‘three bears’ is translated as  $\lambda x [xiong(x) \wedge zhi(x) = 3]$ , that is, it denotes the set of all entities which consist of bear and which contain exactly three atoms. In Krifka’s derivation, the measure word, *zhi* ‘cl’, applies first to the numeral *san* ‘three’, which then combines with the noun *xiong* ‘bear’ to give us the set of objects consisting of three bears<sup>38</sup>. Krifka’s representation is given below and can be read like an “upside down” tree. Syntactic categories are given on the left, and semantic translations are given on the right. The end result is given at the bottom, in this case  $\lambda x [xiong(x) \wedge zhi(x) = 3]$ .

(17)

[<sub>M</sub>zhi],  $\lambda n \lambda P \lambda x [P(x) \wedge zhi(x) = n]$   
 [<sub>Num</sub> san], 3  
 [<sub>MP</sub> san zhi],  $\lambda P \lambda x [P(x) \wedge zhi(x) = 3]$   
 [<sub>N</sub> xiong], xiong  
 [<sub>NP</sub> san zhi xiong],  $\lambda x [xiong(x) \wedge zhi(x) = 3]$

<sup>38</sup> In terms of the semantic analysis of Chinese noun phrases, I adopt Krifka (1989, 1991, 1995) and Chierchia’s (1998) view that classifiers are functions which yield a set of atoms, and do not require numerals as arguments.

Krifka's translation of English mass nouns parallels his translation of Chinese nouns, where the classifier is overt. Thus, a phrase like *five ounces of gold*<sup>39</sup> is translated as  $\lambda x$  [Gold(x)  $\wedge$  oz(x) = 5].<sup>40</sup> In the derivation, the measure word, *ounce*, applies first to the numeral *five*, which then combines with the noun *gold* to give us the set of entities consisting of five ounces of gold.

(18)

[<sub>M</sub> ounces],  $\lambda n \lambda P \lambda x$  [P(x)  $\wedge$  oz(x) = n]

[<sub>Num</sub> five], 5

[<sub>MP</sub> five ounces],  $\lambda P \lambda x$  [P(x)  $\wedge$  oz(x) = 5]

[<sub>N</sub> gold], gold

[<sub>NP</sub> five ounces (of) gold],  $\lambda x$  [gold(x)  $\wedge$  oz(x) = 5]

Given the translation of English mass nouns above, Krifka (1989) can capture the division between constructions such as *five ounces of gold*, which apply to the domain of portions of matter, and those such as *five head of cattle*, which apply to the domain of individuals. This is accomplished in the following way. In the case of an expression like *five ounces of gold*, Krifka asserts that “the measure phrase serves to ‘cut out’ entities of a certain size from a continuum of entities which fall under the head noun” (Krifka 1989, p. 82). On the other hand, in classifier constructions such as *five head of cattle*, the head noun is broken into natural units which depend upon the denotation of the head noun. In these phrases, then, Krifka assumes that the language has a function NU for ‘Natural Unit’. NU applies to the classifiers to yield a measure function compatible with the object lattice. By stating that the measure function must be “compatible with the object lattice”, Krifka captures the fact that classifiers apply to nouns with specific properties. Krifka's derivation of the phrase *five head of cattle* is given below:

<sup>39</sup> Here Krifka assumes that *of* is inserted at the surface.

<sup>40</sup> In Krifka's version, the head noun is “quantized”. Quantized nouns denote objects with precise limits. See Krifka (1989) for details.

(19)

[<sub>M</sub> head],  $\lambda n \lambda x \lambda P [P(x) \wedge \text{NU}(\text{head})(x) = n]$ [<sub>Num</sub> five], 5[<sub>MP</sub> five head],  $\lambda P \lambda x [P(x) \wedge \text{NU}(\text{head})(x) = 5]$ [<sub>N</sub> cattle], cattle[<sub>NP</sub> five head (of) cattle],  $\lambda x [\text{cattle}(x) \wedge \text{NU}(\text{head})(\text{cattle}) = 5]$ 

Krifka also points out that the distinction between singular and plural can be a purely syntactic matter, and does not necessarily match up with semantic number. For example, in English we have plural forms in cases which lack any semantic plurality (e.g. *zero bears*), and in many languages, the singular is used with numerals (Turkish *uc elma* ‘three apple’ vs. *\*uc elma-lar* ‘three apple-plural’). Krifka notes that syntax often forces us to distinguish between mass nouns and count nouns. Another important aspect of Krifka’s (1991) analysis is his proposal (following Carlson 1977) that nouns denote kinds. In English, one of the differences between mass nouns like *wine* and count nouns like *bear* is that the former can be used as directly names of kinds, whereas the latter cannot, for instance *wine is delicious* versus *\*bear is dangerous*. When we examine a language such as Malay, we do not see the same kinds of syntactic distinctions between mass and count nouns.

As we have seen, Krifka extends Link’s theory of classifiers, a critical component in the analysis of languages like Malay. We have also seen that Krifka’s analysis captures differences in the atomicity of the noun by means of the NU operator. This NU operator is important, given the division between nouns which take sortal classifiers and those which take mensural classifiers. I do not adopt these aspects of Krifka’s analysis, as they are not necessary for purposes of this thesis.

### 5.3.3 Kang's (1994) Analysis of Korean Nouns

While Krifka uses classifiers as operators that act as functions compatible with the object lattice, Kang (1994) states that classifiers act as semantic domain shifters. To motivate his claim, Kang (1994) gives a Link-style analysis of Korean nouns. He states that in Korean the denotation of the noun includes both sums and atoms. Kang notes that Korean, like many East Asian languages, does not make a distinction between mass and count nouns, as all appear in the same syntactic structure. He gives the following examples to support his claim.

- (20) Sakwa twu kay  
 Apple two cl.  
 'Two apples'
- (21) Mwul twu can  
 Water two cl.  
 'Two glasses of water'

He also notes that many quantifiers can combine with any noun, shown below.

- (22) Mahun sakwa  
 Many / much apple  
 'Many apples'
- (23) Mahun mwul  
 Many / much water  
 'Much water'
- (24) Cekun sakwa  
 A little / few apple  
 'A few apples'

- (25) Cekun           mwul  
       A little / few water  
       ‘A little water’

Kang analyses Korean “count” nouns as denoting a set that includes atoms and polyatoms because syntactically singular nouns can be used in singular and plural contexts. *Sakwa* denotes \*apple in a Link-style analysis. Korean also has a plural morpheme *-tul* which Kang analyzes as denoting a function which derives a set of plural individuals, or polyatomic entities. In other words *sakwa-tul* denotes \*apple – At.

Because all nouns are neutral for number in Korean, the presence of classifiers is obligatory when counting. Kang explains classifiers as “shifting” the domain of the noun from mass to count. For example, he states that the classifier *can* ‘glass’ shifts the domain of water from a possibly non-atomic semilattice composed of bits of water to an atomic join-semilattice composed of glasses of water. Sortal classifiers, which apply to nouns which denote entities that already contain atoms, apply with minimal effect (i.e. they denote the identity function so that  $f(x) = x$  for all  $x$ ). If the wrong classifier is used, the result of applying the measure function is undefined. For example, \*two pieces of water is semantically anomalous.

This analysis is problematic, however, for the following reason. In order to shift to the count domain from the mass domain, Kang must assume that when a classifier such as *liter* is applied to a noun such as *water*, one ends up with sets of liters of water. There is no unique way to divide a quantity of water up into liters. For example, in a container containing five liters, it is not clear where one liter ends and another begins. Also, one may divide up the container into five liters horizontally, vertically, or perpendicularly. Therefore, liters are not discrete physical entities and their status as atoms as a semilattice is questionable. For this reason, accounts of classifiers which map entities or quantities onto numbers more accurately reflect the effect of the classifier.

Kang's analysis of Korean extends Link's analysis of English to Korean, a language which does not distinguish between mass and count nouns. Kang's analysis therefore provides strong support for my own findings in Malay. However, his analysis of classifiers is problematic as Kang cannot explain why classifiers only occur with numerals in Malay. For these reasons, I reject Kang's analysis of classifiers.

#### 5.3.4 Chierchia (1998) and Classifiers

Chierchia (1998) also investigated the semantic properties of classifiers. He does so by first formalizing the behavior of mass nouns. He argues that mass nouns have a representation parallel to plurals. He states that in the mass domain, minimal parts are not presupposed or foregrounded, and hence the division between singular and plural is neutralized. This is why mass nouns do not take plural morphology – because they are neutral between singular and plural, they already denote sets which include sums, and they do not require plural morphology. Mass nouns also cannot be counted because they do not foreground minimal parts. According to Chierchia, substances which do not have minimal parts readily accessible to our perceptual system (for example, liquids, gases and powders) cannot isolate a suitable set of singularities for the relevant lexical entry to denote. This leads him to the important observation that no language has only count nouns.

Chierchia extends his proposal to classifiers. He argues that there is a difference between classifiers (such as *a grain of*) and measure functions (such as *a liter of*). Classifiers are functions that yield a set of entities rather than a number. Measure functions are functions that yield a number. In Chierchia's analysis, sortal classifiers such as *stack* and *grain* pick out the atoms in the denotation of the noun and map objects onto sets of entities, where mensural classifiers such as *liter* and *pound* pick out arbitrary quantities of the substance denoted by the noun and map these quantities onto numbers.

Chierchia's analysis makes sense for languages such as Chinese, which do not always require numerals with classifiers (in Chinese, classifiers may occur with demonstratives as seen in section 1.5). However, in Malay, classifiers always appear with numerals. It therefore does not benefit us to distinguish between classifier phrases and measure phrases, as all constructions containing classifiers act as "measure" phrases. Chierchia's analysis, then, makes a distinction where none is needed to account for the Malay facts.

I will be assuming Link's formalisation of classifiers as measure functions, and Krifka's derivation of classifier constructions. Link does not give a complete account of the formalization of classifier phrases. Krifka's formalization of classifiers is slightly more complex than my own, although he also assumes classifiers are measure functions. I do not adopt Krifka's notion of NU, which is not necessary for my analysis of Malay. I also assume, contra Link and Krifka, that the measure function, or classifier, combines with the head noun before combining with the numeral. I do this in order to be consistent with the syntax of noun phrases developed in Chapter Two; there is no compelling semantic evidence in favor of either analysis. In the semantic derivations that I am considering all classifiers and measure functions are functions that take two arguments, and as far as the semantics is concerned it does not matter in which order they are combined.

#### **5.4 Classifiers in Malay**

Intuitively, classifiers in Malay, as in other languages, serve to pick a specific quantity of entities out of the set that the noun denotes. The translation of sentences with classifiers will then need to specify how many entities there are, and how they are being measured out. Here I use Link and Krifka's idea that classifiers work as functions which take both the noun and the numeral as arguments. In this case, the classifier is a functor which measures out a specific quantity of the entity being referred to. First, let us take an example such as *tiga liter air* 'three liters of water'.

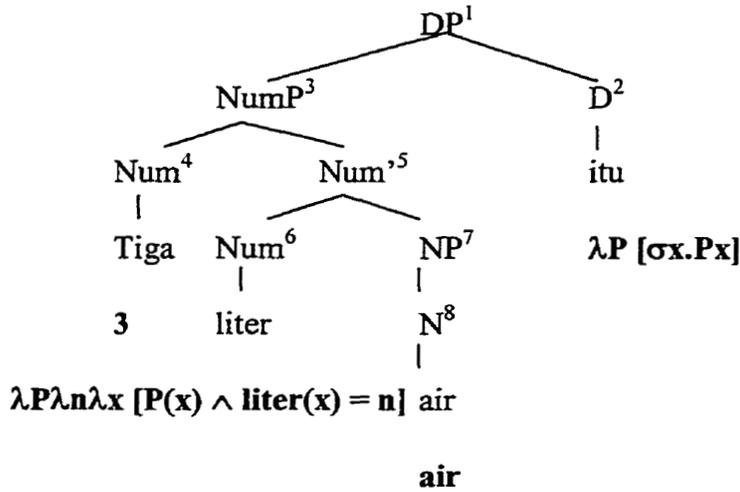


of the noun *air*, which is the set of all portions of matter which are water. In node 5, the measure function takes the noun *air* as an argument. Next, this function is applied to the numeral in node 4. This gives us the set of all portions of matter which are three liters of water, i.e.:  $\lambda x [\text{air}(x) \wedge \text{liter}(x) = 3]$ .

Definite noun phrases with mensural classifiers can be translated in a manner similar to that used in Chapter Four. In this case,  $\lambda x [\text{air}(x) \wedge \text{liter}(x) = 3]$  is the set of all portions of matter consisting of three liters of water.  $\sigma x.Px$  picks out the supremum of this set provided this supremum is itself an element of that set. This will only be the case if there is only one portion of matter which is three liters of water. For example, if I have two containers each containing three liters of water, I have a set consisting of two elements. Each element contains three liters of water, and the sum of all elements of the set is more than three liters of water, it is six liters of water. In this case  $\sigma x.Px$  is undefined. If, on the other hand, I have one container of water which contains three liters of water, I have a set consisting of a single element. The sum of all the elements of this set is therefore three liters of water, and in this case  $\sigma x.Px$  is defined because the supremum of the set is itself an element of the set. In (27), we therefore get an interpretation that picks out the unique entity  $x$  which consists of three liters of water.

(27) tiga liter air itu  
 three liter water the  
 ‘the three liters of water’

**Translation:**  $\sigma x.[\text{air}(x) \wedge \text{liter}(x) = 3]$



- 8) air
- 7) air
- 6)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{liter}(x) = n]$
- 5)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{liter}(x) = n]$  (air)  $\equiv$   
 $\lambda n\lambda x [\text{air}(x) \wedge \text{liter}(x) = n]$
- 4) 3
- 3)  $\lambda n\lambda x [\text{air}(x) \wedge \text{liter}(x) = n]$  (3)  $\equiv$   
 $\lambda x [\text{air}(x) \wedge \text{liter}(x) = 3]$
- 2)  $\lambda P [\sigma x.Px]$
- 1)  $\lambda P [\sigma x.Px]$  ( $\lambda y [\text{air}(y) \wedge \text{liter}(y) = 3]$ )  $\equiv$   $\sigma x.\lambda y[(\text{air}(y) \wedge \text{liter}(y) = 3)](x)$   $\equiv$   $\sigma x.[\text{air}(x) \wedge \text{liter}(x) = 3]$

We can extend this analysis of noun phrases with mensural classifiers to complete sentences, as in Chapter Four. An example of a sentence with this type of classifier is translated the following way:

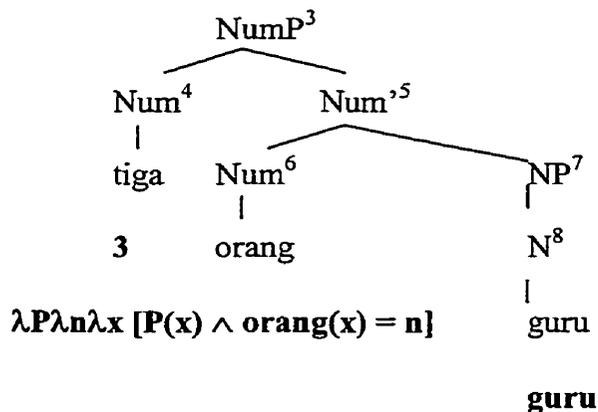
- (28) John minum tiga liter air  
 John drink three liter water  
 ‘John drank three liters of water’  
**Translation:**  $\exists x (\text{air}(x) \wedge \text{liter}(x) = 3 \wedge \text{berat}(x))$

The translation of (28) follows from Chapter Four, where existential closure applies at the sentence level, and the noun phrase contributes a predicate with a free variable. Note that this translation of the sentence yields an interpretation such that there is an entity  $x$  which consists of three liters of water and has the property of being heavy.

We can generalize the above analysis of typical measure functions like *liter* to sortal classifiers, such as *orang* ‘person’. A classifier like *orang* is a measure function which measures out a number of atoms, with restrictions on the properties of the atoms. For example, with *orang* the atoms must be human. We then obtain the following translation for the phrase *tiga orang guru* ‘three teachers’:  $\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$ .

(29) tiga orang guru  
 three cl. teacher  
 ‘three teachers’

**Translation:**  $\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$



8) guru

7) guru

6)  $\lambda P \lambda n \lambda x [P(x) \wedge \text{orang}(x) = n]$

5)  $\lambda P \lambda n \lambda x [P(x) \wedge \text{orang}(x) = n] (\text{guru}) \equiv$

$\lambda n \lambda x [\text{guru}(x) \wedge \text{orang}(x) = n]$

4) 3

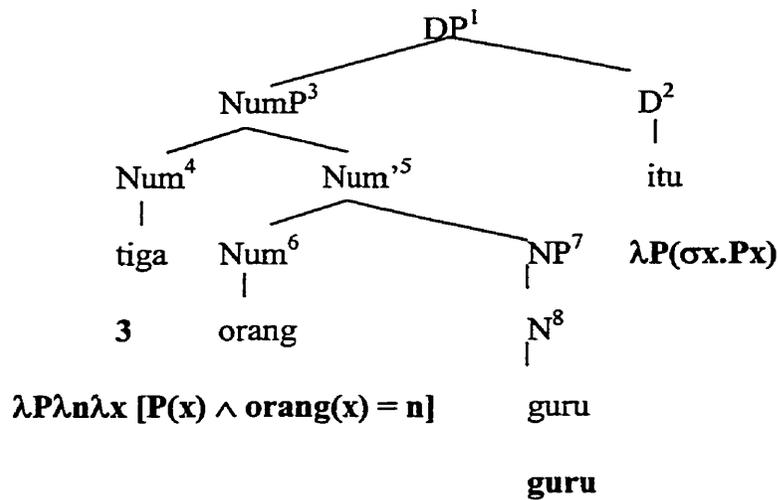
3)  $\lambda n \lambda x [\text{guru}(x) \wedge \text{orang}(x) = n](3) \equiv$   
 $\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$

This translation picks out the set of polyatomic entities consisting of three teachers. It is derived in a manner parallel to example (26). To begin, we have the function denoted by *orang* in node 6. This measure function takes two arguments: first the noun, then the numeral. In nodes 7 and 8 we have the denotation of the noun *guru*, which is the set of all individuals who are teachers. In node 5, the measure function takes the noun *guru* as an argument. Next, the numeral is taken as an argument in node 3. This gives us the set of all entities that consist of three teachers:  $\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$ .

Definite phrases with sortal classifiers can also be translated. In this case,  $\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$  is the set of all entities consisting of three teachers. *Itu* picks out the supremum of the lattice denoted by the NumP provided there is only one entity in the set that consists of three teachers. We therefore get an interpretation that picks out the unique entity *x* which consists of three teachers.

(30) tiga orang guru itu  
 three cl. teacher the  
 'the three teachers'

**Translation:**  $(\sigma x. [\text{guru}(x) \wedge \text{orang}(x) = 3])$



8) guru

7) guru

6)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{orang}(x) = n]$

5)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{orang}(x) = n] (\text{guru}) \equiv$

$\lambda n\lambda x [\text{guru}(x) \wedge \text{orang}(x) = n]$

4) 3

3)  $\lambda n\lambda x [\text{guru}(x) \wedge \text{orang}(x) = n] (3) \equiv$

$\lambda x [\text{guru}(x) \wedge \text{orang}(x) = 3]$

2)  $\lambda P(\sigma x.Px)$

1)  $\lambda P(\sigma x.Px) (\lambda y [\text{guru}(y) \wedge \text{orang}(y) = 3]) \equiv \sigma x.\lambda y[\text{guru}(y) \wedge \text{orang}(y) = 3](x) \equiv$

$\sigma x.[\text{guru}(x) \wedge \text{orang}(x) = 3]$

An example of a sentence with a sortal classifier is translated the following way:

(31) Tiga orang guru berjalan

Three cl. teacher walk

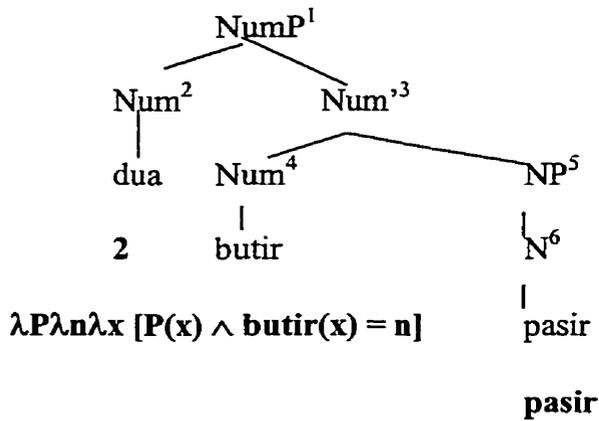
'Three teachers are walking'

**Translation:**  $\exists x (\text{guru}(x) \wedge \text{orang}(x) = 3 \wedge \text{berjalan}(x))$

The above sentence picks out groups of three teachers who are walking. As per usual, existential closure applies at sentence level.

Now that we have seen how typical mensural and sortal classifiers are analyzed, we can derive sentences where the natural unit of a “mass” noun is picked out by a sortal classifier:

- (32) dua butir pasir  
two grain sand  
'two grains of sand'



6) pasir

5) pasir

4)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{butir}(x) = n]$

3)  $\lambda P\lambda n\lambda x [P(x) \wedge \text{butir}(x) = n] (\text{pasir}) \equiv$

$\lambda n\lambda x [\text{pasir}(x) \wedge \text{butir}(x) = n]$

2) 2

1)  $\lambda n\lambda x [\text{pasir}(x) \wedge \text{butir}(x) = n] (2) \equiv \lambda x [\text{pasir}(x) \wedge \text{butir}(x) = 2]$

This analysis is the same as in previous cases.

This analysis of classifiers has the advantage of capturing the fact that classifiers always occur with numerals because the classifier denotes a function which always needs a numeral as an argument. It also has the advantage of treating both sortal and mensural classifiers as simple measure functions. In the next section, we examine cases where the effect of reduplication has the interpretation ‘all kinds of’ rather than ‘more than one’. These cases apply to nouns which are normally thought of as lacking atoms, or minimal parts. We will see that positing a shift from the domain of portions of matter to the domain of kinds accounts for this kind of case.

### 5.5 ‘All kinds of’ Nouns

In some cases, reduplicated nouns have the interpretation ‘all kinds of’ rather than ‘more than one’. This occurs with nouns which are not normally thought of as consisting of discrete, individuatable entities – typical “mass” nouns such as *minyak* ‘oil’, *sup* ‘soup’, and *rumput* ‘grass’. We see this in the example below.

- (33) **Minyak-minyak** masak itu bermutu tinggi  
 Oil-pl cook the quality high  
 ‘The cooking oils are high-quality’  
 all the different varieties of oil are high-quality

We can explain this effect of reduplication in the following way. If the entity denoted by the noun does not have minimal parts, or atoms, then reduplication as we have described it so far, whose function it is to take out the atoms, cannot apply. A semantic shift of domains is therefore forced. In these cases, the semantic shift is from the domain of individuals to the domain of kinds in which the minimal parts each consist of one particular kind. For example, when speaking of kinds of beer, Molson Canadian® would be one kind, Pilsner® another, etc. In this domain it is possible to take out the atoms, and therefore one obtains the interpretation of ‘all kinds of’. I will not be providing a formal

analysis of these kinds of phrases, although Ojeda (1998) provides an analysis of reduplication in Papago that may be an excellent starting point for such an analysis.

## **5.6 Summary**

In this chapter, I showed that there is no distinction between count and mass nouns in Malay, and I extended the analysis of Malay noun phrases to include classifiers. In order to accomplish this, I used Krifka's formalization of classifiers as measure functions, and was able to account for constructions with mensural as well as sortal classifiers. I was also able to capture the fact that in Malay, classifiers always occur with numerals by stating that classifiers denote functions which take two arguments: the noun and the numeral. Lastly, I extended the analysis to account for cases in which reduplication of the noun produces an interpretation of 'all kinds of' by stating that in cases where the entity denoted by the noun does not have atoms, a semantic shift to the domain of kinds takes place.

## Chapter Six –Directions for Further Research and Conclusion

### 6.1 Conclusions

In this chapter a conclusion is given, as well as directions for further research. Along this vein, the role of quantifiers in Malay noun phrases is explored, the typological classification of Malay is discussed as well as the implications of the typological classification for child language acquisition.

In this thesis I have shown that Malay root nouns are neutral for number, even though the language can mark plurality through the use of reduplication. I developed a syntax of Malay DPs in which classifiers and reduplicants are heads of NumP and demonstratives are heads of DP. I explored the semantic representation of the syntactic categories posited and took a detailed look at the use of root versus reduplicated nouns, something other accounts of number-neutral languages have not done. I have shown that the use of reduplicated and non-reduplicated nouns depends on the context of discourse and that Malay demonstratives are beginning to mark definiteness rather than deictic force. I have also explored the Quantity-based implicature associated with non-reduplicated nouns. I have shown that Link's theory of plurals and mass nouns can be extended to Malay, a language which does not distinguish between count and mass nouns. We saw that in languages with number-neutral nouns, nouns have denotations which include both atomic and polyatomic entities and that reduplication removes the atoms from the denotation of the noun. English is different in that root count nouns denote sets of atoms, and plural forms denote sets of polyatomic entities. Extending the observations about root nouns and reduplicated nouns, I examined the distinction between count and mass nouns and found that Malay treats all nouns uniformly in terms of syntax. Malay nouns differ in terms of whether the entity denoted by the noun consists of natural units. The presence or absence of natural units affects the choice of classifier, as sortal classifiers require the existence of natural units, or atoms, and mensural classifiers do not. I presented a formal analysis of classifiers, and in so doing, I presented a compositional semantics of a

language which has obligatory numerals in classifier constructions. This formalization of classifiers was accomplished using Krifka's semantics for classifiers in Chinese and English. Malay is different, however, from the more extensively studied language Chinese. Chinese requires classifiers with determiners as well as with numerals and does not have plural marking.<sup>41</sup> In this way, Malay has provided an interesting perspective on the range of variability possible in classifier languages and this thesis is hopefully the beginning of a more in-depth look at different classifier languages and their characteristics.

## 6.2 Extensions

One interesting area of further research might include an examination of the distribution and semantics of quantifiers in Malay. Quantifiers are normally considered determiners, so a natural extension of the analysis of noun phrases presented in this thesis would be an analysis of quantifiers and their representation in the grammar of Malay.

Two commonly-used quantifiers are *semua* and *setiap*, 'all' and 'every' respectively. Further research is needed on constructions with quantifiers, but I will make some preliminary observations. These quantifiers do not normally occur with numerals. Quantifiers can occur with or without reduplication, and with or without determiners. They do not normally occur with classifiers. Examples of quantifiers with and without reduplication are shown in examples (1) – (5).

- (1) **Semua budak-budak** itu telah mengepong anjing itu  
 All child-pl the already surround dog the  
 'All the children surrounded the dog'

---

<sup>41</sup> Li (1999) argues that *-men* is a plural marker in Chinese. Krifka (1989, 1991, 1995) does not address this possibility in his analysis.

- (2) **Semua buku itu berat**  
 All book the heavy  
 ‘All the books are heavy’
- (3) **Setiap pasukan memenangi satu acara**  
 Each team win one game / event  
 ‘Each team won one game’
- (4) **Setiap buku-buku itu sangat berat**  
 Each book-pl the very heavy  
 ‘Each of the books is individually heavy’
- (5) **Setiap seorang lelaki memenangi satu acara**  
 Each one+cl. boy win one game / event  
 ‘Each and every boy won one event’

From the examples above, we can note that quantifiers do not occur at the right edge of the phrase, and therefore cannot be considered Ds. Nor can quantifiers be of the category Num, as quantifiers can occur with reduplication, and the reduplicant fills the head of Num. Given that quantifiers do not appear to be Ds or Nums, one hypothesis might be that quantifiers are lexical heads and adjoin to the DP or to another maximal projection.<sup>42</sup> Whatever the case may be, it would be interesting to examine the interaction of quantifiers with other lexical items such as determiners, reduplication, and classifiers and to investigate the effect of these interactions on the meaning of the phrase. We can also note that further research might help determine whether it is the case that, as in Korean, quantifiers may be able to occur with all nouns, thereby providing further evidence that Malay does not distinguish between “count” and “mass” nouns. Semantically, an interesting question is what the effect of reduplication of the noun is when there is a quantifier. More generally, we would like to extend the compositional semantics of Malay noun phrases developed in this thesis to quantifiers.

---

<sup>42</sup> It is also possible that quantifiers occur in Spec of Num.

## 6.3 Typological Issues

### 6.3.1 Gil's (1987) Typology of NP Structure

Given the division in the grammar between languages like English, which make use of a count / mass noun distinction, and languages like Malay which do not, one would expect some kind of typology of noun phrases based on the behavior of noun phrases and their components such as quantifiers. David Gil (1987) presents just such a typology of noun phrases based on definiteness, noun-phrase configurationality, and the count-mass distinction. In his view, “whether or not a language possesses an obligatory morphosyntactic strategy for marking (in)definiteness is one of several correlates of a language typology governing NP structure.” (p. 255). Gil splits languages into two types: Type A languages, which distinguish between count and mass nouns, and Type B languages which do not distinguish between count and mass nouns, and can be considered number-neutral. Other characteristics of Type A languages include: obligatory marking of (in)definiteness, obligatory marking of nominal plurality, and the existence of hierarchic interpretations of stacked adjective constructions. Type B language characteristics include: non-obligatory marking of (in)definiteness, non-obligatory marking of nominal plurality, obligatory marking of numeral classification, the existence of adnominal distributive numerals, free NP-internal constituent order, and the existence of stacked adnominal numeral constructions. Gil presents Japanese as a Type B language, and English as a Type A language.

Let us consider Malay according to Gil's typology. The first typological correlate Gil mentions is the obligatory marking of (in)definiteness. Gil asserts that in Type A languages, the distribution of bare nouns is considerably more restricted than that of nouns in construction with articles, and that determiners marking definiteness are obligatory. Although I cannot say with absolute certainty that the distribution of bare nouns in Malay is less restricted than that of nouns with articles, we have seen that bare nouns in Malay appear in structures where they would not be acceptable in English.

- (6) Mereka sedang menyembelih **lembu** untuk majlis keramaianmalam  
 They prog. slaughtering cow for feast night  
 ini  
 this  
 ‘They are slaughtering cow for the feast tonight’

With respect to this correlate, then, Malay patterns as a Type B language.

Secondly, Type B languages do not, according to Gil, obligatorily mark nominal plurality. Although there are restrictions on the use of reduplicated (plural) and non-reduplicated (neutral) forms, the root noun of a word in Malay can be considered as neutral between singular and plural. In this way, then, Malay again patterns with Type B languages.

The third typological correlate involves the use of classifiers. Type B languages treat all nouns as mass nouns, and hence make much less use of nominal plurality markings. We saw this in Chapters One and Three, where the root form of the noun would be acceptable if more than one apple were being referred to. Type B languages also make use of numeral classifiers, as does Malay. Again, Malay appears to pattern with Type B languages.

The fourth correlate involves the existence of adnominal distributive numerals. These are distributive numerals which occur in construction with nouns. An example of an adnominal distributive numeral is shown in (7) from Japanese (Gil 1987, p. 259). English, being a Type A language, does not have adnominal distributive numerals. According to Gil, in Japanese, a Type B language, the numeral-classifier *sansatu* is suffixed with the distributive marker *zutu*, and allows all three English interpretations, shown in (8).

(7) Susumu to Siro-gasansatuzutu no hon-o hakonda  
 Sam and Cyril three-Cl-dist Cop. book-Acc carry-Perf

(8) Sam and Cyril carried the books in threes / three books each / the books three at a time.

I suggest that the prefix *ber-* + reduplication in Malay functions as an adnominal distributive. The example below carries several interpretations, corresponding to the Japanese interpretation of such phrases.

(9) Susan dan Katherine menangat buku **bertiga-tiga**  
 Susan and Katherine carry book in threes  
 ‘Susan and Katherine carried the books in threes / three at a time / three each’

In this case again Malay patterns with Type B languages. The presence of adnominal distributive numerals is possible, according to Gil, because numerals are classified as nominal modifiers rather than determiners, and hence, can distribute over the noun. In Type A languages, numerals are distributives, which cannot be interpreted as functions that satisfy distributivity, and cannot distribute over the noun.

The next correlate Gil discusses is the correlate of “free constituent order”. Type B languages permit a free constituent order within the noun phrase. He gives the following Japanese example (Gil 1987, p. 259):

(10) Susumu-no sansatu no aoi Siro-ga yonda hon  
 Sam-Gen three-Cl. Cop blue Cyril-Nom read-Perf book  
 ‘Sam’s three blue books that Cyril read’

In the above example, we have a nominal head which is modified by a possessor phrase, a numeral, an adjective, and a relative clause. In Gil’s example, the linear order of the four modifiers is flexible, whereas in English the order of modifiers is rigid. In this case,

Malay does not pattern with Type B languages, but rather with Type A languages, as the order of elements in a noun phrase is not free, but rather fixed. We see this in the following ungrammatical phrases, where the order of the constituents in the phrase has been changed.

- (11) tiga ekor ikan merah  
 three cl. fish red  
 ‘three red fish’
- (12) \*tiga ekor merah ikan  
 three cl. red fish  
 ‘three red fish’
- (13) \*tiga merah ekor ikan  
 three red cl. fish  
 ‘three red fish’

The sixth correlate involves stacked numerals. These are constructions in which two numerals can occur adjacent to each other, as adjectives may be stacked in English. Type B languages are said to permit stacked numerals. We see this below in (14), from Japanese which, according to Gil, has “a range of interpretations similar to that of the English *three two-colored books*” (p. 261).

- (14) Sansatu no nisyoku no hon  
 Three-Cl. Cop two-Cl Cop book

Here again, we find that Malay patterns with Type A languages, as stacked numeral constructions are completely ungrammatical.

- (15) \*tiga empat buah buku  
 three four cl. book

The last correlate Gil discusses is the interpretation of stacked adjective constructions, such as *the small powerful engine*. In Type B languages, such a phrase can only be interpreted in one way, that is, as an engine which is small but has a lot of power. In English, however, this sentence has two possible interpretations according to Gil. The first where the engine is powerful relative to engines in general, but small relative to powerful engines; the second where the engine is powerful relative to small engines (p. 262). In Malay, stacked adjectives are not allowed in these kinds of constructions, so it is not possible to test this property.

On the whole, then, we have seen that Malay patterns mostly with the Type B languages, with some exceptions. This is shown more clearly in the following table.

**Table 1 – Typology of Malay Number**

TYPOLOGICAL CORRELATE	TYPE A	TYPE B	MALAY
Obligatory marking of (in)definiteness	+	–	–
Obligatory marking of nominal plurality	+	–	–
Obligatory marking of numeral classification	–	+	+
Existence of adnominal distributive numerals	–	+	+
Free NP-internal constituent order	–	+	–
Existence of stacked adnominal numeral constructions	–	+	–
Existence of hierarchic interpretations of stacked adjective constructions	+	–	N/A

As we can see from our classification of Malay, not all languages can be considered purely Type A or B languages. Gil himself notes the existence of “mixed languages”, which do not fall into either Type A or Type B languages. For example, he classifies Hebrew, Georgian and Russian as mixed languages. As we can see, Malay patterns with Type B languages on the first four correlates, which target number marking and with

Type A parameters on the fifth and sixth correlates, which target the structure of the noun phrase. This being the case, I suggest that these typological correlates target two independent parameters. Still, the correlates examined here suggest that Malay patterns with languages which do not mark their nouns for number. Placing Malay in such a typology clearly shows how Malay noun phrases differ from English noun phrases in terms of an important parameter. It also demonstrates how Malay exhibits properties of a common, yet less understood type of noun phrase.

### **6.3.2 Chierchia's (1998) Typology of NP Structure**

Chierchia (1998) also proposed a semantic parameter of languages which make a count / mass distinction and languages that do not. He states that if a language does not have a count / mass distinction, then it should have the following characteristics. In Chierchia's view, mass nouns are inherently plural. If every noun is plural, then there should be no need for plural morphemes and hence no singular / plural contrast. He also asserts that no indefinite or definite article is necessary, because variants of numerals are covered by the classifiers. Chierchia also predicts the presence of a rich classifier system, and that nouns can occur bare in argument position. He cites Chinese as an example of this kind of language. Malay also appears to have several of these characteristics – it does have a rich classifier system, as seen in Chapter One, section 1.4, and nouns can occur bare in argument position, as seen in Chapter One, section 1.0. However, there are ways in which Malay does not fit quite so nicely into this typology. In Chapter One I argue that reduplication marks the plural in Malay.<sup>43</sup> I also argue in Chapter Three that the demonstratives are coming to mark definiteness only. This type of typology clearly shows the dangers of constructing typological parameters around an examination of very few languages. Although considered a classifier language, Malay lacks many characteristics that Chinese has, and therefore provides an interesting view of classifier languages.

---

<sup>43</sup> Other languages which do not make a count / mass distinction also seem to have plural morphemes, for example, Chinese (Li 1998) and Korean (Kang 1994).

### 6.3.3 Further Typological Questions

Other interesting questions that Malay raises for further typological research include the following. It would be interesting to find out what the correlation is between obligatory plural marking and the presence of classifiers. Sanches (1974) has proposed an implicational universal whereby if a language has classifiers, it will not require plural marking. It would be interesting to see what factors might affect this correlation (e.g. determiners, quantifiers). It would also be interesting to find out if there is a correlation between determiner systems and classifiers. Chierchia (1998) posits that languages with classifiers do not need definite or indefinite determiners. We have seen that Malay provides a counter-example to this claim, but it would be interesting to see if other classifier languages also have this property. Also, the semantic formalization explored in Chapter Four gives rise to an interesting typological question. In Chapter Four, section 4.3, I characterized English as a language where singular noun denotations are sets of atoms and plural noun denotations are sets of sums of atoms. Malay noun roots were proposed to denote sets which included both atoms and sums and reduplicated nouns denote sets which include only sums. A third type of language is theoretically possible where the denotation of root nouns includes sums and atoms, and inflected nouns include only atoms. Maltese (Gil 2000b) and Classical Arabic (Link 1991) appear to be languages which have a 'singulative' operation. It would be interesting to examine the classifier / determiner system of such a language to see what characteristics it may share with languages such as Malay.

### 6.4 Implications for Language Acquisition

As a final question, one might ask what the implications are for child language acquisition. In terms of Gil's typology, one might predict that the two parameters, one based on number representation and the other based on noun phrase configurationality, could be acquired separately. One would therefore expect children to make mistakes in terms of number marking and not in terms of noun phrase configurationality, and vice

versa. One would also expect that, given the syntax developed in Chapter Two, children might acquire lexical categories in Malay before functional categories, as functional categories are largely optional. We therefore expect differences in children acquiring a language like English versus a language like Malay. One difference might be that children learning a language like Malay would learn the plural marker (reduplication) later than children learning English, because in Malay the plural marker is optional. We might also expect children learning Malay to acquire determiners later. Chierchia's (1998) proposal is consistent with this hypothesis. If the count-mass distinction is really a semantic parameter, then we would expect children to make mistakes based on the default setting. As a tentative hypothesis, one may posit that the Malay-type setting is the default, as all languages have mass nouns and not all languages have count nouns. In this case, one would expect English children to start out with the default hypothesis that English is a classifier language. If children do in fact begin with the default hypothesis that English is a classifier language, we might expect them to go through a phase where they treat all nouns as mass nouns. Once, however, they discover that this is not the case, we can surmise that they hypothesize that English makes a count / mass distinction in terms of presence or absence of natural units. In this case, we would expect problems in overgeneralization of plurals to noun like *furniture* and *change* because these nouns are mass nouns, but denote discrete, individual entities. One way to test this would be to look at what types of mistakes children make when learning languages such as English and Malay. We also might expect children to pick up on classifiers which target very salient properties of objects earlier than classifiers which target very abstract properties of objects. Along this vein, we would expect some mensural classifiers to be acquired later than sortal classifiers. Preliminary research in this area with Japanese and English children indicates this may be so (Imai and Gentner 1997).

## References

- Abbi, A. 1985. Reduplicative Structures: A phenomenon of the South Asian Linguistic Area. In Acson, V and Leed, R. (eds.) *Oceanic Linguistics Special Publication No. 20: For Gordon H. Fairbanks*. University of Hawaii Press, Honolulu, p. 159-171.
- Abney, S. 1987. *The English Noun Phrase in Its Sentential Aspect*. Ph.D dissertation, MIT, Cambridge, Massachusetts.
- Allan, K. 1977. Classifiers. *Language*, 53: 285-311.
- Archibald, J. 1996. The Acquisition of Yucatecan Mayan Prosody. In B. Bernhardt, J. Gilbert, and D. Ingram, eds. *Proceedings of the UBC Conference on Phonological Acquisition*. Cascadilla Press, p. 99-112.
- Asher, R. & J. Simpson. 1994. *Encyclopedia of Language and Linguistics*. Oxford, Pergamon.
- Baggaley, V. 1998. *The Syntactic Category of Pronouns*. Unpublished M.A. thesis, University of Calgary.
- Blair, L. 2000. *Arabic / English Bilingual Proficiency in Language Minority Students*. Unpublished M.A. thesis, University of Calgary.
- Boas, F. 1911. *Handbook of North American Languages, Part I*. Bureau of American Ethnology Bulletin 40, Government Printing Office, Washington, D.C.
- Bosveld-de Smet, L. 1997. *On Mass and Plural Quantification: The Case of French des/du-NPs*. Groningen Dissertations in Linguistics 22, University Press, Groningen.
- Bright, W. (ed.) 1991. *International Encyclopedia of Linguistics*. Oxford University Press, New York.
- Bunt, H. 1979. Ensembles and the Formal Semantic Properties of Mass Terms. In F.J. Pelletier (ed.), *Mass Terms: Philosophical Problems*. Synthese Language Library, Vol. 6, p. 249-277, D. Reidel, Dordrecht.
- Carlson, G. 1977. *Reference to Kinds in English*. Ph.D Dissertation, University of Massachusetts, Amherst.
- Chierchia, G. 1998. Plurality of Mass Nouns and the Notion of "Semantic Parameter". In S. Rothstein (ed.), *Events and Grammar*. Kluwer Academic Publishers, Boston, p. 53-103.

- Chierchia, G. & S. McConnell-Ginet, 1996. *Meaning and Grammar: An Introduction to Semantics*. The MIT Press, Cambridge, Massachusetts.
- Chomsky, N. 1986. *Barriers*. The MIT Press, Cambridge, Massachusetts.
- Chung, S. 1976. On the Subject of Two Passives in Indonesian. In C.N. Li (ed.), *Subject and Topic*. Academic Press, New York, p. 57-99.
- Chung, S. 1978. *Case Marking and Grammatical Relations in Polynesian*. University of Texas Press, Austin.
- Cruse, D. 1986. *Lexical Semantics*. Cambridge University Press, Cambridge.
- Darajowidjodjo, S. 1978. *Sentence Patterns of Indonesian*. PALI Language Texts: Southeast Asia. University of Hawaii Press, Honolulu.
- Déchaine, R & M. Wiltschko, 2000. *Decomposing Pronouns*. Unpublished ms.
- Gao, Q. 1994. Chinese NP structure. *Linguistics* 32: 475-510.
- Gil, D. 1987. Definiteness, Noun Phrase Configurability, and the Count-Mass Distinction. In E. Reuland & A. ter Meulen (eds.), *The Representation of (In)definiteness*. The MIT Press, Cambridge, p. 254-269.
- Gil, D. 1994. The Structure of Riau Indonesian. *Nordic Journal of Linguistics* 17: 179-200.
- Gil, D. 1999. Genitives, Number and Indefiniteness: Some Data from English, Singlish, Mandarin and Malay. In F. Plank (ed.), *Agreement Gender Number Genitive & EURO TYP Working Papers, Series 7, Number 23*, The European Science Foundation, p. 109-131.
- Gil, D. 2000a. *Riau Indonesian: A language without Predication*. Presentation given at the University of Calgary, Feb. 28, 2000.
- Gil, D. 2000b. The Maltese 'Collective': A Typological Perspective. In A. Bog and F. Plank (eds.), *The Maltese Noun Phrase Meets Typology*. EURO TYP Working Papers, Series 7, The European Science Foundation, EURO TYP Programme, Berlin.
- Gillon, B. 1991. Towards a Common Semantics for English Count and Mass Nouns. *Linguistic Inquiry* 23: 597-639.
- Ghomeshi, J., R. Jackendoff, N. Rosen, & K. Russell, 2000. *Contrastive Reduplication in English: Is it FOCUS-focus?* Presentation given at the Canadian Linguistic Association. Edmonton, May 28, 2000.

- Grice, H. 1975. Logic and Conversation. In P. Cole and J. Morgan (eds.), *Speech Acts, Syntax and Semantics*, Volume 3. Academic Press, New York, p. 41-58.
- Guilfoyle, E. H. Hung, & L. Travis, 1992. Spec of IP and Spec of VP: Two Subjects in Austronesian Languages. *Natural Language and Linguistic Theory* 10: 375-414.
- Haji, O. 1974a. Malay Noun Phrases. *Language Sciences* 26:12-17.
- Haji, O. 1974b. Classifiers in Malay and Iban. *Anthropological Linguistics* 26:1-17.
- Hassan, A. 1974. *The Morphology of Malay*. Dewan Bahasa dan Pustaka, Kuala Lumpur.
- Heim, I. 1982. *The Semantics of Definite and Indefinite Expressions*. Ph.D dissertation, University of Massachusetts, Amherst.
- Hopper, P. 1988. How Ergative is Malay? In R. McGinn (ed.), *Studies in Austronesian Linguistics*. Monographs in International Studies, Southeast Asia Series No. 76, Ohio University Press, Ohio, p. 441-454.
- Imai, M., & D. Gentner 1997. A Cross-linguistic Study of Early Word Meaning: Universal Ontology and Linguistic Influence. *Cognition* 62: 169-200.
- Kang, B. 1994. Plurality and Other Semantic Aspects of Common Nouns in Korean. *Journal of East Asian Linguistics* 3(1): 1-24.
- Kaufman, D. & M. Aronoff 1991. Morphological Disintegration and Reconstruction in First Language Attrition. In H. Seliger and R. Vago, eds. *First Language Attrition*. Cambridge University Press, p. 175-188..
- Kayne, R. 1994. *The Antisymmetry of Syntax*. The MIT Press, Cambridge, Massachusetts.
- Kiyomi, S. 1995. A New Approach to Reduplication: A Semantic Study of Noun and Verb Reduplication in the Malayo-Polynesian Languages. *Linguistics* 33(6): 1145-1167.
- Krifka, M. 1989. Nominal Reference, Temporal Constitution and Quantification in Event Semantics. In R. Bartsch, J. van Benthem and P. van Emde Boas (eds.), *Semantics and Contextual Expression*. Foris Publications, Dordrecht, p. 75-115.
- Krifka, M. 1991. Four Thousand Ships Passed Through the Lock: Object-induced Measure Functions on Events. *Linguistic Inquiry* 23: 487-520.
- Krifka, M. 1995. Common Nouns: A Contrastive Analysis of Chinese and English. G. Carlson and F.J. Pelletier (eds.), *The Generic Book*. University of Chicago Press, Chicago, p. 398-411.

- Landman, F. 1989. Groups I. *Linguistics and Philosophy* 12(5): 559-605.
- Landman, F. 1991. *Structures for Semantics*. Kluwer Academic Publishers, Boston.
- Lewis, M. 1969. *Sentence Analysis in Modern Malay: with Examples Drawn from Two Plays by Za'ba*. Cambridge University Press, New York.
- Li, A. 1999. Plurality in a Classifier Language. *Journal of East Asian Linguistics* 8(1): 75-99.
- Link, G. 1983. The Logical Analysis of Plurals and Mass Terms. In R. Bauerle et al. (eds), *Meaning, Use, and Interpretation of Language*, de Gruyter, Berlin.
- Link, G. 1991. Quantity and Number. In D. Zaeffer (ed.), *Semantic Universals and Universal Semantics*. Foris Publications, New York, p. 133-149.
- Lonning, J. 1987. Mass Terms and Quantification. *Linguistics and Philosophy* 10: 1-52.
- Lyons, J. 1977. *Semantics* (2 vols.). Cambridge University Press, Cambridge.
- MacDonald, R. 1976. *Indonesian Reference Grammar*. Georgetown University Press, Washington D.C.
- Matthews, P. 1991. *Morphology*. Cambridge Textsbooks in Linguistics, Cambridge University Press, Cambridge.
- Marantz, A. 1982. Re: Reduplication. *Linguistic Inquiry* 13: 435-482.
- Mithun, M. 1984. The Evolution of Noun Incorporation. *Language* 60(4): 847-894.
- Mithun, M. 1988. Lexical Categories and the Evolution of Number Marking. In Hammond, E. & Noonan, M. (eds), *Theoretical Morphology: Approaches in Modern Linguistics*. Academic Press, New York, p. 211-234.
- Moravcsik, E. 1978. Reduplicative Constructions. In J.H. Greenberg (ed.), *Universals of Human Language, Vol. 3: Word Structure*. Stanford University Press, Stanford, p.297-334.
- Ojeda, A. 1998. The Semantics of Collectives and Distributives in Papago. *Natural Language Semantics* 6(3): 245-270.
- Parsons, T. 1979. An Analysis of Mass Terms and Amount Terms. In F.J. Pelletier (ed.), *Mass Terms: Philosophical Problems*. Synthese Language Library, vol. 6, D. Reidel, Dordrecht. p. 137-166.

- Pollock, J-Y. 1989. Verb Movement, UG and the Structure of IP. *Linguistic Inquiry* 20: 365-424.
- Prince, E. 1981. Toward a Taxonomy of Given-new Information. In P. Cole (ed), *Radical Pragmatics*. Academic Press, New York.
- Quine, W. 1960. *Word and Object*. The MIT Press, Cambridge, Massachusetts.
- Quine, W. 1969. Ontological Relativity. In: *Ontological Relativity and Other Essays*. Columbia University Press, New York, p. 26-68.
- Ritter, E. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. In S. Rothstein (ed.), *Syntax and Semantics 25: Perspectives on Phrase Structure*. Academic Press, New York, p. 37-62.
- Rijkhoff, J. 1990. Explaining Word Order in the Noun Phrase. *Linguistics* 28(1): 5-42.
- Sanches, M. 1973. Numerical Classifiers and Plural Marking: An Implicational Universal. *Working Papers on Language Universals* 11:3-22.
- Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. Hartcourt, Brace and World, Inc., New York.
- Sapir, E. 1930. Southern Paiute, a Shoshonean language. Texts of the Kaibab Paiutes and Uintah Utes. Southern Paiute Dictionary. *Proceedings of the American Academy of Arts and Sciences* 65 (1, 2, and 3).
- Schiffrin, D. 1994. *Approaches to Discourse*. Blackwell Publishers, Cambridge.
- Schwarzschild, R. 1996. Pluralities. *Studies in Linguistics and Philosophy, Vol. 61*. Kluwer Academic Publishers, Boston.
- Soh, H.L. 1994. *Tense and Aspect in Malay*. Unpublished M.A. thesis, University of Calgary.
- Stein, M. 1981. *Quantification in Thai*. Ph.D Dissertation, University of Massachusetts, Amherst.
- Tang, C. 1990. A Note on the DP Analysis of the Chinese Noun Phrase. *Linguistics* 28: 337-354.
- Thompson, L. 1965. *A Vietnamese Grammar*. University of Washington Press, Seattle.
- Traugott, E & B. Heine. 1991. *Approaches to Grammaticalization, Vol. 1*. John Benjamins, Amsterdam.

- Travis, L. 2000. *The Syntax of Iteration and Iconicity*. Presentation given at the Canadian Linguistic Association. Edmonton, May 28, 2000.
- T'sou, B. 1976. The Structure of Nominal Classifier Systems. *Austroasiatic Studies 2*: 1215-1248.
- Uhrbach, A. 1988. Topicality and Verbs Encoding Two Arguments in Bahasa Indonesia. In R. McGinn (ed.), *Studies in Austronesian Linguistics*. Monographs in International Studies, Southeast Asia Series No. 76, Ohio University Press, Ohio, p. 455-474.
- Verhaar, J. 1988. Syntactic Ergativity in Contemporary Indonesian. In R. McGinn (ed.), *Studies in Austronesian Linguistics*. Monographs in International Studies, Southeast Asia Series No. 76, Ohio University Press, Ohio, p. 347-367.