Non-Subject Arguments in Indonesian

Simon Musgrave

Submitted in total fulfilment of the requirements of
the degree of Doctor of Philosophy

October 2001

Department of Linguistics and Applied Linguistics
The University of Melbourne

Abstract

The grammatical function subject can be identified reliably in Bahasa Indonesia (Indonesian), but the same is not true of other clause-level nominal constituents in the language. The tests proposed in previous studies for identifying an object grammatical function turn out to be unreliable and inconsistent when the full range of data is considered. This thesis attempts to clarify the problem by examining non-subject arguments in Indonesian in the theoretical framework of Lexical-Functional Grammar. Both the properties associated with the various types of argument and the means by which they are licensed in clauses turn out to be problematic.

Two argument verbs appear in a range of clause types which are related in interesting ways. I argue that it is possible to give a coherent analysis of the system as a whole, treating it as the basic transitive system of the language, when certain relationships are recognised as morphological rather than syntactic. This analysis also reveals similarities between Indonesian and more conservative Austronesian languages such as those of the Philippines. Another type of clause, sometimes referred to as the 'adversative verb' clause type shares an important property with one of the types of transitive clause, that of allowing an oblique argument to appear without a licensing preposition if it is adjoined to the verb which governs it. Another class of verbs, emotion and cognition verbs, has a non-subject argument which can be either a prepositional phrase or a bare nominal, but in this case the bare nominal is not an oblique argument, adjunction is not the relevant syntactic relation. The only possible analysis of these verbs, maintaining the assumption that the transitive system has been identified, is that they have a subject and a secondary object as their arguments. I argue that this analysis is implausible, and that the assumption regarding the transitive system must be reconsidered.

When this is done, I claim that Indonesian is best analysed as having two distinct types of clausal organisation, one in which the major constitutents are subject and verb phrase, and one in which they are subject and predicate. Most verbs can head the predicate constituent in the second type of clause, with morphology playing a crucial derivational role. Evidence from nominalised clauses and from intransitive verbs supports this analysis, but crucial aspects of the analysis resist statement in the framework of Lexical-Functional Grammar, suggesting that the theory is overly-constrained in important areas.

Declaration

This is to certify that:

i. the thesis comprises only my original work except where indicated in the preface,

ii. due acknowledgment has been made in the text to all other material used,

iii. the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Simon Musgrave Date:

Preface

This thesis reports the results of research carried out between 1997 and 2001. The empirical basis of this research consists of a survey of texts carried out by me (the sources used are listed immediately before the other bibliographic references at the end of the thesis), example sentences elicited from speakers of the languages studied, and published material as acknowledged in the body of the thesis. Data on the Sasak language was collected as part of the collaborative Lombok and Sumbawa Research Project (Australian Research Council Large Grant A59803558, Principal researcher: Prof. Peter K. Austin); the other work reported here was my own independent research.

Acknowledgments

My greatest debt is to the speakers of Indonesian, and of the other languages discussed in this thesis, who generously shared their time and knowledge with me. In particular, Umar Muslim was an endlessly patient teacher of both Indonesian and Javanese, and was always willing to share his own insights as a linguist. Katarina Sukamto, Eli Riharti, and Lalu Dasmara all also were very helpful and Waruno Mahdi always answered my e-mails. Lalu Dasmara also helped me with Sasak data, as did Herman Suheri, Syahdan and Yon Mahyuni. I am grateful to Bukhari Daud and Samadi for assisting me with Acehnese data. I am also grateful to Helen McKay for allowing me access to her corpus of Indonesian text (and for employing me).

My supervisor throughout this research has been Peter Austin, and he has made the experience enjoyable by his adroit balance of formal and informal guidance, his enthusiasm for linguistics, and his deep knowledge which he is always willing to share. During my probationary candidature, Lesley Stirling and Dominique Estival were members of my supervisory committee, and I am grateful to them both for helpful comments. The Department of Linguistics and Applied Linguistics at the University of Melbourne is a stimulating and friendly place to work. I am especially grateful to Nick Evans, who always has time (how?) to discuss problems, and always has a suggestion as to how to tackle them; to Rachel Nordlinger for her LFG library; to Anna Margetts for incisive comments; and to Rob Schmittat for technical help (and coffee). I was fortunate to do this research while the department was something of a centre of Austronesian research, and I thank my fellow students Katarina Sukamto, Anthony Jukes, Nick Thieberger and Yusuf Eades for their friendship and support.

Beyond the department, the wider circle of Austronesian scholars were wonderfully welcoming to a novice. I am grateful to all of the following people for many different kinds of assistance: Sander Adelaar, Wayan Arka, Barry Blake, John Bowden, Peter Cole, Susanna Cumming, Novi Djenar, Mark Durie, Michael Ewing, Dan Finer, Margaret Florey, Bill Foley, David Gil, Gabriella Hermon, Nikolaus Himmelmann, Arthur Holmer, Marian Klamer, Waruno Mahdi, Bernd Nothofer, Bill Palmer, Andy Pawley, Stan Starosta, (and a special mention for Mary Ellen Jordan, Austronesianist rtd.),. The LFG community were also very generous; thank you to Avery Andrews, Kersti Borjars, Joan Bresnan, Chris Manning, Jane Simpson and Nigel Vincent.

I am enormously grateful to the government of Australia, through the Department of Education, Training and Youth Affairs, for financial support in the form of an Australian Post-graduate Award which allowed me to work full-time on this thesis. I am also grateful to the Faculty of Arts, The University of Melbourne for a travel grant which enabled me to attend the 8th International Conference on Austronesian Linguistics (Taipei 1997); to the Melbourne Scholarships Office for a Melbourne Abroad Scholarship and to the Department of Linguistics and Applied Linguistics for a travel grant which together enabled me to attend the 7th Austronesian Formal Linguistics Association meeting (Amsterdam 2000) and to visit the University of Manchester. I am also grateful for support from the Lombok and Sumbawa

Research Project (Australian Research Council Large Grant A59803558) which assisted me in attending the 2nd International LFG conference and the 4th Australian Linguistic Institute (Brisbane 1998).

Portions of this work have previously been presented to the following meetings: Austronesian Formal Linguistics Association (Amsterdam 2000); Department of Linguistics Seminar, University of Manchester; the workshop on Voice in Austronesian, 2nd International LFG Conference (Brisbane 1998); the 2nd and 3rd Victorian South East Asian Languages Symposia (1998, 1999); the University of Melbourne Linguistics and Applied Linguistics Post-Graduate Conference (1999, 2000); the Victorian Austronesian Circle; and the Department of Linguistics and Applied Linguistics Austronesian Informal Seminar.

Finally, I could never have completed this thesis without the love and support of my wife, Deborah - thank you for that, and for our two beautiful daughters, Anna and Olivia.

Table of Contents

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
ACC	accusative
ADJ.....ADJ	adjective-forming circumfix
AF	actor focus
APPL	applicative
AS	actor subject
CAUS	causative
CLASS	classifier
COLL	collective
COMP	comparative
COP	copula
DAT	dative
DUP	reduplication
DV	dative voice
EMPH	emphatic particle
EXCL	exclusive
FOC	focus marker
FRAC	fraction
FUT	future
GEN	genitive
HON	honorific
IF	instrumental focus
INCL	inclusive
INT	intensifier
IRR	irrealis
LF	locative focus
LNK	linker
LOC	locative preposition
NEG	negation
NOM	nominative
NOM.....NOM	nominalising circumfix
OBL	oblique
ORD	ordinal
PERF	perfective
PL	plural
POSS	possessive
PROG	progressive
PRT	particle
REAL	realis
REL	relative clause marker
SG	singular
UF	undergoer focus
US	undergoer subject

Glossing Practice

All Indonesian examples are glossed consistently, regardless of the source and spelling is regularised to the post-1972 standard (see section 1.5) except for proper names. With the exception of the allomorphy of prefixes with a final nasal segment, which is described and exemplified in section 1.3, Indonesian morphology is quite transparent. Therefore the morpheme breaks are not indicated in examples. Where the morphological structure of a word is not relevant to the discussion, the morphemes are not separately glossed. Some morphemes are indicated as separate in the gloss line, but are not given a glossing. This occurs in two situations: i) the morpheme in question is the subject of analysis in the surrounding discussion; when the analysis has been established, a suitable gloss is provided in subsequent chapters; ii) the meaning of the morpheme is irrelevant to the discussion. Words which involve reduplication are consistently shown as being of the form base-reduplicant. This is a convenience, and should not be taken to imply any theoretical claim.

Examples from languages other than Indonesian are generally glossed as in the original source. Any additional abbreviations used are then explained in a footnote. The exception to this principle is examples from Tagalog: analyses, and therefore glossing, of Philippine languages vary widely and I have regularised glosses for this language.

1 Introduction

This thesis is an investigation into the properties of non-subject arguments in Bahasa Indonesia (Indonesian) in the framework of Lexical-Functional Grammar. How to identify subjects and what their properties are is generally agreed in studies of Indonesian (one exception to this generalisation is discussed later), but less work has been done on other clausal constituents. This comparative neglect mirrors the situation in syntactic studies more generally: a great deal of work has been devoted to establishing empirically valid and theoretically useful definitions of what it means to call a nominal a subject[1] (see the papers in Li (ed) 1976, especially that by Keenan), but the same attention has not been devoted to other nominal constituents of clauses (but see Plank (ed) 1984 and Dryer 1986). The problems in Indonesian will turn out to be rather more complex than those encountered in many other languages. It will be argued in this work that, in Indonesian, even the non-subject nominal of the most common type of clause with two nominal constituents cannot be described as a (direct) object without raising serious questions, both empirical and theoretical, and that accounting for how different nominal constituents are licensed to appear in clauses poses major challenges. My conclusion will be that, despite the success of LFG as a tool for analysing a wide range of typologically diverse language, the theory as currently formulated is unable to give an adequate account of the facts which will be described.

The remainder of this chapter is organised as follows. Section 1.1 sets out some basic assumptions about nominal constituents of clauses, their properties and the classes into which they can be analysed. Section 1.2 introduces some basic information about Indonesian, and discusses tests which have previously been proposed as diagnostics for subjecthood and objecthood. Section 1.3 introduces enough of the theoretical apparatus of Lexical-Functional Grammar (LFG) for the reader to negotiate the analyses presented later. Section 1.4 presents the account which LFG gives of the licensing of nominal constituents in configurational languages, and discusses some Indonesian data which are contrary to the predictions made by that theory. Section 1.5 discusses the issue of variation in Indonesian and the implications it has for data collection. Finally, section 1.6 is an overview of the remaining chapters.

1.1 Basic assumptions

Two sorts of properties of nominal constituents are of interest in syntactic analysis: *coding properties* and *behavioural properties* (Keenan 1976). Coding properties are manifest formal properties such as case-marking and linear position. Thus, in Latin, the two nominal constituents associated with a transitive verb have different case forms. In most instances, one is in

[1] Throughout, I use the terms *subject* and *object* in an intuitive sense. Where specific grammatical functions are referred to, I use the LFG grammatical function designators SUBJ, OBJ etc.

the case form traditionally termed nominative and is the subject, the other is in the case form termed accusative and is the object:

1. *Puella* *videt* *mensam*
girl.NOM see.3SG table.ACC
'The girl sees the table.'

The case forms fully identify the function of the nominals (in most instances), and word order is relatively free. An alternative strategy is for a nominal to have the same form in all functions, but to be functionally identified by its position. English uses this strategy: subjects precede verbs in unmarked order, and objects immediately follow them. It is possible for a language to use both strategies at once, to greater or lesser extent. Thus, besides linear order, English also has case forms for some pronouns. Coding properties may also be formally marked on another constituent. For example, verbs in Latin (and in English in the case of 3rd person singular present tense) vary in their form dependent on properties of one of the nominal arguments of the clause, the subject. Verbs in other languages, for example Yimas (Foley 1991), show such agreement with more than one nominal.

Behavioural properties are not overt properties of nominals, but are revealed when clauses are manipulated in various ways. For example, in English, the subject nominal of the second of two conjoined clauses can be omitted if it is co-referent with the subject of the first clause:

2. Ellen took the car and drove home.

Non-subject nominal constituents in English do not have this property. Another property of this type is quantifier float. In some languages, quantifiers which are normally part of a nominal constituent can appear elsewhere in the clause, but are still construed as part of the nominal. Where this is possible, it is typically allowed for some nominal constituents but not others, and the possible readings are therefore restricted:

3. Tagalog

Sinusulat	*lahat*	*ng*	*mga*	*bata*	*ang*	*mga*	*liham*
UF.write	all	CORE	PL	child	NOM	PL	letter

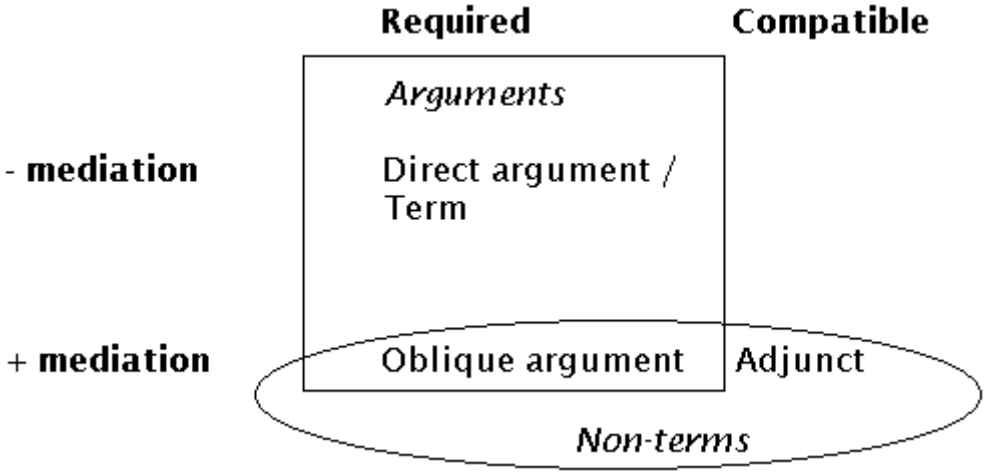
'The/some children write all the letters.'
(NOT: 'All the children write the letters.') (Schachter 1976: ex16b)

Other behavioural properties include whether or not a nominal can be head of a relative clause, whether it can be a controller of the reference of a missing argument in a dependent clause, and whether it can be omitted in a dependent clause (i.e. whether it can be controlled).

There is a consensus across theories that nominal constituents of clauses can be divided into two groups on the basis of their coding and behavioural properties. The most visible of these are coding properties: the first type of nominal is a bare NP in a language without case-marking, or it carries one of the *grammatical cases* in a case-marking language (Blake 1994: 32-34). The second type typically require adpositions in a language without case-marking, they are marked with one of the *semantic cases* in a case-marking language. This division correlates largely, but not perfectly, with

another division: that between nominals which are necessary for the interpretation of the verb and are assigned semantic roles by it, and other nominals which are compatible with the semantic interpretation of the verb (or clause), but which are not required. The correlation breaks down in cases where a nominal is assigned a semantic role by the verb, but nevertheless is coded using semantic case or an adposition. In what follows, I will use the following terminology: *arguments* will refer to the nominals assigned semantic roles by the predicate of a clause; *direct argument* or *term* will refer to an argument that does not require a mediating mechanism such as an adposition; *oblique argument* will refer to an argument for which such a mediating mechanism must be employed. The word *adjunct* will refer to non-arguments, that is elements which are not assigned a semantic role by the predicate, and *non-terms* will on occasion cover both obliques and adjuncts. This terminology is illustrated in the following diagram:

4.



For most theories, this division follows from certain assumptions about how nominal constituents are licensed to appear in clauses. The common view is that predicates have an argument list as a lexical property. This is modelled as argument structure in LFG (Bresnan 2001a: chapter 14), and as the theta grid in Principles and Parameters theory (Haegeman 1991: 41-46). This argument list gives the *valence* of the verb, the number of arguments it syntactically associates with, and interacts with the rules which govern the projection of phrase structure in the language. One result of this interaction, at least in configurational languages, is that there are some arguments whose presence is licensed just by the phrase structure, and there are other arguments whose licensing requires the mediation of some other mechanism.

For example, the English verb *put* has an argument list consisting of two terms and an oblique. This information interacts with the phrase structure rules of the language so that one complement position is projected as sister to the verb, and this can be occupied by an NP without further ado. The NP is licensed to appear there by the lexical properties of the verb and the phrase structure. But the lexical properties of the verb, the fact that it is transitive not ditransitive, only allow one such position to be generated. Phrase structure rules allow the generation of a specifier position in the maximal projection,

and any number of adjunct positions, but these positions do not license arguments in the same way that the complement position does[2]. Some additional mechanism is necessary, and in English the strategy used is to code the argument as a prepositional phrase. These principles account for the difference in grammaticality between the following two sentences:

5. **John put a chair the garden.*
6. *John put a chair in the garden.*

The preposition actually used to mediate the licensing of the argument NP contributes to the interpretation of the clause. There are cases in English where the minimally contentful preposition *of* can be used as a pure case marker, but this is not possible with *put*. This insight also lies behind the distinction mentioned above between grammatical and semantic cases (Blake 1994: 32-34). LFG uses a version of the theory sketched above to account for the licensing of nominals; I will lay this out in detail in section 1.4 after the basics of Indonesian syntax have been described and the apparatus of the LFG theory has been introduced.

1.2 Indonesian syntax

1.2.1 General issues

Bahasa Indonesia is the national language of the Republic of Indonesia and is spoken by about 170 million people, although it is not the first language of the majority of these speakers (Grimes ed. 1996). Further historical details are given in section 1.5, which also discusses the issue of variety in Indonesian and the implications this has for data collection. The phonology of the language is not of interest in this study; the interested reader is referred to Moeliono & Dardjowidjojo (ed. 1988) for information. Orthography is straightforward except for *e* which represents both schwa and a mid-high front vowel. In my examples, following standard practice, no attempt is made to distinguish these. The entries in the dictionary of Echols & Shadily (1994, afterwards E&S) provide this information. In addition, *j* represents a palatal glide, and *c* represents a voiceless affricate. *ng* represents a velar nasal, and if a voiced velar stop follows, *ngg* is written. Following Sneddon (1996), I assume that Indonesian has a standard set of lexical categories: noun, verb, adjective, preposition and adverb. However, the question of whether adjectives are a category distinct from verbs is not clearcut, and is discussed in section 4.1.3, and chapter 5 will argue that the maximal projection of verbs should not always be analysed as VP. The question of whether the functional categories I (or whatever complex of categories is assumed after Pollock (1989)) and D exist in the language is discussed in section 1.3.3.

[2] I ignore here the possibility that the subject is generated in Spec of VP (see references cited by Haegeman (1991: 327, n22). In any case, such proposals agree that the argument is not licensed in that position.

The morphological system of Indonesian is unusual in that it is exclusively derivational[3]. There is no marking for case or gender on nouns, verbs are not marked for tense, aspect or mood, nor is there agreement or cross-referencing. Many verbs can take the prefixes *meN-* and *di-*, traditionally analysed as marking active voice and passive voice respectively. The discussion of chapter 2 will show that this analysis is not correct, and chapter 5 will suggests that these prefixes are in fact derivational also. The prefix *meN-* (and the nominalising prefix *peN-*) end with a nasal segment which assimilates to a following consonant. When the consonant is unvoiced, the nasal fuses with it, except in the case of the affricate *c*. Before a vowel, the form of the prefix is *meng-*, and before a liquid the nasal segment deletes. The following examples illustrate this allomorphy:

7.

Base	**Prefixed form**
anggap	*menganggap*
baca	*membaca*
dapat	*mendapat*
ganggu	*mengganggu*
peras	*memeras*
tulis	*menulis*
kasih	*mengasi*
seluruh	*menyeluruh*
curiga	*mencuriga*
lihat	*melihat*

Aspects of the verbal morphology of Indonesian have been studied in detail, in particular the suffixes *-i* and *-kan* (Arka 1993, Kana 1986, Tampubolon 1983, Voskuil 1996). These two suffixes can derive verbs from noun or adjective roots (examples 8 and 9), and in some cases can derive verbs from other verbs without changing the valence, but with a different meaning (examples 10 and 11). Both suffixes can increase the valence of a verb as applicatives, that is they add a non-subject to the argument array (examples 12 and 13). The result can be a verb with three arguments, none of them introduced by prepositions. *-kan* is also a causative, in which case it adds an agent subject to the argument array[4]. Where the base is not a verb, or is a one-argument verb, the result is a two argument verb (example 9). Where the base is a two argument verb, the derived causative verb has two direct arguments with the original actor appearing as an oblique argument (example 14):

[3] The sole exception to this generalisation is the use of reduplication. This morphological process can indicate plurality when applied to nouns and continued or iterated action when applied to verbs. However, in each case other meanings are possible also (Sneddon 1996: 15-21)

[4] There are occasional instances where *-i* has causative effect. See section 4.1.2.4 for an example.

8. *Pintu dan jendela diwarnainya hijau*
door and window *di*.colour(n.).APPL.3 green
'He painted the doors and windows green.'

9. *Ia menerangkan keadaannya*
3SG *meN*.clear(adj.).CAUS situation.3
'He explained his situation.'

10. *Ia mendapat dia di rumahnya*
3SG *meN*.meet 3SG LOC house.3
'He met her at her house.'

11. *Polisi mendapati bengkel pemalsu uang*
police *meN*.meet.APPL workshop forger money
'Police found the counterfeiter's workshop.'

12. *Dia mengirimi suaminya uang*
3SG *meN*.send.APPL husband.3 money
'She sent her husband money.'

13. *Dia menuliskan ayahnya surat*
3SG *meN*.write.APPL father.3 letter
'She wrote a letter for her father.'

14. *Saya memeriksakan mata ke doktor*
1SG *meN*.examine.CAUS eye to doctor
'I had my eyes examined by the doctor.'

Indonesian morphology as a whole has not yet been analysed satisfactorily. In particular, there are restrictions on the co-occurrence of affixes which are not well understood. One example of this is discussed in section 3.3.1.

The lack of inflectional morphology means that the major coding property available to indicate the function of terms in Indonesian is linear order. The unmarked word order in a clause is for the subject to precede the predicate, and for non-subject arguments to follow the head of the predicate. Given these facts and the lack of other coding properties, I will assume that Indonesian is a configurational language without arguing the position in detail. That is, I assume that there is a subject position which is outside the phrase-structure constituent which includes the positions occupied by non-subject arguments.

Pronouns provide the only exceptions to the generalisation given above regarding coding properties. The third person pronoun has two forms, remnants of a more extensive case-marking system, and one of these has limited distribution which is discussed in section 1.2.2 below. In one clause type, proclitic forms of first and second person pronouns occur, but I argue in chapter 2 that this is a phonological phenomenon. Finally, enclitic pronouns are possible in some positions. The distribution of these forms is important in the argumentation of each chapter in this work, and in chapter 5, I suggest

that their distribution is one of the crucial diagnostics revealing that the Indonesian has two types of clausal structure existing side-by-side.

Coding properties of arguments can offer only limited evidence in the investigation of the status of nominal constituents in Indonesian. Most previous studies have therefore relied more on evidence from behavioural properties. The following sections review some of this work.

1.2.2 Tests of subjecthood

Traditional grammars (McDonald & Darjowidjojo 1967, Sneddon 1996 [afterwards IRG]) and generative studies (Chung 1976a,b, Kana 1986, Arka & Manning, to appear) have agreed that identifying the subject of an Indonesian clause is a straightforward matter, with one clause type being an exception. All agree that the basic word order is subject-predicate, with all major categories possible as heads of the predicate. V is initial in its projection, therefore the word order of verbal clauses is apparently S - V - (O), and in most cases the first noun phrase in a clause can be taken to be subject. As seen in examples 12 and 13 above, two non-oblique arguments can follow the verb. In addition to the derived three argument verbs in the examples above, there are a few basic verbs which take three direct arguments such as *beri* 'give' and *ajar* 'teach'.

There are additional tests available which confirm the word order generalisation for subjects, and enable doubtful cases to be decided. Firstly, the third person singular pronoun has two forms, *dia* and *ia*. The first of these can occur in any environment, but the second is restricted to three possibilities. It can appear whenever it is subject of the clause; it can appear as the agent in a marked clause type discussed in more detail below; and it can appear in the position of object when it is also the subject of a complement clause:

15. *Dia/ia akan menolong kami*
    3SG FUT *meN*.help 1PL.EXCL
    'She will help us.'

16. *Kami akan dia/ia tolong*
    1PL.EXCL FUT 3SG help
    'Us she will help.'

17. *Kami akan menolong dia/*ia*
    1PL.EXCL FUT *meN*.help 3SG
    'We will help her.'

18. *Saya menganggap dia/ia bodoh*
    1SG *meN*.consider 3SG stupid
    'I consider him stupid.'

It is a characteristic of the clause type exemplified in 16 that the agent pronoun must be adjacent to the verb and the positioning of auxiliaries establishes that a clause is of this type. Such structures are thus easily

identified, and in all other cases, the possibility of *ia* occurring indicates that it is in a subject position.

Example 18 introduces the next test which can be used: in structures such as example 18, with certain matrix verbs, the subject of the embedded clause can be treated as object of the matrix clause and can become subject of a passive:

19. *Ali menganggap soal itu beres*
Ali *meN*.consider problem that settled
'Ali believes the problem is settled.'

20. *Soal itu dianggap Ali beres*
problem that *di*.consider Ali settled
'The problem is believed by Ali to be settled.'

This is not possible for any other nominal constituent of the embedded clause:

21. *Mereka menugasi saya melakukan observasi*
3PL *meN*.assign 1SG *meN*.carry.out observation
'They assigned me to carry out observations.'

22. **Observasi ditugasi mereka saya melakukan*
observation *di*.assign 3PL 1SG *meN*.carry.out
(FOR: Observations were assigned by them for me to carry out.')

Therefore, in clauses of this type, the coding properties of the post-verbal pronoun are those of a subject (*ia* is possible), but its behavioural properties are those of a non-subject argument.

The other test which identifies subjects reliably is that it is only the subject position which can be omitted from an embedded clause and have its reference controlled by a constituent in a higher clause. This test is often exemplified with volitive verbs (e.g. Kana 1986: 27), but some of this class of verbs can also be auxiliaries in Indonesian (see discussion in section 4.1.2.7), and to avoid this complication I use examples of control into purpose clauses. Such clauses are introduced by the preposition *untuk* 'for', the subject must be co-referential with some nominal in the matrix clause[5] and it is obligatorily absent:

23. *Dia minta uang untuk (*dia) pergi ke Indonesia*
3SG ask money for go to Indonesia
'He asked for money to go to Indonesia.' (Kana 1986: 28, ex.30)

Even if there is a coreference relation between a nominal of the matrix clause and a non-subject in the embedded clause, the nominal must be present in the embedded clause:

[5] There is some discussion of possible controllers in the following section.

24. *Saya mencari Ali untuk saya membunuh
1SG meN.look.for Ali for 1SG meN.kill
(FOR: 'I looked for Ali in order to kill (him).') (Kana 1986: 28 ex.31)

25. *Saya mencari Ali untuk membunuh
1SG meN.look.for Ali for meN.kill
(FOR: 'I looked for Ali in order to kill (him).') (Kana 1986: 28 ex.32)

26. Saya mencari Ali untuk membunuhnya
1SG meN.look.for Ali for meN.kill.3
'I looked for Ali in order to kill him.'

These tests make it possible to reliably identify subjects in Indonesian. Kana (1986) also claims that relativization is a test for subjecthood. This is discussed in the following section, as Chung (1976a) claims that it is a test for both subjecthood and objecthood.

It was mentioned previously that there was one case in which general agreement as to the identification of subject was not complete. This is the case of the clause type seen in 16 above, which has the patient as left-most nominal, and the agent immediately to the left of the verb. The evidence of coding properties is thus ambiguous, and as example 16 shows the pronoun *ia* can appear as agent in such a clause. In a detailed examination of such clauses, Chung (1976b) demonstrates beyond doubt that it is the patient nominal which has the behavioural properties associated with subject in this clause type. The *ia* test is not relevant, but the other two tests discussed above provide clear evidence for the analysis:

27. Mereka anggap buku ini sudah saya baca
3PL consider book this PERF 1SG read
'They believe this book, I have read.' (Chung 1976b: ex 19)

28. Buku ini dianggap mereka sudah saya baca
book this di.consider 3PL PERF 1SG read
This book is believed by them to have been read by me.'
(Chung 1976b: ex 20)

29. Mereka membeli ikan itu supaya ikan itu dapat
3PL meN.buy fish that so.that fish that can
saya masak
1SG cook
'They bought the fish so that the fish, I could cook.'
(Chung 1976b: ex 28a)

30. Mereka membeli ikan itu untuk dapat saya masak
3PL meN.buy fish that for can 1SG cook
'They bought the fish to be cooked by me.' (Chung 1976b: ex29a)

Rather surprisingly in view of such evidence, later work occasionally takes the question to be still open (e.g. Cartier 1979). Alsagoff (1991, 1992) presents evidence from one dialect of Bahasa Melayu, the national language of Malaysia

which is almost identical to Bahasa Indonesia, that the control test (example 30) is not reliable in this case. This data is discussed in section 2.4.1.

1.2.3 Tests of objecthood

There is far less consensus in the literature as to the status of various non-subject arguments in Indonesian. Chung (1976a) and Kana (1986) both give accounts of what they take to be properties of (direct) objects in Indonesian, but the two accounts differ considerably. Neither is accurate in my opinion, and neither can account for the full range of data examined in this work. Both agree that the object normally immediately follows the verb, although both allow that other possibilities exist: Kana (1986: 28) allows an adverb but not a major clausal constituent to intervene, while Chung (1976a) allows that the postverbal position 'may be changed by scrambling' (42). When a wider range of clause types is considered, positional evidence is of little value. There are several clause types in which a nominal follows the verb, but in each case expression as a prepositional phrase (PP) is also possible. In two of these cases, the nominal is obligatorily adjacent to the verb unless the PP option is chosen:

31.	*Buku*	*itu*	*dibaca*	*(*tampaknya)*	*Ali*
	book	that	*di*.read	apparently	Ali

32.	*Buku*	*itu*	*dibaca*	*tampaknya*	*oleh*	*Ali*
	book	that	*di*.read	apparently	by	Ali
	'The book was apparently read by Ali.'					

33.	*Kantor*	*itu*	*kelihatan*	*(*dengan*	*mudah)*	*polisi*
	office	that	*ke*.see.*an*	with	easy	police

34.	*Kantor*	*itu*	*kelihatan*	*dengan*	*mudah*	*oleh*	*polisi*
	office	that	*ke*.see.*an*	with	easy	by	police
	'That office is easily visible to the police.'						

In the third case, the adjacency requirement is similar to that proposed for object by Kana, but expression as a PP is also possible:

35.	*Saya*	*takut*	*sekali*	*lelaki*	*itu*
	1SG	afraid	very	man	that

36.	*Saya*	*takut*	*sekali*	*dengan*	*lelaki*	*itu*
	1SG	afraid	very	with	man	that
	'I am very much afraid of that man.'					

It will be shown below, using other tests, that the postverbal nominals in examples 31 and 33 are obliques which are realised in an unusual way, and that the postverbal nominal of example 35 is an object of some sort, but not a direct object in the sense intended by either Chung or Kana.

Both these scholars also take the possibility of being made subject in a passive clause to be a test for objecthood[6]. This test correctly rules out the postverbal arguments of examples 31 and 33: 31 is already a passive, and 33 is intransitive. But example 35 is still a problem. There is no passive form of this clause, unless the verb is affixed with the applicative suffix *-i*:

37. **Lelaki itu ditakut saya*

38. *Lelaki itu ditakuti saya*
man that *di*.afraid.APPL 1SG
'That man is feared by me.'

Therefore this test also is not reliable. It has some value, in that in double object structures only the first object can be made subject:

39. *Ali megirimi wanita itu sebuah surat*
Ali *meN*.send.APPL woman that CLASS letter
'Ali sent the woman a letter.'

40. *Wanita itu dikirimi sebuah surat oleh Ali*
woman that *di*.send.APPL CLASS letter by Ali
'The woman was sent a letter by Ali.'

41. **Surat itu dikirimi wanita itu oleh Ali*
letter that *di*.send.APPL woman that by Ali
(FOR: 'The letter was sent the woman by Ali.')

Even this evidence is essentially redundant, as linear order is rigid in such constructions.

Chung claims that three other tests are useful in identifying objects: relative clause formation, reflexive possibilities, and controlling into purpose clauses. Kana argues that:

1) Chung misinterprets the data on relative clause (1986: 25), a conclusion with which I agree, but I will argue that Kana misinterprets certain relative clause data also;

2) Chung's data on reflexivization are wrong but that the test is still useful (1986: 29-31);

and

3) does not discuss the third property.

I will discuss these in reverse order, as the issues raised increase in complexity going in that direction.

The subject property of being potentially a controllee was discussed above with examples from purpose adjuncts. Chung (1976b) claims that deletion of the subject of a purpose adjunct clause is obligatory 'when the

[6] Chung (1976a) takes the possibility of forming a clause of the type of example 18 above as a separate test for objecthood. It is a very reliable generalisation that if the passive with *di-* is possible, then so is this construction: see discussion in chapter 2.

subject of such a complement is coreferential with some NP in the higher clause' (66), but Chung (1976a) claims that deletion is only possible when 'the subject of such a complement is coreferential with the subject or direct object of the higher clause' (48). This latter claim is directly contradicted by evidence from underived ditransitive verbs[7]:

42. *Saya memberi Ali buku itu untuk dibacanya*
1SG *meN*.give Ali book that for *di*,read.3
'I gave Ali the book to read.'
(LIT: 'I gave Ali the book for to be read by him.')

43. **Saya memberi Ali buku itu untuk membacanya*
1SG *meN*.give Ali book that for *meN*,read.3
(FOR: 'I gave Ali the book for him to read it.')

The verb *beri* is an underived ditransitive, and the recipient is the direct object by linear position and would be subject of a *di-* passive clause. But it is the second object which can control the reference of the omitted subject in a purpose clause, and not the direct object. Therefore, although there is clearly an interesting problem to be investigated here, this test cannot be used as a reliable diagnostic for objecthood.

The basis of all reflexive pronouns in Indonesian is *diri* 'self'. This morpheme can occur alone, or with the addition of a possessive pronoun: *diri saya / diriku* 'myself', or with the possessive and the emphatic reflexive *sendiri*: *diriku sendiri* 'myself'. Chung (1976a) and Kana (1986: 29-31) both claim that the possibilities for reflexives in object position are restricted in some way, but they disagree as to the facts. Chung suggests that *diri* plus possessive pronoun can occur in object position, but not elsewhere, while Kana suggests that *diri* with optional possessive pronoun is obligatory in object position when the object is coreferential with the subject, but is optional elsewhere. The crucial example is the following:

44. *Kakak saya membeli sepasang sepatu merah*
older.sibling 1SG *meN*.buy one.pair shoe red
untuk dirinya
for self.3
'My older sister bought a pair of red shoes for herself.'
(Chung 1976a: ex11b, Kana 1986: 31, ex45)

For Chung, this sentence is ungrammatical as it stands and acceptable to some speakers with the addition of *sendiri*, while for Kana it is fully acceptable as given. Neither author remarks on the possibility for reflexives such as *dirinya* to appear in subject position, discussed by Arka & Manning (to appear) and in section 2.3.1.3:

[7] Chung (1976a) also claims that (in Relational Grammar terminology) only initial 1s and 2s can be controllers in this construction, therefore derived ditransitive verbs are not relevant to the discussion.

45. *Dibandingkannya dirinya dengan raja*
di.compare.3 self.3 with king
'He compared himself with the king.'
(LIT: 'Himself be compared by him with the king.' (IRG: 256)

This is also possible for the subject of an embedded clause following a complementizer:

46. *Mereka percaya bahwa dirinya lebih unggul dari*
3PL believe that self.3 more superior from
orang lain
person other
'They believe that they are superior to other people.'
(IRG:206)

The distribution of the various reflexives is thus not a reliable guide to the status of arguments. In addition, my own experience in elicitation is that even the contrasts discussed here are hard to reproduce. In particular, speakers exploit the pragmatic effect of emphasis with *sendiri* frequently and find it hard to separate pragmatic acceptability and grammaticality. Therefore, I conclude that this test also is of limited use in identifying objects, and I will not use it in this study.

The greatest difference between the approaches of Chung and Kana is seen in the question of relativization. The major strategy for forming relative clauses in Indonesian is to omit the head nominal from the embedded clause which is introduced by the invariant particle *yang*:

47. *wanita yang pulang itu*
woman REL go.home that
'the woman who went home'

The question at issue is what grammatical function can be borne by the gapped nominal in the embedded clause. Chung (1976a) claims that both subjects and direct objects can be gapped, while Kana (1986: 24) follows a traditional line in claiming that only subjects may be relativized in this way[8]. Kana argues in detail that Chung has misinterpreted her data, and is correct in this, but her position is also mistaken. Chung bases her claim that objects can be relativized on examples such as the following:

48. *Ikan yang saya masak untuk Ali tidak enak*
fish REL 1SG cook for Ali NEG good
rasanya
taste.3
'The fish that I cooked for Ali didn't taste good.'
(Chung 1976a: ex35a)

As Kana rightly points out (1986: 24-25) an example like this is ambiguous between being an example of the construction seen in example 16 and a

[8] There is also a resumptive pronoun strategy - see chapter 5 for detailed discussion.

regular transitive clause. In the first case, the patient is the subject as discussed in section 1.2.2, and there is no question of an object being relativized. Kana concludes from this that only subjects can be relativized with the gapping strategy, but she does not address a further point raised by Chung. Chung noted that the clause type seen in 16 is restricted to pronoun agents, but that examples parallel to 48 occur with other NPs as agent:

49.	*Kau*	*melihat*	*ikan*	*yang*	*anak*	*itu*	*masak?*
	2SG	*meN*.see	fish	REL	child	that	cook

'Have you seen the fish that the child cooked?'
(Chung 1976a: ex38a)

Both the cited scholars seem to be aware of the data which uncover the true generalisation, but neither makes the necessary step. Thus, Chung comments that:

> *meng-* can appear on the verb of the relative clause if the relativized noun is a subject, but it does not appear if the relativized noun is a direct object (51)

and Kana notes that *masak* is a verb that normally appears without a prefix[9] (1986: 25). The quoted statement from Chung should be extended to include the case of *di-* prefixed verbs: although non-subjects can never be relativized in this case, the prefix must appear when the subject is relativized. The correct statement of the constraint on gapped relative clauses therefore is that a non-subject argument can be gapped in the absence of a verb prefix. This is true for ordinary transitive verbs, all of which can be used without a prefix in a clause type considered by prescriptive grammarians to be substandard, although it is widely used. It is also true of the type of clause exemplified above in 35, as pointed out by Stevens (1970):

50.	*orang*	*yang*	*anak*	*saya*	*lihat*	*itu*
	person	REL	child	1SG	see	that

'the person that my child saw'

51.	*lelaki*	*yang*	*saya*	*takut*	*itu*
	man	REL	1SG	afraid	that

'the man that I fear'

There is no difference in propositional meaning between the two clauses that follow:

52.	*Anak*	*saya*	*lihat*	*orang*	*itu*
	child	1SG	see	person	that

53.	*Anak*	*saya*	*melihat*	*orang*	*itu*
	child	1SG	*meN*.see	person	that

'My child saw that person.'

[9] The normal account is that this verb and a handful of others such as *minta* 'ask', *minum* 'drink' and *mohon* 'ask' have a fossilised prefix included in the root - see section 2.1.

but the non-subject argument can head a gapped relative clause based on example 52 but not one based on example 53. Two interpretations of this data suggest themselves: either the non-subject argument in both of the above clauses is an object, and relativization is not a relevant test for objecthood; or it is a relevant test and the non-subject argument of one clause should be analysed as an object but not the other. Chapter 5 examines this question in detail.

Kana discusses one additional test which she claims is relevant to the identification of objects, one which Chung does not mention, and which is also very germane to the issues raised in the previous paragraph. Reduced forms of the singular pronouns exist[10] which never occur independently, but which attach to a preceding preposition when the pronoun is its complement, to a preceding noun (or some nominal projections) when the pronoun is a possessor, and to some verbs when the pronoun is the non-subject argument. This last possibility is the one of interest here: Kana claims that 'pronominal direct objects may have the full form or the enclitic form' (1986: 29), and that this property is therefore a diagnostic for objecthood. But the facts here closely parallel those relating to relativization, again it is the presence or absence of a verb prefix which is crucial:

54. **Anak saya lihatnya*
child 1SG see.3
55. *Anak saya melihatnya*
child 1SG *meN*.see.3
'My child saw him/her'

Clauses of the type seen in example 35 do not allow enclitics either:

56. **Saya takutnya*
1SG afraid.3
(FOR: 'I am afraid of him.')

Thus this data raises exactly the same question as the relativization data and will be discussed in detail in chapter 5.

This survey shows that none of the tests proposed by Chung (1976a) or Kana (1986), including those proposed by both, is an unequivocal test for objecthood in Indonesian. Rather, close examination of the data shows that some of the tests raise questions about what can legitimately be described as an object in the language.

In the present study, I will rarely discuss the properties of subjects, but where this is necessary I will use the tests discussed in section 1.2.2 which I accept as reliable. As far as non-subject arguments are concerned, I will rely on quantifier float as a test which discriminates between direct arguments, or terms, and all non-terms (this test is introduced and discussed in chapter 2). How the distribution of the various properties discussed above correlates with

[10] In the case of the 3rd person pronoun, the reduced form can have plural reference for some speakers. See chapter 2 for discussion.

the status of various non-subject arguments will be a major focus of the remainder of this study.

1.3 Lexical-Functional Grammar

1.3.1 Types of representation

Lexical-Functional Grammar (LFG) is a theory of grammar developed in the late 1970s by Joan Bresnan and Ronald Kaplan. The theory models language as a set of parallel representations, each of which has to satisfy the constraints appropriate to that type of representation, and which together have to satisfy constraints on the correspondence between types of representation. Foundational material articulating the theory is presented in Kaplan and Bresnan (1982) and this version will be referred to in what follows as *classical LFG*. More recent developments are set out in Bresnan (2001a) and the bibliography of this work provides an up-to-date guide to work within the LFG framework. A brief introduction to the theory is Austin (to appear), and Falk (to appear) is an introductory text.

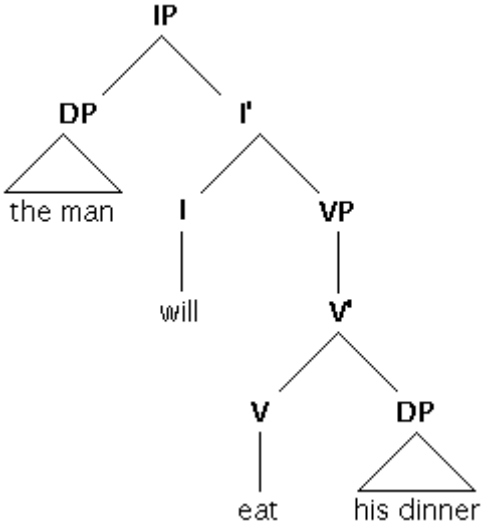
The advantage of using LFG as a framework for syntactic investigation is that it provides a rigorous formalism, the mathematical foundations of which are well worked out, which is also flexible enough to provide useful and insightful descriptions of languages with widely varying typological characteristics. The current work aims to take advantage of the rigour made available by LFG rather than introducing new theoretical developments into the framework, although chapter 5 will suggest some limitations in the theory as currently formulated. This rigour, it is hoped, will clarify the issues to be tackled here and constrain the possible solutions which can be advanced in a way that leads to interesting hypotheses. A wide range of languages have been studied using LFG, therefore a second advantage of using the framework is that the analyses proposed will be stated in a form that allows of cross-linguistic comparison in a simple fashion.

The name of the theory reveals two areas of grammar which are accorded especial importance in the theory. Firstly, a great deal of information is taken to be best treated as lexical information, rather than as part of syntax. One manifestation of this principle is that the principle of lexical integrity is basic: morphological processes are part of the lexicon and the internal structure of words plays no part in syntax. Secondly, grammatical functions (or grammatical relations) are treated as generalisations which capture equivalent mappings between predicate-argument relations and formal expression. These functions are not primitives, as the theory assumes that they can be decomposed (see section 1.3.2.2 below), but they are not reducible to phrase structure configurations or any other specific formal expression. This position is based on the observation that cross-linguistically it is relatively easy to state generalisations in terms of grammatical functions and relatively hard to state the same generalisations in terms of constituent structure or other formal characteristics.

LFG models grammar as a series of parallel representations with independent properties which map onto one another in a principled fashion. Three types of representation will be of interest in the current study. Firstly,

there is the representation of constituent structure, or *c-structure*. C-structures represent the constituency facts of the surface strings of languages using a standard version of X-bar theory[11]. The resulting representations are however rather different from those used by the Principles and Parameters theory (P&P) and its relatives. Empty categories are not a part of LFG c-structures: lexical integrity ensures that only morphologically complete words can be the terminal nodes in c-structure trees, and a principle of economy of expression (Bresnan 2001a: 91) bars the generation of nodes (except terminals and preterminals) which are not required by independent principles such as completeness, coherence and semantic expressivity (see discussion of these principles below). Functional heads and their projections are possible c-structure elements. The following is a partial c-structure for the English sentence *The man will eat his dinner*:

57.



In addition to having all nodes fully expanded, c-structure representations in LFG normally also include functional annotations on each node. These are not part of the c-structure; they specify a part of the correspondence between the c-structure and the associated functional structures. These annotations will be further discussed and exemplified below.

Functional structure or *f-structure* is the representation which includes information about functional relations. Functional here has two senses: this is the representation in which grammatical functions (GFs) are explicitly represented, and the way in which these are modelled can be read as function application in the mathematical sense. These representations take the form of attribute-value matrices (AVMs). An attribute can take an atomic value, or its value can be a complex structure, another AVM, but each attribute must have a unique value[12]. Each lexical head appears in f-structure as the value of a

[11] Trees which do not conform to X-bar theory are allowed in the description of non-configurational languages. The flat structure of such languages is modelled by phrase structure rules of the type $S \rightarrow X^*$, that is, a sentence consists of any number of elements in any order. Unfilled head positions are also possible in some circumstances, see discussion in section 1.3.3.

[12] The contrary is however not the case: different attributes may have the same value.

PRED attribute, along with a list of the GFs which it assigns (if any)[13]. Aside from optional elements (adjuncts, for example), only grammatical functions licensed by the lexical head of an f-structure can appear in it. This requirement is referred to as the *coherence condition*: all GFs must be required by a predicate in the local f-structure. Also, each GF in the f-structure must have a value. For some types of dependency, this value may be obtained from outside the local f-structure by structure-sharing, but the GF attribute nevertheless has a value which can be computed within the local f-structure. This condition is referred to as the *completeness condition*: an f-structure must contain all the GFs which the local predicate requires. In addition to PRED attributes and GF attributes, f-structures also include morphosyntactic features such as tense or number as attributes. The following is the f-structure associated with the English sentence used as an example previously[14]:

58.

$$\begin{bmatrix} \text{PRED} & \text{'eat} < \text{SUBJ, OBJ} > \text{'} \\ \text{TENSE} & \text{fut} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'man'} \\ \text{DEF} & + \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'dinner'} < \text{POSS} > \\ \text{POSS} & \begin{bmatrix} \text{PRED} & \text{pro} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \\ \text{GEND} & \text{m} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

A structure such as this can be interpreted as a collection of mathematical functions in the following fashion. Each complete f-structure is bracketed, and each can be read as a function which assigns values to the attributes it includes. Thus if we call the outermost f-structure in 58 *f1*, we can describe it as a function which returns certain values for various attributes:

59. *f1* (PRED) = 'eat'
f1 (TENSE) = fut
f1 (SUBJ) = *f2*
f1 (OBJ) = *f3*

The process could be repeated for the f-structures modelling the SUBJ (*f2*) and the OBJ *(f3).* LFG assumes that f-structure is the linguistic representation which is most similar cross-linguistically. In other words, the f-structure 58 is assumed to be identical or at least very similar to the f-structure associated with the translation equivalent of the English sentence *The man will eat his*

[13] The value of the PRED attribute is assumed to be some sort of semantic representation of the lexical head. Following common practice, I use the orthographic form of the word in inverted commas as a promissory note for this fuller representation.

[14] The POSS(essor) function is optionally assigned by most English nouns via a lexical redundancy rule.

dinner in other languages, regardless of how widely the c-structure representations might vary across languages.

The third type of representation which will be discussed in this work is argument structure, *a-structure*. This type of representation models the relations between predicates and their arguments. The semantic entry associated with a predicate has a list of the semantic roles assigned by the predicate, but no information about how they are realised. F-structures have information about the GFs assigned by a predicate, but if this information is not part of the predicate's lexical entry (see next section for discussion), some additional representation is necessary which contains the minimum syntactic information necessary for the linking of semantic roles to GFs to be accomplished. A-structure is the additional representation required, and it combines information about the syntactic and semantic prominence of arguments. Terms are more prominent than non-terms, and the a-structure respects this division. Within each division, arguments are ordered by semantic prominence as reflected in the thematic hierarchy (Gruber 1965). The a-structure representation of the sentence used as an example previously would be:

60. eat < x , y >

For comparison, the structure associated with the English verb *put*, which assigns an obligatory oblique GF, would be:

61. put < x , y : z >

where the colon indicates the division between terms and non-terms. Argumentation motivating this type of representation is presented in Manning (1996a) within the framework of Head-Driven Phrase Structure Grammar, but the arguments apply equally to LFG. Other work which makes crucial use of the properties of a-structure includes Arka & Manning (to appear) and Wechsler and Arka (1998). A different version of a-structure is presented by Bresnan (2001a: chapter 14); in this style of structure the division between terms and obliques is not an explicit part of the representation and is instead derived by the effects of the Lexical Mapping Theory (see section 1.3.2.2). The question of which type of representation should be preferred is discussed in sections 2.5.2 and 3.3.2.3. The latter section discusses a case in which the two theories make different predictions, and suggests that the more articulated a-structure is supported by the data in question. In the discussion of LMT below (section 1.3.2.2), I use Bresnan-style representations to reflect the sources accurately.

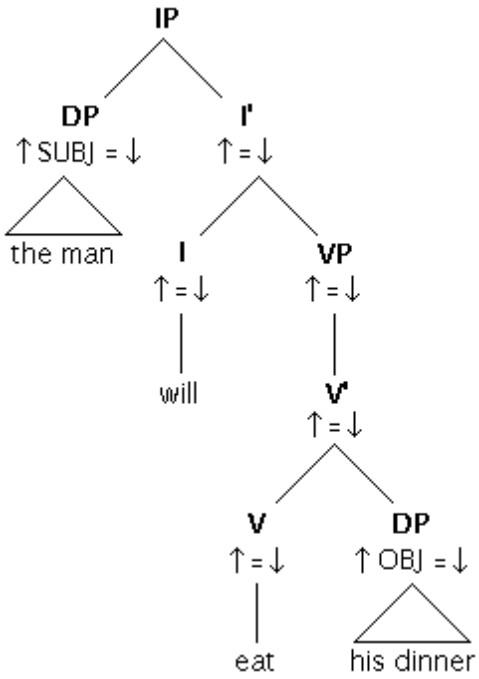
1.3.2 Correspondence principles

The three parallel representations discussed above, as well as others such as semantic structure which are not discussed here, are mapped onto each other by correspondence principles. The ones of interest here are those that map from c-structures to f-structures, and those that map from a-structures to f-structures.

1.3.2.1 C-STRUCTURE TO F-STRUCTURE

Every node in a c-structure must correspond to some f-structure, and this is modelled by annotations on c-structure nodes. These annotations may be introduced by lexical entries or phrase structure rules in special cases, but in general they are provided by default principles (Bresnan 2001a: 102-103). The annotations take the form of equations such as ↑ = ↓, where ↑ can be read as 'the f-structure of my mother node' and ↓ as 'my f-structure'. The equation given in the previous sentence is therefore interpreted as saying 'the f-structure of my mother node is identical to my f-structure'. The annotated version of the tree in 57 is the following:

62.



The equation on the left-most DP node means 'the f-structure of the SUBJ in the f-structure of my mother node is identical to my f-structure'. In this fashion, information from lexical entries is combined into complex f-structures. The process can be given a rigorous mathematical treatment, which will not be set out here (see Bresnan 2001a: 56-60 for details), except to note that if the functional description of a c-structure can be resolved, then there is an infinite set of f-structures which will be compatible with the c-structure, all of which have additional attribute value pairs not mentioned in the functional description. The *minimal* f-structure which is compatible with the description is the linguistically interesting one[15].

For languages in which phrasal categories have endocentric organization, the functional annotations of c-structure nodes are predictable

[15] The account of how f-structures are built up presented here is the standard one. Recent work by Andrews and Manning (1999) has suggested that the use of complete equation as the main tool is misleading. They suggest that more insightful analyses can be given in some cases if only part of the functional information associated with some nodes is passed up to the mother node, rather than assuming total identity of the f-structures. This work does not bear on the issues considered in the present study.

(Bresnan 2001a: 98-109). I take Indonesian to be a language of this type, and I will therefore omit functional annotations from trees in most cases.

1.3.2.2 A-STRUCTURE TO F-STRUCTURE: LEXICAL MAPPING THEORY

The prototypical association of certain semantic roles with syntactic functions, such as that holding between the agent role and the subject function, was already a part of the Sanskrit grammatical tradition (Blake 1994: 65-67). Fillmore (1968) made such associations a part of generative grammar, and much subsequent work has been devoted to analysing which facts about the semantics of verbs are predictors of the linking between semantic roles and syntactic functions (e.g. Carter 1976, Dowty 1991, Foley & Van Valin 1984, Grimshaw 1990, Jackendoff 1990, Tenny 1994). In classical LFG as described in Kaplan and Bresnan (1982), the lexical entries of verbs directly specified the GFs associated with their arguments, but subsequent work has tried to capture linking regularities through a separate module of grammar known as Lexical Mapping Theory (LMT) This theory is first presented in Bresnan & Kanerva 1989 (see also Bresnan 2001a, chapter 14).

In LMT, grammatical functions are treated as bundles of features, the two features in question being ±r(estricted) and ±o(bjective). These features generate a four-celled matrix, giving the four GFs recognised by LFG:

63.

	-r	**+r**
-o	SUBJ	OBL_θ
+o	OBJ	OBJ_θ

It is an implicit claim of this classification that GFs fall into natural classes. Thus, an *unrestricted* GF need not have a semantic role; *restricted* GFs must have a semantic role. Expletive arguments are possible in subject and object position, but not elsewhere and these two GFs therefore constitute a natural class. A *non-objective* syntactic function is one that can be associated with an intransitive predicator. Subjects and obliques fulfil this condition and therefore also constitute a natural class. A further claim is that the four GFs can be ranked in a markedness hierarchy, with negative features taken as unmarked. SUBJ is negatively specified for both features and is therefore least marked; OBJ_θ is positively specified for both and is therefore most marked. The remaining GFs are ranked between these two:

64. Markedness of argument functions (Bresnan 2001a: 309, ex19)
SUBJ > OBJ, OBL_θ > OBJ_θ

This hierarchy plays an important part in the mapping between argument structures and f-structures.

Argument structure (a-structure) consists of a predicator and a list of the arguments associated with it in order of semantic prominence, where semantic prominence is equated with the thematic hierarchy generally accepted since Gruber (1965). The most prominent argument is designated the *logical subject,* and all arguments are annotated with a single feature value which is assigned on the basis of the semantics of the verb. Patientlike arguments receive the feature [*-r*] whilst other arguments receive the feature

[-*o*]. The feature assigned to secondary patients is subject to cross-linguistic variation, but in asymmetric object languages such as English and Indonesian it is [+*o*]. The logical subject is mapped to the subject function unless it is not the first item in the argument list, a situation which can occur in the case of verbs with athematic subjects. Other arguments are mapped to the lowest function on the hierarchy 64 which is compatible with the feature specification they carry.

Comparing the operation of these principles for two English verbs with three arguments each will clarify this scheme. *put* has three arguments, of which only one, the second in thematic prominence, is a patient. Patientlike arguments are associated with the feature [-*r*], and this principle will apply to the second argument of *put*. Other arguments are associated with the feature [-*o*] by default, therefore the complete a-structure of the verb will be:

65. *put* < x, y, z >
[-o] [-r] [-o]

The logical subject is initial in this a-structure, and is therefore mapped onto SUBJ. The remaining roles are mapped to the most marked GFs compatible with their features: in the case of the second argument, this is OBJ, and in the case of the third argument it is OBL_θ. The relevant part of the f-structure associated with the node under which the verb appears in a c-structure will have the following form:

66. PRED 'put < SUBJ, OBJ, OBL_θ >'

give differs in having two arguments which are patientlike. By the Asymmetrical Object Parameter (Bresnan & Moshi 1990), only one of these can be associated with the value [-*r*]. The secondary patient (the one that is least thematically prominent) is associated with the value [+*o*]. The resulting a-structure is:

67. *give* < x, y, z>
[-o] [-r] [+o]

The first two arguments will be mapped to SUBJ and OBJ as before, but the feature associated with the third argument is compatible with the most marked GF, OBJ_θ and it is therefore mapped onto that GF. The resulting partial f-structure is:

68. PRED 'give < SUBJ, OBJ, OBJ_θ >'

The licensing of arguments through the interaction of LMT, f-structure and c-structure will be discussed in section 1.4.

1.3.3 Functional categories in Indonesian syntax

It was mentioned above that functional heads and their projections are part of the vocabulary of LFG c-structures, in common with most syntactic work since the 1980s. But such elements are rather weakly grammaticized in Indonesian. The language has no subject-verb or verb-object agreement, tense and aspect are usually not specified in clauses, number marking of nouns is

rare and definiteness is not marked obligatorily. These facts raise the question of whether it is appropriate to analyse the language as having functional categories, in particular IP as the category of clauses and DP as the category of referential expressions. The purpose of this section is to argue that such an analysis is plausible, and to show how functional categories are handled by LFG.

The argument is easier to make in respect of the clause-level functional categories. Two types of element appear between the subject and the verb in Indonesian clauses, and neither of them can appear in other positions (with the same meaning). These two are negation:

69. *Marisa, kamu* ***tidak*** *mengerti*
Marisa 2SG NEG *meN*.understand
'Marisa, you don't understand.' (SDM: 76)

and expressions of modality:

70. *Aku* ***harus*** *melihat buktinya dulu*
1SG must *meN*.see proof.3 before
'I must see his evidence first.' (SDM: 76)

Expressions of tense and/or aspect also appear in this position, but in at least some cases these words can be used in other clausal positions with similar meaning:

71. *Ia* ***sedang*** *membaca ketika saya datang*
3SG PROG *meN*.read while 1SG come
'He was reading when I came.' (E&S: 487)

72. ***Sedang*** *ia mengucapkan kata-katanya itu isterinya menjerit*
while 3SG *meN*.say word.DUP.3 that wife.3 *meN*.scream
'While he was uttering those words, his wife screamed.' (E&S: 487)

The case of *sudah* is similar: it is used to mark a completed action, but it can also be used as an adverb meaning 'already'[16]. Other adverbs can be placed between the subject and the verb:

73. *Proyek itu* ***tetap*** *akan menjadi milik Candra Surya Abadi*
project that certainly FUT *meN*.become property sign Surya Abadi
'That project will certainly become the property of the Surya Abadi group.' (SDM: 76)

[16] This ambiguity is common in Western Austronesian languages. In Sasak, auxiliary verbs are distinguished positionally and by being clitic hosts, and the class includes *wah* which indicates past action. The same word appears in other positions, and then is translated as 'already' (Austin 2000).

Therefore the status of the temporal and aspectual markers is not entirely clearcut. However, *akan*, glossed FUT above, is clearly an auxiliary verb: it does not occur in other positions[17] and it can be the base of a derived verb *meng-/di-akan* 'aim, strive for' and a nominalisation *keakanan* 'the future'. On this basis, it is reasonable to suppose that there is a position above VP for a head expressing tense and/or aspect, and that other words like *sudah* and *sedang* have dual lexical entries both as adverbs and as functional items which can occupy that position. Various combinations of the three possibilities, including all three, are possible:

74. *Partai itu tidak akan bisa membentuk pemerintahan*
party that NEG FUT able *meN*.form government
'That party will not be able to form a government.'
(IRG: 204)

This suggests that a series of functional head positions is required, as suggested first by Pollock (1989). This point will be important in the argumentation of section 2.2.1.2, but in general I will only include a single projection, labelled IP, in c-structure representations of clauses.

There are two common ways of indicating definiteness in Indonesian nominals. The first is to use the demonstrative *itu* 'that' to indicate that the referent has been mentioned recently. The second is to use the third person possessive clitic *-nya* to indicate that the referent can be understood within the context of interaction, but has not previously been mentioned (IRG: 150-151, Sukamto 1999). The contrast can be seen in the following examples:

75. *Ibu sudah memasak nasi. Nasi itu di*
mother PERF *meN*.cook rice rice that LOC
lemari
cupboard
'Mother has cooked rice. It (LIT: that rice) is in the pantry.'

76. *Kalau mau makan, nasinya di lemari*
if want eat rice.3 LOC cupboard
'If you want to eat, the rice is in the pantry.' (IRG: 151)

The possessive clitic has the same distribution in this usage as it does in its use as a true possessive. This includes appearing closer to the head than demonstratives and relative clauses, and I therefore conclude that it cannot be of the category D. The case of the demonstrative is less clear. *itu* and its complementary *ini* 'this, these' are always the last element in a nominal phrase. This can lead to ambiguity in complex phrases:

77. *Anak yang naik sepeda itu tinggal dekat saya*
child REL climb bike that live close 1SG
'That child riding a bike lives near me.'
OR 'The child riding that bike lives near me.' (IRG: 157)

[17] There is a homophonous preposition, which is not relevant here.

I assume that appearing at the boundary of the phrase (at least most of the time) will be a characteristic of D, therefore these demonstratives are possible candidates. There is a strong reason for doubting that this is the correct analysis, however. It is a reliable generalisation that heads precede complements and other dependents in Indonesian. This is true for verbs and prepositions as can be seen in the preceding examples. It is also true for nouns:

78. *dongeng tentang seorang haji*
legend about one.CLASS haji
'a story about a haji' (IRG: 150)

and for adjectives:

79. *Sulit untuk kita memperoleh bukti*
difficult for 1PL.INCL *meN*.obtain proof
'It is difficult for us to get proof.' (IRG: 183)

And if the position argued for immediately above is correct, it is also true for the functional heads above VP. Therefore, it would be surprising if D were to take its complement in a right-branching structure[18]. On this basis, the only plausible candidates for the category D must occur before the head noun in nominal phrases. This condition restricts the candidates to quantifiers, including numbers, and classifiers.

Classifiers can be eliminated from consideration, because they can only ever occur with a number (IRG: 134):

80. *dua ekor kuda*
two CLASS horse
'two horses'

81. *seekor kuda*
one.CLASS horse
'a horse'

All Indonesian classifiers are originally nouns, and where there is no number associated with them, they revert to their nominal meaning:

82. *ekor kuda*
tail horse
'horse's tail'

This leaves, by elimination, the result that quantifiers are exponents of the category D in Indonesian.

Although this hypothesis is semantically plausible, it is also not without syntactic problems. Chung (1976a) discusses the possibility of quantifiers shifting from the left edge of their nominal phrases to the right edge:

---

[18] Guilfoyle, Hung and Travis (1992: 401) state that: 'There is independent evidence that the determiner (sic) *itu* is only a modifier, not a functional head.' but without giving any details. I assume that what they have in mind is some version of the argument just given.

83. *Semua pemain musik pulang pagi*
all player music go.home morning
'All the musicians left early.'

84. *Pemain musik semua pulang pagi*
player music all go.home morning
'All the musicians left early.'

She notes that this phenomenon is distinct from quantifier float, and that it is associated with some difference in meaning which she is unable to make specific. It should be noted that this possibility also exists for numbers and number/classifier combinations:

85. *Dia mengimpor kursi sebanyak 8000 buah*
3SG *meN*.import chair as.many 8000 CLASS
'He imported as many as 8000 chairs.' (IRG: 140)

However, it is not possible in any of these cases to insert the quantificational element(s) within the remainder of the nominal phrase; the quantifier always must be on one edge of the structure. This is not counter-evidence to the hypothesis proposed, if we assume that the ordering of D and its complement is not rigid.

In chapter 2, I assume the idea originally proposed by Postal (1969) that pronouns are determiners rather than nouns. In association with the analysis argued for above, this assumption predicts that pronouns and quantifiers should not co-occur. There are some possibilities for quantifiers to occur with pronouns, but they are limited in such a way that they are not counter-evidence to the proposed analysis. The only numbers that can occur with pronouns are morphologically complex forms. Bare numbers cannot appear in this environment, only numbers prefixed with *ber-*:

86. *dua orang*
two person
'two people'

87. **dua mereka*
two 3PL
(FOR: 'two of them')

88. *mereka berdua*
3PL *ber*.two
'both of them'

and these numbers do not occur with other nouns (IRG: 58). As discussed previously, morphology in Indonesian is derivational; *ber-* is a prefix commonly associated with intransitive verbs and adjectives. These facts, and its position in structures such as 88, suggest strongly that it is a modifier of some sort in this usage. Some quantifiers can occur with pronouns also. These words precede nouns, except in the case of quantifier shifting discussed above, but they must follow pronouns (IRG: 169-170):

89. *Kami semua harus membuka jalan dahulu*
1PL.INCL all must *meN*.open road before
'All of us must clear a road first.' (IRG: 170)

90. **Semua kami.....*

Whatever the correct analysis of quantifier shifting may be, it is significant that it is obligatory in just this case. The natural interpretation is that it is a mechanism of last resort here which rescues a structure predicted to be impossible on other grounds. Therefore, in the two situations in which pronouns can co-occur with quantifiers, the quantifier does not behave as it does with other nouns. It is plausible to maintain the assumption that quantifiers are generated as functional heads which take NP as their complement, but that the relative ordering of the head and its complement can be reversed under conditions which are not understood[19]. In what follows, I will treat the maximal projection of referential phrases as DP.

I have argued that I and D are functional categories that exist in Indonesian. Following the standard account of such categories, this means that they are compulsory: every clause must be headed by an I and every referential phrase must be headed by D. But as mentioned at the start of this section, functional categories are commonly not instantiated in Indonesian. LFG does not allow null heads in c-structure, therefore it is appropriate to at least sketch here the approach taken to such a situation in the framework. A functional head and the lexical head of its complement map to the same functional structure. Where a functional head is obligatory (for example, if the grammar includes a rule which makes I(nflection) the head of a finite clause), and the functional head is realised morphologically, lexical integrity forces the lexical head to appear in the c-structure position of the functional head. This analysis accounts for many of the structures analysed as examples of head movement in Principles and Parameters accounts (see for example Pollock 1989, Grimshaw 1997). An example of the LFG analysis is the case of so-called verbal-nouns in Welsh. The finite verb is initial in Welsh, giving VSO order when a finite verb is used. But where a finite auxiliary is initial and a non-finite "verbal noun" appears, the word order is Aux-SVO:

91. *Gwelodd Siôn ddraig*
saw. John dragon
'John saw a dragon.'

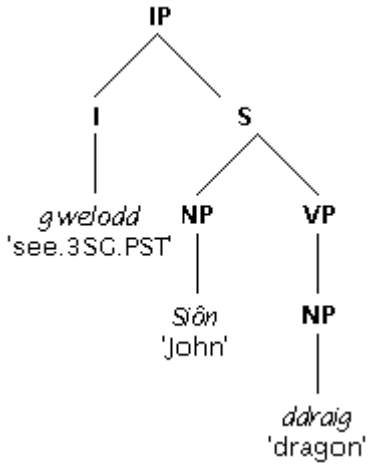
92. *Gwnaeth Siôn weld draig*
do.3SG.PAST John See.VN dragon
'John saw a dragon.' (Sproat 1985: 176)

We can assume that the c-structure of example 92 is that of an IP in which I takes a S consisting of an NP and a VP as complement (see Bresnan 2001a: 127-129 for argumentation). Example 91 can have the identical structure, with

[19] The question of whether any part of a structure such as *sebanyak 8000 buah* should be analysed as the specifier of D, and if so, what linear ordering principles apply is ignored here. Clearly, this an area which needs further research.

the sole exception that the V° position in VP is not present. As the main verb in this example is finite, it must appear clause initially, that is as head of IP. By the structure-function correspondence principles previously discussed, the nodes associated with I, IP, V and VP all map to the same f-structure. Therefore, if lexical material were inserted under V in example 91, the PRED attribute of that f-structure would have two values, a possibility barred by general principles. Therefore only a V node dominating nothing is possible. But such node is not permitted by economy of expression, and it must be omitted all together. The resulting structures are the following (see Bresnan 2001a:129-131 for details):

93.



94.

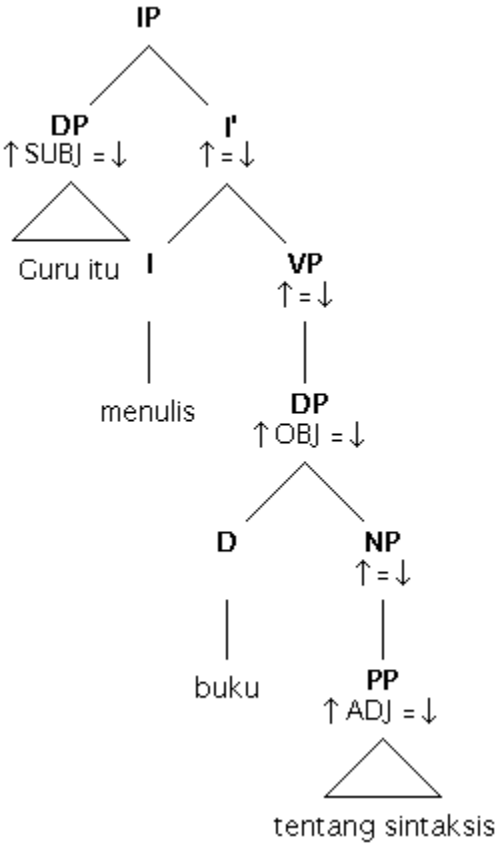
PRED	'see (SUBJ), (OBJ)'	
TENSE	PAST	
SUBJ	PRED	'NAMED-SiÔn'
	PERS	3
	NUM	SG
OBJ	["dragon"]	

(Bresnan 2001a: 131, ex12)

The c-structure head in I is then said to be an *extended head* or the *co-head* of the two c-structure categories (Bresnan 2001a: chapter 7). In the same way, where there is no phonological material realising an obligatory functional category, the lexical head will appear in the c-structure position of the functional head. I propose that, for Indonesian, such structures are associated with a radically underspecified value in the f-structure which is compatible with, and is given content by available contextual information. I will in general omit this attribute from f-structure representations, but it is shown in the following example with the value (...) assigned to the attributes TAM (tense/aspect/mood) in the outer f-structure and DEF (definiteness) in the f-structure associated with the DP[20,21]:

[20] Some non-branching nodes are omitted from the following c-structure.

95.

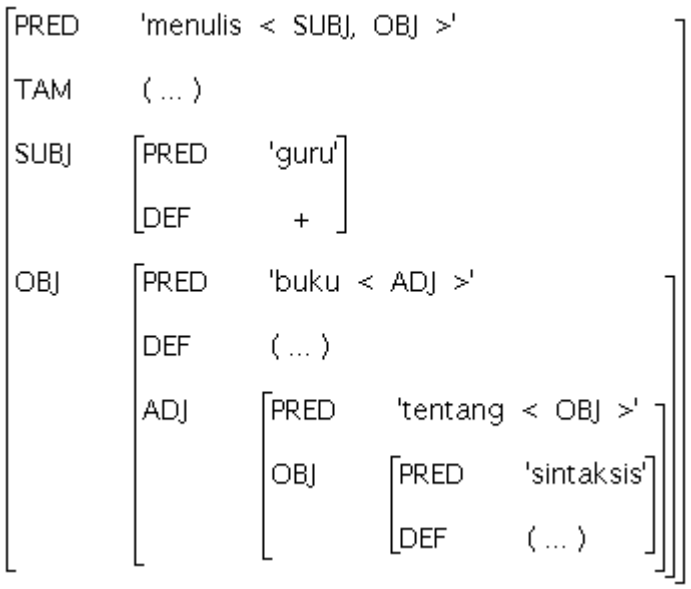


[21] The Indonesian facts perhaps appear closer to those of English, given the similarity of basic word order for the two languages. The accepted LFG analysis of finite clauses in English without auxiliaries is that they are subject-predicate structures, NP, XP, where XP is a VP (Bresnan 2001a: 115). A similar analysis does not seem possible for Indonesian, however, because IP is possible as the XP part of a subject-predicate structure with a non-verbal predicate:

I *Minuman itu harus dingin*
Drink that must cold
'The drinks must be cold.' (IRG: 235)

There is no obvious syntactic distinction between the subject of such clauses and subjects which I analyse as specifiers of IP. Therefore there would seem to be a large class of cases for which two analyses are possible: IP only, or NP-XP (= IP).

96.



97. *Guru* *itu* *menulis* *buku* *tentang* *sintaksis*
teacher that *meN*.write book about syntax
'The teacher wrote a book about syntax.'

I will not pursue this issue further, but I would suggest that the fact that the two functional categories both behave in this way is an important typological fact about Indonesian.

Additional theoretical material relating to LFG is introduced in various parts of this work, as it is required.

1.4 The licensing of arguments

We are now in a position to give a detailed statement of how LFG models the licensing of arguments in a configurational language, and to draw some implications from it. These implications will then be compared with data from Indonesian in order to reveal the nature of the problems to be studied in the following chapters.

A verb has a list of semantic roles associated with it in its lexical entry. LMT provides a set of principles which associate syntactic features with these roles, and the result is an a-structure representation. The mapping principles of LMT then operate on this a-structure and the result is a partial f-structure. These steps are illustrated again for the English verb *put*:

98. *put* V < agent, theme, location > (lexical entry)

99. *put* < x, y, z > (a-structure)
[-o] [-r] [-o]

100. PRED 'put < SUBJ, OBJ, OBL_θ >' (partial f-structure)

In order that the completeness condition be satisfied, the complete f-structure for a clause with this verb must contain a SUBJ, an OBJ and an OBL_θ, and appropriate c-structure positions for them must exist. The X-bar rules of the language and the principles which provide functional annotations on c-

structure nodes must combine to fulfil this condition. I will assume an X-bar theory which is restricted to binary branching, except in the case of second objects, and the annotation principles given by Bresnan (2001a: 102-103). These principles state that the specifier of a functional category is a discourse function. In the case of IP, this allows the grammaticalized discourse function SUBJ to appear as its specifier. The principles also state that the complement of a lexical category is a non-discourse argument function, and this allows the GF OBJ to appear as complement of a verbal head. I assume that ditransitive verbs have two complement positions, their first V' is a ternary-branching node, and that two argument functions can appear in that case. Finally, a principle allows that a maximal projection adjoined to a constituent is optionally a non-argument function. These principles capture the generalisation that arguments are licensed structurally in two configurations: specifier of a functional category, and complement of a lexical category. The last principle also captures the generalisation that obliques and adjuncts have in common the property of not being licensed in this way. The observed properties of a language such as English fall out naturally from this account. SUBJ, OBJ and OBJ_θ are structurally licensed, OBL_θ and adjuncts are not. As adjuncts are not assigned semantic roles by the verb, it is also predictable that the licensing mechanism will make a semantic contribution. Considerations of economy predict that the mechanism used to license non-terms will be identical for both types, and therefore it is predictable that the licensing mechanism for OBL_θ will introduce semantic content also.

If Indonesian is a configurational language without case-marking, the account of the licensing of arguments sketched above will apply, and we can make various predictions. We predict that there will be at least one non-subject argument position where a DP is licensed structurally, and that where only one of these positions exists the DP which occupies it will have the GF OBJ. We predict that other non-subject arguments will appear in different positions and will require some other licensing mechanism, such as prepositions. And we predict that the licensing prepositions will code semantic distinctions between such arguments. The facts of the language cast doubt on all of these predictions.

The first two predictions are thrown into question by data regarding the most common two argument clause types of the language:

101. *Ali* *melihat* *gadis* *itu*
Ali *meN*.see girl that
'Ali saw the girl.'

102. *Gadis* *itu* *dilihat* *Ali*
girl that *di*.see Ali

103. *Gadis* *itu* *dilihatnya*
girl that *di*.see.3

104. *Gadis* *itu* *dilihat* *kemarin* *oleh* *Ali*
girl that *di*.see yesterday by Ali
'The girl was seen (yesterday) by Ali / him.'

Example 101 is a clause of the type described as *active* in most grammars (e.g. IRG), while examples 102-104 are described as *passive*. The subject precedes the verb in all cases. The phrase *gadis itu* is the object in example 101 according to the account developed above, but I have already noted that in a clause which differs only in lacking the verb prefix, this argument has different properties (see discussion of examples 52-56 in section 1.2.3). This is unexpected: we would predict that arguments which are exponents of the same GF will have similar properties. What of the postverbal elements in examples 102-104? In example 104, there is a prepositional phrase, which would be predicted by most theories of voice, and its function is marked by the preposition *oleh* 'by' just as predicted. But example 102 and example 103 do not fit into the picture so easily. We have a bare DP in one case and an attached pronoun in the other: neither has a preposition indicating its function in the clause. Rather than obliques, these have the coding properties of direct arguments, and therefore they are also objects by the theory given above. This in turn would mean that these are not passive clauses under commonly accepted assumptions about the nature of voice alternations. I will argue that these clauses present a more complicated picture still: there is good evidence to show that the post-verbal NP in example 102 is an oblique and that it is licensed by adjacency to the verb rather than by a preposition; there is also evidence that the pronoun in example 103 is a direct argument of the verb, and that it is not licensed by virtue of its syntactic position but by a morphological process. Example 102 already shows that prepositions are not necessary to code functional distinctions between non-terms, but another clause type shows that, in one environment, there is no relation between prepositions and functional distinctions at all. In this case, a variety of prepositions can be used to license a non-subject argument without any associated meaning distinction:

105. *Saya tak ingat akan tuan*
1SG NEG remember about sir
'I don't remember you.' (E&S: 223)

106. *Masih ingatkah Anda pada saya?*
still remember.EMPH 2SG to 1SG
'You still remember me?' (SDPGS: 108)

107. *Saya jadi ingat dengan hit Scot Mackenzie*
1SG thus remember with hit Scot Mackenzie
'Therefore I remember Scot Mackenzie's hit.' (McKay)

In addition, the preposition can be omitted altogether[22]:

[22] For reasons discussed in section 4.1.2.3, this example is ambiguous as to clause type.

108. *Di mana pun, dalam keadaan pun, dia selalu*
LOC where PRT inside situation PRT 3SG always
ingat Marisa
remember Marisa
'No matter where, no matter what situation, he always remembered Marisa.' (SDM: 265)

All of these examples show that the licensing of non-subject arguments in Indonesian is not as straightforward as the theory given above would lead us to expect.

The conclusion that will be argued for in this thesis is that there are two separate systems of clausal organization in Indonesian, each of which licenses non-subject arguments in a different way. Consequently, the properties of the arguments are different in each case. Clause types such as those seen in examples 101-104, and in some other types not yet introduced, will be argued to be typical of one system. I will suggest that this system is related to, though not identical with, the clausal syntax of other more conservative Austronesian languages. The characteristics of this system include that the predicate of the clause can be headed by all major categories, that terms within the predicate constituent are not available to syntactic processes such as extraction, and that there is a tendency for arguments internal to the predicate phrase to be joined to the head in a close relationship. The clause type seen in example 108, on the other hand, is typical of the other, more innovative system. The characteristics of this system include that the second major constituent of the clause must be headed by a verb and that the verb phrase is open to processes such as extraction. The licensing of arguments in the second system proposed above can be accounted for in a theory such as that outlined at the start of this section, but this is not the case for the first system. Additional assumptions are required in order to give an account of the licensing of non-subject arguments when the major constituents of the clause are subject and predicate, as opposed to subject and verb phrase, and I will argue that these additional assumptions are not compatible with LFG as currently formulated.

1.5 The origin of Bahasa Indonesia and the data used

The Indonesian archipelago is an ethnically and linguistically diverse area. The vast majority of the languages spoken in the archipelago are from the Austronesian family, but the variety is nevertheless considerable. The feelings of distinct ethnic identity felt by groups in different regions are evidenced clearly in the recent political history of the Republic of Indonesia and the basis for the nation of Indonesia was a negative emotion. All the various groups involved in the nation-building movement were opposed to Dutch rule, but this did not necessarily imply any further solidarity between the groups. This was seen as a problem from early in the history of the independence movement, and the adoption of a national language was regarded as a crucial unifying move. The native language with the widest geographical distribution and use, although certainly not the greatest number

of speakers, was Malay, which had been the most common trade language of the region for centuries. This was adopted as the basis for a national language, Bahasa Indonesia ('language of Indonesia') by the nationalist youth movement in 1928, and became the official national language when the Republic of Indonesia was proclaimed in 1945.

This language then became the official language of government and of education. The typical Indonesian of earlier generations learned her native language (for example Javanese or Balinese) at home and learned Indonesian at school. Indonesian was therefore a language without speakers for whom it was the first language, although many people could be said to have native speaker competence in the language, having commenced learning it at the age of five or six and having used it daily after that. More recently, increased mobility within the country has meant that marriages between people from different language backgrounds have become more common, and as a result there are now some speakers who have never learnt another language. An additional factor complicating matters is that, as previously mentioned, Malay was a trade language throughout the archipelago for a considerable period of time. Consequently, there are Malay dialects spoken at various places in Indonesia. For example, there are local dialects of Malay spoken in the widely separated cities of Jakarta (on Java) and Ambon (on Seram). There is also variation in the Indonesian spoken in different regions due to the influence of the different first languages spoken (Moeliono 1994, Tanner 1967).

The result of this situation is that there are many varieties spoken which speakers will identify as Indonesian but which may vary widely from one to another. There is a prescriptive standard which is taught in schools and is used as the language of the government and the legal system, but is scarcely heard or read elsewhere. This is the variety which is described by most grammars (e.g. McDonald & Darjowidjojo 1967, IRG), but as a language in use Indonesian is richer. All the resources of the standard are available (if not actually used), but many other possibilities exist. To take a simple example, IRG has extensive discussion of the morphological causative formed with the suffix *-kan* (pp70-78*)*, but does not mention the periphrastic causative formed with the verb *kasi* 'give'. This construction is recognised by Echols and Shadily's dictionary (1994):

109. *Kasi lihat bukunya*
give see book.3
'Let me see the book.' (E&S: 263)

although at the head of the entry for *kasi* they mark it as 'colloquial'[23]. Similar points are made with detailed evidence in respect of the so-called 'bare active' clause type by Voskuil (1996, appendix to chapter 8). Any study of Indonesian which restricts itself to the standard language will therefore miss important phenomena and valuable evidence.

[23] Kana (1986: 77-80) discusses this construction and suggests that it is characteristically eastern Indonesian.

The strategy adopted in the present study has been to use textual surveys to establish the scope of the phenomena under investigation in actual language use, and then to use the information obtained from text examples as the basis for focussed elicitation from native speakers. This is particularly true of the material presented in chapters 3 and 4: chapter 3 takes as starting point several published studies which are themselves based on text surveys and uses them as the basis for further investigation, while chapter 4 is based on my own survey of several texts (full details are given in section 4.1.2.1). Chapter 2 discusses aspects of grammar which are more central, and about which published information is generally reliable. Both the texts used and the judgments provided by my consultants reflect the careful usage of educated Indonesians. Most of the native speakers consulted are of Javanese origin and their judgments might not be replicated by speakers with another first language. The issue of L1 influence on Indonesian is certainly beyond the scope of this study (see Moeliono 1994, Steinhauer 1994).

Bahasa Melayu, the national language of Malaysia, is also based on Malay and is essentially the same as Indonesian. The orthographies of the two languages were not identical, however, as they reflected the phonology to orthography mappings of different European languages. In 1972 the two nations adopted a uniform orthography, which is the current standard for both languages. Some examples in this work are taken from earlier sources, but the orthography has been standardised throughout, with the exception of proper names.

1.6 Overview of the remaining chapters

CHAPTER 2 presents an analysis of the major two argument verb class of Indonesian. The data presented above has shown that there are four clause types in which such verbs can appear (examples 101, 102-104, 16, and 52). Two of these clause types have the more agentive argument as subject, and two have the more patientive argument as subject. In addition, there are restrictions on the nature of the non-subject argument in these last two clause types. These facts have never been explained in a straightforward fashion. The focus of the current study on the properties of non-subject arguments sheds some new light on the problems.

CHAPTER 3 deals with a class of verbs usually described as *adversative verbs*. These are characterised formally by the presence of the circumfix *ke- -an* around the root, and semantically (in almost all cases) by the subject of the verb experiencing adverse consequences as a result of the event:

110. *Tomo kecurian uangnya*
Tomo *ke*.steal.*an* money.3
'Tomo has had his money stolen.'

The non-subject argument in this clause type can be a bare DP as in 110, but sometimes it can be introduced by the preposition *oleh* 'by'. I will argue that this argument is of the same type as the agentive argument in some of the patient subject transitive clauses (such as example 102), and that in each case the argument is an oblique which is licensed by being adjoined to the verb. The *ke- -an* verbs are of interest also because they are naturally treated as

examples of predicate composition: that is, the circumfix should be analysed as a separate predicate which takes another predicate as one of its arguments. But unlike many other phenomena which can be analysed in this fashion, the way in which the governing predicate creates an argument structure is driven by semantic factors, not by the grammatical functions assigned by the governed predicate. The mapping from the a-structure of the complex predicate to grammatical functions also provides an argument in favour of the style of a-structure proposed in Manning (1996a).

CHAPTER 4 is an analysis of the emotion and cognition verbs seen in examples 105-108. I argue that these words have multiple lexical entries, and that only the entries responsible for examples such as 108 are unequivocally verbal. The properties of the non-subject argument are such that it must be considered a direct argument, but other properties of the clause type also make it clear that this type of verb is outside of the transitive system analysed in chapter 2. This conflict forces a choice between two conclusions: either the non-subject argument of emotion verbs is a second object (an OBJ_θ in LFG), a conclusion which is theoretically and empirically problematic, or the assumption that the complex system studied in chapter 2 is the basic transitive verbal system of the language is wrong. On the basis of striking similarities between the emotion verb clauses and the unprefixed agent subject transitive clause type, the second conclusion is adopted.

CHAPTER 5 attempts to motivate the conclusion reached in chapter 4 by an examination of the properties of the two different clausal systems which I propose exist in Indonesian. The interaction of extraction possibilities and the positions that can be occupied by enclitic pronouns is central in this analysis, with the generalisation concerning the distribution of enclitics most naturally stated in terms of the category *predicate*. On the basis of the properties of their non-subject arguments, as analysed in previous chapters, and on the basis of additional evidence, I propose that prefixed verbs and their projections fall within this category and unprefixed verbs do not. Comparison with other Austronesian languages and with proposed reconstructions of Proto-Austronesian syntax shows that the subject-predicate structure is conservative and that the other appears to be an innovation. The chapter concludes by arguing that this insight into the structure of Indonesian resists expression using the current resources of LFG, and suggests what sort of modifications might be necessary.

2 Indonesian Transitive Clauses

Transitive verbs are those which take two term arguments, which in LFG will always be a SUBJ and an OBJ. The class of verbs traditionally treated as transitive in Indonesian can appear in four possible types of clause. In this chapter, I will assume that the possibility of appearing in all four types is the defining characteristic of transitive verbs in the language, and I investigate the status of the non-subject argument in each clause type. I will argue that three clause types always have two direct arguments, and the fourth sometimes does also. The patterns of linking between semantic roles and grammatical functions are not the same for all types, and the word orders associated with the different types vary, therefore this variety cannot obviously be treated within standard views of linking theory or of phrase structure. The analysis developed here is that the variety of the clause patterns, their individual attributes, and the partial complementary distribution between two of them can all be attributed to morphological properties of the language. In particular, I argue that Indonesian makes extensive use of 'small constructions' (Poser 1992), that is, transparent combinations of lexical items which must nevertheless be taken to be part of the morphological component of the language, and which are used as an extra-syntactic means of licensing term arguments.

2.1 The range of transitive constructions

As mentioned above, I will assume that transitive verbs in Indonesian are defined by the possibility of appearing in four different types of clause. Many verbs which have two term arguments can take the prefix *meN-* (where N indicates a morphophonemically conditioned nasal consonant, dropped before liquids, see section 1.2.1for details):

1. *Dia membaca buku itu*
   3SG *meN*.read book that
   S/he read the book.

With a third person actor[1], an alternative prefix *di-* can be used; word order shifts from actor-V-undergoer to undergoer-V-actor and the preposition *oleh* may precede the actor:

2. *Buku itu dibaca (oleh) Ali*
   book that *di*.read (by) Ali
   The book was read by Ali.

In this type of construction, the third person actor often appears as the bound pronoun *-nya* encliticized to the verb:

[1] I use the terms *actor* and *undergoer*, from Role and Reference Grammar (Foley and Van Valin 1984, Van Valin & LaPolla 1997) as shorthand for respectively the more agent-like and the more patient-like arguments of a transitive clause.

3. *Buku itu dibacanya*
book that *di*.read.3SG
The book was read by him.

For many speakers (although not all - see below), the construction exemplified in 2 is not possible with a first or second person actor. In such cases, another construction is used:

4. *Buku itu saya baca*
book that 1SG read
The book, I read.

In this construction, the actor can only be a pronoun (or an address term used as a pronoun substitute), and first and second person singular pronouns are often procliticized to the verb:

5. *Buku itu kubaca*
book that 1SG.read
The book, I read.

The clause types seen in examples 2-5 are all traditionally treated as passives. Finally, in spoken Indonesian it is common to hear sentences such as:

6. *Dia baca buku itu*
3SG read book that
S/he read the book.

Many grammars state that this construction is substandard (e.g. IRG: 'In colloquial usage *meN-* may frequently be dropped .... however ..... its occurrence is required in most natural contexts' [67-68]), or even ungrammatical (e.g. Winstedt 1913: 64), but it is used extensively[2] and it has syntactic properties which differ from the standard construction seen in example 1. In the remainder of this chapter I will use the following purely descriptive terminology: the type of clause exemplified in 1 will be referred to as a *meN*-V clause or construction, the various subtypes seen in examples 2 and 3 will be referred to as *di*-V-PP, *di*-V-DP and *di*-V-*nya* clauses (also generally as *di*-V clauses), both types seen in 4 and 5 will be referred to as Pro-V clauses, and the constructions exemplified in example 6 will be referred to as bare verb constructions. In addition, I will refer to *meN*-V and bare verb clauses collectively as actor subject clauses, and *di*-V and Pro-V clauses will be referred to collectively as undergoer subject clauses.

While in general the possibility of appearing in the full range of these constructions is criterial for assigning a verb to the class of transitives, there are some exceptions and marginal cases which require comment. There are a number of common verbs which have an initial *m-*, including *makan* 'eat',

[2] E. Anderson (1983) reports a survey in which bare verb forms were used in 21% of clauses addressed to strangers, and 67% of clauses addressed to family members. These figures should not be taken too literally - it is impossible to decide in some cases whether a clause has a bare verb or is of the Pro-V type - but they do indicate that the clause type is common in informal speech.

minum 'drink', *minta* 'ask' and *mohon* 'request', which normally appear in actor subject clauses without a prefix. All of these verbs can take the *di-* prefix freely, and indeed can occur with a *meN-* prefix also, although the meaning is usually slightly different in such cases. Thus *makan* simply means 'eat', but E&S give 'consume, use up' as the gloss for *memakan*. Most scholars regard these forms as including a fossilized form of the actor subject prefix (Verhaar 1984a). There are also a number of verbs which regularly take two arguments, but only appear in one type of clause. Almost all of these verbs are in the domain of cognition, typical examples being *mau* 'want' and *ingin* 'wish'. This class of verbs is discussed in detail in chapter 3 of the present work.

Indonesian has a small number of verbs which take three term arguments as a lexical property of the verb root, such as *beri* 'give', *kasi(h)* 'give' and *ajar* 'teach', and these can occur in all four clause types. There are also many two argument verbs which become three argument verbs when suffixed with *-i* or *-kan*[3]. Typical examples of these derivations are:

7. *Ayah mengirim uang kepada saya*
father *meN*.send money to 1SG
'Father sent money to me.'

8. *Ayah mengirimi saya uang*
father *meN*.send.APPL 1SG money
'Father sent me money.'

9. *Pelayan mengambil segelas air*
waiter *meN*.fetch one.glass water
'The waiter fetched a glass of water.'

10. *Pelayan mengambilkan tamu segelas air*
waiter *meN*.fetch.APPL guest one.glass water
'The waiter fetched the guest a glass of water.'

As far as I am aware, it is an exceptionless generalization that all verbs derived with the suffixes *-i* and *-kan* can occur in all four of the clause types discussed above. I will therefore treat ditransitive verbs as a subclass of transitive verbs (as defined here).

Finally, there is a number of intransitive verbs which are ill-formed without the prefix *meN-*. In some cases, the prefix attaches to nouns and adjectives[4] and can be regarded as derivational. Examples are *tepi* 'side' → *menepi* 'move to the side' and *merah* 'red' → *memerah* 'redden, become red'. There is also a group of intransitive verbal roots which require the prefix, such as *menangis* 'cry' and *melompat* 'jump'. In all the cases just discussed, the prefixed verb takes only a single argument and the undergoer subject clause types are therefore not possible. As the prefix appears in clauses in which the

[3] There are also many two argument verbs which include these suffixes in their form.

[4] See chapter 3 for discussion of the question of whether adjective is a lexical category in Indonesian. Here, the term is used pre-theoretically.

actor is subject, the most natural hypothesis would be that intransitive verbs with the *meN-* prefix are unergatives, verbs with only an actor argument. However, the correlation is not perfect. Kana (1986: 55) states that 'most of the prefixed verbs are unergative, while some of the prefixed verbs are unaccusative.' A crucial test for Kana is the effect the suffix *-kan* has when added to an intransitive verb[5]. The suffixed verb has an additional argument which is an actor in some cases and an undergoer in others. Kana (1986: 55-68) takes the first class to be unaccusatives and the second to be unergatives. Intransitive verbs with the *meN-* prefix occur in both classes:

11. *Anak itu menangis*
child that *meN*.cry
'The child is crying.'

12. *Anak itu menangiskan kucing yang hilang*
child that *meN*.cry.APPL cat REL lost
'The child cried over the lost cat.' (Kana 1986: 61, ex39a)

13. *Penyakit malaria menular di Irian Jaya*
illness malaria *meN*.spread LOC Irian Jaya
'Malaria is spreading in Irian Jaya.' (Kana 1986: 65, ex61b)

14. *Nyamuk menularkan penyakit malaria*
mosquito *meN*.spread.CAUS illness malaria
'Mosquitoes spread malaria.' (Kana 1986: 65, ex61a)

I leave open the question of whether any or all of these intransitive possibilities can be or ought to be included in the account of the prefix *meN-* developed below.

The availability of so many constructions for transitive verbs raises a number of issues. Firstly, the linking of arguments to grammatical functions is not straightforward. There are two clause types with the actor argument linked to the subject grammatical function, and two with the undergoer linked to subject. At first sight, this suggests redundancy which should be explained. Also, what is apparently the same verb form, the bare verb, can occur in two types of clause, those seen in examples 4 and 6, with opposite linking[6]. This also requires explanation. Secondly, the constituency pattern of the Pro-V construction is different to that of all the others. This clause type has both arguments of the transitive verb before the verb[7] unlike the other three clause types. This variation must be accounted for within a constrained theory of phrase structure. Finally, the constraints on the nature of the argument linked

[5] For further discussion, see section 4.3.1.2.

[6] Both of these examples can be exemplars of both of these clause types, although there is a clear difference in intonation when either the subject is postposed in example 4, or the non-subject is topicalized in example 6 (Chung 1978a).

[7] This discussion assumes neutral intonation. The subject of a Pro-V clause can follow the verb with distinctive intonation (Chung 1978a). For discussion of topicalization from *meN*-V and *di*-V clauses, see below.

to subject for the undergoer subject constructions must be explained. The fact that pronouns are treated differently from other noun phrases suggests an explanation in terms of the animacy hierarchy. There is some evidence for animacy-based splits in the treatment of actors in Western Austronesian languages (Musgrave 2000), but the split in Indonesian will be shown not to coincide with the split between the two types of undergoer subject clause, and I will argue that the possible types of actors in *di*-V and Pro-V constructions are conditioned by morphological properties of the language. Before exploring these questions, it is first necessary to make it clear that the prefix *di-* is part of the same paradigm as the prefix *meN-*, that is, a verbal prefix, and that it is not a proclitic pronoun, as has sometimes been argued (Cartier 1979, Guilfoyle, Hung and Travis 1992, Shibatani 1985).

2.2 The Status of *di-*

2.2.1 *di-* as pronoun

Pronominal clitics are possible in the Pro-V construction for first and second person singular actors, see example 5 above. The first and second person proclitics are *ku-* and *kau-* respectively. These are transparent reductions of the free pronouns *aku* and *engkau*. The normal third person singular free pronoun is *dia*, therefore there is some plausibility to the suggestion that what appears to be a prefix, *di-*, should instead be considered to be a proclitic. This line of argument has the advantage of reducing the two undergoer subject constructions to one: all undergoer subject clauses are Pro-V clauses on this account. A recent presentation of this view is that of Guilfoyle, Hung and Travis (1992). I summarise their position below, and then turn to the arguments against the analysis of *di-* as a proclitic. Firstly, various problems arise in assigning a phrase-structure position to *di-* if it is a pronoun, and secondly, how semantic roles are associated with syntactic arguments is problematic. In addition, the historical evidence does not support the view that *di-* derives from *dia*.

2.2.1.1 THE ARGUMENTS OF GUILFOYLE, HUNG AND TRAVIS

In an influential paper, Schachter (1976) reported that Tagalog (Austronesian, Philippines) does not always assign to a single DP all the properties normally associated with subjects. Philippine-type languages allow a variety of non-actor subjects, and in such constructions the syntactic subject does not have all the properties that might be expected. For example, the subject is the only DP available for relativization in such clauses, but it is not the only possible antecedent for reflexives. Indeed, the subject can itself be a reflexive anteceded by the actor. Two DPs in a single clause have subject properties, and as Schachter points out, it is not obvious which should be called subject.

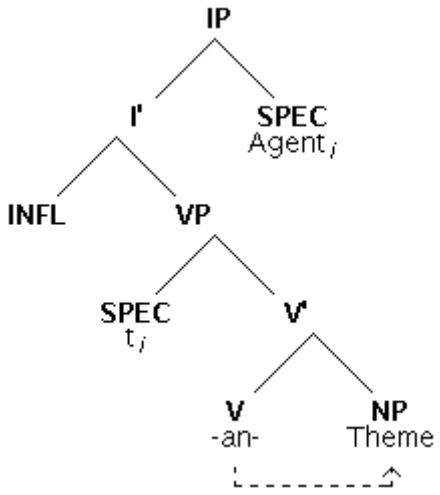
Guilfoyle, Hung and Travis (1992, afterwards GHT) is an attempt to capture the split in subject properties noted by Schachter in a configurational model (Principles & Parameters Theory). The crucial assumption in this account is that the agent (or more exactly, the more agentive argument of a transitive verb, what I am calling actor) is base-generated in Spec of VP and

that this position can remain occupied even if some other argument moves to Spec of IP, the canonical subject position, at s-structure. Two structural subject positions exist, and can be filled simultaneously and this explains the fact that subject properties can be shared between two DPs. The s-structure position of DPs depends, as is usual in such analyses, on the case-marking properties of the verb and its associated functional heads. The analysis is originally motivated for Malagasy and Tagalog, where all verbs are affixed, the affixes varying with the semantic role of the subject DP. GHT claim that each affix case marks a DP in a different d-structure position[8], and that whichever DP is left without case-marking moves to Spec of IP in order to satisfy the case filter (Chomsky 1981). Thus the Malagasy infix *-an-*, which identifies the actor as the subject of the clause, is analysed by GHT as an assigner of case to the undergoer. This leaves the actor with no case-marking, and it must move to Spec of IP[9]:

[8] Semantic roles and d-structure positions are assumed to be related in a straightforward way, along the lines of Baker's (1988) Uniformity of Theta Assignment Hypothesis.

[9] I do not deal with the GHT's account of case-marking in my main discussion, but it is worth noting some problems with this aspect of the analysis. Firstly, it seems to claim that predicates in the languages discussed have no intrinsic case-assigning properties, a claim which is contrary to the assumptions of most grammatical theories (but see Foley 1998 for a related view). The claim also encounters empirical problems: for example Tagalog recent perfective verbs have no subject-identifying morphology, but they nevertheless license arguments, and the surface coding of those arguments is the same as that given to non-subject core arguments of normal verbs (Kroeger 1993: 50). Secondly, the analysis involves additional assumptions that GHT do not make clear: that the Extended Projection Principle holds and that therefore *something* has to move to Spec of IP in any clause; and that there is some 'last-resort' case-assignment process that deals with obliques after subject selection. This second assumption is necessary because, in both Tagalog and Malagasy, NPs which would otherwise appear as obliques can become subject. If these NPs had inherent case, assigned by the oblique determiner *sa* in Tagalog for example, then no motivation for the movement to Spec of IP would exist. But any such rescue mechanism seems an *ad hoc* manoeuvre, particularly in Indonesian where obliques are introduced by prepositions with semantic value. Thirdly, GHT equate the Austronesian clause types which do not have an actor as subject with the passive of languages such as English. But the morphology which characterises these clauses assigns case in the GHT account, whereas the morphology which characterises passive in current accounts (e.g. Jaeggli 1986) is analysed as *absorbing* a theta-role and therefore preventing case from being assigned. Finally, verbal morphology of the Philippine type has traditionally been viewed as directly reflecting the NP selected as subject (the title of Blake (1906) is "Expression of case by the verb in Tagalog"). The GHT analysis is counter-intuitive in that the verbal morphology is analysed as having a case relation with some NP, but never with the subject.

15.

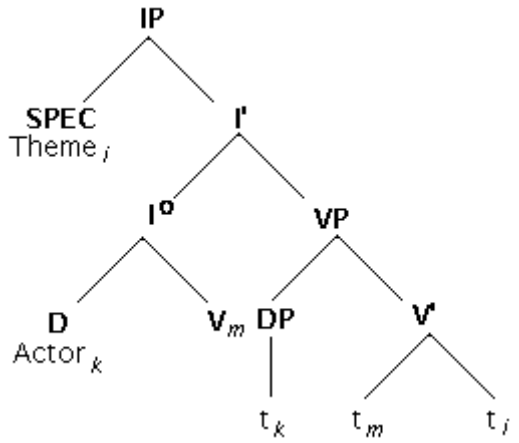


(GHT: ex5a, dotted arrow indicates case-marking)

In languages such as Tagalog and Malagasy, the claim that different affixes case mark DPs in different d-structure positions forcing movement to Spec of IP for other DPs is reasonably straightforward as all verbs (see footnote 9 for an exception) have affixes. However, in the case of Indonesian (and Bahasa Melayu, they are treated together by GHT), the analysis is made more complex by the existence of unaffixed verbs.

The analysis for *meN-* verbs is straightforward: the prefix enables the verb to case mark the undergoer argument and the actor argument must move from Spec of VP to Spec of IP to receive case there, just as in the tree above, example 15, *modulo* the ordering under IP. GHT claim that a unified analysis is possible of the undergoer subject constructions, treating them all as passives in which the actor argument is generated as D° in the position Spec of VP, and is licensed (case-marked) by being adjoined to V°, which has moved to I:

16.



When the actor is first or second person, D° is realized as either the full pronoun, or as a proclitic. When the actor is third person, D° is realized as either one of the pronouns *dia* or *mereka*, or as the proclitic *di-* which can be co-indexed with a post-verbal NP, the complement of the actor pronoun determiner. This DP remains in its original position and the surface order is a result of the movement of the verb to I°. In all cases, the undergoer argument does not receive case in its d-structure position as sister of V°, and must move

to Spec of IP to pass the case filter. On GHT's account, the proclitic *di-* is a realization of the features for third person (see discussion of what this might mean below), and this accounts for the restrictions on the use of non-third person agents with *di-* verbs (cf. IRG:249-250 especially footnote 2). GHT note that liberal speakers allow *di-* verbs with a non-third person agent introduced by *oleh* in clauses similar to example 2 above. Their theory predicts that such speakers are no longer using *di-* as a spellout of third person features. The morpheme is not a pronoun in this case, and a postverbal actor cannot therefore be licensed by being co-indexed with the pronoun *di-* adjoined to the verb in I. The prediction then is that such speakers will not allow any bare DP agents with *di-* verbs. GHT provide no data on this point, and the prediction is hard to test without examining a substantial corpus of spoken Indonesian as speakers reject *di*-V clauses with post-verbal actors such as *oleh saya* and such constructions are avoided in writing. In the absence of such data, I would however doubt the accuracy of the prediction for the reason that a very common use of *di*-V clauses is with a third person enclitic actor (see example 3), and it would be surprising to discover that GHT's liberal speakers did not use such constructions.

GHT's analyses of both the clause they term active (*meN*-V clauses) and those they term passive (*di*-V and Pro-V clauses) are thrown into doubt by the existence of bare verb constructions, which are not considered by GHT. Clauses such as example 6 above show that Indonesian verbs can license (case-mark) an undergoer in the absence of the prefix *meN-*. This fact goes against the GHT account, where this prefix case-marks the undergoer *in situ*. If the bare verb can case-mark its internal argument, why is the prefix needed at all? A similar line of argument can be advanced against the GHT analysis of Pro-V clauses: if a bare verb can case-mark the undergoer inside VP, why does it have to move to Spec of IP when the actor is a proclitic?

2.2.1.2 THE PHRASE-STRUCTURE POSITION OF *DI-*

There are two possible positions as to the status of the reduced pronouns *ku-* and *kau-*: either they are true clitics, occupying their own phrase structure positions and phonologically dependent on the verb to their right; or they are attached to the verbs in the lexicon, in the fashion argued for French object clitics by Miller (1992). No issue arises for the type of Pro-V clause in which the actor appears as a full pronoun because in such cases pronoun substitutes are freely available and these do not constitute a closed class. Indonesian speakers command multiple exponents of first and second person pronouns, but these can also be freely substituted with kin terms, address terms and proper names in certain contexts[10]. In the following text example, the speaker, who is (or believes himself to be) the hearer's uncle, uses the kin term *oom* 'uncle' as a pronoun substitute:

[10] GHT (p398) state that the actor of a Pro-V clause 'is restricted to being a pronoun, clitic or proper name'. This is not quite correct: proper names can only be used in this type of clause as pronoun substitutes referring to the speaker or the hearer. A proper name with 3rd person reference is not possible.

17. *Oom tidak mau merusak wanita yang Oom*
uncle NEG want *meN*.damage woman REL uncle
cintai sekaligus Oom puja
love.APPL as.well.as uncle worship
'I do not want to hurt the woman I love and worship.'
(SDM: 113)

Where a proper name is used as a pronoun substitute, this can consist of more than one phonological word, a two part name for example, or a title and a name:

18. *Buku itu sudah Prof Jones baca, bukan?*
book that PERF Prof Jones read NEG
'Have you read that book?' (addressed to Prof. Jones)

Such possibilities seem to rule out an analysis in which all pronouns in the Pro-V construction are treated as phonologically dependent on the following verb, with the apparent cliticization of the reduced forms only an orthographic convention.

The forms which are orthographically combined with the verb fail most of the standard tests for lexical attachment. They do not undergo or trigger morphophonological processes[11], there are no arbitrary gaps in their distribution and there are no suppletive forms. The sole criterion that might suggest a lexical origin is that co-ordination of following verbs is noticeably worse for the reduced forms than for the free forms:

19. *?Buku itu saya beli dan baca*
book that 1SG buy and read

20. **Buku itu kubeli dan baca*
book that 1SG .buy and read

21. *Buku itu kubeli dan kubaca*
book that 1SG .buy and 1SG.read
'I bought and read the book.'

However, it is not clear that co-ordination is a reliable test for the status of clitics; even Miller's French data show inconsistencies (see Sadler 1998 for discussion). Therefore, I will assume that it is preferable to treat all actors in the Pro-V clause type in a unified fashion as clitics having their own distinct phrase structure positions[12].

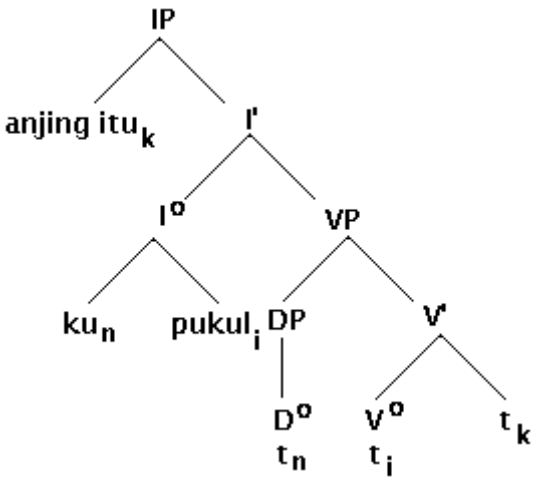
GHT assume that all actors in Indonesian originate as specifiers of VP. This position is not case-marked and therefore the actor always has to move to some other position to satisfy the case filter. Where the verb is prefixed

[11] This test is not crucial. No Indonesian prefix with a final vowel shows any morphophonemic variation.

[12] Some of the arguments developed regarding the position which *di-* then is considered to occupy do not apply if this assumption is incorrect. However, the other arguments developed in following sections do still apply.

with *meN-*, which case-marks the undergoer, the actor moves to the specifier of IP, the canonical subject position. Where there is no *meN-*, some other move is required. *di-* is a pronoun, therefore all undergoer subject clauses have a pronoun actor. Following Postal (1969), GHT treat pronouns as determiners, and the actor pronoun is therefore originally the head of the DP which fills the specifier position in VP. This head moves to I° and by joining the verb there it satisfies the case-marking requirement, giving the surface order Pronoun - Verb:

22.



(GHT example 34)

If there is a post-verbal actor, then it is the NP complement of the pronoun D° which remains in its original position.

The claim that the pronoun-verb combination finishes in the position of I° is not straightforward however. GHT do not give any examples in which anything that might be an overt I appears, nor do they discuss this possibility, but functional elements can appear between the subject and the verb in Indonesian. When this occurs, the position to be occupied by the pronoun-verb combination is problematic. Three types of functional element can appear between the subject and the verb in Indonesian: negation, temporal/aspectual markers and modals. Leaving aside the question of what category any of these should be assigned to, it is a fact that in Pro-V clauses, all of these precede the pronoun actor:

23.

Buku	itu	tidak	akan	bisa	saya	baca
book	that	NEG	FUT	possible	1SG	read

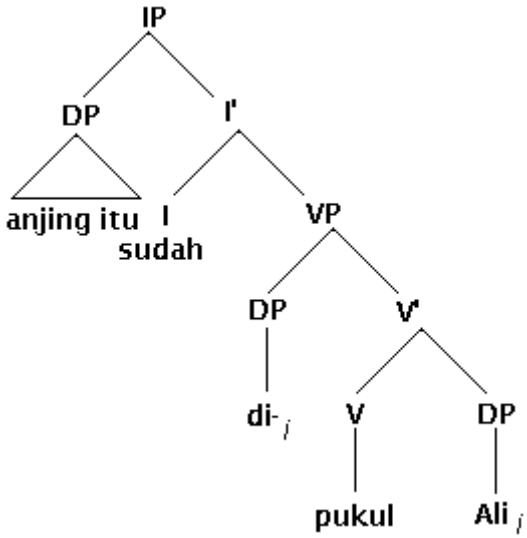
'I will not be able to read the book.'

All these elements might be accommodated by assuming that IP is not a single category, but rather can be analysed as a series of functional heads (Tense, Negation etc,) and their projections (Pollock 1989). But to account for an example such as 23, GHT are committed to a claim that there is some additional functional head above VP in Indonesian which is never overt, and which is required for their analysis to succeed. It is hard to think of plausible candidates for such a head. Some agreement head (Agr) might appear to be a possibility, but on closer inspection the idea fails. The necessary head must be the lowest head in the I complex, and this position is not compatible with AgrS. AgrO could occupy this position, but GHT explicitly characterise this

construction as a passive, which presumably means that the verb is intransitive and therefore AgrO cannot be present[13]. Also it cannot be the case, on GHT's arguments, that in a case such as example 23 the verb remains in its d-structure position. If this were so, the relation between the actor pronoun and the verb would still be the d-structure configuration assumed by GHT, a configuration in which they argue case-marking requirements cannot be met. The position of the preverbal pronouns, including *di-*, is therefore not adequately accounted for by GHT.

The position of the pronoun actor is a problem also for a theory which does not permit movement such as LFG. The problematic cases for such a theory, however, come in clauses without any functional projections. As the tree in 24 shows, where there is some functional head above VP, two specifier positions are available which can be occupied by the subject nominal and the actor pronoun:

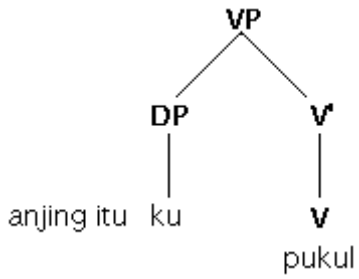
24.

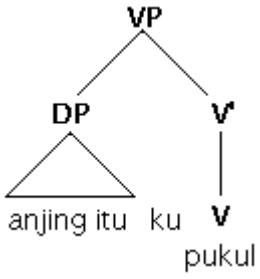


anjing	*itu*	*sudah*	*dipukul*	*Ali*
dog	that	PERF	*di*.hit	Ali

In the absence of such a head, it is by no means obvious where the actor pronoun should be placed. The structure can be taken to be a VP only (see trees in 25 below), in which case only a single specifier position is available for one of the arguments. Whichever nominal is analysed as occupying the available specifier position, the other is left without a home:

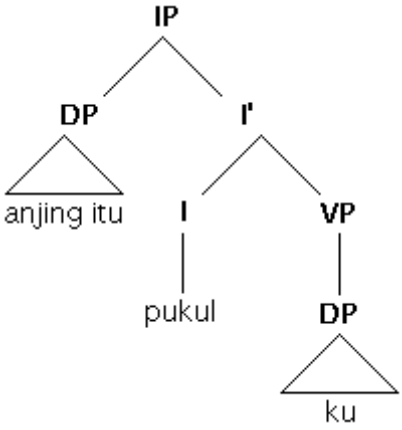
[13] GHT argue that passivization has no effect on theta assignment, but that it does prevent the verb from assigning accusative case. My argument assumes that this position rules out the possibility of AgrO occurring.

25. a 

b 

anjing *itu* *kupukul*
dog that 1SG.hit

Another possibility is that the tree in example 25b is correct as regards the position of the subject and the verb, and that the actor pronoun is a complement of the verb in a left-branching structure. However, Indonesian has heads robustly preceding complements for all major categories and this analysis is therefore implausible. LFG also allows lexical heads to appear in the position of functional heads, leaving the projection of the lexical head headless in the phrase-structure tree (although not in the functional structure) (Bresnan 2001a: Chapter 7 and section 1.3.3). Such an analysis would allow the verb to take the pronoun actor as a complement, but would produce the wrong linear order:

26. 

In the case where there is no functional head above VP in the clause, a theory such as LFG can offer no account of the phrase structure of Pro-V clauses, if we assume that the pronoun actor occupies an independent node and we exclude the possibility of a null head and the possibility of two specifiers in a single projection.

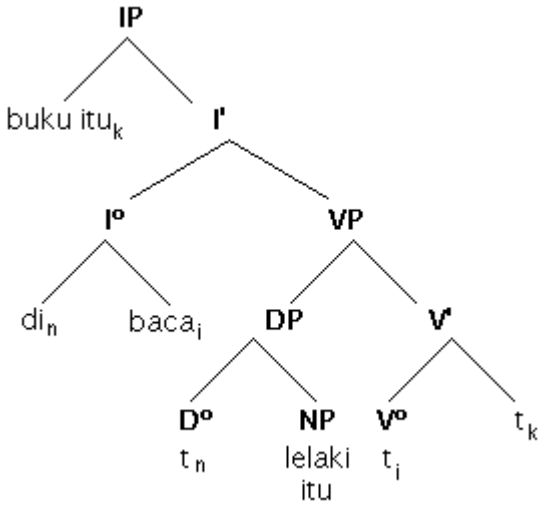
It should be noted that the argument made in this last paragraph is true for Pro-V clauses in general, whether or not *di-* is a pronoun. If *di-* is claimed to be a pronoun, then this is an argument against that position. If *di-* is not a pronoun (as I claim), then this is a more general problem for the analysis of Pro-V clauses, one to which I return below. There I will argue that the insight of GHT, that the pronoun actor and the verb are in a close relation, is accurate, but that the nature of the relationship is morphological rather than syntactic.

2.2.1.3 ACTOR DOUBLING

As discussed in the previous section, GHT's analysis assumes that an actor occurring immediately after the verb in a *di*-V clause is the complement of the D head *di-* which has undergone head movement. This claim raises various problems which are discussed in this section. The discussion concentrates on the case in which the postverbal actor is a pronoun, as these are the cases in which the problems are most evident. However, I take it that similar problems would arise in the case of DP postverbal actors, if the analysis were pursued in sufficient detail.

GHT assume that pronouns are generated as D°, and that the complement of D is NP:

27.



buku *itu* *dibaca* *lelaki* *itu*
book that *di*.read man that
'The book was read by the man.' (GHT example 36)

These assumptions lead to the prediction that a postverbal actor cannot be a pronoun. The pronoun is a D, and projects a DP, while the D head which has moved to become *di-* before the verb requires a DP complement. This prediction is not true: the pronouns *-nya* (see example 3) and *mereka* are both possible as postverbal actors, as are proper names, which GHT also treat as Ds[14]:

28. Buku itu dibaca mereka
book that *di*.read 3PL
'The book was read by them.'

29. Buku itu dibaca Ali
book that *di*.read Ali
'The book was read by Ali.'

One solution to this problem might be to argue that D as a functional projection can take a range of categories as its complement (following the

[14] GHT acknowledge the possibility of proper names as post-verbal actors (n.30) and suggest that they can be optionally analysed as NPs, but they ignore the case of pronouns.

reasoning of Abney 1987), and that one possible complement is another DP. This would allow two D heads to be generated in one maximal projection, but would raise another problem, that of how two referential expressions can be generated in a single DP. If the reference of the two is identical, the structure would be redundant and if the reference of the two is not identical, the semantics must be incoherent. GHT's analysis is already open to this criticism in that they go against the spirit of Postal's work on which they rely. A central point of Postal's (1969) paper is that the noun projection associated with a pronoun determiner can never appear precisely because the determiner is referential. The same criticism applies even more strongly if two Ds are posited.

Another possible solution might be to claim that the pronoun *di-* is the true argument and that any postverbal actor is an adjunct linked to the pronoun only by coreference, that is to adapt the analysis of non-configurational languages proposed by Jelinek (1984). This solution also can be rejected. Firstly, Jelinek's proposal is intended to explain a cluster of properties which co-occur in the languages referred to as non-configurational, three of which are crucial: free word order, discontinuous constituents and null anaphora. Indonesian does not have free word order, nor does it allow discontinuous constituents. It is more tolerant of omitted arguments than English, but Jelinek (1984) argues with respect to Chinese that this phenomenon can occur in configurational languages and should be accounted for outside of sentence grammar. There is therefore no reason to assume that Indonesian is a systematically non-configurational language. Secondly, there is no reason to assume that it has non-configurational characteristics in just the one clause type. Two predictions which would be made by the proposed analysis fail, and one of these also undermines a strength of the GHT analysis as it stands. If *di-* is a true pronominal argument which can be optionally doubled by an adjunct, then it would be predicted that a clause with no other exponent of actor could have a reading with a specific third person actor. This is not true; such clauses have a reading similar to that of agentless passive clauses in English[15]:

30. Barang ini akan dikirim ke Jepang
goods this FUT *di*.send to Japan
'These goods will be sent to Japan.'
NOT: 'These goods will be sent to Japan by him/her/them.'

Secondly, Jelinek's analysis makes it a characteristic of adjuncts doubling pronominal arguments that they have no fixed position in the clause (word order is free). If the postverbal actor were such an adjunct, it would be predicted that it could appear in any position. This is not true. The postverbal actor must be immediately adjacent to the verb, unless it is introduced by the preposition *oleh*. This fact is accounted for in GHT's analysis, but modifying

[15] In sequences of action clauses with the same actor, it is common in Indonesian discourse for the non-initial clauses to be *di*-V clauses without an actor. In such cases, the actor is specific but I assume that such possibilities are accounted for outside of sentence grammar as previously discussed.

their account in the fashion discussed here would lose this advantage as the postverbal actor would no longer be generated in Spec of VP, an A-position. I discuss the possibility of a similar approach in the LFG framework below, and reject it for similar reasons.

The problem of actor doubling is equally awkward in other frameworks. For example, an analysis in LFG would be unable to assign a well-formed functional structure to a clause with a post-verbal actor. Whatever grammatical function is assigned to the pronoun *di-* should also be assigned to the post-verbal actor for the interpretation to come out right. But LFG has a strict condition which bars a single attribute, such as a grammatical function, having two values. No coherent f-structure is possible as demonstrated for parallel facts from Irish in a detailed analysis by Andrews (1990). One possible solution in LFG would be to claim that the pronoun carries the grammatical function, and that the post-verbal actor is licensed by being the exponent of some discourse function (Bresnan and Mchombo 1987). However, the behaviour of the post-verbal actor is not at all what might be expected on this analysis: occurring in non-typical linear position, especially at the edge of a clause, is characteristic of discourse functional DPs, but a postverbal actor must always be adjacent to the verb. It can be clause-final, but this is accidental: it happens only when there is not other material in the clause to follow. It would also be unexpected for an argument assigned a discourse function to be expressed as a reduced pronoun.

The reduced pronouns *ku-* and *kau-* behave exactly as might be expected given these considerations, no other exponent of the actor role can occur in the clause. But if *di-* is really a reduced form of the pronoun, parallel to *ku-* and *kau-*, then we would predict that it should also be possible for these pronouns to be doubled after the verb. All such possibilities are ungrammatical:

31. **Buku itu kubacaku*
book that 1SG.read.1SG

32. **Buku itu kubaca saya*
book that 1SG.read 1SG

33. **Buku itu kubaca oleh saya*
book that 1SG.read by 1SG
(FOR: 'I read the book.')

This might be accounted for on the grounds that the 1st and 2nd person pronouns are fully referential in context, whereas a 3rd person pronoun may need additional specification. However, this argument cannot be maintained because for all speakers *-nya* is a possible post-verbal actor but it is not more specific than the pronoun *dia*, the supposed source of *di-*. Canonically, *dia* has only human (or at least animate reference); some speakers allow it to have inanimate reference also, but in all such cases its reference is singular. For *-nya* on the other hand, most speakers allow inanimate reference and many allow plural reference. Therefore, if *di-* is a reduced form of *dia*, post-verbal *-nya* does not make the reference more specific. The fact that *ku-* and *kau-* do not allow a post-verbal actor must be attributed to some other reason. This

asymmetry between *ku-* and *kau-* and *di-* cannot be explained if *di-* is taken to be a true pronoun. It can be explained if *di-* has less referential content than the true pronouns, and I discuss this possibility below.

Although examples 4 and 5 give two structures for the Pro-V clause type, and although the reduced pronoun is conventionally written as one word with the verb (these forms were often referred to formerly as 'conjugated verb forms', see for example Gonda 1949), there is little evidence of any syntactic difference between the two possibilities (apart from the co-ordination data discussed above). Therefore, we would also predict that *dia* as actor in a Pro-V clause could be doubled with a post-verbal actor. As pointed out by Arka and Manning (to appear), this prediction is also not correct (I omit the possibility of doubling with *-nya* which is ruled out below on other grounds):

34. **Orang itu dia lihat Ali*
man that 3SG see Ali

35. **Orang itu dia lihat oleh Ali*
man that 3SG see by Ali
(FOR: 'Ali saw that person.')

Thus in addition to the theoretical arguments for not analyzing *di-* as part of the same paradigm as *ku-* and *kau-*, there is direct morphosyntactic evidence to support this position. There is also evidence that *di-* has significantly similar properties to the verb prefix *meN-*, and this is discussed in the following section.

One response to the evidence discussed in this section might be to claim that *di-* does indeed originate from *dia* and was at some stage a pronoun in its own right, but that it is now an agreement marker, that is, its referential content has become reduced. The split in actor possibilities between the two undergoer subject clause types would still be accounted for, and the arguments made above would be irrelevant. However, this possibility can also be eliminated. As mentioned previously, 1st and 2nd person actors do occur in *di-*V clauses on occasion, typically in prepositional phrases headed by *oleh*[16]. This already casts doubt on the interpretation of *di-* as an agreement marker, but further evidence renders it untenable. If *di-* were an agreement marker, we would predict that its effects would be most obvious in *di-*V clauses without an overt actor. But it is precisely in such cases that it is possible to have an implied 1st or 2nd person actor (Kaswanti Purwo 1988)[17]:

36. *Kau ini hanya cewek yang diurus*
2SG this only girl REL *di*.take.care.of
'As for you here, it's only girls that you have on your mind'
(LIT: 'You here, it's only girls that are taken care of.')
(Hilman & Boim 1992: 80)

[16] Another possibility is discussed in section 2.4.3.2.

[17] Jelinek (1984) discusses examples of this type in non-configurational and pro-drop languages. Her discussion is not relevant to Indonesian, for the reasons previously given.

Therefore, the weaker hypothesis that *di-* is synchronically an agreement marker also fails.

2.2.1.4 SEMANTIC ROLE ASSIGNMENT

In general, modern syntax assumes a bi-unique mapping between semantic roles and syntactic arguments. Where arguments have been advanced that this relationship should be loosened, this has typically been in the direction of allowing an argument to have more than one semantic role (e.g. Jackendoff 1987 and elsewhere). An analysis which claims that *di-* is a pronoun also claims that the opposite type of relationship is true: that more than one argument can have the same semantic role. Intuitively, this possibility seems more pernicious, and major syntactic theories have made greater efforts to exclude it.

GHT do not discuss this issue, so I am forced to speculate how they might deal with it. The theta criterion must be satisfied at d-structure in the framework GHT adopt. This seems straightforward; the DP which is in the specifier of VP at d-structure is assigned the theta role Agent. But what happens at s-structure? It is perfectly possible for no post-verbal actor to appear in a *di*-V clause:

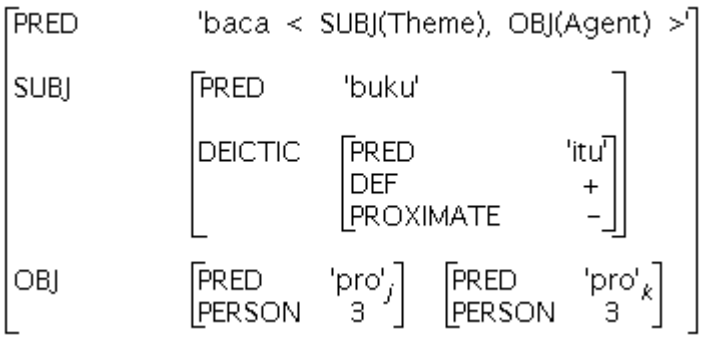
37. *Barang ini akan dikirim ke Jepang*
goods this FUT *di*.send to Japan
'These goods will be sent to Japan.'

This would suggest that the pronoun *di-* carries the theta role, and for a consistent analysis we would expect this to be the case even when there is a post-verbal actor. The fact that the pronoun is the head of the category to which the theta role was originally assigned might be taken as an additional argument for this position. But the post-verbal actor certainly contributes to the interpretation of the clause: what then is the relationship between it and the preverbal pronoun? Certainly the classic Principles and Parameters answer to this sort of problem is not available to GHT. The two positions which apparently are assigned the same theta role do not form a chain, the categories are not identical and the binding relations are not admissible. The argument applies *a fortiori* if a chain including the trace of D is considered. The question posed above is a reflex of the more general question raised above: how is it possible for two referential expressions to be generated under one DP node?

The problem is equally awkward in other frameworks. For example, an analysis in LFG would be unable to assign a well-formed functional structure to a clause with a post-verbal actor. Whatever grammatical function is assigned to the pronoun *di-* should also be assigned to the post-verbal actor for the interpretation to come out right. But LFG has a strict condition which bars a single attribute, such as a grammatical function, having two values. No coherent f-structure is possible if *di-* is a pronoun, that is if it has a PRED value. This feature in LFG is part of the lexical entry of any item that has a semantic representation associated with it. For example, the clause in

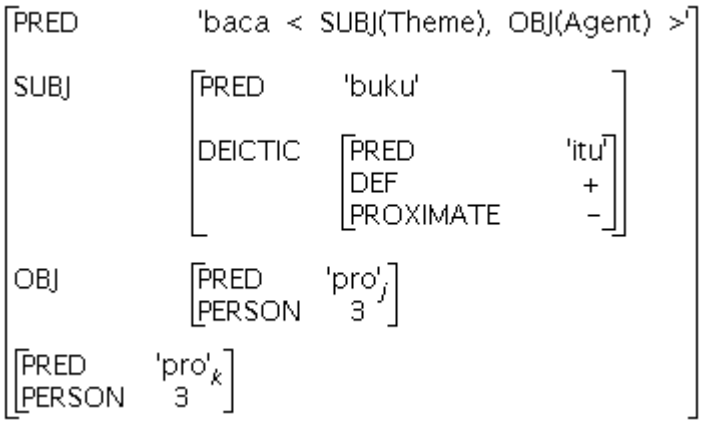
example 3, *Buku itu dibacanya,* would be associated with an f-structure along the lines of 38, if we assume that *di-* is a pronoun[18]:

38.



The object grammatical function has two values here, which is not permitted (Kaplan and Bresnan 1982: 181). The alternative is that one of the pronouns is not the value of any attribute in the f-structure:

39.



This f-structure is equally impossible: if the pronoun is not assigned a function in some f-structure, it cannot be semantically interpreted.

2.2.1.5 HISTORICAL EVIDENCE

The question of the historical origin of *di-* and the question of its synchronic analysis are logically independent. The prefix could have originated as a reduced form of the pronoun *dia* and have undergone reanalysis to a verb prefix at some point; equally it could have some other origin and be currently analysed by speakers as a pronoun (although the evidence presented here shows this to be highly improbable). However, as the proponents of the position that *di-* is a pronoun typically use the similarity in shape of the prefix and the free pronoun as an argument without adducing historical evidence (Kana 1986 is an honourable exception), it is appropriate to briefly consider the available material.

---

[18] This example uses *-nya* as post-verbal agent because it is unequivocally a term argument - see discussion below, section 2.3.1.3. I later argue that in such cases, *-nya* is morphologically attached to the verb. The impossibility of the analysis is only more apparent in this case: the verb form includes a pronoun in its morphological representation and the function cannot take any other value (cf. a closely parallel analysis of Irish subject morphology, Andrews 1990).

The historical literature offers two sources besides *dia* for *di-*: the prefix *ni-* which was a passive or object-oriented marker in Old Malay, and the preposition *di* used as an agent marker. The first position has been argued by earlier linguists as reflected in De Casparis (1956) and by Hopper (1979). However, the Old Malay prefix *ni-* was compatible with first and second person actors as well as third persons. Therefore, if this is the source of *di-*, some account of how its usage came to be restricted must be provided, and this has not yet been done as far as I am aware. In addition, this proposal has to account for the shift from a nasal to an oral consonant. This can be done in the case of the Old Malay prefix *mar-*, which has as its contemporary reflex *ber-*, but the arguments which apply in that case do not apply to *ni-* (Adelaar 1992b, p.c.). The second source is better supported in the literature. In at least one Malayic dialect, Salako, it is still obvious that *di* is a pre-nominal marker of agency which can also precede the verb, and a similar phenomenon is attested in Minangkabau (Adelaar 1992a, b).

Another argument against deriving *di-* from the free pronoun *dia* is offered by Wolff (cited by Kana 1986:102) who suggests that it entered Indonesian from Javanese[19].The possible reasoning behind this suggestion leads to another issue. There are actually two 3rd person singular pronouns in Indonesian, *dia* and *ia*. The second of these only occurs in writing and is restricted to preverbal positions (simplifying slightly: see discussion in section 1.2.2), typically subject of a clause or agent in a Pro-V clause. The existence of both these forms is the remnant of an older system in which the personal pronouns had special oblique forms: *daku* for 1st person, *dikau* for 2nd person with *dia* as the 3rd person oblique form. So in older forms of the language, the *ia* form would have been used as actor of Pro-V clauses, making the derivation of *di-* from a free pronoun less plausible. Javanese has a set of verb prefixes used in undergoer subject clauses, and the prefix indicating 1st person actor is *tak-* or *dak-*. The origin of Wolff's suggestion therefore would seem to be an assumption that the initial consonant was more widespread in Javanese than in Indonesian, and that it is therefore a more plausible source for the Indonesian form[20].

In his monograph on Proto-Malayic, Adelaar (1992a:162) does not use *di-* as the basis for any reconstruction, although in later work (Adelaar 1992b) he accepts a prepositional origin for the prefix and it is possible to say with reasonable confidence that a pronominal origin for *di-* is unlikely. To the extent that this conclusion is relevant to synchronic analysis then, historical evidence agrees with the arguments presented here.

2.2.1.6 SUMMARY

The preceding sections have shown that serious theoretical problems arise when *di-* is analysed as a pronoun, that there is direct morphosyntactic evidence that it behaves differently to the two reduced pronouns *ku-* and *kau-*, and that the historical evidence does not support the free pronoun *dia* as the

[19] This discussion is necessarily speculative: Kana gives only the bare claim made by Wolff.

[20] This point was made to me by Sander Adelaar (p.c.)

source of *di-*. There is, however, evidence that *di-* shares properties with the verb prefix *meN-*, and to I now turn to this.

2.2.2 *di-* as verb prefix

This section will examine the morphosyntactic evidence that *di-* is a verb prefix parallel to *meN-*. Much of the evidence hinges on properties of post-verbal arguments in clauses with prefixed verbs, and this introduces an additional complication. Recent work by Arka and Manning (to appear, afterwards A&M) has argued that the post-verbal actor in a *di-*V clause is only a term (a direct argument) when it takes the form of the attached pronoun *-nya*. Noun phrase post-verbal actors are non-terms. The evidence for this claim will be discussed in detail in section 2.3.1.3; the crucial fact is that *-nya* can antecede an anaphor in subject position of a *di-*V clause. But no other post-verbal actor can do this:

40. **Dirinya tidak diperhatikan Haris*
self.3 NEG *di*.look.after Haris

41. *Dirinya tidak diperhatikannya*
self.3 NEG *di*.look.after.3
'(S)he didn't take care of her-/himself.'

Accepting this data, and Arka and Manning's arguments, means also accepting that a direct comparison of the non-subject arguments of transitive verbs in *meN-*V and *di-*V clauses is not always legitimate. Therefore, where this is a consideration, I use the second object of a *di-*prefixed ditransitive verb as the comparison. Such arguments are terms on the basis of a separate test (quantifier float).

2.2.2.1 POSTVERBAL ATTACHED PRONOUNS

Attached pronouns occur only with verbs with a prefix, *meN-* or *di-*. This generalisation is independent of the status of the non-subject argument because, as mentioned above, I assume that the attached pronoun is still a direct argument following a *di-* prefixed verb. With a *meN-* verb, the full range of attached pronouns is possible, *-ku* for 1st person singular, *-mu* for 2nd person and *-nya* for 3rd person:

42. *Ali melihatku/-mu/-nya*
Ali *meN*.see.1SG/2/3
'Ali saw me/you/(s)he).'

43. *dia tidak dapat melakukannya lagi*
3SG NEG able *meN*.act.APPL.3SG further
'he couldn't do it again' (SDM:70)

As seen above (example 3) and in the following text example, the pronoun *-nya* can attach to a *di-* verb also:

44. *Tanah itu diwarisinya dari kakeknya*
land that *di*.heir.APPL.3SG from grandfather.3SG
'That land he inherited from his grandfather.' (SDM:245)

The other two attached pronouns do not occur, as *di*-V constructions are restricted to 3rd person actors.

In contrast, no bare verb form in Indonesian can have a pronoun attached to the right. In addition to the Pro-V clause type, there is also an actor subject bare verb clause type in Indonesian (introduced in example 6):

45. *Ali baca buku itu*
Ali read book that
'Ali read the book.'

This type of structure is common in spoken Indonesian and in writing which emulates colloquial speech, although it is not part of the prescriptive standard (see Voskuil 1996, Appendix to chapter 8 for discussion). The two types of clause with bare verb forms are distinguished by the positioning of temporal and modal auxiliaries which occur to the right of the actor in bare verb clauses and to its left in Pro-V clauses. An attached pronoun following the verb is impossible in both cases:

46. **Akan saya lihatnya*
FUT 1SG see.3

47. **Saya akan lihatnya*
1SG FUT see.3
(FOR: 'I will see him.')

Further evidence for this generalisation, that only prefixed verbs can host attached pronouns, comes from a class of quasi-verbal predicates. These are predicates which denote emotional and cognitive states and which do not take the prefixes *meN-* or *di-*. They normally take a single direct argument, the experiencer, with the stimulus appearing as a prepositional phrase. Some of these words also allow the possibility of the stimulus appearing as a bare DP, but it cannot be an attached pronoun unless a transitive verb is derived using an applicative suffix[21]:

[21] Pronouns can be attached to some prepositions, and therefore the stimulus can be an attached pronoun when it is within a prepositional phrase: *Saya sayang padanya*. Chapter 4 is an extensive discussion of this class of words.

48. *Saya* *sayang* *pada* *Siti*
1SG pity to Siti

49. *Saya* *sayang* *Siti*
1SG pity Siti

50. **Saya* *sayangnya*
1SG pity.3

51. *Saya* *menyayanginya*
1SG *meN*.pity.APPL.3
'I pity Siti / her.'

Therefore, *di-* patterns with *meN-* in this respect: they both license an attached pronoun. I argue below (section 2.5.2) that the relation between the host verb and the pronoun varies depending on the prefix. This does not affect the descriptive generalisation discussed here, but it does suggest that it is a matter of phonological realization, rather than a syntactic property.

2.2.2.2 EXTRACTION AND TOPICALIZATION

Both *meN-* and *di-* block extraction of a non-subject argument. This fact, combined with the fact that verbs with these prefixes are the most common in written Indonesian, has led to the received wisdom that Indonesian permits only subject extraction. This is incorrect; at least two types of clause allow extraction of the non-subject argument and crucially both of these have unprefixed verbs. The data in this section uses relative clauses throughout to exemplify the extraction possibilities. In the prescriptive standard, questions are formed as equational clauses with the question word as predicate and a headless relative clause as subject:

52. *Siapa* *yang* *kauundang?*
who REL 2.invite
'Who did you invite?' (LIT: Who is the one you invited?')

Therefore, the discussion of questions falls under the discussion of relative clauses for this register. There are also some wh- *in situ* possibilities which would needlessly complicate the discussion, and I ignore these here.

A non-subject argument can head a relative clause with an emotion predicate construction:

53. *orang* *yang* *saya* *suka* *itu*
person REL 1SG like that
'the person that I like'

Similarly, a non-subject argument can be extracted with the bare verb actor subject construction previously discussed. But extraction of the actor from a Pro-V clause is totally impossible, giving rise to the following paradigm[22]:

[22] Example 47 is possible only if the name *Badu* is used as a pronoun substitute. *Badu* cannot be a 3rd person in this case; see discussion in section 2.2.1.2.

54. (Bare verb, subject extraction)
Inilah Badu yang sudah baca buku itu
this.PRT Badu REL PERF read book that
'This is Badu who has read the book.'

55. (Bare verb, non-subject extraction)
Inilah buku yang Badu sudah baca
this.PRT book REL Badu PERF read

56. (Pro-V clause, subject extraction)
Inilah buku yang sudah Badu baca
this.PRT book REL PERF Badu read
'This is the book that Badu has read.'

57. (Pro-V clause, non-subject extraction)
**Inilah Badu yang buku itu sudah baca*
this.PRT Badu REL book that PERF read
(FOR: 'This is Badu who has read the book.')

These examples demonstrate that Indonesian, unlike say Tagalog, does not have a blanket ban on the extraction of non-subjects. Possible reasons for the impossibility of example 57 will be discussed below.

Verbs prefixed with *meN-* cannot have their non-subject argument extracted:

58. **Itulah buku yang saya membaca*
that.PRT book REL 1SG *meN*.read
(FOR: 'That is the book that I read.')

The same apparently applies to *di*-V clauses:

59. **Ali yang buku itu dibaca*
Ali REL book that *di*.read
(FOR: 'Ali who read the book')

But I have argued above that bare DP actors in *di*-V clauses are not direct arguments of the verb, and there is evidence that this type of extraction is limited to direct arguments in Indonesian[23]. Therefore, the comparison between example 58 and example 59 is not legitimate. The crucial data relate to the possibility of extracting the second object of a ditransitive verb. Such objects can be shown to be direct arguments of the verb, and I will assume that they are VP constituents in both *meN*-V and *di*-V clauses. If verb prefixation makes VP an island for extraction, we would predict that such second objects would not be extractable for either type of clause, and this is the case:

60. **buku yang Ali memberi saya itu*
book REL Ali *meN*.give 1SG that
(FOR: 'the book that Ali gave to me')

[23] See chapter 5 for discussion.

61. **buku yang saya diberi Ali itu*
book REL 1SG *di*.give Ali that
(FOR: 'the book that was given to me by Ali')

Again, the evidence suggests that verbs in *di*-V clauses behave in the same way as verbs prefixed with *meN*-.

In addition to extraction being impossible from the environment of a *meN*- prefixed verb, topicalization also fails from such structures:

62. **Buku ini Badu sudah membaca*
book this Badu PERF *meN*.read
(FOR: 'This book, Badu read.')

As might be expected, given the requirement that bare DP actors appear adjacent to the verb, the identical topicalization from a *di*-V clause is also bad. But if *oleh* is used, the structure is perfectly acceptable:

63. **Badu buku ini dibaca*
Badu book this *di*.read

64. *Oleh Badu buku ini dibaca*
by Badu book this *di*.read
'By Badu, this book was read.'

The considerations previously discussed suggest that in this case also the true comparison should be with the second object of a ditransitive. Topicalization of such an argument is not possible from a clause with either a *meN*- verb or a *di*- verb:

65. **Buku itu Ali memberi Siti*
book that Ali *meN*.give.3 Siti
(FOR: 'That book, Ali gave Siti.')

66. **Buku itu Siti diberi Ali*
book that Siti *di*.give Ali
(FOR: 'That book, Siti was given by Ali.')

Topicalization of the non-subject argument from a clause with a bare verb form is fully acceptable:

67. *Film itu saya mau lihat*
film that 1SG want see
'That film I want to see.'

68. *Gadis itu saya suka sekali*
girl that 1SG like very
'That girl I like a lot.'

Topicalization is not possible for the non-subject argument of a Pro-V clause, but again I would suggest that this is for independent reasons to be discussed below. The contrast between bare verbs and verbs with prefixed morphemes is again the crucial point: *meN*- and *di*- behave in the same way.

Voskuil (1996: 192-195) claims that *meN-* and *di-* have an additional common syntactic property: they both license a zero anaphor as their non-subject argument. He offers the evidence reviewed above of contrasts between the acceptability of left-dislocations from *meN*-V clauses and from parallel bare verb clauses to support his claim. However, when a headless relative clause is part of an equational predication, the constraint does not apply:

69. *Surat ini saya yang menulis(nya)*
letter this 1SG REL *meN*.write
'As for this letter, it is me who wrote (it).' (Voskuil 1996: 193 ex.32)

Voskuil claims that this construction is possible because it does not involve movement leaving a trace, but rather there is a null resumptive pronoun in the original position of the left-dislocated DP. Given this assumption, tests which distinguish between traces and null pronouns (*pro*) are available. Firstly, non-referential DPs can be moved but cannot be left-dislocated, because *pro* is referential:

70. *Tidak satu buku pun saya akan baca*
NEG one book EMPH 1SG FUT read
'Not one single book I shall read.' (Voskuil 1996: 194 ex.37)

71. **Tidak satu buku pun saya yang menulis*
NEG one book EMPH 1SG REL *meN*.write
(Voskuil 1996: 194 ex.38)

Secondly, locally bound anaphors can be moved but not dislocated because anaphors and DP traces are both subject to Principle A of the binding theory (Chomsky 1981), but *pro* is a pronoun subject to Principle B and therefore in complementary distribution with anaphors:

72. *Dirinya sendiri Ali tidak percayai*
self.3 self Ali NEG believe.APPL
'Himself Ali doesn't trust.' (Voskuil 1996: 195 ex.39)

73. **Dirinya sendiri Ali yang mempercayai*
self.3 self Ali REL *meN*.believe.APPL
(Voskuil 1996: 195 ex.40)

All the evidence which Voskuil gives for these claims involves *meN*-V clauses, but the arguments developed here would lead us to predict that *di*-V clauses should behave in a parallel fashion. However, this is not the case. Such constructions with a *di*-prefixed verb in the relative clause are consistently bad:

74. **Ali surat ini yang ditulis*
Ali letter this REL *di*.write
(FOR: 'As for Ali, it was this letter that was written (by him).')

75. **Ali surat ini yang ditulisnya*
Ali letter this REL *di*.write.3
(FOR: 'As for Ali, it was this letter that was written (by him).')

Voskuil's claim that this possibility is in some sense dependent on the possibility of a resumptive pronoun means that the status of the actor should not be relevant here. The resumptive pronoun will always be a term, just as the undergoer of a *meN*-V clause is. Leaving aside this point, we note that in fact this construction is not possible also with the second object of a ditransitive verb:

76. **Buku itu Ali yang diberinya*
book that Ali REL *di*.give.3
(FOR: 'As for that book, it was Ali who was given it by her.')

I will suggest a reason for this difference between *meN*- verbs and *di*- verbs below.

2.2.3 Summary

The preceding sections have presented a variety of arguments designed to show that *di*- is not a proclitic pronoun, a reduced form of *dia*. The most detailed analysis along these lines, that of GHT, has been shown to be flawed. Both historical and morphosyntactic evidence have been presented to show that any such analysis is incorrect, and finally it has been shown that such an analysis is not compatible with basic principles of LFG. On the other hand, there is morphosyntactic evidence which shows that *di*- has properties in common with *meN*-, which is unequivocally a verb prefix[24]. I will therefore proceed on the basis that *di*- is a verbal prefix, in paradigmatic relation to *meN*- which has never been claimed to be anything but a prefix, to my knowledge. The next topic for discussion is then the syntactic function of these prefixes.

2.3 Prefixed Verbs: *meN*-V and *di*-V clauses

2.3.1 The status of non-subject actors

The constructions with a *di*- prefixed verb are described as passives in many grammars, although reservations are often expressed. The following from IRG is typical:

> Some grammars of Indonesian do not use the terms active and passive, instead using terms such as subject and object construction or subjective

[24] Many Austronesian languages have a prefix indicating that the actor is the subject of a clause which consists of a nasal consonant only (e.g. Balinese, Sasak, Javanese). In informal usage, such a prefix is sometimes used in Indonesian. For example, the actor subject form of the verb *tulis* 'write' would be *nulis* rather than *menulis*. Umar Muslim (p.c.) points out that in this case, the verb behaves like the bare verbs discussed in the preceding sections, and not like a prefixed verb. This suggests that *meN*- is not a unitary morpheme at some level of analysis (see Benjamins 1997).

> and objective focus[25]. This is because there are differences between the constructions in Indonesian and the active and passive in European languages such as English, both in structure and function. Nevertheless, there are also important similarities, and the relationship between the two constructions is often similar to the relationship between active and passive in English, allowing the same terms to be used to describe them. (p247)

The differences commonly noted between this construction and the English passive are that the Indonesian construction is used much more commonly, and that it often cannot be translated by the passive voice in English. The second of these reasons is irrelevant from a linguistic point of view; the first is intriguing and there is a valuable discourse-oriented literature which examines these issues (for example Cumming 1991, Hopper 1988, Kaswanti Purwo 1988 and McCune 1979). However, there are language-internal reasons why the construction should be more common in Indonesian. Extraction (relativisation and question formation) is largely restricted to subjects in Indonesian, not including adjuncts, apart from some exceptions which will be discussed later (and see also chapter 5). Therefore, the *di*-V construction is very commonly used to feed extraction and might be expected to be more common. The data collected by E. Anderson (1983) suggests that politeness and social distance are also relevant considerations. Discourse studies are not always as exact as might be desired in specifying whether such factors are controlled in their figures, but it seems likely that in narrative at least, there is still a tendency for the *di-* construction to be used in non-initial clauses of event sequences[26].

There are more serious syntactic reasons to doubt that all the *di-* constructions are passive. I assume here that passive is a lexical process which derives an intransitive verb from a transitive verb, the subject of the new verb being thematically equivalent to the undergoer of the transitive verb, and the lexical process being formally marked on the intransitive verb. The *di-* construction does not clearly fulfil either of these criteria. The prefix on the verb changes from *meN-* to *di-*, but it would be perverse to describe either of these as formally unmarked in relation to the other. Unmarked verbs occur in Indonesian (see examples 4, 5 and 6 above), and both *men*-V and *di*-V constructions are formally marked relative to them. But by another, and in this case more compelling criterion, the *meN*-V construction must be considered as the basic construction of the language: a correct sentence with this verb form can be given for any pair of arguments and any register. Both the *di*-V construction and the Pro-V construction seen in example 4 impose restrictions on the actor argument, and the bare verb construction seen in example 6 is not useable in formal Indonesian. *meN*-V clauses are distributionally unmarked, and must be taken to be the basic clause type of the language.

---

[25] The confusion of levels implicit in this terminology is endemic to the literature on Indonesian (SM).

[26] Helen McKay (p.c.) reports that in a corpus of contemporary Indonesian journalism, the *di-* construction is used with a frequency comparable to English passive. However, this genre is influenced by English in both vocabulary and structure.

In terms of syntactic transitivity, *men*-V and *di*-V constructions do not contrast clearly either. I take syntactic transitivity to mean the ability of a verb to license two direct arguments; in this case either type of verb apparently meets the criterion. With *men-* verbs, the postverbal argument always occurs without a preposition, and with *di-* verbs, the post-verbal A only needs a preposition when it is not adjacent to the verb:

77. *Dia* *menjemput* *saya*
3SG *meN*.meet 1SG
'S/he met me.'

78. *Saya* *dijemput* *Ali*
1SG *di*.meet Ali
'I was met by Ali.'

79. *Kami* *dibawa* *ke* *bioskop* *oleh* *ayah*
1PL.EXCL *di*.take to cinema by father
'We were taken to the cinema by Father. '
(IRG:247,248,259)

But the arguments of A&M, introduced above (section 2.2.2) show that this data is deceptive; the bare DPs in examples such as 78 are not in fact direct arguments. Only the attached pronoun *-nya* is a term in *di*-V clauses. *di-* verbs **can** license two direct arguments, *men-* verbs **always** do[27]; there is an asymmetry between the two, but a weak one. Transitive verbs are sometimes defined as those which can be passivized. On this criterion, the Indonesian pattern is still unclear; given the symmetry between the two constructions just discussed, it is not impossible to consider *men-* verbs as the passive (or rather antipassive) of *di-* verbs. In other Austronesian languages such as Balinese and Tagalog, the undergoer subject clause type has the formally unmarked verb, and an ergative-antipassive analysis has been proposed for such languages by various scholars (Artawa 1999, Blake 1988, De Guzman 1988).

Various accounts have been offered in the literature as to how the actor argument of a *di*-V construction should be analysed, and I now discuss three of these before presenting my own conclusions.

2.3.1.1 THE GENITIVE AGENT HYPOTHESIS

All DPs in Tagalog and a number of other Philippine languages have an obligatory case marker or determiner. The marker used for non-subject direct arguments *ng* [vAN] is identical, or at least homophonous, with the marker used for dependents within DPs:

[27] Excepting of course the small number of intransitive *men-* verbs such as *menangis* 'weep, cry'.

80. *iniabot* *ng* *manggamot* *sa* *sundalo*
UF.REAL.within.reach GEN IRR.AF.medicine OBL soldier
ang *itlog*
NOM egg
'The physician handed the egg to the soldier.' (Bloomfield 1917: 30)

81. *bata'* *ng* *dalaga*
child GEN young woman
'child of the young woman' (Himmelmann 1991: 8)

The glossing practice in the above examples is quite commonly followed by Philippinists, with the marker uniformly glossed as GEN(itive). Whether this common marking should be taken to indicate a common function is the subject of some debate. Two questions must be distinguished here: (i) are all non-subject arguments treated the same regardless of thematic role; and (ii) are (some) dependents of verbs functionally identical to dependents of nouns.

The first question can be answered definitely only for Malagasy, where the two types of dependent relations are distinguished morphologically for both non-subject clausal arguments and DP-internal arguments (Mahdi 1988). In Tagalog, there is no formal difference amongst non-subject clausal arguments; the actor of an undergoer subject clause and the undergoer of an actor subject clause are both marked with *ng*, and when an oblique argument is made subject, both actor and undergoer are marked by *ng*. Some researchers have reported differences in syntactic behaviour dependent on thematic role (e.g. De Guzman 1997), but the data are hard to interpret, given the effects of thematic prominence in Tagalog grammar noted by Schachter (1976). And the most detailed study of Tagalog syntax published in recent years (Kroeger 1993: 47) concludes that non-subject agents are direct arguments (terms) just as non-subject undergoers are.

The second question addresses the issue of whether the same functional relation can obtain between a verb and its dependent and a noun and its dependent. The formal identity of case marking could indicate that the functional relation between the two elements is the same in each case, or it could be only a case of available formal resources being deployed in two distinct contexts where there is no possibility of ambiguity. It should be noted that the issue here is distinct from the problem posed by such syncretisms as that of ergative and instrumental case in many Australian languages (Blake 1994: 25-26) where two clause-level functions are involved. If the two functions are taken to be identical, this would lead to a view of Tagalog clause structure and DP structure as very similar, if not identical. Naylor (1980) has argued that this is indeed true for Tagalog, that *ng* simply marks that an attributive relation of some sort holds between the head and its dependent. She carries these views on to their logical conclusion in a later paper (Naylor 1995) which argues that all Tagalog clauses are equational predications which specify the identity between two nominal elements. On this view, all non-subject arguments are within DPs, and the case marking reflects this.

Voskuil (1996, Ch.7) attempts to construct a theory of case for Western Austronesian languages which assumes that there is a clause-level genitive

case, and that this is distinct from a clause-level accusative case. Genitive is the case of non-subject actors on this analysis, which allows Voskuil to maintain that undergoer subject clauses are passive and have a different case array from active clauses. This analysis is dubious on several grounds. Firstly, the distinction between genitive and accusative case is formally justified only in Malagasy. In the other languages Voskuil considers, Tagalog and Indonesian, the distinction does not exist in form, and he gives no evidence and no arguments to support the distinction in these languages. Secondly, the evidence presented by Kroeger (1993) for Tagalog, and that presented by Arka and Manning (to appear) as well as the data presented below for Indonesian, all go to show that (some) non-subject agents in these languages are direct arguments. Thirdly, Voskuil himself builds extensive argumentation on the fact that the non-subject argument in *meN*-V and *di*-V clauses have identical properties. This is surprising if the two are different types of arguments, appearing in different though formally indistinguishable cases. Finally, Voskuil claims that there are parallel case marking possibilities and mechanisms in clauses and DPs, but he does not allow that the similarity might be deeper than that, he disavows an analysis in terms of nominalisation as discussed above. But his position is not logically satisfactory in this respect. For all these reasons, I cannot accept Voskuil's analysis.

2.3.1.2 AGENT INCORPORATION

Another response to these problems is to claim that the *di*-constructions are not a unitary phenomenon. One study along these lines is Myhill (1988), who argues that when the actor argument is introduced by a preposition, the *di*-V-PP construction is a straightforward passive, but that when the actor argument is a postverbal DP (*di*-V-DP), it is incorporated into the verb. The evidence supporting this claim is that the bare DP must be adjacent to the verb, and that it tends to be non-specific. There are several weaknesses in Myhill's proposal. The most immediate one is that a postverbal actor need not be a single word. This must cast doubt on the notion that such nouns are incorporated as, across languages, consistently only single nouns are incorporated, as acknowledged by Myhill, citing Mithun (1984). Myhill suggests two responses to the problem: either the entire DP is incorporated, or only the head noun is incorporated (p120). In defence of the first possibility, Myhill claims that 'prototypical noun incorporation involves only one word', but that this should not rule out the possibility of whole phrases being incorporated. But he presents no evidence in support of this position[28], and the source on which he relies for the claim of prototypicality (Mithun 1984) does not make any such claim on my reading, and indeed has no examples of more than one word being incorporated. If the alternative

[28] Myhill's statistical evidence of factors which condition the use of the construction is open to essentially the same criticism. He examines a number of factors which predict the occurrence of the construction, none of which is categorical. This is as expected in one direction, i.e. not every clause which meets the conditions for incorporation will actually use it, but is highly surprising in the opposite direction, i.e. not every clause which exhibits incorporation meets the conditions.

analysis is correct, there should exist some syntactic evidence that the remnant of the DP is a separate constituent; Myhill presents no such evidence.

Mithun (1984) argues that an incorporated noun is non-referential, and that therefore 'it is not marked for definiteness or number' (849). Myhill's own examples show that the actors he claims are incorporated do not meet this criterion:

82. *Api yang kuning menyala itulah yang dipandang*
fire REL yellow *meN*.burn that.PRT REL *di*.look.at
keduanya
COLL.two.3
'The burning yellow fire was what was being looked at by the two of them.' (Myhill 1988: ex18)

In this example, the actor argument is quantified: there are exactly two people referred to. The argument is also specific, with the enclitic *-nya* used to mark definiteness (see section 1.2.3). Indonesian bare DP actors clearly also fail this test for incorporation.

Further problems emerge with Myhill's use of Mithun's 1984 study as support for his position. The part of Mithun's work which Myhill sees as particularly relevant to the Indonesian case discusses the use of incorporation to systematically background known information. There is little doubt that *di-* V-DP clauses in Indonesian can be used for this purpose, even more so when the actor argument is the clitic pronoun *-nya*, but Mithun sees this as the typical manifestation of type 3 incorporation in her scheme. This is explicitly stated to occur typically in polysynthetic languages, and to imply the existence of type 1 and type 2 incorporation in the language also. Indonesian is not a polysynthetic language and has little trace of type 1 incorporation; there are a few lexicalised verb plus noun combinations, but the process is not productive and such verbs almost always occur without any prefix (Verhaar 1984a:48-9)[29]. Type 2 incorporation, in which the argument structure of the clause is modified, does not exist in Indonesian. Even if the construction Myhill analyses as incorporation is treated as type 1, the combination of verb and noun as an intransitive predicate, expected properties are lacking, for example there is no tendency for common combinations to be lexicalised. This is certainly related to another problem in Myhill's account: the implausibility of agent incorporation. Just as verb-agent idioms are much less common than verb-patient idioms, lexicalised agent plus verb units are unlikely to occur because we conceive of agents as more independent from actions than patients (Dowty 1991). Again, although Myhill claims that there is no theoretical impossibility in the analysis, it clashes with an implicational hierarchy established by Mithun which has patient as the most likely semantic role to be incorporated. Indonesian has only limited evidence of patient incorporation, and even less sign of incorporation of other roles. Myhill's analysis thus suggests either that Mithun's hierarchy is wrong, or Indonesian does not follow what is otherwise a well-supported universal. A

[29] See section 5.3.5 for discussion of another clause type that has been claimed to involve incorporation.

possible escape would be to claim that incorporation works on a syntactic basis rather than a semantic basis, and that objects are the primary target. Then Indonesian *di-* verbs might be analysed as ergative constructions with the agent as object, and the incorporation would be as expected. However, this involves two additional assumptions (that *di-* verbs with DP actors occur in ergative clauses, and that the agent of an ergative construction is an object [see Manning 1996a for discussion]) and contradicts an explicit statement of Mithun's, that her generalisations are independent of the syntactic systems of different languages. For these reasons, I reject Myhill's proposal[30].

2.3.1.3 TERMS AND NON-TERMS

Arka and Manning (to appear, afterwards A&M) also argue that *di-* verb constructions do not form a single group, but they separate the constructions differently. Based on evidence from binding phenomena, A&M claim that *di*-V-*nya* clauses should be distinguished from the other two possibilities[31]. The crucial evidence depends on the question of whether the agent can bind a reflexive in subject position. On a purely configurational account of binding (Chomsky 1981), this should be impossible. However, a theory which allows thematic prominence a role in binding might make a different prediction, and it is well known that in other Western Malayo-Polynesian languages binding is constrained by thematic rather than syntactic prominence (Schachter 1976). The theory advanced by Manning (1996a,b) posits that binding relations are computed on a syntactic level of argument structure, which respects the division between terms and non-terms, with thematic prominence ordering the arguments within these two divisions (see discussion in section 1.3.3). Thus, in Manning's theory, if an agent is a term, it will be able to bind a reflexive in any other argument position (even subject), but if the agent is a non-term, it will not be able to bind any term. The crucial data is as follows:

83. *?*Dirinya* *tidak* *diperhatikan* *Amir*
self.3 NEG *di*.care Amir
Himself was not taken care of by Amir. (A&M: ex29a)

84. *Dirinya* *tidak* *diperhatikannya*
self.3 NEG *di*.care.3
S/he didn't take care of her/himself (A&M: ex25a)

In example 83, the DP agent *Amir* cannot bind a reflexive in subject position and therefore it is not a term. The argument structure of the clause is:

85. <dirinya : Amir >

where items left of the colon are terms and those on the right are not. On the other hand, in example 84 the reflexive in subject position is acceptable and

---

[30] Clynes (1995) makes a claim similar to that of Myhill for Balinese, but for actor subject clauses: 'Indefinite nominal complements of Actor Pivot verbs are incorporated into the verb to form a syntactically intransitive unit' (196). For the reasons discussed in the main text, this is a typologically more plausible claim.

[31] Myhill 1988 does not deal with the enclitic *-nya* at all.

by Manning's theory the agent clitic must be a term. The argument structure in this case is:

86. < -nya, dirinya : >

The actor is a term and as the most thematically prominent term is first on the list. Further support for this analysis comes from the fact that in the case of a ditransitive verb, an agent DP cannot bind a reflexive theme argument:

87. *Amir*$_i$ *diperlihatkan* *Ayah*$_k$ *foto* *dirinya*$_{i/*k}$
Amir *di*.show father photo self.3
Amir$_i$ was shown the picture of himself$_{i/*k}$ by father$_k$. (A&M: ex29c)

In this example, the grammatical subject is the only possible binder of the reflexive and the argument structure of the clause must be:

88. < Amir, foto dirinya : Ayah >

A&M do not offer an example where the agent is introduced by *oleh*, but they would clearly predict that such an agent would not be able to bind a reflexive in any term position.

Additional evidence supports the analysis of A&M. Indonesian permits floating quantifiers, and this test reliably distinguishes between direct arguments, and obliques and adjuncts (terms and non-terms). The quantifiers *seluruh* and *semua* 'all' both normally precede the noun which they quantify, but can on occasion be moved to the end of the clause, sometimes with the addition of the suffix *-nya*. Keenan (1976) points out that if any type of DP allows quantifier float (QF), it will be subjects, and this is true for Indonesian:

89. *Pertahanan-tahanan* *telah* *kabur* *seluruhnya*
prisoners.DUP PERF flee all.3
'The prisoners all ran away.' (Chung 1976a: n.30)

QF is also possible from both the second argument of a *meN*-V clause, and from the third argument of a ditransitive verb[32]:

90. *Saya* *mukul* *anak-anak* *itu* *kemarin* *semuanya*
1SG hit child.DUP that yesterday all.3
'I hit all the children yesterday.'

91. *Saya* *memberinya* *hadiah* *itu* *dulu* *semua*
1SG *meN*.give.3 present that before all
'I gave her all the presents before.'

But QF is not possible from an uncontroversial oblique, that is a DP within a PP:

[32] The quantifiers which float can also be shifted to the left of the NP they modify in a separate phenomenon (see Chung 1976b and discussion in section 1.3.3). This process is not sensitive to the grammatical function of the NP. Therefore in examples 90 and 91 and parallel examples an adverbial appears between the leftmost NP and the quantifier in order to ensure that it is QF that has occurred.

92. *Orang-orang Sasak datang dengan anak-anaknya*
man.DUP Sasak come with child.DUP
semuanya
all.3
'All the Sasak people came with their children.' NOT
*'The Sasak people came with all their children.'

The generalisation seems to be that quantifiers can float from direct arguments (terms) but not from oblique arguments. On this test, the non-subject agent of a *di*-V clause is not a term:

93. **Ali dipukul mereka dulu semua*
Ali *di*.hit 3PL before all
(For: 'Ali was hit by all of them previously.')

This is consistent with A&M's binding evidence.

The evidence then shows that some clauses with *di-* verbs have two term arguments (those with *-nya* as actor), while the rest have only one term, the subject. It is therefore wrong to call *di-* a marker of passive voice: some clauses with *di-* verbs are passive, but others are not. The correct generalisation, as A&M point out, is that *di-* indicates that the subject of the clause is an undergoer. I now examine some implications of this claim.

2.3.2 Lexically specified linking

As is well known, the Philippine languages such as Tagalog allow arguments with a wide variety of semantic roles, even non-selected arguments, to be subject of the clause:

94. *Bumili ang lalake ng isda ng pera sa*
AF-buy NOM man GEN fish GEN money OBL
tindahan
store
'The man bought fish with money in the store.'

95. *Binili ng lalake ang isda ng pera sa*
UF-buy GEN man NOM fish GEN money OBL
tindahan
store
'The man bought the fish with money in the store.'

96. *Binilhan ng lalake ng isda ng pera ang*
LF-buy GEN man GEN fish GEN money NOM
tindahan
store
'The man bought fish with money in the store.'

97. *Ipinambili* *ng* *lalake* *ng* *isda* *ang* *pera* *sa*
IF-buy GEN man GEN fish NOM money OBL
tindahan
store
'The man bought fish with the money in the store.'

Recent analyses (Foley 1998, Kroeger 1998, Sells 1998) in constraint-based formalisms have treated such alternations as the result of derivational morphology. Verbs (or precategorial roots in Foley's analysis) only specify which semantic arguments must be terms in their basic lexical entries. As can be seen above, these arguments have only two possible realizations: either they are subjects and marked with *ang*, or they are non-subjects and marked with *ng* (phonologically [ŋaN]). All other nominals are marked with *sa* except when they are subjects.

This type of system lacks two typical phenomena of voice systems. There is no unequivocal demotion of the actor in undergoer subject clause types, and non-terms are not promoted initially to a non-subject function, that is there is no applicative process. Theories of the linking of semantic roles to syntactic argument positions are based on a correspondence between two hierarchies, a semantic one, the thematic hierarchy, and a syntactic one in which subject is considered to the most salient function available. In the unmarked case, the two hierarchies are aligned, that is, the most thematically prominent argument is linked to the subject syntactic function[33] (see discussion in section 1.3.3). Passive can be accommodated within this system simply: the most thematically prominent argument is made unavailable for linking to a core syntactic function and the thematically most prominent remaining argument is linked to subject. This assumption is made fully explicit by, for example, Bresnan and Kanerva (1989), who take it as basic to their theory that agents can be subjects or obliques, but never any other syntactic function. But such an account will not work for a language like Tagalog. In all the clause types in which the actor is not subject, it is apparently still a core argument (see Kroeger 1993 for extensive discussion of the status of non-subject actors). That is, it is still available for linking to the subject function, yet it can be passed over. This argument can be extended by considering the clause types in which more peripheral arguments are subject (the examples above by no means exhaust the possibilities in Tagalog). Leaving aside causatives, in which semantic considerations would lead us to expect the added argument to become subject in any case, the cross-linguistically typical pattern for adding an extra argument (applicative) to a verb is that the new argument becomes a non-subject direct argument, which can then become subject via passivization. But in Tagalog, the first possibility does not exist: there are no added non-subject arguments, and the second stage is obligatory: all added arguments are subjects. Voskuil (1996) claims that all applicative processes in Tagalog trigger obligatory passivization, but without providing any explanatory account of why this should be so. Such an account can preserve the generalisations of linking theory, but only by the

[33] I ignore the possibility of an ergative system for simplicity.

addition of unmotivated processes. In contrast, the analyses mentioned previously take the opposite path and accept that conventional linking theories do not apply to Philippine-type languages. Then the simplest account is that each verbal affix in the voice system specifies directly the linking of one semantic role to the subject function.

Clearly, Indonesian does not have a full-fledged Philippine-type system (see Thomas 1978 for an attempted analysis in these terms), but the conclusion reached in the previous section, that *di-* indicates that the undergoer is linked to subject, suggests that the Indonesian system is related. Four arguments can be advanced to support this position. The first has already been made: that *di-* is not a passive marker, but something else. The second is that there are restrictions on choice of subject in Tagalog (and other Philippine languages) which are dependent on inherent properties of the argument. While the restrictions in Indonesian oppose pronouns to other nominals and third person to the other two, the Tagalog restriction is related to specificity. In general, Philippine subjects are always specific, and specific undergoers must be subject (Adams & Manaster-Ramer 1988). Nevertheless, in both cases the inherent properties of the nominal are important. The third is that the existence of the bare verb, actor subject construction discussed previously raises questions about the function of *meN-*. A pair of sentences such as the following are equivalent in propositional meaning:

98. *Ali akan melihat ibunya*
Ali FUT *meN*.see mother.3
99. *Ali akan lihat ibunya*
Ali FUT see mother.3
'Ali will see his mother.'

The linking of arguments to syntactic functions also is identical; *Ali* is the subject in each case and *ibunya* is object. Therefore it is not immediately obvious what the verb prefix *meN-* actually accomplishes in a clause like example 98. It has the syntactic effects discussed above in section 2.2.2, but it seems unlikely that the prefix is used in order to make enclitics possible or to prevent extraction of non-subject arguments. If the prefix is part of a voice system (using the term rather loosely), then general principles of economy of expression (Bresnan 2001a: 91, Chomsky 1991) would suggest that it is unnecessary[34]. This conclusion is also supported by the fact that the linking of actors to the subject grammatical function is the cross-linguistic default (excluding ergative languages), and we would therefore expect that pattern to be formally unmarked. Indeed, this is the most plausible interpretation of the clause type seen in example 99; the linking of arguments to syntactic functions is the unmarked, default option. The prefix *meN-* then must carry some additional information. *meN-* forms a paradigm with *di-*, and we have argued that *di-* specifies the linking of the undergoer argument to subject.

[34] The much lower frequency of use of verb prefixes in most varieties of spoken Indonesian suggest that economy is driving language change in this case.

Therefore, we can conclude that *meN-* specifies the opposite linking, of actor to subject[35].

The other argument that supports the analysis is that the syntax of Indonesian clauses with prefixed verbs is similar to the syntax of Tagalog clauses. The most obvious example of this is the restriction on relativization possibilities. As discussed in section 2.2.2.2, a verb prefix in Indonesian blocks extraction of any argument except the subject. The same is true of Tagalog, only the nominative argument can be extracted from a clause:

100. *isdang ibinigay ng lalake sa bata*
fish.LNK UV.PERF.give GEN man OBL child
'the fish which was given to the child by the man'

101. *batang binigyan ng lalake ng isda*
child.LNK PERF.give.DV GEN man GEN fish
'the child who was given fish by the man'

102. **isdang nagbigay ang lalake sa bata*
fish.LNK AV.PERF.give NOM man OBL child
(FOR: 'the fish which the man gave to the child')

103. **isdang binigyan ng lalake ang bata*
fish.LNK PERF.give.DV GEN man NOM child
(FOR: 'the fish which the child was given by the man')

(Examples from Kroeger 1993: 23-24)

The descriptive generalisation here is that no nominal marked with *ang* (or the equivalent suppletive pronoun forms) can appear within the relative clause, and it is intuitively clear that this fact is related to the fact that the verbal morphology uniquely identifies the semantic role associated with that argument. Various proposals have been made as to the nature of the link between these two phenomena (Nakamura 1998, Sells 1998, Voskuil 1996); the issues are taken up in detail in chapter 5 of this work. At this point, I only draw attention to the fact that the correlation exists and that it appears in Indonesian in exactly the place we would predict given the view of the verb prefixes developed here. Indonesian allows extraction of non-subject arguments, but only when a verb prefix is not present.

These arguments lead me to propose the following tentative lexical entries for the prefixes *meN-* and *di-*:

104. *meN-* prefix ___V_{tr} < x, y, ... >
|
SUBJ

[35] I ignore here the fact that there is a small group of intransitive verbs which have the *meN-* prefix.

105. *di-* prefix ___V_{tr} < x, y, ... >
|
SUBJ

Actor and undergoer are not concepts which have any formal status in LFG, therefore the linking stipulation introduced by the prefixes is stated in terms of a-structure. Indonesian verbs with three direct arguments always have the thematically least-prominent argument as second object, therefore these entries only need refer to the first two items in the argument array. The representation 105 makes no statement about the status of the actor when *di-* is prefixed to a transitive verb. I take up this issue in the sections 2.4.3.2-4.

2.4 Pro-V clauses

Before presenting my analysis of the Pro-V clause type (sections 2.4.2 - 4), I first briefly consider an alternative proposal which, on the basis of data from Bahasa Melayu (Malaysian) treats this clause type as having a pragmatic function TOPIC which is distinct from the grammatical function SUBJ.

2.4.1 TOPIC and SUBJECT as distinct functions

An analysis of the Pro-V construction is offered by Alsagoff (1991, 1992). On this account, the construction is neither a passive, nor an ergative construction, but rather one in which the pragmatic function TOPIC and the grammatical function SUBJ are assigned to different arguments. Alsagoff denies that the undergoer DP is the subject, treating the actor as such instead. The crucial data supporting this account comes from Malay (essentially the same language as Indonesian). Some speakers of Malay will not allow a Pro-V construction as the embedded clause in an obligatory control construction. Cross-linguistically, it is true that the controlled position is always the grammatical subject in such structures, and this is also true for *meN-* and *di-* verbs:

106. *Ali menyuruh Mariam [_ memukul doktor itu]*
Ali *meN*.ask Mariam [_ *meN*.beat doctor that]
'Ali asked Mariam to beat the doctor'

107. *Ali menyuruh doktor itu [_ dipukul oleh Mariam]*
Ali *meN*.ask doctor that [_ *di*.beat by Mariam]
'Ali asked the doctor to be beaten by Mariam'
(Alsagoff 1991: ex3a, 4a)

Both of these sentences are grammatical for all speakers. But for some Malay speakers a Pro-V structure is impossible in the lower clause:

108. *Ali menyuruh doktor itu [_ saya periksa]*
Ali *meN*.ask doctor that [_ 1SG examine]
'Ali asked the doctor to be examined by me' (Alsagoff 1991: ex8a)

For speakers of what Alsagoff (1991) calls Dialect A, this sentence is fully grammatical, but for speakers of Dialect B, it is not possible. For both dialects,

there is no reading in which the doctor is the agent of the examining. Alsagoff argues that one of the arguments of the embedded clause must be subject and there must be some reason why it cannot be controlled. The answer to this problem, Alsagoff proposes, is that the clause-initial DP in Malay (and Indonesian) is always TOPIC, prefixed verbs select one or other argument of a transitive verb as both subject and TOPIC, but the Pro-V construction dissociates the two: the patient in clause-initial position is TOPIC (and OBJ) and the immediately preverbal pronoun is subject. Speakers of the restrictive dialect only allow control of positions which are both TOPIC and subject, whilst speakers Dialect A allow any TOPIC to be controlled. The analysis of clause-initial DPs as TOPIC is motivated by the fact that only this DP cannot be questioned using a wh-*in situ* strategy, the addition of the clitic *-kah* to the questioned element:

109.a **Mariamkah memukul doktor itu tadi?*
Mariam.*kah* *meN*.beat doctor that just now

b *Mariam memukul doktor itukah tadi?*
Mariam *meN*.beat doctor that.*kah* just now
'Was it the doctor that Mariam beat just now?'

110. *Mariam memukul doktor itu tadikah?*
Mariam *meN*.beat doctor that just now.*kah*
'Was it just now that Mariam beat the doctor?"
(Alsagoff 1991:10a-c)

Alsagoff's argument is that questioned elements are focussed, FOCUS and TOPIC are incompatible, therefore the inability of the initial DP to be questioned in this way shows that it is TOPIC.

There are a number of problems with this analysis. Firstly, and as discussed in chapter 1, the majority of evidence seems to favour the view that the immediately preverbal pronoun is not grammatical subject. Chung (1976b) reviews a large body of evidence in detail (e.g. control of adverbial clauses), and concludes that the undergoer DP is subject in this construction, and most subsequent studies have agreed with this conclusion. I would place particular weight on one configurational fact: in all other clause types, the subject precedes any temporal or modal operator:

111. *Ali sudah membaca buku itu*
Ali PERF *meN*.read book that
'Ali has read the book.'

112. *Buku itu sudah dibaca Ali*
book that PERF *di*.read Ali
'The book has been read by Ali.'

113. *Saya sudah baca buku itu*
1SG PERF read book that
'I have read the book.'

114. *Buku itu sudah saya baca*
book that PERF 1SG read
'The book, I have read.'

The patient in the Pro-V construction (example 114) is in the same position as the subject of each of the other three constructions, while the agent is in a different position. The patient of a Pro-V construction can be right-dislocated for discourse effect without affecting the relative position of agent and temporal/modal operators[36]:

115. *Bisa kami terbangkan layangan itu*
able 1PL.EXCL fly kite that
'We can fly the kite' (Chung 1978a:16b)

Such examples provide a clear contrast with the bare verb clause type seen in example 113. In a footnote (1992: 163,n51), Alsagoff claims that there is no restriction as to the relative ordering of auxiliaries and actors in Pro-V clauses. This statement contradicts many published sources (McDonald & Dardjowidjojo 1967: 237, Moeliono & Dardjowidjojo ed. 1988: 280-281, IRG: 249, Voskuil 1996: 59-60) and native speaker judgments, and suggests that Alsagoff does not differentiate between the Pro-V and bare verb clause types in her data[37].

In order to generate the correct word order, Alsagoff proposes phrase structure rules which make the specifier of IP the topic position, with S as the complement of I (1992: 133-134). The position of SUBJ is then immediate daughter of S. These rules are not descriptively adequate. Firstly, they do not account for the possibility of a Pro-V clause without an overt I; in such a case there are two DPs to the left of the verb and there is no obvious structural analysis which allows these two positions (see detailed discussion in section 2.2.1.2). Secondly, if the ordering of actor and auxiliary is not crucial, as claimed by Alsagoff, then it is predicted that two DPs, the TOPIC and the SUBJ, should be able to appear to the left of an auxiliary. Again, the phrase structure rules as presented do not allow this possibility. Such clauses are possible, but they differ in their intonational properties from Pro-V clauses. Thus the Pro-V clause example 116 is spoken as a single intonational unit, but example 117 requires two intonation units with a break after the initial DP[38]:

---

[36] In the absence of temporals, modals or negation, Pro-V constructions with postposed subjects and bare verb constructions are indistinguishable. Disambiguation is primarily by sentence stress patterns; see Chung 1978a for details.

[37] Example 5b from Alsagoff (1991), *Saya periksa doktor itu*, is starred, but is acceptable to Indonesian speakers subject to the qualifications in the discussion of example 5 above. Whether this is a difference between Bahasa Indonesia and Bahasa Melayu, or a weakness in Alsagoff's data I have not been able to establish yet.

[38] The initial DP of example 117 is not a TOPIC by Alsagoff's test: it can be questioned in situ with *-kah*.

116.	*Buku*	*itu*	*sudah*	*saya*	*baca*
	book	that	PERF	1SG	read
117.	*Buku*	*itu*	*saya*	*sudah*	*baca*
	book	that	1SG	PAST	read

'The book, I read.'

This evidence is consistent with treating a structure such as example 117 as a left-dislocation structure and therefore distinct from the Pro-V clause of 116.

There is an additional empirical argument against Alsagoff's analysis of Pro-V clauses. She treats the actor as subject in such clauses, but this would then contradict a generalisation about subjects in Indonesian: that subjects cannot be reduced pronouns. As exemplified several times (e.g. example 5), the actor in a Pro-V clause can be one of the reduced pronouns *ku-* and *kau-*. The undergoer in this clause type cannot be a reduced pronoun, and shares this property with the subject argument of the other three transitive clause types, not to mention all other clause types in the language (this issue is discussed further in section 5.1.2). It might be possible to rescue Alsagoff's position by restating the generalisation in terms of TOPIC: TOPIC cannot be a reduced pronoun. But this formulation has lesser empirical coverage than the alternative. In example 116, the reduced pronoun *ku-* can replace *saya*, but this substitution is not possible in example 117. In conjunction with Alsagoff's overall account, the proposed constraint makes no prediction about this. In both cases, *saya* is not TOPIC, and can potentially be reduced. An additional statement about linear adjacency must be included in the grammar to account for the contrast. But the formulation in terms of SUBJ, with the assumption that these clauses are distinct structures, gives the correct prediction without the additional stipulation. In one case, *saya* is not a SUBJ (example 116) and reduction is possible, while in the other case it is a SUBJ and cannot be reduced (example 117). Alsagoff's account could be further modified to recognise the bare verb clause type as distinct from the Pro-V clause type, but this would mean abandoning a central claim of her theory, that the selection of TOPICs is accomplished by the verbal prefixes.

For the reasons just discussed, Alsagoff's analysis can be rejected for Indonesian, at least. The speakers who allow control into Pro-V clauses accept examples in which the subject (Chung's not Alsagoff's) position in the embedded clause is controlled, as expected. The question raised by Alsagoff's data from the second dialect is not whether this is a subject position, but why some speakers do not allow it to be controlled. I have no explanation for this phenomenon.

2.4.2 Syntax and morphology in Pro-V clauses

Salient aspects of the previous discussion can be summarised in a set of descriptive generalisations. In undergoer subject clauses:

1) only pronoun actors can be non-subject terms

2) no pronoun actor (except *mereka*) can be a non-subject non-term

3) something occurs immediately before the verb in all undergoer-subject clauses

4) only *-nya* can express a non-subject actor term after the verb

The two items which require specific reference in these statements, *mereka* and *-nya*, can reasonably be taken to be special cases. Therefore, my strategy will be to idealise the data and ignore the inconsistencies introduced by these elements. After proposing an analysis of the regularised system which results, I will examine the additional assumptions which are necessary to account for the irregularities also. Thus, the system to be analysed can be characterised by the following three statements. In undergoer subject clauses:

1) only pronoun actors can be non-subject terms

2) no pronoun actor can be a non-subject non-term

3) something occurs immediately before the verb in all undergoer-subject clauses

My first step will be to briefly recapitulate and discuss the crucial evidence for each of these generalizations.

2.4.3 Pronoun actors and termhood

2.4.3.1 ONLY PRONOUN ACTORS CAN BE NON-SUBJECT TERMS

The crucial evidence in this case is the binding patterns analysed by Arka and Manning (to appear). An undergoer subject can be an anaphor, some version of the reflexive *diri*, anteceded by the actor in any Pro-V clause:

118. *Diri saya saya serahkan ke polisi*
self 1SG 1SG surrender to police
'I surrendered myself to the police.' (A&M ex.16a)

and in a *di*-V clause when the actor is represented as the post-verbal attached pronoun *-nya* (example 84). A post-verbal actor which is an DP or within a PP cannot antecede a reflexive in subject position (example 83). If we accept Arka and Manning's version of binding theory which is well-motivated in various publications (Manning 1996b, Manning and Sag 1999, Wechsler and Arka 1998), then the best interpretation of these facts is that, in undergoer subject clauses, the actor is a term when it is a pronoun (ignoring post-verbal *mereka*) but not otherwise.

2.4.3.2 PRONOUN ACTORS CANNOT BE NON-SUBJECT NON-TERMS

Leaving aside *mereka*, the test case for this statement is the 3rd person singular pronoun *dia*. This pronoun is fully grammatical as actor in a Pro-V clause, indeed this possibility was one of the factors leading to the hypothesis that *di-* was also a pronominal element. Typical statements of prescriptive rules for Indonesian grammar claim that *di*-V clauses are possible with any 3rd person actor, therefore we would predict that a clause such as the following would be possible:

119. *?Buku itu dibaca dia*
book that *di*.read 3SG
'The book was read by him/her.'

But I cannot recall ever seeing a clause like this in written Indonesian, and some speakers reject it outright as ungrammatical. One speaker, however, offered a particularly interesting comment on this example. He said that he could imagine uttering such a clause in a context in which the actor was very contrastive, but that in such a context he could imagine also using a 1st or 2nd person actor as a post-verbal actor. As noted previously, 1st and 2nd person actors do appear as non-terms in prepositional phrases headed by *oleh* in casual speech, and this possibility is acknowledged in some grammar books. But the possibility of a 1st or 2nd person pronoun directly following a *di-* prefixed verb would not be countenanced.

Even in the cases in which a pronominal actor occurs with *oleh*, the actor is contrastive. In the words of Poedjosoedarmo (1986):

> If the speaker wishes to put information focus on the agent of the passive verb, the agent must be moved to a position where the rules [of information structure SM] allow it to receive focus (p13)

Poedjosoedarmo gives the following example:

120. *Itu dibuat oleh saya*
that *di*.do by 1SG
'That was done by me.' (Poedjosoedarmo 1986: 13)

Therefore, it would seem that this generalisation should be amended to refer only to topical pronoun actors. Pronouns are topical by default, presumably, therefore the cases which limit the generalisation in this way are few and far between, but they do exist.

2.4.3.3 COMBINING THESE TWO GENERALISATIONS

The combined effect of the two descriptive generalisations just discussed reduces to a single constraint:

121. (Topical) Pronoun actors must be terms.

This constraint can be viewed in several ways. It might be considered to be a statement specific to the grammar of Indonesian which is typologically strange, and an historical explanation might be proposed (see Himmelmann 1996, van den Berg 1996 for studies of Sulawesi languages along these lines). One alternative is to view the constraint as part of an optimality-theoretic system (Prince & Smolensky 1993), that is as a defeasible constraint which is part of the grammar of all languages but whose effects are more or less apparent depending on how it is ranked with respect to other constraints and how it interacts with them[39]. This constraint would always be satisfied by actor subject clauses, and would be satisfied in all clauses in languages with ergative organization, if we accept that such languages have truly transitive clauses (see Manning 1996a for discussion). The one type of clause which would breach the constraint is the undergoer-subject clause of a language with nominative-accusative orientation, that is the passive. If we assume that

[39] The formulation of the constraint in OT terms would be opposite in polarity: *PRO actor NON-TERM - incur a mark for any pronoun actor which is not a term.

the standard view of passive is correct and that this can be reconstructed in terms of OT (see Aissen 1999, Legendre, Wilson and Smolensky 1993, and Sells, to appear, for discussion), then no syntactic means exists for the constraint to be satisfied. The actor is removed from the list of terms associated with a verb by the lexical process which derives passive verbs so it is never available for syntactic expression as a term. However, this leaves open the possibility that a language might have a morphological means of expression for pronoun actors. Transparently, morphology precedes mapping of arguments to syntactic functions (LMT or whatever other theory we assume); lexical processes with morphological consequences like applicative and passive affect the mapping. Therefore in principle a language might be able to satisfy constraint 121 without violating mapping theory. I will argue below that Indonesian is such a language, and I will also argue that once this possibility is allowed, the constraint on the possible means of expressing pronoun actors turns out to be artefactual. The apparent constraint arises as a result of the interaction of the morphological resources of the language with the general principles of Economy of Expression.

2.4.4 Morphology of undergoer subject clauses

2.4.4.1 A PREVERBAL MORPHOLOGICAL SLOT

The prefix *di-* is phonologically and orthographically part of the verb, always immediately to its left. And the pronoun actor of a Pro-V clause cannot be separated from the verb either. This is seen most clearly from the interaction of the restriction on possible actors in the Pro-V clause type and the relative ordering of auxiliaries and preverbal nominals. As discussed previously (section 2.2.1.2), a proper name can be used as a pronoun substitute in a Pro-V clause, but a 3rd person reading is impossible in that context. Thus in 122, where the proper name appears between the auxiliary and the main verb, it can only be read as a pronoun substitute:

122. *Mobil saya akan Ali beli*
car 1SG FUT Ali buy
'You will buy my car.' (addressed to Ali)

Where the order of the proper name and the auxiliary is reversed, the natural reading is to give the name 3rd person reference; the pronoun substitute reading is marginal:

123. *Mobil saya Ali akan beli*
car 1SG Ali FUT buy
'My car Ali will buy.' (possibly: 'My car you will buy.' [addressed to Ali])

Also, if an auxiliary intervenes between pronoun and main verb, the pronoun has subject properties, such as being a controlled function in a purpose clause, but this is not true for the pronoun in a Pro-V clause. As topicalization of the object of a bare verb clause is not allowed in a subordinate clause, the word order of the two structures is different, but the contrast is evident nevertheless:

124. *Dia datang ke Indonesia untuk dapat lihat saya*
3SG come to Indonesia for able see 1SG
'She came to Indonesia to be able to see me.'

125. *Dia datang ke Indonesia untuk dapat saya lihat*
3SG come to Indonesia for able 1SG see
'She came to Indonesia to be able to be seen by me.'

In both examples, the controlled subject would appear before *dapat*, but the difference in word order results in a reversal of semantic roles of the controlled subject. When the pronoun immediately precedes the verb, the undergoer is subject, confirming the generalisation stated above.

This generalisation is again close to the insight which persuaded earlier scholars that *di-* must be pronominal. The arguments presented in section 2.2 show that this is not the case, but the arguments presented there also showed that there is no natural analysis of the phrase structure position of the actor pronoun in Pro-V clauses. Therefore, we are forced to examine the possibility that the combination of pronoun and verb is morphological. In fact, *ex hypothesi*, we should already embrace this position. The existence of the bare verb, actor subject clause type eliminates the possibility that the verb in Pro-V clauses is bare in the same sense, as it is impossible that the same verb form should allow linking of either argument to the subject function. Syntactic processes cannot access the linking of arguments to grammatical functions, therefore some morphology must mediate the linking in Pro-V clauses and the pronoun is the only possible candidate.

Two problems arise immediately. Firstly, structures with both a pronoun (or proclitic) and a verb prefix are totally ungrammatical:

126. **Akan saya melihat Ali*
FUT 1SG *meN*.see Ali
(FOR: 'I will see Ali.')

127. **Ali akan saya dilihat*
Ali FUT 1SG *di*.see
(FOR: 'Ali will be seen by me.')

Assuming that the pronoun and verb combination is morphological, and that information about the mapping of arguments to syntactic functions is conveyed by the preverbal material in *meN*-V, *di*-V and Pro-V clauses[40], this fact might be taken as showing that there is a single slot which can be filled by elements from two different paradigms. This would be an awkward analysis, but it is not hard to show that it need not be correct. It must be true that both verb prefixes and immediately preverbal pronouns contribute information about linking of arguments, because when neither is present a default linking is adopted. If preverbal pronouns (or perhaps the morphological position they occupy) carry the information that the undergoer

[40] Of course, in Pro-V clauses, information about the semantic content of an argument is also present.

is linked to the subject function, then it is unsurprising that the prefix *meN-* and a preverbal pronoun cannot co-occur, because they would carry conflicting information. The co-occurrence of *di-* and a pronoun would not involve a conflict of information, but it would mean that the same information was given twice, and it is reasonable to assume that this possibility is therefore ruled out on grounds of economy.

At this point, it is possible to see that Indonesian undergoer subject clauses do show a split dependent on the semantic properties of arguments, that is a animacy hierarchy split of the type described by Silverstein (1976). Pronoun actors must be terms (taking *mereka* as a special case) and non-pronoun actors cannot be terms. There is an overlay of discourse-functional effects which cloud the issue, but essentially it is the semantics of the actor which drive this split. But this semantic split does not coincide with the split between the two clause types, because *di*-V-*nya* clauses have a pronominal actor which is a term. The formal split between the two clause types is governed by the morphological resources of the language, not the nature of the actor argument. Traditional formulations which describe the two clause types as overlapping in the case of 3rd person actors are not fine-grained enough.

The second and more serious problem is that I have argued previously that pronoun actors in the Pro-V clause type are not affixes (section 2.2.1.2). Therefore, the analysis will depend on developing a view of morphology in which that module of grammar is not limited to word-formation processes. I turn to this question in the next section.

2.4.4.2 THE LIMITS OF MORPHOLOGY

In previous sections, I have argued that there is no obvious syntactic analysis of the Pro-V construction and section 2.2.1.2 argued that the evidence for treating even the reduced pronouns in the Pro-V clause type as affixes was not convincing. Therefore a morphological analysis would seem at first to be also out of the question. However, two recent studies (Poser 1992, Sadler 1998) have presented strong arguments that the morphological component of grammar is not identical to the word formation processes (including affixation), but that some processes of combining separate words should be considered as morphological rather than syntactic

The notion that a word form derived by a restricted rule will pre-empt the derivation of another form by a more general rule has a long history (see Andrews 1990 and Poser 1992 for references). The idea was formalised by Kiparsky as an instance of the Elsewhere Principle (Kiparsky 1982) and can be precisely stated in a unification grammar framework, as shown by Andrews 1990 (see also Blevins 1995). Poser (1992) develops in detail the insight that a similar effect can obtain between a word form and a phrasal construction, that is, the expression of a morphological category may exhibit a complementary distribution between a synthetic and a periphrastic form[41]. For example,

[41] This insight is also present in previous work. Andrews (1984) already analyses English adjective forms along the same lines as Poser, and Di Sciullo & Williams (1987) also note this case.

English adjectives form comparatives and superlatives in one of two ways, either by affixation of *-er* and *-est*, or by combination with *more* and *most*. The choice is determined by prosodic factors which are irrelevant here. What is relevant is that the two strategies are in complementary distribution:

128.

big	bigger	*more big
small	smaller	*more small
good	better	*more good
funny	funnier	*more funny
silly	sillier	*more silly
childish	*childisher	more childish
regal	*regaler	more regal
damaging	*damaginger	more damaging
symmetric	*symmetricer	more symmetric
vivacious	*vivaciouser	more vivacious

(Poser 1992 ex. 24)

Another example discussed at length by Poser is that of the Japanese periphrastic verb construction with *suru* 'do'. In this construction, the verb is preceded by a noun (often a loan word) to create a complex verb:

129.

denwa suru	to telephone
sanpo suru	to take a walk
doraibu suru	to drive
nokku suru	to knock

(Poser 1992 ex. 1,2)

Japanese also has a productive process of forming deverbal nouns:

130.

Verb Stem	**Derived Noun**	**Gloss**
ir	iri	parching
kari	kari	borrowing
mamor	mamori	protection
oyog	oyogi	swimming
sabak	sabaki	judgment

(Poser 1992 ex. 9)

We might then predict that it would be possible to form a deverbal noun, and then to use that noun as part of a periphrastic verb. But in fact, such derivations are systematically blocked:

131.

*iri suru	parch
*mamori suru	protect
*oyogi suru	swim
*sabaki suru	judge

(Poser 1992 ex. 13)

Poser argues that this is a blocking effect: the original verb stem exists and therefore the periphrastic form is not possible. And if this is the case, he further argues, then blocking is not restricted to the word formation processes of the lexicon. He notes that it is necessary to constrain such possibilities in some way, and suggests two limiting factors. Firstly, only morphological categories, categories that are expressed affixally in some

language, can be expressed by morphological processes. Secondly, the cases he examines[42] can all be considered as examples of 'small' categories, that is, a phrasal category which dominates only zero-level categories. So the Japanese periphrastic verbs are of category V°, dominating a N° and another V°. At least in the case of the English adjective data, this second condition is not adequate as comparative measure expressions admit of phrasal expansion (e.g. *exactly three times more expensive / exactly three times dearer*. See Bresnan 1973, Bresnan 2001b). It should also be noted that a unification-based account such as that of Andrews (1990) covers all the data discussed by Poser. Andrews' theory is discussed in more detail in section 2.5.2; see also Bresnan (2001b) for further discussion of blocking in an optimality-theoretic system.

The elements immediately in question, those that are used in the Pro-V clause type fall under both of Poser's conditions. Non-subject pronoun arguments can be expressed by affixes in many languages such as Yimas[43]:

132. *pu-ka-tay*
3PL.O - 1SG.A - see
'I saw them.'

133. *pu-ŋa-tay*
3PL.A - 1SG.O - see
'They saw me.' (Foley 1991: 196)

Therefore this is a morphological category by Poser's first criterion. Poser's second criterion is also satisfied if we accept that GHT are right about the category of the preverbal pronoun in Pro-V clauses. I argued above that their analysis was incorrect as to the prefix *di-* being of the same category, and also as to the position in which the pronoun is generated. But I see no reason to doubt their claim that the pronoun (or pronoun substitute) is of category Dº. Whatever the category of the pronoun is, it is clear that it must be a zero-level category. It cannot be modified *in situ* and it cannot even support a floated quantifier:

134. *Buku itu mereka baca semua*
book that 3PL read all
POSSIBLY: 'They read all the books.' BUT NOT: 'They all read the book.'

The pronoun is a head which does not project, and it must be a head of a different type from nouns in general as they cannot appear in the preverbal position in this clause type (except as pronoun substitutes). The other generally accepted head with nominal characteristics is D and this therefore is the logical choice[44]. Two analyses of the range of items which can appear

[42] Including some Basque data not discussed here.

[43] Glossing as in source: A and O are used following Dixon (1979) to identify respectively the more agent-like and more patient-like arguments of a transitive verb.

[44] See section 1.3.3 for argument supporting the claim that Indonesian has the functional category D. In the event that the claim is not valid, an alternative might be to claim that pronouns are the only type of nominal which can appear without projecting additional

seem possible: either the lexical item consists of the features associated with the pronominal meaning, and address terms and proper names are contextually determined phonological realizations of those features; or a lexical rule allows any expression which can be used to address an individual to also be a pronoun substitute. Nothing seems to hang on the choice between these two.

Sadler (1998) uses arguments similar to those of Poser to support an analysis of certain clitics in Welsh. The salient facts show strong similarities to the Indonesian facts discussed here. Thus for Welsh Sadler claims:

> A pronominal argument *almost always* fails to appear solely in canonical argument position, but is (obligatorily) expressed by means of either cliticisation or pronominal incorporation. (p25)

This is exactly parallel to the claim developed above, that for the relevant arguments (non-subject pronoun actors), Indonesian does not allow realization in canonical argument position, but only permits a morphological expression. Sadler also notes that a clitic or incorporated pronominal can never be doubled by a full lexical argument, also exactly paralleled by the Indonesian facts[45]. She shows that there is no good evidence to treat the Welsh clitics as affixes, but that there is no satisfactory phrase-structure analysis for their positioning, again a parallel with the Indonesian facts. In order to explain the properties mentioned above, Sadler analyses the clitics as Dº heads adjoined to the Xº head of their projection. The crucial move made explicit in Sadler's analysis is that the adjoined element in such a structure can express an argument function[46].

Welsh clitics certainly do express argument functions: object when attached to V, subject when attached to I and possessor when attached to N, and if the lexical adjunction analysis is accepted the remaining questions are how the argument function is assigned and how the possible assignments are constrained. On the first point, Sadler notes that the same clitic forms can express different functions depending on the item they are attached to, the clitics do not have different case forms. This means that argument functions are not associated with the clitic forms in the lexicon, as is the case for dependent-marking languages with varied case forms. The clitics are not agreement markers either, they cannot be doubled by free nominals. Therefore head-marking is not being used to identify argument functions either. The remaining possibility in LFG is that the argument function is assigned configurationally, by functional annotation of a c-structure schema:

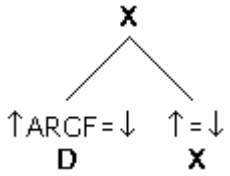
---

structure. But this claim would presumably be rooted in the semantics of specificity and would imply the further claim that bare nouns appeared with a null element of some kind which supply the necessary additional semantics. Such an analysis is a notational variant of that which invokes D as a category.

[45] Welsh allows a copy pronoun in some cases, but does not share this property even with the closely related Irish Gaelic (Andrews 1990, Sadler 1998: 27)

[46] It is not clear to me whether this move is implicit in Poser's analysis of Japanese periphrastic verbs. He calls such structures 'incorporated' which suggests that the light verb is not assigning an argument function to the associated nominal, but this is never stated definitely.

135.



Example 135 is to be interpreted as stating that a D° can be adjoined (morphologically) to a lexical head, and be assigned an argument function in the f-structure associated with the head. This leads naturally to the second question: are there constraints on the argument function which can be assigned by any given lexical head? Sadler suggests that only local argument functions can be assigned in this way, local in the sense that the argument's position in c-structure is within the maximal projection of the head. Sadler's precise formulation is:

136. A language may make use of a lexical adjunction structure to express argument functions. The argument functions which can be expressed are limited to those generated within the X^{max} projection of the c-head X° by the endocentric mapping principles. (p25)

Bresnan (2001a: 98-109) gives a full discussion of endocentricity; the effect of the principles is to allow the possibilities noted above: a clitic attached to V is an object, one attached to I is a subject and one attached to N is a possessor.

Again, the Indonesian data shows striking similarity to the Welsh facts. Indonesian is not a case-marking language. There are two instances in the language of lexical items having different forms depending on their syntactic environment. One is the 3rd person singular pronoun *dia/ia* discussed briefly in section 2.2.1.4, a relic of a prior stage in which the language had some case-marking. The other is the 2nd person bound pronoun which has the form *kau-* in the Pro-V clause type and the form *-mu* in all other positions. In other words, one form occurs only in the case in which a pronoun can be attached to the left of its host, the other form occurs in the other cases where the pronoun attaches to the right of the host. This should be regarded as a phonological phenomenon as there is no consistent relation between form and grammatical function in this case. In the most comparable cases, that of *-mu* attached to a *meN-* prefixed verb, and *kau-* in a Pro-V clause, the attached pronoun is a non-subject in both, and can be treated as OBJ in each also. The correlations which exist then are with semantic role and with linear position. The semantic role of agent (actor) is also assigned by the preposition *oleh*, and it can host the pronoun *-mu* but not *kau-*:

137. *Hal itu kedengaran olehmu*
matter that overhear by.2
'That matter got overheard by you.'

138. **Hal itu kedengaran olehkau*

Therefore linear position is the conditioning factor, and the difference can be treated as phonological. As noted previously, Indonesian also lacks verbal agreement of any kind. As discussed in section 2.2.1.3, and in this section, doubling of attached pronouns is never possible. The pronouns do not reference arguments, they are arguments.

Using the work of Poser and Sadler as a basis, an analysis of actor pronouns in Indonesian as morphological elements is very plausible. In the following section, I present such an analysis, beginning with Pro-V clauses.

2.5 An analysis

2.5.1 Pro-V clauses

I adopt the following assumptions which have been argued for above:

1) Pronouns are generated as D°.

2) Adjunction of zero-level categories is a possible morphological operation, at least for V° and D° in undergoer subject clauses.

Assumption number 2 needs more comment. As discussed above (section 2.4.4.1), the preverbal pronoun codes the linking of undergoer to subject as well as the pronominal meaning. This must be the case, as a bare verb with no morphological dependents has default linking of arguments to grammatical functions i.e. actor is linked to subject. Therefore a morphological rule is required rather than a statement that the adjunction is possible[47]:

139. Pro-V Adjunction

$$V \rightarrow \quad D^{\circ}_{\text{actor}} \qquad V_{tr}$$
$$\uparrow \text{ARGF} = \downarrow \qquad \uparrow = \downarrow$$
$$[\text{PRED 'pro'}] \qquad < x,\ y,\ \ldots. >$$
$$\Big| \quad \text{SUBJ}$$

It is not necessary for this rule to specify what argument function the pronoun will be assigned[48], if we allow that the actor in an ergative construction is an object. Sadler's principle (136 above) predicts that only an argument function generated within the maximal projection of the c-head will be possible. The c-head is V, and object is the grammatical function canonically generated within its X^{max}. Given that the linking for subject is specified, this is not a point of major interest, but it is preferable to use the less stipulative formulation, particularly as this accords with a generalisation which has some validity cross-linguistically[49].

---

[47] The case of ditransitives is handled as noted previously: see section 2.3.2.

[48] I assume that the actor in an ergative construction can be an OBJ: see Bresnan & Kanerva (1989, n32) and Manning (1996a) for various opinions on the question.

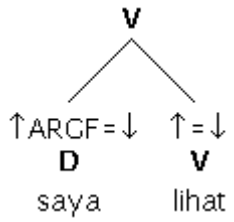
[49] It is not clear to me how this principle operates when there is no auxiliary in a clause and V is an extended head occupying I in c-structure (see section 1.2.3). In such a case, the endocentric mapping principles which Sadler (1998) relies on could give two results: if the head is treated as V, the OBJ function will be available, but if it is treated as I, the SUBJ function will be available. I assume that lexical categorisation takes precedence in such a situation.

The effect of this rule is as follows. A transitive verb has a lexical entry such as the following (I use only actor and undergoer as shorthand for the semantic roles assigned by the verb):

140. *lihat* V_{tr} 'see' < ACTOR, UNDERGOER >

The morphological rule 139 applies, giving rise to the following structure and its associated partial f-structure:

141.

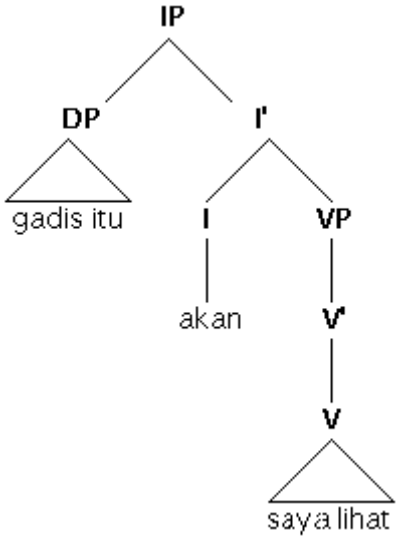


142.

$$\begin{bmatrix} \text{PRED} & \text{'see < x, y >'} \\ \text{SUBJ} & [\ ]_y \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERSON} & 1 \\ \text{NUMBER} & \text{sg} \end{bmatrix}_x \end{bmatrix}$$

This verb can form a part of a larger structure, following general principles, with the following result:

143.



144.

$$\begin{bmatrix} \text{PRED} & \text{'see < x, y >'} \\ \text{TENSE} & \text{FUT} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'girl'} \\ \text{DEF} & + \\ \text{PROXIMATE} & - \end{bmatrix}_x \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERSON} & 1 \\ \text{NUMBER} & \text{sg} \end{bmatrix}_y \end{bmatrix}$$

This analysis accounts for various properties of the Pro-V clause type in a natural fashion:

1) Treating the pronoun and verb unit as a morphological combination resolves the issue of the phrase structure position of the pronoun, explains why the two cannot be separated and why the pronoun is a non-projecting category.
2) It was noted previously (section 2.2.2.2) that we might expect the non-subject argument of Pro-V clauses to be extractable as the verb is not prefixed. Under the present analysis, the possibility of extraction is ruled out by the fact that the verb is not really bare, it has morphological material associated with it; and by the fact that the non-subject argument is realized morphologically and is therefore not visible to syntactic operations.
3) Allowing an argument function to be associated with a morphological element explains the binding theoretic properties of the Pro-V clause type, while also constraining the possibility of non-subject actor terms in accord with the data (see also the following section for binding in *di*-V-*nya* clauses), and in accord with general considerations.

The data on co-ordination in this type of clause (discussed in section 2.2.1.2) also fit in with the current analysis. It is impossible to co-ordinate two pronouns as actors:

145. **Anjing itu saya dan dia pukul*
dog that 1SG and 3SG strike
(FOR: 'I and he hit the dog.')

This possibility is ruled out on two grounds. Firstly, co-ordination is a part of syntax and therefore is not available as a part of a morphological process, and secondly, for any combination of pronouns there is a single pronoun which could be used instead and the co-ordinate structure is therefore ruled out on grounds of economy of expression. Thus in example 145, *saya dan dia* 'I and he' could be replaced by *kami* '1st person plural exclusive'. As for the co-ordination of two verbs with a single actor pronoun, this would seem to be ruled out again by the non-availability of co-ordination in morphological processes. The marginal acceptability of such structures for some speakers can be attributed to the transparency of the intention.

2.5.2 Adding *-nya* to the picture

The qualification which *-nya* added to the system, noted at the start of section 2.4, was that it is the only actor pronoun which can occur after the verb (*modulo* the effect of contrast and focus discussed previously). It is natural to associate this difference between *-nya* and the other pronouns (still excepting *mereka*) with the fact that it is the only pronoun that all speakers use for non-human reference (previously discussed in section 2.2.1.3). Prescriptively, the 3rd person pronouns *dia* and *mereka* have only human reference, although in informal usage this distinction is not always maintained. But such usage is not universally accepted and is seen by most commentators as innovative (see Kaswanti Purwo 1984 and IRG: 167-168). We might then revise the rule for Pro-V adjunction (139 above) to include the constraint that the adjoined pronoun must include the feature [HUMAN +] in its matrix.

But the fact that *-nya* following a *di-* prefixed verb codes a term suggests that perhaps rule 139 needs more far-reaching revision. The null hypothesis must be that the same morphological process operates in the case of both preverbal and postverbal actor pronouns, that is that the morphological adjunction process does not have ordering specified. Unless there is strong evidence to the contrary, we must adopt this position, which would lead in turn to the necessity of separating rule 139 into two components. One of these would express the generalisation that a personal pronoun adjoined to the left of a transitive verb results in an undergoer subject verb, while the other would express the generalisation that any pronoun adjoined to an undergoer subject verb is assigned an argument function.

The evidence shows that *-nya* following a *di-* prefixed verb is in a morphological relation to the verb, whereas *-nya* following a *meN-* prefixed verb or a preposition is a true clitic in a syntactic relation with its host. The previous discussion has established that the actor in a *di*-V clause is a term when it is realized as *-nya*, but not when it is realized as a DP adjacent to the verb. This is a surprising result if *-nya* is a clitic in the commonly understood sense, that is, a syntactic terminal in its own right which is phonologically

dependent on another word. If *-nya* is a clitic in this sense, then we would assume that it would be a constituent of the same category as any other nominal (regardless of whether we label this as NP or DP) and it would be in the same phrase structure position relative to the verb. Therefore, the difference between the term, in the one case, and the non-term, in the other case, would be unexplained. We could stipulate that a pronoun in that position was a term, but even this *ad hoc* solution is not sufficient, as the pronoun *mereka* can occur in the position (see example 28), but it is not a term:

146. *Sendiri*$_{*i/j}$ *diperhatikan* *mereka*$_i$
self *di*.look.after 3PL
'They$_i$ look after themselves$_j$.'

Co-reference is not possible in this case; the actor cannot antecede a subject reflexive pronoun.

Additional evidence as to the status of *-nya* in *di*-V clauses comes from co-ordination. I suggested above (section 2.1.1.2) that co-ordination possibilities did not provide a reliable test to discriminate between clitics and affixes. However, in this case, I believe that the evidence is useful because the judgments seem to be more definite than those for co-ordination in Pro-V clauses, and because the co-ordination data are not the only evidence available. The binding data already establish that *-nya* in a *di*-V clause has a different status to other actors in such clauses. But quantifier float data shows that such actors are not terms in the same way that full DP arguments are. Some speakers will allow *-nya* to have plural reference (see also section 2.1.1.3):

147. *Saya* *sudah* *membacanya*
1SG PERF *meN*.read.3
'I've read it/them.' (IRG: 167)

But speakers who allow plural reference for the pronoun will not allow a floated quantifier to modify it:

148. *?(*)Buku* *itu* *dibacanya* *dulu* *semua*
book that *di*.read.3 previously all
'All the books were read by him/her/them (for some speakers).'
NOT: 'The book was read by all of them.'

This is parallel to the data for preverbal actor pronouns discussed in the previous section.

Attached pronouns such as *-nya* occur in three contexts in Indonesian: following prefixed verbs, following some prepositions, and as possessors within DPs. Following a *meN-* prefixed verb or a preposition[50], these pronouns behave like standard clitics, and co-ordination with a following DP is possible:

[50] I do not consider possessors here. For further discussion of attached pronouns, see chapter 5

149. *Saya mencintaimu dan ibumu*
1SG *meN*.love.APPL.2 and mother.2
'I love you and your mother.'

150. *Saya pergi dengannya dan Ali ke kampung.*
1SG go with.3 and Ali to village
'I went with him and Ali to the village.'

But such coordination is not possible with *-nya* following a *di-* prefixed verb:

151. **Siti dilihatnya dan ibunya*
Siti *di*.see.3 and mother.3
(For: 'Siti was seen by her and her mother.'

Co-ordination for bare DP actors following *di-* prefixed verbs is fine, even when the first co-ordinated element is a pronoun:

152. *Buku itu dibaca guru dan mahasiswa*
book that *di*.read teacher and student
'The book was read by the teacher and by the student.'

153. *Buku itu dibaca mereka dan kita*
book that *di*.read 3PL and 1PL.INCL
'The book was read by them and by us.'

This evidence shows again that *-nya* attached to a *di-* prefixed verb is not a clitic and that it has a special status. Again, it is parallel to the data for preverbal actor pronouns and again it supports the argument that these pronouns are zero-level elements morphologically attached to the head verb.

An alternative analysis would be that *–nya* represents two homophonous pronominal forms, one strong and one weak in the sense of Cardinaletti & Starke (1996). The impossibility of co-ordination is a diagnostic for deficient pronouns for Cardinaletti and Starke, and the *–nya* which attaches to *di*-prefixed verbs would be deficient by this test, whereas the *–nya* which appears elsewhere would not. However, pronouns which are not deficient are assumed to be phonologically free in this scheme, and *–nya* always requires a host on its left: it cannot occur on its own. Such an analysis would also ignore the parallels between the preverbal pronoun of the Pro-V clause type and *–nya* following a *di*-prefixed verb. The pronouns of the Pro-V clause are strong by another of Cardinaletti & Starke's tests, in that they can have only human reference, but the *–nya* following a *di*-prefixed verb would be weak. This analysis therefore cannot be accepted[51].

An additional argument for accepting this proposal is that it allows us to account for the contrast between *meN*-V clauses and *di*-V clauses with respect

[51] It is possible that there is some distinction between human and non-human reference for the clitic *–nya*. The relevant examples above (examples 149 and 150) have human referents for the clitic, and co-ordination may fail in similar cases where *–nya* has non-human reference (for speakers who allow this). I have not checked this point, and in any case it is not clear to me whether the point is relevant given the considerations discussed above.

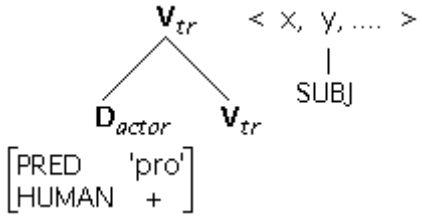
to topicalization from cleft structures. In section 2.2.2.2, I repeated evidence reported by Voskuil (1996) showing that topicalization of the non-subject argument was possible from within a headless relative clause (example 69 repeated):

69. *Surat ini saya yang menulis(nya)*
letter this 1SG REL *meN*.write
'As for this letter, it is me who wrote (it).' (Voskuil 1996: 193 ex.32)

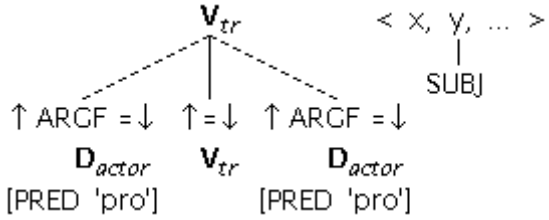
This structure is possible only when the verb prefix is *meN-*, with *di-* it is impossible. Voskuil links the possibility of this structure to the possibility of a resumptive pronoun within the relative clause, claiming that there is a resumptive pronoun in all cases, but that it is not always overt. Regardless of the correctness of Voskuil's account, the basic observation is intuitively correct: the resumptive pronoun and the extraction possibilities are interdependent. In section 2.2.2.2, I suggested that as the attached pronoun was possible in the parallel *di*-V clause, it was surprising that the topicalization failed. But in light of the theory developed in this section, this apparent paradox is resolved. *-nya* following a *di-* prefixed verb is not an attached pronoun, it is a morphological element. Therefore the two cases are not parallel. The pronoun is attached to the verb morphologically, and the resultant lexical structure is opaque to syntactic processes such as topicalization.

In light of the discussion immediately above, I propose the following two rules which replace rule 139 above and also generate *di*-V-*nya* clauses:

154. Undergoer Subject Verb Morphological Rule



155. Undergoer Subject Verb Pronoun Adjunction

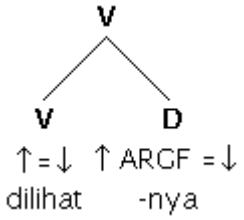


Representation 154 is intended to mean that a transitive verb which has a personal pronoun left adjoined to it will have the linking of undergoer to subject specified at the dominating V node, and representation 155 is intended to mean that if a V node which has the specified linking of undergoer to subject dominates an adjoined pronoun, the pronoun is

assigned an argument function by the verb[52]. Pro-V clauses receive the same analysis as that given in the previous section, as in that case the two new rules have the same effect together as the single rule 139.

The analysis of *di*-V-*nya* clauses is now straightforward also. The actor is not a personal pronoun, therefore rule 154 does not apply. But the prefix *di-* can be joined to the verb, resulting in a verb with the linking of undergoer to subject specified (see 105). The principles forcing a topical pronoun actor to be a term still apply, and rule 155 makes this possible, resulting in the following structures:

156.



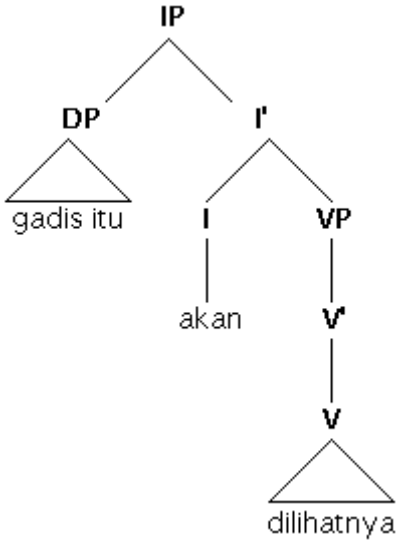
157.

$$\begin{bmatrix} \text{PRED} & \text{'see } < x, y > \text{'} \\ \text{SUBJ} & [\ ]_y \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERSON} & 3 \end{bmatrix}_x \end{bmatrix}$$

The structure of a complete clause is:

[52] Clearly, there is some redundancy between these two rules and this suggests that a deeper generalisation is being missed. I return to the question in chapter 5.

158.



159.

$$\begin{bmatrix} \text{PRED} & \text{'see < x, y >'} \\ \text{TENSE} & \text{FUT} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'girl'} \\ \text{DEF} & + \\ \text{PROXIMATE} & - \end{bmatrix}_y \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERSON} & 3 \end{bmatrix}_x \end{bmatrix}$$

This analysis accounts for the binding facts, the co-ordination facts and the topicalization facts. Note that LMT predicts that the actor should never be an OBJ: it has the feature [*-o*], which should prevent it ever being mapped to any GF except SUBJ and OBL_θ. This prediction may be falsified by languages with ergative-absolutive alignment (see footnote 48 and references cited there). In any case, the analysis presented here for both Pro-V and *di*-V-*nya* clauses apparently contradicts this prediction. It is possible, though, to maintain the analysis of verb prefixes proposed here, and LMT, given certain assumptions. LMT must only apply after all morphological processes: this assumption is well-motivated and not problematic. Given this assumption, a *di*-V-*nya* verb has both its arguments mapped to GFs when it is morphologically complete, and LMT is irrelevant, no further mapping remains to be accomplished except in the case of ditransitive verbs. I assume that in such cases, the feature [*+o*] will be assigned to the third argument, as the assignment of features is dependent on the semantic status of arguments. The ditransitive verb has two patient-like arguments, regardless of how these are realized, and features can be assigned accordingly. The mapping to OBJ_θ follows from the principles of LMT.

The case of a *di*-prefixed verb which does not have *-nya* attached to it is more complex and the most general account depends on assuming one of the versions of a-structure introduced in section 1.3.1 rather than the other. I have argued above that *di-* does not mark passive, but its effects are nevertheless similar to passive. Passive blocks the mapping of the most thematically prominent argument to the SUBJ GF; mapping follows afterwards

by general principles. In the case of an undergoer subject verb, the SUBJ GF cannot map to the most thematically prominent argument because that GF has been assigned lexically. That argument has the feature [-*o*] by general principles, and as SUBJ is not available, the only compatible GF is OBL_{θ}. This account follows naturally within the version of a-structure presented by Bresnan (2001a: Chapter 14), in which the distinction between term and non-term is not marked. For this version of a-structure, a ditransitive verb has three arguments which can be realized in various ways depending on the interaction of specific features of lexical items and general principles. However, if the alternative conception of a-structure is assumed, this account will not work. If the number of term arguments is specified in a-structure, then a ditransitive verb with the prefix *di-* and without *–nya* as actor has a different valence to the base verb: it has only two terms rather than three. In this case, Bresnan's version of a-structure allows the more general account and is to be preferred. However, in section 3.3.2.3, I discuss a case where the data seem to favour the other conception of a-structure (that of Manning 1996a).

An account incorporating the assumptions just discussed is possible. But in chapter 5, I will suggest an alternative. The lexical entries for the verb prefixes (examples 104 and 105) and the morphological rule for pronoun adjunction (example 155) already implicitly claim that LMT is not relevant for the linking of an argument to the SUBJ function as this mapping is lexically specified. In chapter 5 I will extend the argument and suggest that, in the case of prefixed verbs, LMT may not be relevant for the linking to other functions either.

Free pronouns appearing after a *di*-prefixed verb would be non-terms, and are therefore barred by the constraint 121. But the constraint does not bar the appearance of 1st or 2nd person attached pronouns following a *di-*prefixed verb. Two related questions therefore must be answered: why are forms such as *dia lihat* and *dilihatnya* both possible, and why is *saya lihat* possible, but not *dilihatku*? Intuitively, the answer to both questions will appeal to some blocking principle. *saya lihat* and *dilihatku* would have identical possibilities for the reference of the pronoun, therefore one blocks the other, but *dia* and *-nya* do not have identical reference, and both are possible. This intuition can indeed be captured in a precise fashion. The principle of morphological blocking has been stated formally in LFG by Andrews (1990):

160. Suppose the structure S has a preterminal node P occupied by a lexical item l_1. and there is another lexical item l_2 such that the f-structure determined by the lexical entry of l_1 properly subsumes that determined by the lexical entry of l_2, and that of l_2 subsumes the f-structure associated with P in S (the complete structure, after all unifications have been carried out). Then S is blocked.

One f-structure, F, subsumes another f-structure, F', if the information in F is a subset of the information in F', if F' is more specific. Therefore this principle states that a lexical item is blocked from appearing in a structure if there is another lexical item which is more specific, but which does not alter the total information content of the structure. The most natural account of *-nya* is to

say that its lexical entry includes the information that it is a pronoun, and that it has the value '3' for the attribute PERSON. This would make it compatible with both singular and plural, and human and non-human antecedents[53]. However, if this approach is adopted, the principle above (160) predicts that a form such as *dia lihat* will block *dilihatnya*. The f-structure associated with *dilihatnya* would be:

161. *di*-V clause: *dilihatnya*

PRED	'see' < SUBJ, OBJ >'
SUBJECT	(undergoer)
OBJECT	[PRED 'pro'; PERSON 3]

while that associated with *dia lihat* would be:

162. Pro-V clause: *dia lihat*

PRED	'see' < SUBJ, OBJ >'
SUBJECT	(undergoer)
OBJECT	[PRED 'pro'; PERSON 3; NUMBER sg; HUMAN +]

The structure 161 is less specific than the structure 162: 161 subsumes 162. Therefore *dia lihat* is a potential blocker of *dilihatnya* but not vice versa. The f-structure associated with the total structure in which *dilihatnya* appears will include (minimally) additional information about the SUBJ, but any such structure will not be subsumed by the structure 162 which includes additional information about the OBJ. Therefore, in accordance with Andrews' principle, the structure which includes *dilihatnya* is not blocked: both structures are possible.

Different considerations apply in the case of the 1st and 2nd person pronouns. The f-structures associated with, for example, *saya lihat* and *dilihatku* are identical:

163.

PRED	'see' < SUBJ, OBJ >'
SUBJ	(undergoer)
OBJ	[PRED 'pro'; PERSON 1; NUMBER sg]

The Morphological Blocking Principle stipulates that one item may block another only if it is properly subsumed by it; in this case the two items

[53] Conservative speakers would have values specified for one or both of these other attributes in their lexical entries.

(trivially) subsume each other, but no relation of proper subsumption exists. Andrews (1990, n.12) suggests that in such a situation free alternation is possible. But that is not the case here; some other factor is at work. I assume that the explanation is to be sought in economy of expression. One item, *saya lihat* consists of a pronoun and a verb stem, morphologically composed, while the other consist of the same verb stem with an affix and a pronoun, all three morphologically composed. There is no reason not to suppose that economy of expression may be a principle in choosing between alternative morphological expressions of the same material, as well as in choosing between morphological and syntactic expressions of identical material. And economy would clearly prefer *saya lihat*, the observed result. The possibility of focussed pronouns appearing after *di-* prefixed verbs also fits plausibly into this analysis. The constraint 121 is waived in the case of a non-topical pronoun, and the addition of the information that the pronoun was focussed would remove the blocking relation between the two possible expressions.

2.5.3 Adding *mereka*

The analysis developed to this point predicts that a personal pronoun actor, except when strongly contrastive, will only appear in one type of undergoer subject clause, the Pro-V type. *-nya* is the only pronoun which can appear in a *di*-V clause because it is the only one whose feature matrix is not possible in a Pro-V clause and which is therefore not blocked from being expressed in another way. However, the 3rd person plural pronoun *mereka* can appear in both clause types, contrary to prediction:

164.	*Buku-buku*	*ini*	*sudah*	*mereka*	*baca*
	book.DUP	this	PERF	3PL	read
165.	*Buku-buku*	*ini*	*sudah*	*dibaca*	*mereka*
	book.DUP	this	PERF	*di*.read	3PL
	'They have read these books.'				

There is no obvious syntactic reason why this should be the case, and the only explanation I can offer is a historical one.

mereka was not part of the original Malay pronoun system, but was borrowed from Javanese at some point. Zoetmulder (1982) suggests that the Javanese *marika* was a demonstrative, also used as a 3rd person pronoun. But Adelaar (1992b: 125) cites two early nineteenth century sources on Malay which both have *marika* as a noun. It is possible then that the word was borrowed into Malay as a noun which soon came to be used as a pronoun substitute. As a noun, the word would be able to appear as actor in *di*-V clauses (see following section for discussion), and as a pronoun substitute it would also be able to be used in Pro-V clauses. The nominal meaning was lost at some stage, but the postverbal possibility was left as a relic.

This scenario gains some indirect support from the usage of pronoun substitutes in modern Indonesian. Kinship terms are used frequently as pronoun substitutes in Pro-V clauses, as previously discussed:

166. *Buku itu ibu baca*
book that mother read
'I / you read the book.'

The same words are freely available as actors in *di*-V clauses, like any other noun. But in such cases, pronoun substitutes can still have a 1st or 2nd person reading, despite the general constraint against 1st or 2nd person actors in *di*-V clauses:

167. *Buku itu dibaca ibu*
book that *di*.read mother
'Mother read the book.' OR
'I read the book.' OR
'You read the book.'

This evidence shows that there is an ambiguity as to the status of pronoun substitutes within the overall system, and this gives some plausibility to the scenario sketched above. Of course, real credibility will depend on detailed examination of the use of *mereka* in Old Malay texts, work which is yet to be done.

2.5.4 Bare DP actors in *di*-V clauses

A&M argue, and I have adopted this position, that the bare DP actor of a *di*-V clause is a non-term, that is an oblique argument. The surprising consequence of this position is that an DP which is obligatorily adjacent to a verb is not a direct argument of it. This seems to be contrary to standard views of phrase structure. The actor behaves as though it were within the immediate projection of the verbal head, a position normally reserved for direct arguments, but A&M's analysis of the binding facts, and the quantifier float data discussed above, show that it is not a direct argument. Note also that with (derived) ditransitive verbs in the *di-* construction, the actor DP occurs inside the second object:

168. *Amir membelikan saya baju*
Amir *meN*.buy.APPL 1SG shirt
'Amir bought me a shirt'

169. *Saya dibelikan Amir baju*
1SG *di*.buy.APPL Amir shirt
'I was bought a shirt by Amir' (A&M:28b)

I assume that the theme argument in 168 and 169 (*baju*) is a VP constituent. Therefore there are two possible interpretations of such structures. Either the theme is a direct argument and the structure has an oblique argument closer to the head of the VP than a direct argument, or the theme is also an oblique argument. The first alternative is problematic within standard theories of phrase structure[54], but it must be correct because the QF data rule out the

[54] This is especially true for a monostratal framework such as LFG; the surface configuration in such a theory cannot be derived by movement operations.

second alternative. As shown above (example 91), the second object of a ditransitive verb can float a quantifier, and can therefore be considered a direct argument. The account of passive which I am assuming here predicts that the passive of a ditransitive verb will be a transitive verb with both the non-subject arguments of the active remaining direct arguments. This prediction is borne out in Indonesian, with the original second object still able to float a quantifier in the *di*-V clause:

170.	*Ali*	*diberi*	*mereka*	*buku*	*itu*	*dulu*	*semua*
	Ali	*di*.give	3PL	book	that	before	all
	'Ali was given all the books by them before.'						
	(NOT: 'Ali was given the book(s) by all of them before.')						

This phenomenon is not affected by the status of the agent:

171.	*Ali*	*diberinya*	*buku*	*itu*	*dulu*	*semua*
	Ali	*di*.give.3	book	that	before	all
	'Ali was given all the books previously.'					

Thus the first possible interpretation noted above must be correct: in such constructions a direct argument, subcategorized for by the verb, is further from the verb than a non-term argument.

The relationship between a bare DP actor and its verb is unusual in another respect. The adjacency requirement between the two is absolute; if anything intervenes, then the preposition *oleh* is required before the actor:

172.	*Film*	*itu*	*dilihat*	*dulu*	**(oleh)*	*Umar*
	film	that	*di*.see	before	by	Umar
	'That film was seen previously by Umar.'					

The behaviour of the two types of prefixed verbs is different in this respect. With *meN-* prefixed verbs, there is no alternative way of expressing the undergoer[55] and material can intervene between the verb and the object:

173.	*Saya*	*membeli*	*sekarang*	*buku*	*itu*
	1SG	*meN*.buy	now	book	that
	'I bought the book just now.'				

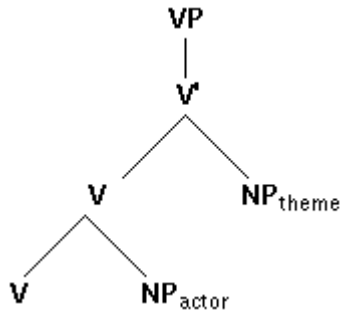
These facts show that the actor of a *di*-V clause is licensed in some fashion only if it is adjacent to the verb (or part of the verb morphologically - see previous section), otherwise the preposition is needed to license the DP.

The most satisfactory account of these facts would seem to be that a *di-* prefixed verb gives the possibility for a non-term argument DP to be right adjoined to it when that DP has the semantic role of actor. This proposal gives the right word order result, and also suggests a reason for the rigidity of the word order. It also gives a structure in which the verb and the actor are not in

[55] This was not the case in Old Malay, where all transitive verbs could use the preposition *akan* to introduce their object (Cumming 1991).

a head complement relation, but in which the object of a ditransitive verb is the complement of a head:

174.



A similar proposal has been made by Cole and Hermon (2000) to account for word order facts in Toba Batak. In fact, Toba Batak exhibits rigid ordering in both actor subject and undergoer subject constructions, with the non-subject argument appearing between the verb and the subject in each case, and the second object of ditransitives appearing outside the subject[56]:

175. *Mangalean* *buku* *si-John* *natoari* *tu-si-Mary*
ACT-give book HON-John yesterday to-HON-Mary
'John gave the book to Mary yesterday.'

176. *Dilean* *si-John* *sada* *buku* *tu-si-Mary*
PASS-give HON-John one book to-HON-Mary
'One book was given to Mary by John.'

In the words of Cole and Hermon:

> Both the Direct Object and the Passive Agent appear to be "frozen" to the verb. Not only can no element intervene between either of them and the verb, but neither can undergo any movement rule.

The verb prefixes involved in the two languages (Indonesian and Toba Batak) are clearly cognate, and it is therefore unsurprising that other properties of the languages should be similar. Note however that Indonesian only has these properties in the *di*-V clause type; example 173 above shows that word order in *meN*-V clauses is less rigid[57].

At this stage, I am not able to offer a non-stipulative account of this phenomenon. Therefore, I only propose a language-specific rule as follows:

---

[56] Glossing as in Cole and Hermon (2000): ACT - active, PASS - passive, HON - honorific.

[57] Besides the word order facts, Cole and Hermon's proposal is intended to account for binding facts in Toba Batak which, like Tagalog and Indonesian, allows non-subject actors to bind anaphors in subject position. The present work assumes that binding relations are not solely dependent on phrase structure configurations, and this aspect of Cole and Hermon's work is not relevant here. Keenan (2000) makes a proposal related to those discussed here, that actor and verb form a constituent in some Malagasy clauses.

177. Actor adjunction (Indonesian)

$$V \rightarrow \quad \underset{\uparrow = \downarrow}{\text{di-V}} \qquad \underset{\uparrow \text{OBL}_\theta = \downarrow}{\text{DP}_{\text{actor}}}$$

The data and arguments presented in section 2.5.2 establish that there is a difference in the relation of *di-* prefixed verbs and following *-nya* and a following DP. Although I have argued that the actor is adjoined to the verb in each case, the nature of the adjunction is different. In one case, the two elements are joined morphologically, in the other case syntactically, and the different properties of the two types of actors follow from this difference. The possibility of licensing oblique arguments by adjunction to the governing head will be discussed further in section 3.3.3.2, where it will become clear that the possibility is more generally available in Indonesian.

This concludes my discussion of *di*-V clause types. I assume that the type of clause in which the actor occurs in a prepositional phrase is a standard passive clause type, and does not need additional discussion here. However, I will revisit this assumption and other aspects of the syntax of *di*-V clauses in chapter 5, where I consider an alternative view of the structure of transitive clauses in Indonesian.

2.6 Summary

In this chapter, I have presented an analysis of the common clause types for verbs with two direct arguments in Indonesian. The fact that there are four clause types in which such verbs can appear raises various problems which have not been satisfactorily dealt with previously. Here, I have argued that *di-* is a verbal prefix rather than a pronoun and that it is in paradigmatic opposition to the prefix *meN-*. Not all clauses with the prefix *di-* are intransitive and it is therefore inappropriate to describe it as a passive marker. Rather, it signals that the undergoer of the clause is linked to the subject grammatical function. By symmetry, the prefix *meN-* must therefore signal the linking of actor to subject, rather than marking active voice in the sense usually understood. This conclusion is reinforced by the existence of a clause type with a bare verb form in which the actor is subject; this type is the cross-linguistic default linking option, an active voice.

Consideration of the two types of undergoer subject clauses forces the conclusion that pronoun actors must be terms in such clauses, except when they are contrastive (focussed), and that the only means by which this expression of arguments can be achieved is morphological. Indonesian therefore has a highly unusual transitive system. It has features (specified linking to subject) which are related to Philippine languages, it has only a limited place for features which are the norm elsewhere (default linking of actor to subject), and it has the possibility of undergoer subject clauses with two direct arguments. Normally, one would describe such clauses as ergative. But as has been argued in detail here, this pattern is only possible because of certain morphological properties of the language, and there is good evidence that the actor arguments in these 'ergative' clauses are invisible to syntactic operations. The clause types in question are not 'morphologically ergative' in the usual sense of that term - there is no distinctive form associated with the

actor of these clauses. But it seems equally inappropriate to describe this as 'syntactic ergativity', when the ergative argument is apparently not syntactically represented. I leave the choice of appropriate terminology to the reader.

The central concern of this thesis is to elucidate the status of non-subject arguments and the manner in which they are licensed. The assumption that the system analysed here is that of transitive verbs in Indonesian carries with it the assumption that the non-subject argument of actor subject clauses is an object, OBJ in LFG terms. However, the data in section 2.2.2 has shown that there are significant differences in the properties of this argument in the two types of actor subject clause: in one case, the object can be an attached pronoun but it cannot head a gapped relative clause, and in the other case, an attached pronoun is not possible but relativization is. This suggests that the assumption made above may not be reliable, and further data to be examined in chapter 4 will show that this is indeed the case. The non-subject arguments of undergoer subject clauses present a more complex picture. They can be direct arguments if realized morphologically, and as obliques they can be licensed in two ways, by a preposition or by adjunction to the verbal head. Chapter 3 will present data for another type of clause which shows that this second possibility is more widespread in Indonesian, and can be the only possible means of realizing an oblique argument in one case. Undergoer subject forms of transitive verbs therefore use various strategies for licensing their non-subject arguments. Some of these are unusual cross-linguistically, and do not fit easily into the theory of argument realization set out in chapter 1. It is tempting to propose that there is some correlation between these facts and the fact that the subject argument of these verbs is directly specified by morphology, rather than being selected by the operations of LMT (except in the case of the bare verb actor subject clause type). This possibility will be pursued in chapter 5, when a fuller picture of Indonesian syntax is available.

3 *ke- -an* verbs

This chapter deals with a group of predicates in Indonesian which are derived from other words by the addition of the circumfix *ke- -an*. This affix has several functions in Indonesian, but the syntactic behaviour of the words under consideration is distinctive and shows that they are verbs. The semantics of these verbs is also distinctive, almost always implying some bad effect on the subject:

1. *Tomo kecurian mobil*
   Tomo *ke*.steal.*an* car
   'Tomo's car was stolen.' (IRG: 120)

For this reason, these verbs have often been referred to as adversative verbs (e.g. Darjowidjojo 1978) or as adversative passives (e.g. Kana 1986) or accidental passives (Junus 1971)[1]. A clause such as example 1 has two nominals which meet the *prima facie* criterion for being direct arguments in Indonesian: they are not introduced by prepositions. The discussion of chapter 2 has already shown that this criterion is an oversimplification, and evidence presented below will show that the initial DP is a subject, but that the second DP does not have the properties expected of any type of direct argument in Indonesian.

The first part of this chapter is a descriptive account of this type of verb, based largely on published sources. Of particular interest here is the issue of whether there are semantically-based sub-groups which can be identified within the group of verbs and, if so, whether these correlate with differences in syntactic behaviour. Section 3.2 examines previous accounts of the syntax of this clause type, particularly the analysis of Kana (1986, chapter 6) who treats them as clause unions in a Relational Grammar framework. Section 3.3 presents a new analysis, arguing that the predicate *ke- -an* exhibits highly unusual properties in the way in which a new argument structure is created from that of the base predicate with which it combines. This process is sensitive to the semantic relations between the base predicate and its arguments, and takes no account of the grammatical functions assigned by the base. The status of the bare DP second argument which some *ke- -an* verbs allow (see example 1) is also investigated in detail, and I conclude that it is an oblique argument. I claim, on the basis of this analysis, that the possibility of adjoining an oblique argument to the verb which licenses it is not restricted to the agent in undergoer subject clauses.

[1] The assumptions which I make in this work, described in chapters 1 and 2, lead me to reject these labels as inappropriate. These verbs are not the result of a lexical process which blocks the thematically most prominent argument from linking to a direct argument function and therefore they are not passive clauses.

3.1 A description of the *ke- -an* verbs

3.1.1 Functions of the circumfix *ke-...-an*

The most common function of the *ke- -an* affix in Indonesian is to derive nouns, usually abstract. Thus *baik* 'good' has the nominalisation *kebaikan* 'goodness', and *raja* 'king' has the derived noun *kerajaan* 'kingdom'. This type of derivation is very productive. In section 4.1.2.8, such nominalizations of emotion and cognition predicates are discussed; of the 41 words for which data is presented, 31 have *ke- -an* nominalizations. A text example is the following:

2. *Susila merasa dan berpendapatan, usahanya itu*
Susila AS.feel and hold.opinion effort.3 that
semata-mata akan menimbulkan ***keamanan****,*
solely FUT AS.float.CAUS NOM.safe.NOM
kesentosaan*,* ***kemakmuran*** *dan* ***kecerdasan***
NOM.quiet.NOM NOM.rich.NOM and NOM.shrewd.NOM
rakyat.
populace
'Susila felt and believed that only that effort of his would bring about the safety, peace, prosperity and shrewdness of the people.'
(Iskandar: 1946: 61, quoted in Teeuw 1977)

This circumfix also derives two types of adjectives. There is a derivation of limited productivity, mainly restricted to colour words, with the meaning 'possessing to some degree the characteristic indicated by the base' (IRG: 50). The base is reduplicated in this derivation, thus *biru* 'blue' gives *kebiru-biruan* 'bluish' and *ilmu* 'knowledge' gives *keilmu-ilmuan* 'quasi-scientific'. This second example shows a tendency noted by Sneddon (IRG: 50) for these derivatives to have a negative connotation. The following is a textual example:

3. *Meskipun jalan kendaraan itu amat cepat - maklum,*
although road vehicle that very fast know
abang-abang supir ***kegila-gilaan*** *- tapi terasa*
older.brother.DUP driver ADJ.insane.DUP.ADJ but felt
olehku lambat juga.
by.1SG leisurely only
'Although the car's passage was very fast - you should know the chauffeur was a very crazy - but it felt slow to me.'
(Iskandar 1952: 94, quoted in Teeuw 1977)

The second type of adjectival derivation takes adjectival bases and intensifies them: *kecil* 'small' gives *kekecilan* 'too small'. This function has limited productivity in Indonesian. Sneddon ascribes this derivation to Javanese influence (IRG:50): the affix also exists in Javanese and is fully productive in this function in that language[2].

[2] Teeuw (1977) does not mention this function at all. This may be due to the fact that his data are all drawn from the work of an Indonesian writer of a previous generation, who

It is not easy to draw a boundary between the various functions of the circumfix *ke- -an* described above, or between at least the last of them and the verbal(izing) function which is the focus of this chapter. IRG (50) gives the following as an example illustrating the last function of the affix mentioned in the previous paragraph:

4. *Kamu kebanyakan uang*
   2 *ke*.much.*an* money
   'You have too much money.'

The structure of this clause suggests strongly that it is not headed by an adjective. Adjectives in Indonesian do not generally take nominal arguments in this fashion (see discussion in section 4.1.3), and the clause seems to be semantically compatible with the adversative verb pattern: on this account the meaning would more accurately be 'You suffer from much money'. Other sources include this word as a verb (Junus 1971), and indeed in a later paper, Sneddon uses this example to demonstrate the spread of the pattern to include derivations from new bases (Sneddon 2000).

None of the sources mentioned thus far gives explicit criteria by which the verbal words formed on this pattern can be distinguished from the words belonging to other categories. Teeuw (1977) quotes the following passage to show the extent to which categorisation is flexible even within a short passage:

5. *India telah* ***kehilangan*** *seorang pemimpin besar. Dan*
   India PERF *ke*.lost.*an* one.CLASS leader large and
   kehilangan *bangsa India itupun terasa oleh*
   ke.lost.*an* nation India that.EMPH felt by
   Indonesia sebagai ***kehilangan*** *sendiri.*
   Indonesia as *ke*.lost.*an* self
   'India has suffered the loss of a great leader. And that loss of the Indian nation is felt as its own loss by Indonesia.'
   (Iskandar 1952: 425, quoted in Teeuw 1977)

The structures in this passage show at least how to differentiate verbs from nouns. Each of the instances of the word *kehilangan* in this passage is followed by a DP, however the syntactic relations involved are different. In the first instance, *kehilangan* is the predicate of the clause: it is preceded by the auxiliary *telah*. In the second instance, *kehilangan* heads the subject DP, and the third *kehilangan* heads the complement of the preposition *sebagai*. In the second and third cases, the following DP could be replaced by the bound pronoun *-nya* functioning as a possessive, but this is not possible in the first case:

6. **India telah kehilangannya*
7. *Dan kehilangannya terasa ....*

---

was a native of Minangkabau. IRG (50) notes that the function in question is not considered standard by all educated speakers.

The problem of distinguishing adjectives and verbs remains. This is not a simple problem in Indonesian, and is discussed in detail in section 4.1.3 in relation to another group of words. Some of the crucial tests to be discussed there are morphological, and are not applicable in this case as the *ke- -an* circumfix does not co-occur with other morphology characteristic of adjectives. The syntactic tests discussed in section 4.1.3 also are not relevant in this case. Even the apparently clearly adjectival *ke- -an* forms do not occur in adverbial phrases with *dengan*:

8. **Dia berkata dengan kemanisan*
   3SG speak with *ke*.sweet.*an*
   (FOR: 'He spoke too sweetly.')

These forms also cannot directly modify a noun; the relativizer *yang* is always necessary:

9. *gue *(yang) kemanisan itu*
   cake REL *ke*.sweet.*an* that
   'the oversweet cake'

This evidence suggests that the *ke- -an* forms traditionally regarded as adjectives (IRG, Teeuw 1977) are in fact intransitive verbs. This should be borne in mind by the reader, but is not an important fact in relation to the arguments presented here. What is important is that there are a number of forms with the affix *ke- -an* which can have two nominal arguments and that this is impossible for Indonesian adjectives except where they have the comparative morpheme *se-* prefixed to them:

10. *orang setinggi pohon*
    person COMP.high tree
    'a person as tall as a tree'

11. **orang tinggi pohon*

The discussion of this chapter will take it that the *ke- -an* forms which can take two nominal arguments are unequivocal verbs. The discussion will refer to single argument *ke- -an* verbs from time to time, but always with the implied qualification that the categorial status of these forms is less secure.

3.1.2 Subgroups of *ke- -an* verbs

Various attempts have been made to divide the *ke- -an* verbs into sub-groups on the basis of their semantics, their syntactic behaviour, their derivational history or some combination of these factors. The differences in methodology are no doubt largely responsible for the differences in results: Junus (1971) identifies seven classes of *ke- -an* verbs, Dardjowidjojo (1978, afterwards SSAKI) identifies six and Sneddon (2000, afterwards ALKV) finds three, of which the one has four subclasses and another two[3]. These classifications are all based largely on semantic factors (see discussion below).

[3] IRG identifies six subclasses which correspond fairly closely to the groups in Sneddon (2000). The addition of the higher-level grouping is new in the later work.

Hiorth (1976) considers the syntactic environment to be more important and identifies three classes: *ke- -an* verbs which are followed by an agent, those which are followed by a DP which is not an agent, and intransitive verbs. This is similar but not identical to the classification adopted by Kana (1986). This author is primarily interested in the grammatical relations assigned by these verbs, and therefore only identifies three classes on the basis of whether there is a second argument, and whether it is a DP or a PP. Junus bases his classification on the transformations which are possible for each class of verbs he identifies[4], and while in some cases this method reveals important facts, it is often obscure and the empirical base is lacking in detail. The following discussion is based mainly on the classifications of Dardjowidjojo and Sneddon which are presented in more detail and with better supporting evidence. I draw attention to parallels with Junus's work where relevant.

The classification of SSAKI is based on the semantic relation between the verb and its associated nominals. That of ALKV is primarily based on the semantic relation between the verb and its subject. There is therefore some overlap in the two classifications, the most significant being the separation of the small group of verbs which are related to perception verbs:

12.

Root	**Gloss**	**ke- -an verb**	**Gloss**
dengar	hear	*kedengaran*	can be heard, audible
dapat	discover	*kedapatan*	be found out, detected
lihat	see	*kelihatan*	can be seen, visible
tahu	know	*ketahuan*	be found out, discovered

Junus (1971) also identifies these as a group. The perceived entity is always the subject:

13. *Rumahnya kelihatan dari sini*
house.3 *ke*.see.*an* from here
'His house can be seen from here.' (ALKV: ex32)

ALKV notes that the perceiver is always implicitly present even when unexpressed, and that these verbs can take a subjectless clause as complement if the subject is human:

14. *Dia kedapatan menggelapkan uang negara*
3SG *ke*.discover.*an* AS.dark.CAUS money state
'He was caught embezzling state funds.' (ALKV: ex29)

The perceiver is often introduced by the preposition *oleh* 'by', but this is only obligatory with pronouns and when the perceiver follows a complement clause. This class is not semantically uniform. *kedapatan* and *ketahuan* carry the implication that the event has adverse consequences for the subject if it is human, whereas *kelihatan* and *kedengaran* have no such implication. They have instead an abilitative meaning (ALKV: 524), and can be followed by an adjectival complement.

[4] Junus uses the term *transformation*, in the sense of Harris (1965) rather than that of Chomsky (1957 and elsewhere).

The other area of significant overlap between the two classifications is in the treatment of verbs with a single argument. ALKV treats all of these in a large group which includes all the verbs having a subject with the thematic role of patient and which suffers adverse consequences. This group also includes a set of verbs with two arguments, and the single argument verbs are subclassified according to the category of their root. SSAVI recognises two classes of one argument verbs and does not group them together, nor does it link either or both classes with any two argument verbs except in the overall grouping. The second group which is discussed in SSAVI may have only a single member, *kehujanan* 'caught in the rain' derived from the root *hujan* 'rain', which can be either a noun or a verb. ALKV treats this as a one argument verb based on a noun, while SSAVI takes it to be based on an argumentless ambient verb.

The remaining verbs, that is the two argument verbs which are not perception verbs, are where the classifications of ALKV and SSAVI diverge. ALKV identifies two groups. Firstly, there are two argument verbs with a subject which suffers adverse consequences from the action referred to by the verb such as *keracunan* 'be poisoned by/with' (group 1d). In this case, the second argument can be either the agent of the action:

15. *Mas Begong kemasukan setan*
brother Begong *ke*.enter.*an* devil
'Mr Begong got possessed by a devil.' (SSAVI: ex6)

or it can be the medium through which the action was carried out:

16. *Saya keracunan air basi*
1SG *ke*.poison.*an* water spoiled
'I was poisoned by bad water.'

In the classification of ALKV, this group is on the same level as all the one-argument verbs. Together they form the first-level class of verbs with patient subjects, and some verbs occur in both one-argument and two-argument groups. The second class of two-argument verbs is a first-level group. The subject is not a patient but rather is the possessor of the other nominal. The subject suffers adverse consequences as a result of something happening to the second argument:

17. *Ali kematian ayahnya*
Ali *ke*.die.*an* father.3
'Ali suffered the death of his father.' (ALKV: ex20)

Clauses with verbs of this type can be (partially) paraphrased by a clause containing the root verb and the two nominals in a single DP as possessor and possessed:

18. *Ayah Ali mati*
father Ali die
'Ali's father died.' (ALKV: ex21)

This relationship was also noted by Junus (1971). ALKV divides this class into two subclasses on the basis of the thematic role of the second argument. In one subclass, this argument is affected by the action as in example 13, and is termed a patient in ALKV (group 2a). In the other subclass, the second argument is transferred by the action and is therefore termed a theme by ALKV (group 2b). All the verbs of this second class denote stealing:

19. *Saya* *kecurian* *sepeda*
1SG *ke*.steal.*an* bicycle
'I had my bike stolen.' (ALKV: ex22)

The partial paraphrase for this subclass uses an undergoer subject verb in the 'agentless passive' clause type. Again the possession relation is made explicit in the paraphrase:

20. *Sepeda* *saya* *dicuri*
bicycle 1SG US.steal
'My bike was stolen.' (ALKV: ex23)

This group is not identified at all by Junus (1971).

SSAVI also identifies three groups of two-argument verbs, other than the perception verbs, but they do not correspond exactly with the ALKV groupings. For SSAVI, the subject of all three groups is a patient. In one case, the second argument is an agent; this group includes words such as *kemasukan* (see example 15 above) and corresponds to part of ALKV's group 1d. For the other two groups identified by SSAVI, both arguments are claimed to be patients[5]. The difference between the two groups lies in the nature of the verb. In the first case, the verb denotes a process and in the second it denotes an action. The second group corresponds closely to ALKV's group 2b, but the examples given in SSAVI suggest that its first group corresponds to part of ALKV's group 1d. Nothing in SSAVI's classification corresponds to ALKV's group 2a, but ALKV's group 1d is not a unitary group in the classification of SSAVI. SSAVI makes no claim for any relationships between the six groups it identifies, whereas ALKV's classification is hierarchical. Figure 3.1 sets out these relations in a graphical format.

[5] SSAVI's terminology is based on that of Chafe (1970) and the extension of term *patient* may not be identical to that of the same term in ALKV.

ALKV			SSAVI
1 a	*kelaparan* 'suffer hunger'	*kelaparan* *kehujanan salju*	1 5[1]
b	*kehujanan* 'be rained on'		
c	*ketiduran* 'oversleep'		
d	*kejatuhan* 'be fallen on'		2
	kemasukan 'be possessed by'		6
2 a	*kematian* 'suffer the death of'		
b	*kecurian* 'have stolen'		3
3	*kedengaran* 'be audible'		4

Table 3.1 - Classifications of *ke- -an* verbs

1. SSAVI treats *kehujanan salju* 'to be snowed on' as a compound verb. ALKV argues against this position.

The analysis to be developed below will treat *ke- -an* verbs as complex predicates, and will argue that the semantic relations between the base predicate and its arguments are crucial to determining the argument structure of the *ke- -an* verb. Therefore it is important to retain all the distinctions which can be observed in the data, and SSAVI's observation that some two argument *ke- -an* verbs take an agent as the second argument will be significant. On the other hand, the category of the base of one argument verbs is not important; if the base assigns a thematic role it must be compatible with that assigned by the other part of the complex predicate and if the base assigns no role, only that assigned by the other predicate will be present. I therefore assume the following set of possibilities[6]:

21.

Number of Arguments	**Argument 1**	**Argument 2**
1 argument	Patient	
2 argument	Patient	Theme
	Patient	Agent
	Possessor	Patient
	Possessor	Theme
	Theme	Perceiver

One fact is made clear in this tabulation: although almost all *ke- -an* verbs imply that the subject is affected by the event, only two groups of the *ke- -an* verbs have thematic profiles which one would normally consider typical of passive verbs: the third and the last of the possibilities listed in 21. Therefore, in addition to the syntactic reasons for rejecting passive as an appropriate label for these verbs, it is also not particularly appropriate on semantic grounds.

3.2 Analyses of clauses with *ke- -an* verbs

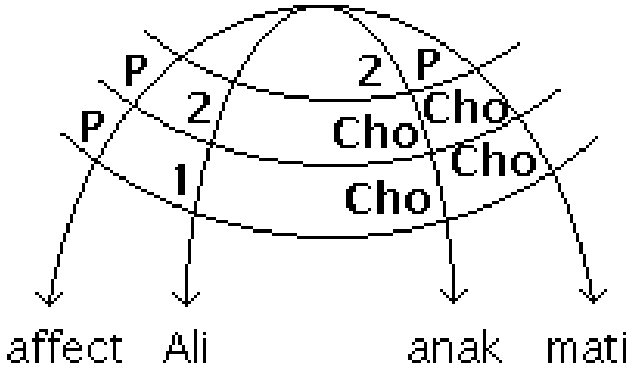
Only two of the sources discussed in the previous section go beyond offering a taxonomy of *ke- -an* verbs, and in each case the analysis that is offered is based on the idea that these verbs have a dual nature. The framework employed by SSAVI treats the relation between the verb and each of its arguments as semantically distinct. The verb need not have identical semantic characteristics in its relation to each argument. Therefore, for each class except the single argument verbs, SSAVI has two semantically distinct verbs in the representation of each clause. One of these is the root and the other is the root plus the circumfix. For example, SSAVI analyses the verb *kejatuhan* 'to be fallen on' as consisting of the verb *jatuh* which is a process and has a patient which is affected by the process (the object which falls), and the verb *kejatuhan* which is a state and has also a patient which is the entity adversely affected by the event. This analysis makes no claim about the syntactic relations between the verb(s) and the arguments; only semantic relations are considered.

[6] Patient here is an entity directly affected by the event and Theme is an entity which participates in the event without causing it or being affected by it.

3.2.1 The analysis of Kana (1986)

Kana (1986, chapter 6) takes essentially the same idea but presents it formally within the framework of Relational Grammar. Clauses with *ke- -an* verbs have a single verb in their surface string, but have two predicates in some level of representation. In RG terms, they are examples of clause union. The circumfix *ke- -an* is a predicate in its own right in this analysis, having a meaning something like 'befall, affect adversely' (Kana 1986: 185). The root of the verb and its arguments constitutes the inner predicate sector, and *ke- -an* is the outer predicate, initializing relations in a later stratum. The relational diagram for a typical clause is as follows[7]:

22.



23. *Ali kematian anak*
Ali AFFECT-die child
'Ali had a child die on him.' (Kana 1986: 196, ex2' [original gloss])

The predicate *mati* 'die' is initialized in the innermost stratum. It is an unaccusative verb (according to Kana - see discussion below and in section 4.3) and therefore it also initializes a 2 arc. In the next stratum, the union predicate *ke- -an* is initialized, forcing the base predicate *mati* into chômage. The union predicate also initializes a 2 arc and the Stratal Uniqueness Law (Perlmutter and Postal 1983) does not allow two arcs to bear the same relation in a single stratum. Therefore the 2 initialized by the base predicate is also forced into chômage. The Final 1 Law (Perlmutter and Postal 1983) requires every clause to have a 1 arc (a subject) in its final stratum, and this forces the 2 initialized by the union predicate to be revalued in the final stratum.

Kana exhaustively treats the various arrays of relations which are possible in the initial stratum given her assumptions. There are seven possibilities, two of which are unattested in Indonesian according to Kana. One of these, that a single nominal heads two arcs in the initial stratum (multiattachment in RG terminology) does not occur in the language at all according to Kana and this possibility is ignored. Another possible initial stratum, with an unergative verb apparently does not exist in Indonesian either. However, this claim rests on Kana's analysis of intransitive verbs (1986 chapter 3), which is problematic. I return to this issue below in section 3.2.2.2

[7] This discussion follows that of Kana 1986 in assuming that clause unions are monoclausal, multi-predicate structures (Davies and Rosen 1988), rather than multiclausal structures. As far as I am aware, nothing in the discussion hinges on this choice.

Base Predicate Type	Initial Relations	Relation to union 2[1]	Final relation	Example
Predicate only	Ø		Ø	*kehujanan*
Unaccusative	2	= 2^U	1	*ketiduran*
	2	≠ 2^U	2-Chômeur	*kematian*
Transitive[2]	(1) 2	≠ 2^U	2-Chômeur	*kecurian*
Transitive	1 2	2 = 2^U	1-Chômeur 1	*kedengaran*

Table 3.2 - Types of *ke- -an* clause according to Kana (1986)

1. 2^U = union nominal, initial 2 of union stratum
2. (1) = unspecified 1 argument, semantically present but not expressed

(and see also section 4.3). The argument initialized by the union predicate *ke- -an* will always be the final 1, therefore the interest of Kana's analysis is in the fate of any arguments initialized in the inner stratum. The possibilities are summarised in Table 3.2 (I subdivide one of Kana's categories for clarity). Various aspects of Kana's analysis are made obvious in this presentation. Firstly, an argument of the base predicate can only be the final 1 when it is identified with the 2 initialized by the union predicate. Any other argument of the base predicate will be a chômeur. Secondly, it is claimed that two types of chômeur occur in the final strata of various clause types. The evidence for this claim is examined in section 3.2.2.1.

There are three areas in which Kana's analysis is problematic. Firstly, the claim that there are two possible second arguments which occur with *ke- -an* verbs is dubious. Secondly, the claim that unergative verbs cannot be the predicate of the innermost stratum is unmotivated and apparently contradicted by the data. And finally, Kana makes no attempt to capture the fact that in many cases the subject of the *ke- -an* verb is in a possession relation with the second argument. I now deal with each of these problems in turn.

3.2.2 Problems with Kana's analysis

3.2.2.1 THE SECOND ARGUMENT

Kana divides the properties of the second argument of two-argument *ke- -an* clauses into two groups (1986: 200-201). The first group includes two properties, that the position of the argument is postverbal, and that it has no preposition. These are properties of 2s in transitive clauses for Kana. The second group of properties includes the fact that the second argument cannot be a bound pronoun:

24. **Ali kematiannya*
Ali *ke*.die.*an*.3
(FOR: Ali had him die.' (Kana 1986: 200, ex30)

and the various properties associated with the fact that these verbs do not participate in the transitive system described in chapter 2. Thus, they cannot occur with the prefix *meN-*:

25. **Ali mengematian anaknya*
Ali AS.*ke*.die.*an* child.3
(FOR: Ali had his son die on him.') (Kana 1986: 200, ex31b)

or any of the other clause types discussed in chapter 2. On this basis, Kana suggests that these arguments should be taken to be 2-chômeurs: they have some of the properties expected of final 2s, but they also lack some crucial properties. There are however some *ke- -an* verbs which take the preposition *oleh* before their second argument (actually only three verbs are listed by Kana for this group: *kelihatan*, *kedengaran*, and *kedapatan*). This preposition is one means of expressing the actor in an undergoer subject clause. Such actors are 1-chômeurs in Kana's analysis (1986, chapter 4) and the second arguments of these *ke- -an* verbs must be the same.

For reasons discussed in chapter 1, I do not consider some of Kana's evidence on this point relevant, and on some other points her data is incomplete. The discussion of section 2.2.2.1 has shown that distribution of bound pronouns is conditioned by complex factors and it is not clear yet whether *ke- -an* verbs should be expected to allow them. This point will be clarified in chapter 5, when the distribution of bound pronouns is established. Postverbal arguments not introduced by a preposition may be direct arguments in Indonesian, but are not necessarily so (see argumentation in section 2.5.4). The most important syntactic tests are whether or not the DP can float a quantifier, and whether it can be extracted to head a relative clause. The second arguments of all two-argument *ke- -an* verbs fail these tests. For each type of two-argument verb, the second argument cannot float a quantifier:

26. **Saya kecurian mobil-mobil itu kemarin semua*
1SG *ke*.steal.*an* car.DUP that yesterday all
(FOR: I had all those cars stolen yesterday.)

27. **Saya kematian anak-anak kemarin semua*
1SG *ke*.die.*an* child.DUP yesterday all
(FOR: I suffered the death of all my children yesterday.'

28. **Mobil itu kelihatan polisi kemarin semua*
car that *ke*.see.*an* police yesterday all
(FOR: The car was visible to all the police yesterday.)

And for each type, the second argument cannot head a relative clause:

29. **mobil yang saya kecurian*
car REL 1SG *ke*.steal.*an*
(FOR: 'the car that I had stolen')

30. **anak yang saya kematian*
child REL 1SG *ke*.die.*an*
(FOR: 'the child that I had die')

31. **polisi yang mobil kelihatan*
police REL car *ke*.see.*an*
(FOR: 'the police to whom the car was visible')

The second argument also cannot be fronted in the constructions discussed in section 2.2.2.2:

32. **Mobil itu, saya kecurian*
car that 1SG *ke*.steal.*an*
(FOR: 'That car, I had stolen (from me).')

33. **Mobil itu saya yang kecurian*
car that 1SG REL *ke*.steal.*an*
(FOR: That car, it was me who had it stolen.)

This evidence suggests that not only is the second argument of these verbs not a direct argument, but also that it is not available to any syntactic processes.

Kana's description of the verbs which according to her have a 1-chômeur as their second argument is also misleading. Her statement is the following:

> For examples such as 5 [an example with *kedengaran*. SM], the postverbal nominal is preceded by *oleh*, the regular marker of 1-chômeurs in passives.

But the preposition is not obligatory in all cases. It is obligatory if the second argument is a pronoun, or if the second argument does not immediately follow the verb (ALKV: 524-525):

34. *Dia ketahuan (oleh) polisi mencuri barang itu*
3SG *ke*.know.*an* (by) police AS.steal goods that
OR: *Dia ketahuan mencuri barang itu oleh polisi*
3SG *ke*.know.*an* AS.steal goods that by police
'He was caught stealing those goods by the police.' (ALKV: ex33

35. *Mobil itu kedengaran *(oleh) saya*
car that *ke*.hear.*an* by 1SG
'The car could be heard by me.'

These conditions have a great deal in common with those that restrict the appearance of actor DPs in undergoer subject clauses - see section 2.5 for details. It is thus by no means as clear as Kana claims that 1-chômeurs and 2-chômeurs display different syntactic behaviour. It is also the case that *oleh* may be used with the second argument of some other *ke- -an* verbs. ALKV has the following example:

36. *Apalagi mereka pun keduluan oleh Speed*
furthermore 3PL EMPH *ke*.precede.*an* by Speed
'Furthermore, they were beaten to it by Speed.' (ALKV: ex36)

and Verhaar (1984b) gives examples with *keracunan* 'be poisoned', *kebanjiran* 'be flooded' and *ketinggalan* 'be left behind'[8].

One additional point can be made about the behaviour of the second arguments of *ke- -an* verbs. It is only the perception verb group which allow the second argument to be moved away from the verb at all. For the other two-argument types, the intervention of any material (except for *oleh* in the cases discussed above) between the verb and the second argument is impossible:

[8] Verhaar (1984b) claims that in the case of *ketinggalan* the preposition is possible because the second argument is an experiencer. For him, this verb then groups with *kelihatan* and the perception verbs. The suggestion seems semantically unconvincing.

37. **Saya kecurian dulu sepeda*
1SG *ke*.steal.*an* before bicycle
(FOR: 'I had my bike stolen previously.')

38. **Dia kedatangan dulu ibunya*
3SG *ke*.come.*an* before mother.3
(FOR: 'He was visited unexpectedly by his mother previously.')

Again, this is reminiscent of the behaviour of actor DPs in undergoer subject clauses. To conclude the discussion of this point, the evidence shows that the second argument of the *ke- -an* perception verbs has different properties to those of the other verbs, but these properties are not so different as to allow us to conclude on syntactic grounds that the arguments have a different status.

3.2.2.2 RESTRICTIONS ON BASE VERBS

Kana claims that only unaccusative intransitive verbs are possible as the base predicate with which *ke- -an* combines (1986: 208-211), although she can find no motivation for this restriction. There are however apparent counter-examples to this claim. SSAVI classifies the verbs *kedatangan* 'be visited´and *kemasukan* 'be entered, be possessed (by a devil)' as having an agent as second argument, and they are apparently derived from the intransitive verbs *datang* 'come, arrive' and *masuk* 'enter, go into' which intuitively have agentive subjects. Kana bases her classification of Indonesian intransitives on morphological criteria (1986, chapter 3, see also discussion in section 4.3) and notes that some Indonesian unaccusative verbs, including the two under discussion here, seem to have actor arguments. Relational Grammar after Rosen (1984) allows such mismatches, and Kana does not consider them as invalidating her hypothesis. The two most important tests that Kana uses for identifying unaccusatives are *-kan* suffixation (1986: 55-68) and *peN- -an* nominalizations (1986: 68-77)[9]. *datang* and *masuk* do display behaviour similar to that of Kana's other unaccusatives on these tests. Both use *-kan* as a causative suffix rather than an applicative:

39. *Polisi mendatangkannya*
police AS.come.CAUS.3
'The police had him brought in.' (E&S: 132)

40. *Ia memasukkan barang di lemari*
3SG AS.enter.CAUS goods LOC cupboard
'She put the things in the cupboard.' (E&S: 364)

And *masuk* has a derived nominal with the *peN- - an* affix:

[9] The capital N in *peN-* indicates an underspecified nasal segment which exhibits the same allomorphy as that in the actor subject prefix *meN-*. The two prefixes are undoubtedly linked historically.

41. *pemasukan* 'entering, registration, import, introduction'

Against this, both of these verbs allowed agent-oriented adverbs to occur with them:

42. *Dia datang ke Australia dengan sengaja*
3SG come to Australia with deliberate
'He came to Australia deliberately.'

I will argue below that this type of *ke- -an* verb is in fact derived from a transitive base predicate, with deletion of an applicative suffix in the morphological component. On this basis, Kana's claim that there are no unergative verbs which serve as base for *ke- -an* derivations is correct, but still unmotivated. A semantic reason will be suggested for the constraint in section 3.3.1.

3.2.2.3 POSSESSOR SUBJECTS

Another weakness in Kana's treatment is also revealed by Table 3.2. The two cases in which the 2 argument of the base predicate is not identical to the argument initialized by the union predicate *ke- -an* coincide with the cases in which the final subject is the possessor of the entity affected by the base predicate. The generalisation is that there are two possible derivations for a base predicate which takes a patient or theme argument. Either the undergoer of the base is identical to the undergoer of the union predicate and it will be the final subject, or it is not identical to the union nominal, in which case it must be the possessor of that entity. Kana discusses the possibility of a possessor ascension analysis for the *ke- -an* verbs of this type, but rejects it because it cannot be used for all the *ke- -an* verbs (1986: 225-226). But in the cases where a possessor is the final subject, the relationship between the two entities is an entailment of the *ke- -an* verb. For example, with the verb *kecurian* 'to have stolen', it is entailed that the stolen object was in the possession of the final subject at the time of the stealing event. It is possible to utter the following sentence:

43. *Saya kecurian uang itu, tapi bukan uang saya*
1SG *ke*.steal.*an* money that but NEG money 1SG
'I had the money stolen, but it was not my money.'

but it must still be the case that the money was stolen from the speaker; it was in their possession at the relevant time. Kana's analysis does not capture this fact about the semantic content of some *ke- -an* verbs. However, the generalisation just noted makes it possible to delimit precisely the cases in which possession is a relevant factor, and to characterise them semantically. In every case in which the base predicate takes a patient argument, the final subject of the *ke- -an* verb is affected by the event denoted by the base, either directly by being identical to the original patient, or indirectly by being the possessor of the original patient or theme. It is clearly desirable for an analysis of these verbs to capture this generalisation.

3.3.3.4 SUMMARY

This section has argued that Kana's (1986) analysis of *ke- -an* verbs as clause unions, whilst based on a correct intuition, is deficient in three respects. Firstly, the evidence for the existence of two distinct grammatical relations among the second arguments of *ke- -an* verbs is not reliable. While there is a definite semantic difference between the second arguments of various verbs, they should nevertheless be given a unified syntactic treatment if possible. Secondly, the syntactic evidence for the status of certain verbs as unaccusative is not convincing enough to override the semantic intuition that a verb like *datang* 'come' has an agent argument. And lastly, Kana's analysis misses important semantic generalisations about the correspondence between the argument array of the base predicate and that of the derived verb. In the following section, I propose an analysis in the LFG framework which attempts to remedy all these defects.

3.3 A new analysis

The analysis presented here adopts the insight that *ke- -an* verbs are the result of the composition of two predicates. The principle of lexical integrity is upheld by LFG, and therefore the composition of the base predicate and *ke- -an* must be a lexical process in an LFG analysis, rather than being a syntactic operation as in the Relational Grammar analysis (see Alsina 1997 and Andrews & Manning 1999 for LFG treatments of complex predicates). Nevertheless, the essential insight is the same. The discussion of the previous section suggests a better sub-classification of the *ke- -an* verbs, based on the semantics of the two predicates and of their arguments, may reveal more accurate semantic generalisations which reveal how that composition operates. Section 3.3.1 presents arguments that the cases where the nature of the base predicate is not clear can be resolved by assuming that derived verbs with the applicative suffix *-i* can be the base for *ke- -an* derivation, but that the suffix is deleted in the process. Section 3.3.2 sets out in detail how the argument structure of *ke- -an* verbs is derived from that of the base predicate on semantic principles, including a discussion of the complex case of the verb *ketinggalan* 'be left behind'. This section concludes with a proposal for the lexical entry of the predicate *ke- -an*. Section 3.3.3 examines the syntax of clauses with *ke- -an* verbs, paying particular attention to the question of what grammatical function is assigned to the second argument of these verbs. I agree with the explicit characterisation of the *ke- -an* morpheme given by SSAVI and Kana (1986) as having a meaning of 'undergo, affect' with a strong implication that the effects are adverse[10]. I also agree with Kana that the *ke- -an* morpheme constitutes a predicate in its own right, and that it takes an undergoer argument as its subject[11]. We would immediately predict from general principles that such a predicate ought to be intransitive: by virtue of its semantics, this predicate has an argument from the bottom of the thematic

[10] The other sources cited accept some version of this position implicitly.

[11] The existence of the perception group requires that the statement be in this form. If these verbs are excepted, a stronger claim can be made: *ke- -an* takes an affected patient as its subject.

hierarchy, therefore any second argument associated with it would necessarily be more thematically prominent and would be expected to link to the subject grammatical function. The evidence discussed in section 3.3.3 will confirm this prediction, showing that the second argument must be analysed as an OBL_θ in the LFG theory of grammatical functions.

3.3.1 What are the base predicates?

In order to analyse the relation between the argument structure of the base predicates which derive *ke- -an* verbs and the final argument structure, it is first necessary to establish exactly what the base predicate is in at least some cases. Where the base takes no arguments, as in the case of nouns, there is no issue. Nor is there a problem in the case of one argument base predicates such as adjectives and a few verbs. However, what is not immediately clear is whether certain *ke- -an* verbs derive from a root word, or whether there is an intermediate derived predicate which feeds *ke- -an* affixation. It is always the root which is affixed with *ke- -an*, but I will argue here that there is evidence in at least one group of verbs that an intermediate stage is involved. The argument is clearest for the several *ke- -an* verbs which have alternative argument-taking properties. For example, the root *racun* 'poison' is a noun and derives a *ke- -an* verb which has only a subject. This is as would be predicted: a root which takes no arguments combines with another predicate which takes at least a subject, and the result is a one-argument verb:

44. *Saya* *keracunan*
1SG *ke*.poison.*an*
'I was poisoned.'

But this verb can also appear in a clause with two arguments:

45. *Dia* *keracunan* *permen* *basi*
3SG *ke*.posion.*an* candy spoiled
'He was poisoned by spoiled sweets.' (ALKV: ex16)

The semantics of the one argument clause are exactly parallel to those of other clauses with *ke- -an* verbs based on nouns. A literal rendering might be 'He was adversely affected by poison', just as *kemalaman*, from the noun *malam* 'night' means 'be adversely affected by night (be overtaken by night)' and *kesetanan*, from the noun *setan* 'devil', means 'be adversely affected by a devil (be possessed by an evil spirit)'. But the meaning of example 45 is rather different and does not correspond in any obvious way to a structure with a noun base. There is no compound noun structure meaning 'spoiled sweet poison', and even if there was, then it would be expected to be the input to the *ke- -an* affixation because, at least as a nominalising circumfix, *ke- -an* can attach outside compounds:

46. *terus* *terang* *keterusterangan*
straight direct candour, straightforwardness
tidak *jelas* *ketidakjelasan*
neg clear obscurity

There is, however, a derived verb *rancuni* 'poison someone or something' which has an argument structure closely related to that of *keracunan*:

47. *Permen basi meracuninya*
candy spoiled AS.poison.APPL.3
'Spoiled sweets poisoned him.'

Similar alternations can be observed with the roots *hujan* 'rain' and *banjir* 'flood', the data for *hujan* being especially enlightening. The root can be either a noun or an ambient verb which takes no arguments. It also serves as the base for a *ke- -an* verb which can occur with one argument or with two:

48. *Kami kehujanan*
1PL.EXCL *ke*.rain.*an*
'We were caught in the rain.'

49. *Dia kehujanan batu*
3SG *ke*.rain.*an* stone
'He was showered with stones.' (ALKV: ex14)

The two-argument structure in this case cannot be based on a compound noun. The compound *hujan batu* does exist, but it means 'hail'. Additionally, the two-argument structure is almost impossible with a literal meaning because as a base predicate *hujan* takes no subject (Indonesian does not have expletive pronouns):

50. *Hari ini hujan*
day this rain
'It's raining today.'

Therefore, there is no potential candidate for the second argument position except the rain itself, which would be a redundant way of expressing the same meaning as example 48:

51. **Kami kehujanan hujan*
1PL.EXCL *ke*.rain.*an* rain
(FOR: We were rained on by the rain.)

But, as with the root *racun*, there is a derived verb which has an argument structure identical to that of *kehujanan*:

52. *Batu-batu menghujaninya*
stone.DUP AS.rain.APPL.3
'Stones rained on him.'

In fact, all the verbs which ALKV lists in its group 1d, verbs with two arguments of which the first is a patient, appear to be derived from verbs with the suffix *-i*, with one exception discussed in section 3.3.1.2.

There is independent evidence that morphological material can be deleted in Indonesian when a sequence of affixation occurs. For example, the adjective *tinggi* 'high' can be the base for a causative verb *tinggikan* 'raise, heighten'. There is also a derived noun *peninggian* (*peN-tinggi-an*) and

although this appears to take the root adjective as its base, the meaning is 'heightening, raising' which is clearly based on the causative verb even though the suffix *-kan* is not retained. A parallel example is provided by the adjective root *pancar* 'scattered'. The causative verb is *pancarkan* 'broadcast' and the nominalisation is *pemancaran* (*peN-pancar-an*) 'action of broadcasting'. This morphological process is very common in Indonesian, particularly in nominalizations of verbs. IRG (28) lists several examples for the prefix *peN-*, which derives nouns indicating the performer of an action[12]:

53.

Base	Gloss	Caus V	Gloss	Nominal	Gloss
bangkit	rise	*bangkitkan*	generate	*pembangkit*	generator
satu	one	*persatukan*	unify	*pemersatu*	unifier
tenang	calm	*tenangkan*	make calm	*penenang*	sedative
kendali	rein	*kendalikan*	control	*menara pengendali*	control tower

There are also cases where a derived nominal may be ambiguous between a derivation from an unsuffixed verb and a suffixed verb (IRG: 40):

54.

Verb	Gloss	Nominal	Gloss
tembak	shoot (s.o.)	*penembakan*	shooting (of s.o.)
tembakkan	fire (a bullet)	*penembakan*	firing (of a bullet)

and in at least one case, either of two suffixed verbs could be the base, but neither suffix appears in the derived nominal:

55.

Verb	Gloss	Nominal	Gloss
terangi	illuminate	*penerangan*	lighting
terangkan	explain	*penerangan*	explanation

(The root *terang* means 'clear, bright')[13].

A similar point is made by ALKV in relation to two *ke- -an* verbs, *kegelapan* 'overtaken by darkness' and *kepedasan* 'suffer from food which has too much chili'. These two verbs are apparently derived from adjectives, *gelap* 'dark' and *pedas* 'hot'. However, the other *ke- -an* verbs which are derived from adjectives can be replaced by the root adjective without drastic change of meaning:

56. *Dia* *kelaparan*
3SG *ke*.hungry.*an*
'He suffers from hunger.'

57. *Dia* *lapar*
3SG hungry
'He is hungry.'

[12] Some causative formations include the prefix *per-*. *persatukan* is such a verb, see IRG: 98-103 for details.

[13] See also footnotes 6 and 24 of chapter 4 for discussion of other relevant examples.

This is not the case for *gelap* and *pedas*:

58. *Saya* *kegelapan*
1SG *ke*.dark.*an*
'I was overtaken by darkness.'

59. *?Saya* *gelap*
1SG dark
'I am dark.'

For this reason, ALKV suggests that the *ke- -an* verbs are derived from nominalizations of the adjectives, also formed with *ke- -an*, *kegelapan* 'darkness' and *kepedasan* 'heat' with deletion of the doubled affix. It is not therefore implausible to suggest that the applicative suffix *-i* can be deleted from a verb when that verb is the input to the formation of a *ke- -an* verb, and this hypothesis does account for the argument structure which is available to the predicate composition process. The parallel between some *ke- -an* verbs and undergoer subject verbs with the suffix *-i* has been noted by previous studies (ALKV, Cartier 1978, and SSAVI although in that case the relevance of the parallel is denied). The difference between the two possibilities is the adversative meaning which is present in the *ke- -an* clause, otherwise the two are propositionally similar:

60. *Dia* *kedatangan* *wartawan*
3SG *ke*.come.*an* journalist
'He was unexpectedly visited by a journalist.' (ALKV: ex11)

61. *Dia* *didatangi* *wartawan*
3SG US.come.APPL journalist
'He was visited by a journalist.' (ALKV: ex12)

No doubt this is one reason for the term 'adversative passive' being applied to the *ke- -an* clause type.

The arguments of this section have been directed to establishing that one group of *ke- -an* verbs are derived from transitive verbs with the applicative suffix *-i*, even though this suffix does not appear in the derived form. Establishing this has three consequences for the discussion of the following section. Firstly, the issue of whether there are single argument verbs with an agent which function as the base for *ke- -an* verbs is irrelevant: the dubious cases actually involve a base transitive predicate. Secondly, the *ke- -an* verbs in question are the result of the composition of the predicate *ke- -an* and a predicate with two arguments. And thirdly, the second argument of the base predicate has a particular character. While generalisations about the functions of verbal morphology in Indonesian are very hard to maintain, it is possible to say that the applicative suffix *-i* is most often associated with the promotion of a location to the object function (for detailed examination of the question, see Tampubolon 1983 and Voskuil 1996). This is clearly true of the majority of the verbs in ALKV's group 1d. It follows from this, that the objects of transitive verbs with the *-i* suffix are not conceptualised as affected entities. It is certainly possible for them to be affected by the event, but this is

incidental. Their essential role is as the location of the event. A clear example of this can be seen with the verb *jatuhi* 'fall on'. It can be used with an object which, on the basis of real world knowledge, can be said to be affected:

62. *Kelapa itu menjatuhi kepala saya*
coconut that AS.fall.APPL head 1SG
'The coconut fell on my head.'

But another lexical item can be substituted in the object position, and the apparent entailment of being affected is not present:

63. *Kelapa itu menjatuhi atap rumah saya*
coconut that AS.fall.APPL roof house 1SG
'The coconut fell on the roof of my house.'

Given that the roof in question is of some robust material, we do not assume from such an utterance that the roof was affected by the event. This fact about verbs with the suffix *-i* will be important in the following discussion.

3.3.2 Properties of the *ke- -an* predicate

3.3.2.1 DERIVING THE ARGUMENT STRUCTURE

The unusual properties of the predicate *ke- -an* are best seen by considering first the verbs for which a new participant is added to the event, the possessor of the affected entity. There are two groups of these verbs. One group has intransitive bases, the other transitive bases, but in each case the subject of the *ke- -an* verb is a person in some relation to the entity affected by the event such that the person is also affected. The relation can always be syntactically expressed by a possessor-possessed structure, although the actual relationship may not be of that nature[14]. The entity affected by the event denoted by the base predicate is inanimate for almost all these verbs. the exceptions are *kematian* 'suffer the death of' and *keguguran* 'have a miscarriage'[15]. The entity affected by the base predicate in these cases is no longer animate as a result of the event; therefore the same generalisation does in fact apply to these two verbs also. Therefore we can say that the subject of these *ke- -an* verbs is the entity most affected by the event, where a human who experiences adverse effects is treated as more affected than an inanimate patient or theme. I shall refer to the added argument in the following discussion as an *extended patient*.

The *ke- -an* verbs which have an extended patient as subject are an example of external possession in the sense of Payne and Barshi (1999). These scholars define an external possession construction as one in which:

[14] This qualification is intended to cover the case of kinship relations in particular.

[15] *keguguran* never has an expressed second argument. ALKV argues (p521) that the semantic relation between the two arguments of *kematian* is parallel to the relation between the subject of *keguguran* and an understood second argument *janin* 'fetus'. The discussion assumes this position.

> a semantic possessor-possessum relation is expressed by coding the possessor as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum. (3)

External possession constructions with an adversative meaning are attested in at least one other Austronesian language, Tukang Besi (Donohue 1999). Shibatani (1994) argues that external possession should be seen as part of a wider spectrum of constructions, including ethical dative and adversative constructions, which have in common the feature that they treat a participant that is not a semantic argument of the verb as a syntactic argument. He suggests that the non-thematic argument is integrated into the construction on a pragmatic basis, and that some readings, such as inalienable possession, are more easily achieved than others. *ke- -an* verbs do not fit into such an account well. Although some *ke- -an* verbs introduce an argument which is not semantically related to the base predicate, the account presented here claims that such arguments are semantically selected by the *ke- -an* predicate itself: it is a one-argument predicate which takes a patient as its subject. Almost all *ke- -an* verbs are adversative in meaning, and the connection between this type of construction and external possession pointed out by Shibatani is important. It is therefore unsurprising that the one type of verb should allow the other type of meaning in some cases, but the possibility is more syntacticized in Indonesian than in some of the other languages discussed by Shibatani (see for example the comparison of Spanish and Japanese, examples 28 and 29 and related discussion). A case where the nature of the possession relation between two possibilities is significant is discussed in section 3.3.2.3, although even in that case the proposal given here accounts adequately for the contrast.

There is a clear contrast between the *ke- -an* verbs formed from transitive verbs which have affected patients and those which, as argued in section 3.3.1, are based on transitive verbs whose second argument is a location. Where the base predicate is an *-i* suffixed verb, there is no possibility of the affectedness being extended to a possessor. This can be understood by recalling the argument presented in section 3.3.1 to the effect that affectedness was not entailed for the object of such verbs, but it is entailed for the subject of *ke- -an* verbs (with the exception of the perception group). This semantic requirement of the *ke- -an* predicate is exhausted in converting the non-affected second argument of a transitive verb into an affected subject. On the other hand, where there is a fully affected argument already available, the semantic force of the *ke- -an* predicate can be used to add a new affected argument (the *sufferer* in the terminology of ALKV). The proposal here is that the argument structure of *ke- -an* verbs is established on semantic grounds; the most affected entity in the situation described will be selected as subject of the *ke- -an* verb.

The crucial data which confirms that the argument expression of *ke- -an* is semantically driven comes from two verbs both meaning 'to be addicted', *kecanduan* and *ketagihan*. The base forms of both these verbs take the addicted person as subject:

64. *Ia telah mencandu minum kopi*
3SG past AS.addicted drink coffee
'She has become addicted to coffee.' (E&S: 103)

65. *Saya tagih rokok*
1SG addicted cigarette
'I am addicted to smoking.' (E&S: 540)

The *ke- -an* verb has the same subject, the addicted person:

66. *Kamu sudah kecanduan cara roman yang*
2SG past *ke*.addicted.*an* way romantic REL
dilihat di film
US.see LOC film
'You are addicted to the romantic ways which you have seen in the movies. (E&S: 104)

67. *Ia ketagihan madat*
3SG *ke*.addicted.*an* opium
'He is addicted to opium.' (E&S: 541)

This would be highly surprising if the *ke- -an* verb selected its subject on the basis of the grammatical functions assigned by the base predicate. In all other cases where the base predicate has two arguments, the subject of the *ke- -an* verb is either the undergoer or a human related to the undergoer. But in this case, the subject of the *ke- -an* verbs is the actor of the base predicate. This is, however, unsurprising if the subject of the *ke- -an* verb is selected on semantic grounds. For these two verbs, the actor of the base predicate is the entity which suffers adverse effects from the event (more accurately, the state) which the base predicate denotes.

This semantic theory allows us to give a principled account of the complete data. Firstly, there are bases which have no argument of their own. In this case, the derived verb has a single argument which is affected by the referent of the base, for example *malam* 'night' and *kemalaman* 'overtaken by darkness'. Secondly, the base can have a single argument which is the theme of a state, typically an adjectival base. The derived verb has the same single argument with the added meaning that it is affected by the state, for example *lapar* 'hungry' and *kelaparan* 'be affected by hunger'. Thirdly, the base can be a verb with only an affected patient argument. In this case, *ke- -an* can add an argument, an extended patient. Fourthly, the base can be a transitive verb with an affected patient. *ke- -an* can add an extended patient in this case, but it only allows the expression of two arguments and the agent remains unexpressed. This fact can be interpreted in two ways. Either it is a requirement of the *ke- -an* predicate that if the base predicate has a patient, this must be expressed in the final argument structure, or the arguments of the base predicate are ordered with respect to affectedness, and *ke- -an* takes its arguments in order from the lowest on this scale. The only evidence that seems to bear on this question is that two verbs can alternate between expressing only a patient as subject, and having an extended patient as

subject and the original patient as second argument. These are *kebakaran* 'catch fire' and *ketinggalan* 'be left behind':

68. *Rumah saya kebakaran*
house 1SG *ke*.burn.*an*
'My house caught fire.'

69. *Amat kebakaran rumah*
Amat *ke*.burn.*an* house
'Amat's house was burned down.'

70. *Buku saya ketinggalan di perpustakaan*
book 1SG *ke*.leave.*an* LOC library
'I left my book in the library.'
[LIT: My book was unintentionally left in the library.] (ALKV: ex8)

71. *Saya ketinggalan buku di rumah*
1SG *ke*.leave.*an* book LOC house
'I left my book at home.' (ALKV: ex27)

This data is inconclusive but probably tends to support the first interpretation, expression of a patient of the base predicate is part of both structures. The fifth possibility is that the base is transitive, but the undergoer argument is not coded as affected, that is, it is one of the *-i* suffixed verbs discussed above. In this case, the *ke- -an* changes the semantics of the base predicate by making its undergoer an affected patient, and the possibility of an extended patient is not available. Similar arguments apply to two of the perception verbs, *ketahuan* and *kedapatan*. Both of these can plausibly be derived from transitive verbs with the suffix *-i*. *dapat* has various meanings as a root, mainly modal, but also including 'be found', but this is an intransitive verb. *dapati*, on the other hand, is a transitive verb meaning 'find, discover'. Similarly, the root *tahu* means 'know, be cognizant of' and *ketahui* means 'find out, discover'[16]. Both *ketahuan* and *kedapatan* 'always indicate something negative or unfavourable for [the subject]' (AKLV: 524).

The remaining two perception verbs, *kedengaran* 'be audible' and *kelihatan* 'be visible' are the only two *ke- -an* verbs for which the semantic account is less convincing. The root in each case forms an unsuffixed transitive verb:

72. *Dia mendengar suara ibunya*
3SG AS.hear voice mother.3
'She heard her mother's voice.'

73. *Kami sudah melihat film itu*
1PL.EXCL PERF AS.see film that
'We've seen that movie.'

[16] The added morpheme *ke-* in the derived verb is due to the fact that this root is treated as monosyllabic.

The perceived entity is thus treated as a patient rather than a location, as far as verbal morphology is concerned, and in this case we would predict that an extended patient should be possible with the *ke- -an* verbs. This is not the case:

74. **Saya kelihatan mobil*
1SG *ke.see.an* car
(FOR: My car was visible.)

It is unlikely that these two verbs are actually based on *-i* suffixed forms, as E&S do not list a *dengari* at all. Therefore, we are forced to conclude that these two verbs behave like those with a non-affected undergoer even though they do not have the morphology that this behaviour is usually associated with. I assume that this oddity is not unrelated to the fact that the semantic effect of *ke- -an* affixation for these two verbs is different from all other cases. For these verbs, the derived verb has an abilitative rather than an adversative meaning. There is still a semantic distinction between a *ke- -an* verb clause and the closest paraphrase, but it is not the same distinction as in the cases previously discussed:

75. *Musik itu saya dengar*
music that 1SG hear
'The music was heard by me.'

76. *Musik itu kedengaran oleh saya*
music that *ke*.hear.*an* by 1SG
'I could hear the music.' (ALKV: ex34)

I offer no account here of the factors which cause these differences.

This semantic account makes it easy to understand why there are no *ke- -an* verbs with unergative intransitives as base. By definition, such verbs do not have an affected patient in their thematic structure. But the predicate *ke- -an* requires an affected entity as its subject, and such an argument can only be added when the base has no arguments, or when an extended patient is part of the situation. Unergative intransitives meet neither of these conditions. If such a verb were combined with the *ke- -an* predicate, there would be no possible subject which fulfilled the semantic restrictions imposed by the predicate and therefore the lexical composition process must fail.

3.3.2.2 WHAT GRAMMATICAL FUNCTIONS?

3.3.2.2.1 Subject

All the sources mentioned thus far agree on one point, that the DP which normally precedes the verb and whose referent is the entity which suffers adversely as a result of the event, is a subject. The tests described in section 1.1 all point to this conclusion. Thus, this position can be occupied by the third person pronoun *ia*:

77. *ia* *telah* *kecurian* *barang-barangnya*
3SG PERF *ke*.steal.*an* article.DUP.3
'his luggage was stolen from him' (Hiorth 1976: 78)

The DP in this position can be extracted and can head a relative clause:

78. *bayi* *yang* *ketularan* *penyakit* *dari* *para* *pedagang*
baby REL *ke*.infect.*an* disease from PL trader
'a baby that caught a disease from salesmen' (Hiorth 1976: 77)

It is also be the case that this position can be controlled in a complement clause. Some examples are potentially ambiguous:

79. *Aku* *mimpi* *kedatangan* *seorang* *asing*
1SG dream *ke*.come.*an* someone alien
'I dreamt that an unknown person visited me.' (Hiorth 1976: 76)

The gloss provided by Hiorth suggests that this is an example with a complement clause. But it is also marginally possible to construe *kedatangan* here as a noun, meaning 'arrival'. The gloss for the sentence would then be 'I dreamt of the arrival of an unknown person'[17] This particular case is further complicated by the fact that the data discussed in section 4.1 include no example of the verb *mimpi* taking a complement clause; on the other hand they also do not include unequivocal examples of it taking a bare DP complement. Elicited examples can resolve the question easily. The word *kehujanan* has no nominal reading; it only means 'be rained on', and it can occur in a complement clause without an independent subject. Note also that *kehujanan* is negated by *tidak* in the following example, rather than by the nominal negator *bukan*:

80. *Saya* *senang* *tidak* *kehujanan.*
1SG happy NEG *ke*.rain.*an*
'I am happy not to be rained on.'

The adversely affected argument of a *ke- -an* verb can control the reference of the subject of a predicative complement, and this provides additional evidence that such arguments are indeed the subject of their clause. Data which reinforces this argument is presented in section 3.3.2.2.2.

3.3.2.2.2 The status of the second argument - predicative complements

It was argued in section 3.2.2.1 above that the evidence for treating the second arguments of *ke- -an* verbs as being of more than one type was less than compelling. I will therefore assume that all such arguments have the

[17] Native speaker judgment is that this meaning would be expressed by the following sentence:

Aku *memimpikan* *kedatangan* *orang* *asing*
1SG AS.dream.APPL NOM.come.NOM person alien
'I dreamt of the arrival of a stranger.'

same syntactic status, and turn to the question of what exactly that status is. LFG recognises only four grammatical functions that a verb can assign to a nominal argument: SUBJ, OBJ, OBJ_θ and OBL_θ. The data presented in section 3.3.3.1 show that the preverbal argument is the SUBJ of a clause with a *ke- -an* verb. General principles of unification-based grammar prevent an attribute, such as a grammatical function, being assigned more than once in the relevant structure, therefore a clause can only have one SUBJ and the second argument of the *ke- -an* verb is not it. OBJ and OBJ_θ are both terms and it is a property of terms in Indonesian that they can launch floating quantifiers. As discussed in section 3.2.2.1, the second arguments of *ke- -an* verbs do not have this property, or any other property associated with term arguments except for appearing without a preposition. I therefore conclude that these arguments are not either OBJ or OBJ_θs. The only possibility which remains is that the second argument of a *ke- -an* verb is an OBL_θ. This conclusion is supported by evidence from the *ke- -an* verbs which take a predicative complement as an argument.

A number of verbs which serve as the base for *ke- -an* derivations can appear with a complement clause predicated of their subject (an XCOMP). This property is inherited by the derived verb, but in some cases there is a significant change to the control relation which establishes the reference of the missing subject. Some of the base verbs which allow predicative complements have the subject position of the complement controlled by the matrix object, and some have it controlled by the matrix subject. But in all cases, the subject of the derived *ke- -an* verb controls the reference of the subject of the complement. *ke- -an* verbs are formed by a lexical process, and this process affects the control relation, therefore this fact demonstrates that these are examples of lexically induced functional control (Bresnan 1982). It is a consequence of the LFG theory of control that 'the controllers of lexically-induced functional control relations must change under the lexical operations on functional assignments' (Bresnan 1982: 401). This clause type therefore provides additional evidence that the second argument of *ke- -an* verbs is a non-term, that is an OBL_θ.

There are base verbs which can take a SUBJ and an XCOMP as arguments, and in this case there is no change in the control properties as a result of the composition with the *ke- -an* predicate. For example, the verbs *tagih* 'addicted to' and *candu* 'addicted to' are of this type. By a lexical redundancy rule (Bresnan 1982), the following control equation is added to the lexical entries of these verbs: (↑SUBJ) = (↑XCOMP SUBJ). This allows clauses such as the following:

81. *Ali tagih membaca koran*
Ali addicted AS.read newspaper
'Ali is obsessed with reading the paper.' (E&S: 540)

When this verb is combined with the *ke- -an* verb, the subject of the base is the affected entity in the situation and therefore is also the subject of the derived verb. Therefore the control equation is unchanged:

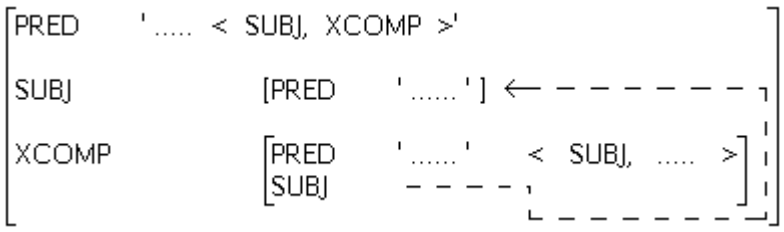
82. *ke- -an* V 'undergo < x : P < ..... > , (y), (XCOMP) >'
tagih V 'addicted < x : XCOMP >'
(↑SUBJ) = (↑XCOMP SUBJ)
ketagihan V 'suffer addiction < x : XCOMP >'
(↑SUBJ) = (↑XCOMP SUBJ)

The resulting verb appears in a clause such as the following:

83. *Dia* *ketagihan* *berjudi*
3SG *ke*.addicted.*an* gamble
'He is addicted to gambling.' (ALKV: ex40)

For both the base verb and the derived verb, the f-structure of the clause is of the following type:

84.



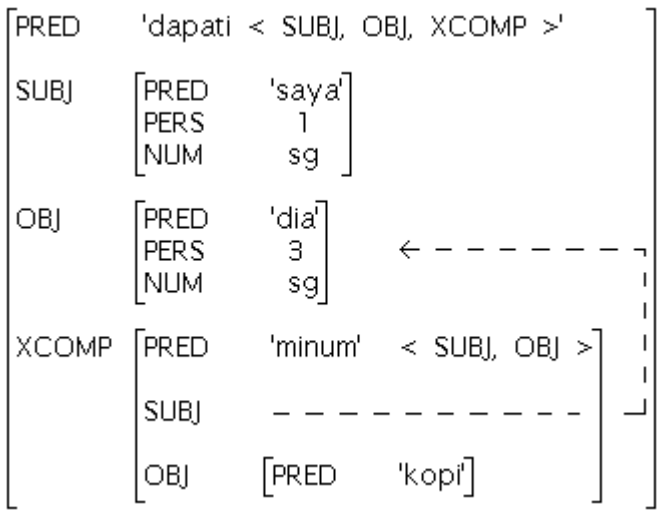
The subjects of the two clause nuclei are identical.

The other *ke- -an* verbs which allow a complement clause are derived from verbs for which the XCOMP subject is controlled by the matrix object. Where the XCOMP appears in the lexical entries of these verbs, it carries the annotation (↑OBJ) = (↑XCOMP SUBJ). This allows a clause such as the following:

85. *Saya* *mendapatinya* *minum* *kopi*
1SG AS.find.3 drink coffee
'I found him drinking coffee.'

with the associated f-structure:

86.



But where this lexical entry has served as the base for a *ke- -an* verb, the control properties of the derived verb are different. In this case, the subject of the derived verb controls the reference of the subject of the complement:

87. *Dia kedapatan menggelapkan uang negara*
3SG *ke*.find.*an* AS.dark.CAUS money state
'He was caught embezzling state funds.'

The *ke- -an* verb must have the control equation (↑SUBJ) = (↑XCOMP SUBJ) with an f-structure of the type illustrated in 84. As noted above, it is a consequence of the LFG theory of control that 'the controllers of lexically-induced functional control relations must change under the lexical operations on functional assignments' (Bresnan 1982: 401). From this, it can be predicted that lexically-induced functional control relations will change under operations such as passive. This is indeed the case for a verb such as *dapati*: when the undergoer-subject form *didapati* is used, control is by the SUBJ, the OBJ of the actor –subject form. the following example only has a reading in which Ali is the coffee-drinker[18]:

88. *Ali didapati minum kopi oleh Siti*
1SG AS.find.3 drink coffee by Siti
'Ali was found by Siti drinking coffee'

The other type of functional control relation allowed by the theory is *constructionally-induced* functional control. As the terminology implies, the relevant control equation is introduced as a functional annotation of a c-structure node in this case. Any such annotation cannot be affected by lexical processes. Therefore we can conclude that a verb such as *dapati* can lexically induce functional control.

Additional properties of functional control can be derived from the LFG theory of control, two of which are relevant to this discussion. Firstly,

[18] The account of Pro-V clauses and *di*-V-*nya* clauses developed in Chapter 2 suggests that the control properties of a verb such as *dapati* might be ambiguous in such clause types. This is a question for further research; related Balinese data are discussed by Arka & Simpson (1998).

constraints on the lexical encoding of semantically restricted functions ensure that they can never be functional controllers (Bresnan 1982)[19]. And secondly, there is an unmarked hierarchy for choice of controller. If the lexical entry of a verb includes the function XCOMP, the controller will be the OBJ if there is one and it will be the SUBJ otherwise. In some cases, it is possible for a *ke- -an* verb to have two nominal arguments and an XCOMP also. If it were the case that the second nominal argument was a term, we would predict that it would control the reference of the subject of the complement clause. But this is not the case; the subject of the *ke- -an* verb controls the complement subject in these cases also:

89. *Dia ketahuan polisi mencuri barang itu*
3SG *ke*.know.*an* police AS.steal goods that
'He was caught stealing those goods by the police.' (AKLV: ex33)

This is therefore additional strong evidence that the second argument of a *ke- -an* verb is an oblique (and also that the preverbal argument is a subject). An alternative explanation might be that the *ke- -an* predicate itself has a lexical specification that its subject will be the controller for any open function which it might take as an argument. But this account would miss the generalisation that an XCOMP is only possible with a *ke- -an* verb if it is possible with the base verb: taking such a complement is a property of the base verb which is inherited in the lexical process which forms *ke- -an* verbs.

3.3.2.3 A LEXICAL ENTRY FOR *KE- -AN*

The predicate *ke- -an* takes as arguments a subject, another predicate with its complete argument structure, and optionally an oblique argument and a clausal complement with a controlled subject position:

90. *ke- -an* V 'undergo < x : P < ..... > , (y), (XCOMP) >'

Following Alsina (1997), I assume that the lexical process of predicate composition involves the identification of arguments of the embedded predicate with arguments of the inclusive predicate. For *ke- -an*, the possibilities for identifying arguments are constrained by the following principles:

[19] Bresnan (1982) treats OBJ2 as a semantically unrestricted function, while Bresnan and Kanerva (1989) and later work treat OBJ_θ as a semantically restricted function. This has consequences for the theory of control as set out in Bresnan (1982), but they are not relevant to the current discussion. Bresnan (1982) gives OBJ2 as the first choice for controller if that function is assigned by a verb; I omit this possibility from the discussion.

91. 1) The argument obligatorily associated with *ke- -an* is a patient and the most affected entity in the situation must be identified with it.

2) A more animate entity in a close relation (typically possessor) to an affected patient is an extended patient, and is the most affected entity in a situation.

3) An affected patient or theme must be identified with an argument of the inclusive predicate.

I assume also that some general constraint against discarding information ensures that all the arguments of the base verb that can be realised will in fact be realised. The control properties of the XCOMP argument fall out from general principles; these were discussed in section 3.3.2.2.2.

The working of the principles set out in 91 can be seen from the following examples. The noun *malam* 'night' has no arguments, therefore the lexical entry for the derived verb *kemalaman* will be the following:

92. *kemalaman* V 'be affected by night < x >'

The subject function for *ke- -an* is obligatory, therefore a new argument is added, and the meaning of that predicate guarantees that this argument will be a patient. The (derived) verb *jatuhi* has the following lexical entry (subscripted numbers are added to assist keeping track of the identity of arguments):

93. *jatuhi* V_{tr} 'fall on < x_1, y_2 >'

The second argument is a location not an affected patient, therefore principles 91.2 and 91.3 do not apply. The location is nevertheless the most affected entity and it therefore becomes subject. The oblique slot of the *ke- -an* verb is available to realise the first argument of the base verb, and the result is:

94. *kejatuhan* V 'be fallen on' < x_2 : y_1 >

The verb *mati* 'die' takes a single argument which is an affected patient. Therefore principle 91.2 is applicable: a more animate entity, in this case a live person, in a possessor-like relation to that patient is an extended patient in this situation and the most affected entity. By principle 91.1, this person will be the subject and the oblique slot will be available to realise the argument of the base predicate:

95. *mati* V_{itr} < x_1 (possessed by y_2) >
96. *kematian* V < x_2 : y_1 >

The other classes of verbs described previously can be derived in similar fashion.

The a-structure of *ke- -an* verbs provides evidence that deriving a verb's valence from the interaction of its a-structure and the principles of LMT does not work in all cases. If the lexical entry of *ke- -an* does not represent the fact that it is an intransitive verb with a single direct argument, the version of a-structure presented in Bresnan (2001a) and LMT make two incorrect

predictions, the first of which also depends on the analysis presented in chapter 2. Firstly, where the base predicate has an agent and a patient as arguments, the a-structure of the complete predicate would be the following, with linking to semantic roles included for clarity:

97.

	(agent)	(patient)
ke-V-*an*	< x,	y >
		|
		SUBJ

This configuration meets the conditions for morphological adjunction of a pronoun as a direct argument: the verb has lexically specified linking of its second argument to the SUBJ GF (chapter 2, example 153). But this is not possible, as seen in example 24. The a-structure therefore cannot be as shown in 97. Secondly, in the case where an extended patient and the patient of the base predicate are the arguments of the *ke- -an* verb, the a-structure would be:

98.

	(extended patient)	(patient)
ke-V-*an*	< x,	y >
	|	
	SUBJ	

Indonesian has second objects, and I have assumed that these are treated in the same way as in English. Therefore, we would expect the second patient-like argument in this configuration to be assigned the feature [*+o*], and be mapped to the GF OBJ_θ, giving a verb with the GFs SUBJ and OBJ_θ. But the evidence presented in section 3.2.2.1 has made it clear that the second argument of such verbs is not a term, therefore not an OBJ_θ. Neither of these cases is problematic if the information that *ke- -an* has a single term argument is represented in the a-structure. Only one term position is available, and principle 91.1 specifies which argument will be mapped to it. Any other argument can only be realised as an oblique. I take this to be important evidence in support of the version of a-structure advocated by Manning (1996a).

3.3.2.4 THE CASE OF *KETINGGALAN*

ALKV groups the verb *ketinggalan* 'be left behind' with the verbs discussed in the section 3.3.1.1, those that have two arguments of which the first is a patient. But the range of clause types in which this verb can appear shows that this classification is unsatisfactory. Verhaar (1984b) treats *ketinggalan* as belonging to the same group as the perception verbs. This is semantically implausible, but *ketinggalan* can occur with a complement clause with controlled subject, which is a point of similarity. In fact, close examination of the data shows that this verb can appear in almost every clausal structure attested for *ke- -an* verbs. The account developed above forces the conclusion that this can only occur if there is more than one predicate which is serving as the base for *ke- -an* derivation, and this section will argue that this is indeed the correct solution.

The most immediate and striking contrast is between the following two possibilities. Firstly, *ketinggalan* can occur with an extended patient as subject, but also allows the true patient to occur in that position:

99. *Saya ketinggalan buku di rumah*
1SG *ke*.leave.*an* book LOC house
'I left my book at home.' [LIT: I suffered the leaving of my book at home.] (ALKV: ex27)

100. *Buku saya ketinggalan di perpustakaan*
book 1SG *ke*.leave.*an* LOC library
'I left my book in the library.'
[LIT: My book was unintentionally left in the library.] (ALKV ex8)

The same verb can be used in a structure parallel to example 99, but without the possibility of the alternative:

101. *Saya ketinggalan bis*
1SG *ke*.leave.*an* bus
'I missed the bus.' (LIT: I was left behind by the bus.) (ALKV: ex17)

102. **Bis saya ketinggalan*

This contrast might be argued to be due to the factors discussed by Shibatani (1994). The relation between a book and its possessor, and a bus and one of its passengers, even when we can say *my bus*, is sufficiently different that one case might allow the possessor to be integrated into the scene and the other might not. However, in this second case, the second argument is an actor: it is the bus that leaves and thereby affects the subject. Following the arguments made previously, an extended patient is only possible when the base predicate has an affected patient as one of its arguments, and therefore the base predicate for examples 99 and 100 must be of this type. It will have the entity left behind as patient. On the other hand, the base predicate in example 101 will not have an affected patient in its argument structure. An appeal to pragmatic factors is not necessary, even in this case.

There are two possible bases for the *ketinggalan* seen in examples 99 and 100. The intransitive verb *tinggal* means 'stay, be left', it plausibly has a patient as its sole argument, and the meaning is compatible. On this account, the literal meaning of example 99 would be 'I suffered because the book stayed at home.' The alternative is that the causative verb *tinggalkan* is the base. This verb is normally translated as 'leave behind', the more literal meaning is 'cause to stay, cause to be left'. This is a transitive verb with an agent, the causer, as well as the affected patient inherited from the root verb, and if this were the base predicate for example 99, it would mean that the subject was an agent as well as an extended patient. This contradiction is sufficient reason to reject this analysis, and to prefer the first alternative. This means, in turn, that, despite the translation, the attribution of agency to the subject of this sentence is an implicature only. There is a structure in which *ketinggalan* does seem to have *tinggalkan* as the base verb. It is possible to specify the agent as the second argument:

103. *Sebagian barangnya ketinggalan olehnya di hotel*
part goods.3 *ke*.leave.*an* by.3 LOC hotel
'He had (unintentionally) left part of his luggage at the hotel.' (ALKV: ex39)

In this case, the causative verb is the only possible base, with deletion of the derivational morphology as argued for previously.

It remains to account for the *ketinggalan* seen in example 101. I argued above that this verb cannot have an affected patient in the argument structure of its base predicate. Therefore neither the intransitive *tinggal* nor the causative *tinggalkan* are possible bases. The obvious alternative is to look for an *-i* suffixed verb which would make this *ketinggalan* quite parallel to the other verbs in ALKV's group 1d (example 101 appears in the discussion of that group in the source). There is a verb *tinggali*, but it means 'live in, stay at' or 'bequeath', none of which seems relevant here. However, *tinggalkan* has a second meaning, which is 'leave (a location)':

104. *Ia meninggalkan negerinya*
3SG AS.leave.APPL? country.3
'He left his country.' (E&S: 578)

The questions as to how exactly this verb is derived, and whether the *-kan* suffix in this case is an applicative are beyond the scope of this work. We only note that there are issues to be explored here. For present purposes, it is sufficient to note that this use of the verb *tinggalkan* seems to be the base which we are looking for. The second argument is a location rather than an affected patient, and this brings this *ketinggalan* into line with the other verbs with a location as second argument, and gives a plausible reading for example 83 in which the subject is (metonymically) a location: 'I was affected because the bus left my location'.

This discussion does not exhaust the possibilities of *ketinggalan*. This verb also allows a complement clause:

105. *Petani tidak ketinggalan menghadapi kesamaan*
farmer NEG *ke*.leave.*an* AS.across.APPL *ke*.same.*an*
selera terhadap aneka komoditas sebab
appetite towards variety commodity because
kemampuan beli petani sudah demikian hebat
ke.able.*an* buy farmer already like.that intense
'Farmers have not been left behind in facing the similarity of demand for a variety of commodities because the buying capacity of farmers is already so great.' (McKay)

This follows from the general principles already discussed. The verb *tinggalkan* can take a clausal complement with object control:

106. *Saya meninggalkan Ali menulis surat-surat itu*
1SG AS.leave.CAUS Ali AS.write letter.DUP that
'I left Ali writing those letters.'

As expected, when a verb with this argument structure undergoes predicate composition with *ke- -an*, the patient of the base becomes subject and the new verb has subject control of the complement subject function.

3.3.3 Obliques and adjunction

It is true in general in Indonesian that requiring a preposition is definitional for oblique arguments, and as discussed in section 3.2.2.1, the second argument of a *ke- -an* verb is rarely introduced by a preposition. It was also remarked in that section that these second arguments had properties that were reminiscent of those of actor DPs in undergoer subject clauses. Section 2.5.4 assumed rather stipulatively that an actor DP in a clause with a verb with the undergoer prefix *di-* was an oblique argument, even though the preposition *oleh* was not always used to introduce it. We can now re-examine this case, and in comparing it with the case of *ke- -an* verbs, offer a more complete account.

There are three significant similarities between the two cases. Firstly, in order to appear without a preposition the DP must be adjacent to the verb. This fact about word order, particularly in relation to ditransitive verbs was the essential motivation for the suggestion in section 2.5.4 that such DPs are adjoined to the verb. Secondly, in the *di*-V clause type, it is always possible to include the preposition *oleh* 'by', which marks agency[20]. And in *ke- -an* clauses the preposition *oleh* can also introduce the second argument in some cases. This is rather rare in this clause type, but, if we consider where it is possible and where it is not, the explanation is straightforward. The discussion in ALKV (pp524-525) and the examples given in section 3.2.2.1 make it clear that the preposition can be used whenever the second argument is treated as an actor. Note that it is necessary to use the concept actor here rather than the more restrictive agent, as verbs such as *kelihatan* 'be visible' allow *oleh* with their second argument. But this is not surprising, since in its transitive form, a verb such as *lihat* can use *oleh* to mark the actor in undergoer subject clauses:

107.	*Pak Bomo*	*dilihat*	*kemarin*	*oleh*	*Ali*
	Mr Bomo	US.see	yesterday	by	Ali
	'Mr Bomo was seen yesterday by Ali.'				

The surprising fact about the use of *oleh* with *ke- -an* verbs is that it cannot be used to introduce the second argument of the verbs which SSAVI identifies as having an agent as second argument. SSAVI notes this fact as odd: that non-agentive base predicates such as *lihat* 'see' and *tahu* 'know' derive verbs which allow *oleh*, but the apparently agentive *masuk* 'enter' does not:

---

[20] It is very clear in Indonesian that this is the primary function of the preposition *oleh,* which does not have the locational meaning which *by* has in English. Its only other meaning is 'because'; this is not semantically unrelated and is available in rather different structures, e.g.

(i)	*Aku*	*merasa*	*kesal*	*oleh*	*sikapnya*	*itu*
	1SG	AS.feel	angry	because	attitude.3	that
	'I felt angry at his attitude.'					

108. *Setan masuk ke Mas Bengong*
devil enter in brother Bengong
'The devil went into Begong.'

109. *Mas Bengong kemasukan (*oleh) setan*
brother Bengong *ke*.enter.*an* by devil
'Bengong got possessed by a devil.' (SSAVI: ex6)

The undergoer subject form of the verb *masuki*, which I claim is the base predicate of *kemasukan* is perfectly grammatical with *oleh*:

110. *Mas Bengong dimasuki oleh setan*
brother Bengong US.enter.APPL by devil
'Bengong was possessed by a devil.'

This problem is left to future research, and for the present I will assume that *oleh* is possible with any *ke- -an* verb whose second argument is an agent. The remainder of the *ke- -an* verbs with two arguments have a patient or a theme as their second argument, and in these cases there is no alternative realisation with a preposition. But, across languages, the semantic notion affected patient is consistently associated with a structural position, (direct) object (Tenny 1994) to the extent that it would be surprising to find a preposition which coded those semantics. Certainly Indonesian does not have such a preposition. Therefore, for the majority of two argument *ke- -an* verbs the means for coding the second argument in a prepositional phrase do not exist.

The third similarity between actor DPs in undergoer subject clauses and the second arguments of *ke- -an* verbs is in the way in which pronouns are treated in the two environments. In both cases, a pronoun adjacent to the verb is dispreferred. Where the prepositional realisation is available for the second argument of a *ke- -an* verb, it is compulsory for pronouns. And elsewhere, pronouns are very rare. The only verb which the native speakers I consulted would accept with a pronoun as second argument was *keduluan* 'be preceded by':

111. *Saya keduluan dia*
1SG *ke*.precede.*an* 3SG
'I was preceded by him.'

Similarly, pronouns are generally impossible as actor in a *di*-V clause, except in the case of *mereka* which is argued to be an anomaly in section 2.5.3 and where they are strongly focussed, but they are possible with the preposition *oleh*. This is certainly true for a third person pronoun:

112. *Saya dijemput oleh dia /olehnya*
1SG US.meet by 3SG by.3
'I was met by him.' (IRG: 248)

First and second person pronouns are generally unacceptable as actor in a *di*-V clause, but they are judged to be more acceptable after the preposition *oleh* than immediately following the verb. Thus IRG (249 n2) mentions the possibility of a phrase such as *oleh saya* appearing in a *di*-V clause, but not

the alternative. In the context of his discussion, this omission clearly means that such structures are not possible at all.

The evidence discussed here makes it clear that whatever syntactic process accounts for the behaviour of actor DPs in *di*-V clauses should also cover the behaviour of the second arguments of *ke- -an* clauses. The account offered in section 2.5.4 allowed the adjunction of an actor DP, which was an oblique, to a particular morphological class of verbs. These restrictions can now be seen to be too constrained: another class of verbs allows adjunction and arguments with a variety of semantic roles can adjoin. However the restriction of the adjunction process to oblique arguments can be maintained. Adjunction to the verb is not possible in the case of non-terms which are not semantically selected by the predicate, that is, adjuncts. There is no possible alternation between structures such as examples 113 and 114 parallel to that shown in example 34:

113. *Siti pergi ke Surabaya*
Siti go to Surabaya
'Siti went to Surabaya.'

114. **Siti pergi Surabaya*
Siti go Surabaya

The correct generalisation then is that an oblique argument which is assigned a semantic role by a verb can be realised by adjunction to that verb. Rule 177 of chapter 2 was the following:

177. Actor adjunction (Indonesian)

$$V \rightarrow \underset{\uparrow = \downarrow}{di\text{-}V} \quad \underset{\uparrow \, OBL_\theta = \downarrow}{DP_{actor}}$$

This can now be stated more generally:

115. Oblique adjunction (Indonesian)

$$V \rightarrow \underset{\uparrow = \downarrow}{V} \quad \underset{\uparrow \, OBL_\theta = \downarrow}{DP}$$

Subsequent discussion in chapter 4 and chapter 5 will show that this statement is too general, but the statement is now less stipulative than that offered in chapter 2. The majority of two-argument *ke- -an* verbs are unusual in that their second argument is a patient and no prepositional encoding of this role is possible. Therefore their only possible realisation is adjoined to the verb.

The general impossibility of pronouns as actors in *di*-V clauses was attributed to a constraint barring non-term pronoun actors in chapter 2. This is not a factor in the distribution of pronoun second arguments for *ke- -an* verbs, and this fact raises the question of whether the account in chapter 2 is correct, or incomplete. The fact that pronouns with *oleh* are much better following *ke- -an* verbs than in *di*-V clauses suggests that the similarities may be coincidental, and the fact that at least one *ke- -an* verb can occur with a pronoun adjoined to it would seem to confirm this. It is however possible that

adjunction structures including pronouns are avoided. A restriction of this type might be related to a difference in referentiality between pronouns and other nominals. However, I argued in section 2.3.1.2 that Myhill's (1988) account of the actor in *di*-V clauses as an incorporated noun fails, in part, because such actors are fully referential. It is therefore hard to see exactly how any such constraint would work. Such a constraint would explain the behaviour of the *ke- -an* verbs, but would not undermine the account of *di*-V clauses given in chapter 2, although the effects observed there might then be overdetermined by the grammar.

3.4 Conclusion

This chapter has presented an analysis of *ke- -an* verbs in Indonesian, often described as adversative verbs. I have argued that these verbs are derived by a lexical process of predicate composition, and that the argument structure of the derived verb is organised on unusual semantic principles. The predicate *ke- -an* selects the most affected entity in a situation as its subject. The notion of 'most affected entity' includes a human in a close relation with the affected patient of the base verb.

Examination of the syntax of clauses with *ke- -an* verbs has shown that many of these verbs have a non-subject argument which is an oblique, but which does not necessarily have the typical coding properties of an oblique. That is, this argument is not typically introduced by a preposition. This conclusion has forced us to extend the idea introduced in section 2.5.4, that Indonesian allows the actor in an undergoer subject clause to be adjoined to the verb. The second argument of many *ke- -an* verbs has similar properties to those of the actors discussed in chapter 2: they cannot be separated from the verb and they can, in some cases, alternate with expression as a prepositional phrase. Where this alternative form of expression is not possible, I argued that this is due to the lack of a semantically appropriate preposition. In the following chapter, I examine a case where a group of verbs alternate between having a prepositional phrase and a noun phrase as their non-subject arguments. But in this case, I will argue that the noun phrase is not adjoined to the verb, but is in fact a direct argument even though the verbs are not a part of the transitive system analysed in chapter 2.

4 Emotion/Cognition Predicates

This chapter describes and analyses a class of predicates in Indonesian which are semantically associated with two arguments, but which do not behave syntactically in the same fashion as the verbs described in Chapter 2. Almost all of the predicates in this class denote emotional or cognitive states, a class of predicates which cross-linguistically often stand apart from the prototypical transitive pattern of a language (Belletti & Rizzi 1988, Grimshaw 1990, Kachru 1990). As the whole class of predicates has received comparatively little attention in the literature (but see Kana 1986: 283–289 and Stevens (1970)), section 4.1 argues that membership of the class in Indonesian depends on the predicate taking its second argument (the stimulus of emotion or cognition) as a prepositional phrase, and provides a description of the properties and patterns of use of these predicates. A subclass of emotion predicates also allow the stimulus to be a DP in a construction described by Stevens (1970) as 'pseudo-transitive', and this group will be the focus of investigation in section 4.2, which attempts to answer the question: what grammatical function is assigned to the second argument of such clauses? Section 4.3 recasts the analysis of section 4.2 in terms of LFG and section 4.4 discusses some theoretical and empirical issues which arise from this analysis.

4.1 A class of predicates

4.1.1 Criterion of membership

I stated above that I take the property of coding the second argument as a PP to be the appropriate defining property for the class of predicates under consideration:

1. *Sampai* *Retno* *lupa* *pada* *dirinya*
   until Retno forget to self.3
   'Until Retno forgot herself.' (PYD: 187)

2. *Jadi* *rakyat* *boleh* *mual* *dan* *bosan* *dengan* *sang* *PM*
   then populace may feel.queasy and bored with HON PM
   'Then the people may feel dissatisfied and bored with the honourable PM.' (McKay)

Both Stevens (1970) and Kana (1986), while acknowledging the existence of the prepositional construction, assume that appearance of a predicate in the 'pseudo-transitive' construction is criterial. There are two reasons for rejecting this approach.

Firstly, there are Indonesian predicates which are clearly of the same semantic type as those considered by Stevens and Kana but which cannot occur in the 'pseudo-transitive' construction. For these predicates, the prepositional construction is obligatory, not optional. An example is the predicate *puas* 'satisfied, content':

3. *Dan mereka juga sudah puas dengan kasus*
and 3PL also PERF content with case
wanita yang hilang itu?
woman REL lost that
'And were they also satisfied with the case of the woman who disappeared?' (SDPGS: 128)

4. **Dan mereka juga sudah puas kasus wanita yang hilang itu?*

Other predicates which behave like this are *cinta* 'love' and *setia* 'loyal':

5. *Aku tetap cinta padamu*
1SG constant love to.2
'I still love you.' (E&S:118)

6. **Aku tetap cinta kamu.*

7. *kakek-kakek yang setia padaku*
grandfather .DUP REL loyal to.1SG
'old relatives who are loyal to me.' (PYD:92)

8. **kakek-kakek yang setia aku*
(but OK as 'my loyal old relatives'

In other respects, such predicates share many properties with the 'pseudo-transitives': they typically derive verbs which participate in the system analysed in chapter 2 when an applicative suffix is added:

9. *Kurasa dia mencintaimu*
1SG.feel 3SG AS.love.APPL.2
'I feel that he loves you.' (PYD:39)

And some of them can take a non-finite clause as complement, often with a controlled subject position:

10. *Sedang kita sendiri hanya puas menjadi*
while 1PL.INCL INT.self only content become
pengikut-pengikutnya
follower.DUP.3
'While only we ourselves are content to become his followers.' (McKay)

All of these properties are discussed in more detail in following sections. My point here is that there is a clear overlap in properties between predicates which only ever occur in the prepositional construction and the 'pseudo-transitive' predicates. The most satisfactory approach therefore is to treat the 'pseudo-transitives' as a subset of the more general class.

The second reason for preferring this approach is that the prepositional construction is widespread across the Austronesian language family. Many languages in the Western Austronesian group have this feature, but the 'pseudo-transitive' construction is less common. Examples of such languages are Acehnese, Balinese, Sasak and Tagalog.

Acehnese has many verbs which occur in the prepositional construction, and which are formally intransitive. This language has a classic fluid-S system (Dixon 1994:78–83). Transitive verbs have obligatory proclitics cross-referencing the actor; the undergoer is optionally cross-referenced by an enclitic[1]. Some intransitive verbs allow either type of coding for their argument, with a corresponding semantic difference and this group includes some emotion verbs[2]:

11.	*gopnyan*	*hana-*	*inseueh-*	*geuh*	*keu-*	*lôn*
	he/she	NEG-IN	feel compassion	3	DAT	1

'He/she has no sympathy towards me.' (Durie 1985:4-13:49)

12.	*gopnyan*	*hana-*	*geu-*	*inseueh*	*keu-*	*lôn*
	he/she	NEG-IN	3	feel compassion	DAT	1

'He/she feels no sympathy towards me (by choice).' (Durie:1985:4-20:49)

Example 11 makes it clear that this verb is intransitive: otherwise a proclitic would be required. Durie suggests that the preposition *keu* is omissible (1985:61), and gives an example:

13.	*h'an-*	*galak-*	*geuh*	*(keu)-*	*u-*	*muda*
	NEG	like	3	(DAT)	coconut	young

'He doesn't like young coconut.' (Durie 1985: 4-83, 62)

My own research suggests that the preposition is not generally omissible, and is obligatory when the following DP is animate. However, the speaker with whom I worked spoke a different dialect to that described by Durie, and this variation might be due to this difference. Acehnese shares the prepositional construction and the 'pseudo-transitive' construction with Indonesian, and there is some evidence that the prepositional construction is basic.

Balinese also has semantically similar verbs which introduce the stimulus argument with a preposition, and Artawa, Artini and Blake (1997) refer to these verbs as 'two-place intransitives'. They include *tresna* 'to love', *inget* 'to remember' and *demen* 'to like':

14.	*cerik-cerik*	*jani*	*lebihan*	*sing*	*demen*	*teken*	*jaja roti*
	children	now	more	NEG	like	with	cake

Children today more and more do not like cakes
(Artawa, Artini & Blake 1997:4 ex8)

The sources which I have consulted do not give any examples of 'pseudo-transitive' constructions, but it is possible to make the stimulus a direct argument of the verb, and the verbal morphology shows that its argument structure is changed to allow this[3]:

---

[1] Some dialects lack enclitics altogether (Bukhari Daud, p.c.).

[2] Glossing as in the source: IN = inchoative.

[3] Glossing as in source: DEF = definite, AD = advancement.

15. *Tiang* *demen* *gati* *ajak* *rok-e*
1SG like much with dress-DEF
'I like the dress.'

16. *Rok-e* *demen-in* *tiang*
dress-DEF like-AD 1SG
'I like the dress.' (Artawa, Artini & Blake 1997:13 ex32)

The additional verbal morpheme in example 16 which Artawa *et al.* gloss as AD(vancement) is clearly parallel to the Indonesian applicative in example 9 above, and also probably cognate to it (see Artawa 1999: 43-45). Note also that when the verb becomes transitive in example 16, the actor appears in the postverbal position required for a Balinese transitive verb without a nasal prefix (Artawa and Blake 1997). On the available data, it seems clear that the prepositional construction can be considered basic in Balinese.

Tagalog also has a group of verbs which behave in a similar fashion. In the simplest case, the experiencer argument is marked with the nominative case marker *ang*, and the stimulus is marked by the oblique case marker *sa*:

17. *na-gulat* *ang* *babae* *sa* *boyfriend* *niya*
AS-surprised NOM woman OBL boyfriend 3SG.POSS
'The woman was surprised at her boyfriend.' (Martin 1996: 252 ex39)

The case marking possibilities of Tagalog mean that the language does not allow the 'pseudo-transitive' construction. And if a non-core argument is promoted, it must become subject. Therefore the only alternative construction is one in which the stimulus is subject while the experiencer is still a core argument (i.e. a transitive construction):

18. *k-in-a-gulat-an* *ng* *babae* *ang* *boyfriend* *niya*
US-surprised CORE woman NOM boyfriend 3SG.POSS
'The woman was surprised at her boyfriend.' (Martin 1996:252 ex40)

The verbs which behave in this fashion are semantically similar to those in Indonesian and Balinese. Additional evidence as to the spread of this construction within the Austronesian family is not hard to find. For example, the so-called 'middle verbs' of Polynesian languages have much in common with this construction (Chung 1978b: 216-234, Seiter 1978).

A very small number of predicates do not fit the general pattern. These predicates do not have transitive forms (except with applicative suffixes), but they also never occur in the prepositional construction. The predicates in this group include *mau* 'want', *punya* 'own, possess', *tahu* 'know' and *perlu* 'need'. A further noteworthy feature of this group is that these verbs seem able to become auxiliaries very easily. Note the position of *mau* relative to the actor pronoun in the following example:

19. *Nasi itu mau dia makan*
rice that want 3SG eat
'That rice, he wants to eat it.' (Waruno Mahdi, p.c.)

In many cases, *mau* has an almost purely aspectual meaning and can occur with non-sentient subjects:

20. *Rumah itu mau roboh*
house that want collapse
'The house is about to collapse.' (Waruno Mahdi, p.c.)

The predicate *ingin* 'want, desire' should probably also be included with this group.

21. *Ingin dia makan, nasi itu*
wish 3SG eat rice that
'He wishes to eat it, that rice.' (Waruno Mahdi, p.c.)

Kana (1986: 284) has an example of this predicate in the prepositional construction, but I have no such examples in my data and E&S does not give such an example. It may be significant that the preposition in Kana's example is *akan*; this preposition is rarely encountered in most registers of contemporary Indonesian (see discussion in section 4.1.2.2 below) and has a distinctly formal or archaic feel. I would therefore suggest that *ingin* may previously have behaved in the same way as the main group of emotion predicates, but that it is moving (or has moved) to the anomalous group. Quite plausibly, the other predicates in this anomalous group have followed a similar path. Several of these predicates can also be used as auxiliaries, but lacking the prepositional construction is neither a necessary nor a sufficient condition for being used as an auxiliary. On the one hand, *punya* and *tahu* are not used as auxiliaries, and on the other hand *suka*, which retains the prepositional construction, is (see further discussion in section 4.1.2.7 below).

4.1.2 Range of constructions

Most of the predicates under consideration here can appear in a range of constructions. In addition to the prepositional construction and the 'pseudo-transitive' construction, most predicates form a transitive verb with an applicative suffix, and a small group form transitive verbs directly without an applicative. Transitive verbs are also derived from some predicates with the causative suffix *-kan*. Many predicates can also take a complement clause as an argument. There are various possibilities in this regard: a full clause with a complementizer, a full clause with no complementizer, or a reduced clause with the subject position controlled. Finally, non-predicative uses are possible for some of these words: some appear as nouns and some as nominal modifiers[4]. The following sections give examples and brief discussion of each of these constructions in turn. For each predicate for which I have a textual

[4] The question of to what lexical category the words being investigated should be assigned is discussed in detail in section 4.1.3.

Word	Gloss	DP	Applic	Caus	C + S	Ø + S	Control
bangga	proud		(-kan)	+			
benci[1]	hate	+			+		+
bimbang	anxious		-kan	+			+
bosan	bored	+		+	+		+
cemas	worried						+
cemburu	jealous		(-i)	(+)			
cinta	love		-i				
curiga	suspect		-i	+	+		
ganggu	upset						
gemar	fond of		(-i)				+
gemas	annoyed		(-i)	+			
gila	insane	+	-i	+			
heran	surprised		-kan	+	+		+
ingat[2]	remember	+	-kan	+	+	+	+
jemu	fed up			+			
jengkel	annoyed			+			+
kangen	long for	+	(-i)	+			
kasih	love		-i				
kasihan	pity	+	-i	(+)			+
kecut	afraid			+			
kenal[3]	know s.o.	+	-i	+			
kuatir	afraid	+	(-i)	+		+	
lupa	forget	+	-kan		+		+
malu	ashamed			(+)			
marah	angry	+		+	+		
mimpi	dream	+	-kan				
mirip	resemble	+					
peduli	care about	+	-kan			+	
percaya[4]	believe	+	-i	(+)	+	+	+
prihatin	concerned		-kan	+			+
puas	satisfied			+			+
sadar	aware		(-i)	(+)	+		
salut	respect						
sayang	pity, love	+	-i/-kan	+			+
senang	like	+	-i	+	+	+	+
setia	loyal		(-i)				
simpati	sympathetic						
suka	like	+	-i	+	+	+	+
takut[5]	afraid	+	-i/-kan	+	+	+	+
tega	have the heart to						+
yakin	sure	+	-i	+	+	+	+

Table 4.1 – Indonesian Emotion and Cognition Words

Occurrence in various clause types

Table 4.1 – Legend and notes

DP	Can have a bare DP as complement
Applic	Takes the specified suffix to derive a transitive verb
Caus	Takes the suffix *-kan* to derive a causative verb
C + S	Can be followed by a full clause following a complementizer
Ø + S	Can be followed by a full clause without a complementizer
Control	Can be followed by a clause with a controlled subject position

A plus sign in any column indicates that at least one textual example of the clause type has been recorded. A bracketed plus sign indicates that native speakers believe the clause type is possible, but that no textual examples have been recorded.

In the column **Applic,** the entries indicate which applicative suffix, if any, is used with the word. Bracketing has the same meaning as in all the other columns: a native speaker attests the usage but no text example has been collected.

Notes:

[1] For discussion of whether *benci* has an applicative formation with *-i*, see section 4.1.2.4.

[2] There is a transitive verb derived directly from *ingat* without an applicative suffix. Therefore the examples with an DP complement are dubious.

[3] There is a transitive verb derived directly from *kenal* without an applicative suffix. Therefore the examples with an DP complement are dubious.

[4] There is a transitive verb derived directly from *percaya* without an applicative suffix. Therefore the examples with an DP complement are dubious.

[5] The suffix *-i* can also derive a causative with *takut*. See section 4.1.2.4 for discussion.

example of the prepositional construction, Table 4.1 shows which constructions it can appear in.

4.1.2.1 THE DATA

The data on which this survey is based are drawn from the following sources:

1) Two contemporary novels, one set in Jakarta (SDM) and one set in Surabaya (PYD). I have not attempted to analyse any variation which might be caused by this difference, as other factors are not controlled for. In particular, PYD is more 'slangy' in tone. However, the central characters in each novel are educated Indonesians, and I assume that these sources have made some attempt to accurately portray the speech of such persons.
2) One translation of an English novel into Indonesian (SDPGS). This source has certainly introduced some distortions into my data in terms of relative frequencies of certain constructions. For example, Miss Christie's characters are fond of adding the expression 'I dare say' to their comments. This is translated as *Saya berani katakan...*, and I am sure that the appearance of the predicate *berani* with reduced complement is consequently over-represented in my data.
3) Examples drawn from a corpus of contemporary Indonesian journalism collected by Helen McKay.
4) Example sentences provided by E&S. This is the standard Indonesian-English dictionary, and its usage is probably conservative. Most of the examples which I have using the preposition *akan*, which is no longer common in spoken Indonesian come from this source.
5) Native speaker judgments. The two speakers consulted are both educated Indonesians of Javanese background.

Full bibliographic details of the sources described in 1,2 and 4 are provided immediately before the reference section of this work.

Pending the discussion of the lexical categorization of the words under discussion, I use the neutral term *predicate* throughout this section. I refer to the two semantic roles associated with these predicates as *experiencer* and *stimulus*.

4.1.2.2 PREPOSITIONAL CONSTRUCTION

A wide range of prepositions occurs in this construction. At least historically the most widespread in usage is *akan*, which marks the second argument of all two place verbs in classical Malay (Cumming 1991: 42-45). It is therefore theoretically possible with all the emotion predicates. However, it is not commonly used in non-formal registers of Indonesian now. Occasional examples do occur

22. *Retno sangat takut akan perpisahan dengan Sandy*
Retno very afraid about separation with Sandy
'Retno was very afraid of parting from Sandy.' (PYD: 183)

The most common prepositions used are *dengan* 'with' and *(ke)pada* 'to, at'. A native speaker has informed me that she was taught, as a prescriptive rule, that in this construction *dengan* should be used with non-human nouns following, and *kepada* with humans. The data shows this pattern as a trend, but it is certainly not observed as an absolute rule:

23. *Dan orang yang paling tidak puas dengan*
and person REL important NEG content with
keputusan itu
decision that
'And the person who is least satisfied with that decision' (SDM:28)

24. *saya percaya ia tidak kenal dengan orang*
1SG believe 3SG NEG know with person
semacam itu
certain kind that
'I don't believe he knows anyone like that.' (SDPGS: 245)

25. *Barangkali ia hanya bosan pada hotel itu*
maybe 3SG only bored to hotel that
'Maybe he just got bored with the hotel.' (SDPGS: 101)

26. *Dia sangat benci padaku.*
3SG very hate to.1SG
'She really hates me.' (SDM:266)

27. *Marilah kita memperkuat tekat untuk tetap*
please.EMPH 1PL.INCL AS.support resolve for constant
setia kepada aspirasi
loyal to aspiration
'Let us strengthen our commitment to remain loyal to the ideals.' (McKay)

28. *Marisa sangat takut kepadanya*
Marisa very afraid to.3
'Marisa was very frightened of him.' (SDM:259)

There is a colloquial equivalent of *dengan, sama* 'together with', and this also is used in this construction. All the examples I have in my database have human stimuli:

29. *Aku kangen sama dia.*
1SG long for with 3SG
'I long for her.' (PYD:33)

The preposition *tentang* 'about' can occur with some predicates, typically cognitive rather than affective ones. However, the only textual examples which I have for this preposition come from the source translated from

English. Most of these appear to be direct translations of the English expression 'sure about ...':

30. *Anda yakin tentang hal itu?*
2SG sure about matter that
'Are you sure about that matter?' (SDPGS: 147)

In that source, there are also examples with *heran* 'surprised, astonished' and *peduli* 'attend to, care about'. In each case, direct translation from English seems likely. These examples suggest that although this preposition is not ungrammatical with some cognitive predicates, the usage is probably not natural for native speakers. Table 4.2 shows the possible co-occurrences of predicates and prepositions, and the data presented there reinforce this point. The distribution of *tentang* is restricted in comparison to the other five prepositions listed there, suggesting again that this preposition is only used in translation of English (and perhaps Dutch) expressions.

Three other prepositions occur in isolated examples: *ke* 'to', *atas* 'on'[5] and *terhadap* 'about, toward':

31. *Dia malah marah ke Presiden Peru Fujimori*
3SG instead angry to President Peru Fujimori
'Instead, he was angry with Peru's President Fujimori.' (McKay)

32. *Saya senang sekali atas konsolidasinya*
1SG happy very on consolidation.3
'I am very happy about the merger.' (McKay)

33. *organisasi yang peduli terhadap permasalahan*
organization REL care about set.of.problems
kependudukan
demography
'an organization which cares about demographic problems' (McKay)

Note that all of these examples are from journalism; it is possible that contemporary usage allows a wider range of prepositions in this construction.

There are two generalisations which can be made about the use of prepositions in this construction. Firstly, the general locative preposition *di* is never used, and the only spatial concept which plays any role in this construction is movement towards. More recent usage may be altering this picture, with the appearance of *atas* and *terhadap* which can mean 'across from' or 'opposite'. This fact will be discussed further in section 4.2 in relation to Kana's (1986) claim that the stimulus in this construction is a Location. This is a fact about Indonesian; other Austronesian languages use locative prepositions in their parallel construction. For example, in Sasak (Menó-mené variety), the locative preposition *léq* is used with emotion predicates:

[5] Strictly speaking, *atas* is a locational noun and it normally occurs together with the general locative preposition *di*, thus *di atas* 'on top of'. By a transparent reanalysis, *atas* has come to be used as a preposition in its own right. A similar process can be seen with *dalam* 'inside', from *di dalam* 'at the inside of'.

Word	Gloss	akan	dengan	kepada	pada	sama	other
bangga	proud	+	+	+	(+)	(+)	atas
benci	hate	+	+	+	+	(+)	
bimbang	anxious	+	?	?	?	?	
bosan	bored	(+)	+	(+)	+	(+)	
cemas	worried	+	(+)	(+)	(+)	(+)	
cemburu	jealous	(+)	(+)	+	+		
cinta	love	(+)	(+)	(+)	+	+	terhadap
curiga	suspect	+	(+)	(+)	+	(+)	
ganggu	upset	(+)	+				
gemar	fond of	+	(+)		(+)		
gemas	annoyed	(+)	(+)	(+)	(+)	(+)	
gila	insane	(+)	+		+	(+)	tentang
heran	surprised	(+)	(+)	(+)	+	(+)	tentang
ingat	remember	+	+	(+)	+	+	(tentang)
jemu	fed up	?	+	?	?	?	
jengkel	annoyed	(+)	+	(+)	(+)	(+)	
kangen	long for	(+)	(+)	(+)	(+)	+	
kasih	love	(+)	(+)	(+)	+	(+)	
kasihan	pity	(+)	(+)	(+)	+	(+)	
kenal	know s.o.	(+)	+	(+)	+	(+)	
kuatir	afraid	(+)	(+)	(+)	(+)	(+)	
lupa	forget	(+)	+	(+)	+	(+)	di
malu	ashamed	(+)	(+)	(+)	(+)	+	
marah	angry	(+)	+	(+)	+	(+)	ke
mimpi	dream	(+)					(tentang)
mirip	resemble	(+)	+			(+)	
peduli	care about	+	+	(+)	+	(+)	tentang, terhadap
percaya	believe	+	+	+	+	(+)	
prihatin	concerned	+	(+)	(+)	(+)	(+)	atas
puas	satisfied	(+)	+		(+)	(+)	atas
sadar	aware	+	(+)		(+)	(+)	(tentang)
salut	respect	+	+	(+)	(+)	(+)	
sayang	pity, love	(+)	(+)	+	+	+	
senang	like	+	+	(+)	+	(+)	atas terhadap
setia	loyal	(+)	(+)	+	+	(+)	terhadap
simpati	sympathetic		(+)	(+)	+	(+)	
suka	like	+	+	+	+	(+)	
takut[5]	afraid	+	(+)	+	+	(+)	
tega	have the heart to		(+)	(+)	+	(+)	
yakin	sure	+	+		(+)	(+)	tentang

Table 4.2 – Indonesian Emotion and Cognition Words

Co-occurrence with various prepositions

34. *Takutk* *léq* *acòng*
afraid.1SG LOC dog
'I am afraid of dogs.'

This fact also leads to the second generalisation, that the semantic contribution of the preposition is not important.

This second generalisation is illustrated by Table 4.2 which shows the co-occurrence possibilities for each predicate in the database and six prepositions. Five of the prepositions (*tentang* being the exception) can each occur with almost every predicate, the gaps which do appear seeming to be quite idiosyncratic. For many predicates, examples occur with more than one preposition, and in one case, there are examples with five prepositions (see also example 33 above):

35. *Saya* *tidak* ***peduli*** ***akan*** *perkataannya*
1SG NEG care about NOM.word.NOM.3
'I don't care what he said.' (E&S: 415)

36. *Mereka* *tidak* ***peduli*** *sedikit* *pun* ***tentang*** *dia*
3PL NEG care a.little EMPH about 3SG
'They didn't care even a little about him.' (SDPGS: 184)

37. *Moskow* *seolah* *tak* ***peduli*** ***dengan*** *segala* *macam*
Moscow as.if NEG care with all type
reaksi *di* *dunia.*
reaction LOC world
'Moscow apparently doesn't care about any sort of reaction from the world.' (McKay)

38. *Mereka* *seperti* *tak* ***peduli*** ***pada*** *Kepala*
3PL as NEG care to head
Bulog *Beddu* *Amang*
distribution.agency Beddu Amang
'They don't seem to care about the chief of the distribution agency, Beddu Amang.' (McKay)

Such evidence suggests very strongly that there is no semantic dependency between the preposition and its complement in this construction, and that the stimulus is assigned its semantic role directly by the predicate. If a predicate in the database occurs with more than one preposition, the two used are most likely to be *dengan* and *(ke)pada*. There is no common semantic ground between these two, and they are not mutually substitutable in other constructions. *(ke)pada* is the preposition normally used for the recipient in act of giving, but *dengan* is impossible in such a context:

39. *Ia* *memberikan* *uangnya* *kepada* *adiknya*
3SG AS.give.APPL money.3 to brother.3
'He handed over his money to his younger brother.' (E&S: 73)

40. *Ia* *memberikan* *uangnya* *dengan* *adiknya*
3SG AS.give.APPL money.3 with brother.3
'He handed over his money with(/*to) his younger brother.'

dengan is typically used to indicate a co-participant in an action, and *pada* cannot have this meaning either:

41. *Ia* *pergi* *dengan* *ayahnya*
3SG go with father.3
'She went with her father.' (E&S: 138)

42. *Ia* *pergi* *pada* *ayahnya*
3SG go to father.3
'She went to(/*with) her father.'

As the above examples also show, in other contexts there is no basis for the tendency (and prescriptive rule) noted above for *dengan* to occur with inanimate stimuli.

In English, some predicative adjectives display similar behaviour:

43. *John was angry **with** her.*

44. *John was angry **at** her.*

The tendency for a distinction between the prepositions appropriate for human stimuli and those appropriate for non-humans is also seen (at least to some extent) in English:

45. *Mary is annoyed with/?about Bill.*

46. *Mary is annoyed about/?with the traffic*

Example 43 is possible with the preposition *about*, but the most natural reading is that Mary is annoyed about some action preformed by Bill, not by Bill himself. The use of *with* in example 44 seems less felicitous, and creates an anthropomorphising effect.

The general principle is that when a predicate semantically selects a complement without licensing it as a direct argument, then the licensing element must be semantically compatible with the selected role. In the case where a language does not have a licenser, such as an adposition or case-marker, which is semantically specific for the role selected, any other linker which is compatible can be used. Thus, neither English nor Indonesian has a preposition which specifically codes the semantic role 'stimulus of emotion', and predicates which select this role without licensing it can use any other preposition which is compatible (and this does not include purely locational ones). Where a preposition is specific, it is the obligatory linker as for agents in both languages (*by* in English, *oleh* in Indonesian). The intermediate case is a verb like *put* which semantically selects a location as one of its arguments without licensing it directly. In consequence, any compatible licenser is acceptable.

A small number of predicates occur with one of the intransitive verb prefixes *ter-* and *ber-* in this prepositional construction. The *ter-* prefix usually has a meaning of involuntary activity, typically translated by an English passive:

47. *Ardi teringat pada wajah seekor.*
Ardi *ter*.remember to face CLASS(animal)
'Ardi was reminded of the face of an animal.' (SDM: 16)

48. *Saya berpikir tentang pulau Jawa.*
1SG *ber*.think about island Java
'I think about Java.'

These examples are not surprising: these two prefixes are generally agreed to occur with intransitive verbs and they therefore fall under the same general pattern as that discussed thus far in this section. What is surprising, is that occasional examples appear of transitive verbs with a following preposition:

49. *Aku sendiri tak mengerti denganmu*
1SG self NEG AS.understand with.2
'I myself don't understand you.' (PYD: 194)

which can be compared with a more normal usage for this verb:

50. *Ia mengerti maksud surat itu*
3SG AS.understand intention letter that
'She understands the meaning of that letter.' (E&S: 160)

This verb is clearly in the same semantic field as the predicates under discussion, and it would seem that the prepositional construction exerts an influence. Occasional examples also occur in which the predicate is derived as a transitive verb with an applicative suffix, but a preposition also appears before the stimulus:

51. *Yang penting saya sangat mencintai Sandy*
REL important 1SG very AS.love.APPL Sandy
dan ***menyayangkan atas*** *semua kejadian itu*
and AS.pity.APPL on all event that
'What is important is that I love Sandy very much and that I regret all these events.' (PYD: 200)

The issue of where applicative verbs in Indonesian can or should be followed by a preposition is complex (see Chung 1976a for some discussion). The confusing cases seem to be limited to three place predicates, where the third argument may have a preposition even when a derived verb is used. Derived two-place verbs generally behave as straightforward transitive verbs, and the above example (51) is atypical.

4.1.2.3 BARE NOUN-PHRASE COMPLEMENTS

Many of the emotion and cognition words which appear with a prepositional phrase complement can also occur with a bare DP complement.

This clause type is discussed in detail in section 4.2, therefore I only provide examples here to illustrate the construction:

52. *Kamu* ***lupa*** ***rumahku?***
2SG forget house.1SG
'You have forgotten my house?' (SDM: 96)

53. *Mau dijadikan uang, saya* ***takut*** ***konduite*** ***jelek***
want US.cause money 1SG afraid report bad
'Wanting money to be made, I am afraid of a bad report.' (McKay)

54. *Cuma aku tidak* ***suka*** ***omongannya*** *yang sok suci,*
only 1SG NEG like word.3 REL as.if pure
'I just don't like his speeches that take the moral high ground.'
(PYD: 87)

55. *Saya nggak* ***senang*** ***wartawannya***
1SG NEG like journalist.3
'I don't like the journalist.' (McKay)

Not surprisingly, many clauses of this type are like the examples just given in having 1st or 2nd person pronouns as experiencer. But other types of nominal are possible in this semantic role, including non-pronominal 3rd persons:

56. *Dua manusia sesama jenis bercumbuan di pojok*
two human equal kind flirt LOC corner
jalan, tanpa peduli keramaian sekitar
road without care NOM.busy.NOM around
'Two human beings of the same sex flirt at the street corner, without caring about the surrounding bustle.' (McKay)

57. *Kalau boleh memilih, Ardi lebih suka Marisa*
if permitted AS.choose Ardi more like Marisa
bersikap selembut penampilannya
ber.attitude as.soft appearance.3
'If he had a choice, Ardi would have preferred Marisa to be as gentle as her appearance.' (SDM : 27)

This clause type is therefore not a type of Pro-V clause with non-standard word order.

Speakers report that there is no semantic difference between the clause type with a preposition, and that which lacks it. The following examples are adjacent clauses from dialogue:

58. *Papa sayang padamu, Risa.*
father love to.2 Risa
'Father loves you, Risa'.

59. *Risa juga sayang Papa.*
Risa also love father
'Risa loves Father too.' (SDM: 293)

This is the moment in the story at which the heroine is reconciled with her step-father, and the author's intention is to portray a mutuality of emotion. Even in this context it is acceptable for one character to use one clause type and for the second character to respond with the other type. In section 4.2, I argue that the DP complement in clauses such as example 59 are direct arguments, while the PP in examples like 58 contains an oblique argument. The lack of semantic differentiation is therefore somewhat surprising. Where there is a choice between coding an argument as an oblique or as a direct argument, the direct argument is usually more fully implicated in the action. A well-known example of this phenomenon in English is the conative alternation:

60. The hunter shot the deer.

61. The hunter shot at the deer.

where example 60 entails that the bullet hit its target, whereas example 61 does not. Similar remarks apply to the argument alternation possible with verbs of the *spray/load* class (Hall 1965, see also Dowty 1991: 587 and references cited there). The Indonesian case is different according to the evidence presented here, and I would suggest that this difference is linked to the fact that, as discussed in the previous section, the preposition makes no semantic contribution. This question is discussed further in section 4.1.3.2.

4.1.2.4 APPLICATIVES AND CAUSATIVES

The majority of the emotion/cognition predicates derive transitive verbs with one of the applicative suffixes *-i* and *-kan* (*-kan* also derives causatives, discussed later in this section). Typical examples for each type are:

62. *Kurasa dia* ***mencintaimu***
1SG.feel 3SG AS.love.APPL.2
'I feel that she loves you.' (PYD: 39)

63. *Aku senang kau sudah bisa* ***melupakan*** *wanita*
1SG happy 2SG PERF able AS.forget.APPL woman
itu, San
that Sandy
'I am happy that you were able to forget that woman, Sandy.' (PYD: 195)

Kana (1986: 287) implies that *lupa* is the only predicate which takes the *-kan* suffix in this construction, but examination of Table 4.1 shows that this is not the case. Applicative derivation results in a verb which can participate fully in the system described in Chapter 2. Both verb prefixes are possible and are used to make arguments accessible for processes such as relativisation:

64. *Kamu punya ibu yang* ***menyayangimu***.
2SG have mother REL AS.love.APPL.2SG
'You have a mother who loves you.' (SDM: 84)

65. *Untuk menemui anak yang begitu* ***disayanginya***.
for AS.meet.APPL child REL so US.love.APPL.3
'In order to meet a child who is so loved by her.' (SDM: 208)

The pronoun agent construction is possible; note the relative position of the pronouns and the negation in example 66 and auxiliary in example 67:

66. *Apakah ada sesuatu yang* ***tak*** ***kau*** *sukai*
what.EMPH be something REL NEG 2SG like.APPL
tentang Nicky?
about Nicky
'Is there something you don't like about Nicky?' (PYD: 82)

67. *Itulah yang* ***telah*** ***Anda*** *lupakan*
that.EMPH REL PERF 2SG forget.APPL
'That is what you have forgotten.' (SDPGS: 314)

The bare verb actor subject construction is also possible; note the position of the adverb *sangat* immediately before the verb *sayangi* in the following example:

68. *Yakni, bahwa salah seorang dari yang hilang*
that.is that one one.person from REL missing
itu adalah seorang yang kami - dari divisi -
that COP one.person REL 1PL.EXCL from division
sangat kenal, ***sangat*** ***sayangi***
very know very love.APPL
'That is, the fact that one of those who were missing is a person who we - from the division - know so well, love so much.'
(Simatupang 1993: 117, quoted in Voskuil 1996: 212)

The data examined here gives little support to theories which differentiate the two applicative suffixes semantically (Tampubolon 1983, Kana 1986, Voskuil 1996). Although there is a tendency for predicates with more cognitive meanings to take the *-kan* suffix and for more emotional meanings to occur with the *-i* suffix, the division is not clear cut. For example, it is not obvious why *curiga* 'suspect' and *percaya* should behave more like emotion words than *bangga* 'proud' or *peduli* 'attend to, care about'. There are also at least two predicates which can derive transitive verbs with both applicative suffixes. One is *sayang* 'pity, love' which has the derivatives *sayangi* 'love' and *sayangkan* 'regret, deplore'. Although the two derived verbs have different meanings, the thematic relations between the verb and its arguments are essentially the same in each case. Even more complex is the case of *takut* 'fear' which derives transitives with both suffixes, with both derivatives also having a causative reading:

69. *Hanya Tuhan yang sewajarnya kita* ***takuti***
only God REL truly 1PL.INCL fear.APPL
'We need fear only God.' (E&S: 544)

70. *Tapi yang* ***ditakutkan*** *oleh Mustika bukanlah wanita*
but REL US.fear.APPL by Mustika NEG.EMPH woman
yang sedang dikejar-kejar oleh Sandy
REL PROG US.chase.DUP by Sandy
'But what Mustika feared was not the woman who was being chased by Sandy...' (PYD: 30)

71. *Ia* ***menakuti*** *dan membentak orang bawahannya*
3SG AS.fear.CAUS and AS.snap.at person subordinate.3
'She intimidated her subordinates and spoke harshly to them.' (E&S: 544)

72. *berarti ia melangkah ke dalam bahaya yang*
meaning 3SG AS.step in inside danger REL
menakutkan
AS.fear.CAUS
'meaning that he stepped into a terrifying danger.' (SDPGS: 274)

It is not common to find the *-i* suffix used as a causative, but this data show that it is possible. These particular examples raise a problem in that the same verb forms can have the experiencer linked to either grammatical function, suggesting that two lexical entries would be required for each verb form.

As shown in Table 4.1, causative derivations are possible with many predicates, of both the emotion and cognition types:

73. *Yang lebih* ***menjengkelkan*** *lagi, mertuanya ikut*
REL more AS.angry.CAUS again parent-in-law follow
campur tangan.
mixed hand
'What was more annoying, her in-laws were involved.' (SDM: 110)

74. *Ia selalu menasehati dan* ***mengingatkan***
3SG always AS.advice.APPL and AS.remember.CAUS
dirinya berhati-hati
self.3 careful
'He always advised and reminded himself to be careful.' (PYD: 167)

Indonesian[6] systematically lacks lexical causatives in this semantic field. English verbs of this type such as *remind* (example 74) and *frighten* (example 72) are always translated by derived Indonesian verbs. A possible restriction on the use of the *-kan* suffix as a causative is discussed further in section 4.3.1.2.

[6] And other Austronesian languages, as far as my research has gone.

4.1.2.5 TRANSITIVE VERBS

A few predicates can form transitive verbs without a derivational suffix. The most common of these is *ingat* 'remember':

75. *Sandy masih melamun mengingat kehangatan*
Sandy still AS.muse AS.remember warmth
tubuh Nicky
body Nicky
'Sandy still daydreamed remembering the warmth of Nicky's body.' (PYD: 142)

The other common form is *benci* 'hate':

76. *Salahkah kalau dia membenciku?*
mistake.EMPH if 3SG AS.hate.1SG
'Is it wrong if he hates me?' (SDM: 229)

Some native speakers claim that in such clauses *benci* is an applicative verb with *-i*, but there is no evidence that this is true, either synchronically or diachronically (Sander Adelaar, p.c.).

If a predicate forms a simple transitive verb, clauses in which the bare form occurs with a bare DP stimulus cannot be counted as examples of the construction considered in section 4.2. The clause could be a transitive clause with a bare verb form in such cases, and I do not include verbs with simple transitive forms as participating in the DP stimulus construction.

4.1.2.6 INTRANSITIVE USES

Most of the predicates under discussion can be used in a clause with one argument:

77. *Ia tetap menjadi suami yang setia*
3SG constant AS.become husband REL loyal
'He certainly became a loyal husband.' (SDPGS: 70)

78. *Pasti ia pergi karena takut*
certainly 3SG go because afraid
'Certainly he left because he was afraid.' (SDPGS: 57)

In such examples, the predicate behaves like an adjective. (The issue of what lexical category or categories these predicates should be assigned to is discussed in 4.1.3 below.) But with many predicates, an objectless emotion or cognition cannot be construed, and in such cases a second argument is clearly implied:

79. *aku sudah nggak ingat lagi*
1SG PERF NEG remember again
'I did not remember (it) any more.' (PYD: 60)

80. *dan Mustika nampak manja dan bersandar ke*
and Mustika apparently clinging and lean to
Sandy yang nampak tak begitu suka
Sandy REL apparently NEG like.that like
'and Mustika wanted to cling to and lean on Sandy who appeared not to like (it).' (PYD: 12)

Such examples show that many of these predicates do require two arguments in their semantic representation.

4.1.2.7 CONSTRUCTIONS WITH COMPLEMENT CLAUSES

Many emotion and cognition predicates occur in clauses in which the second argument is a complement clause. In fact, for some of the most common predicates this is the most common type of appearance: this is true for *suka* 'like' and *senang* 'happy, like' and also for *yakin* 'sure', although the data studied may not be representative for this predicate (see discussion in section 4.1.2.1 above). Three subtypes of this construction occur.

Firstly, the complement can be a full clause introduced by a complementizer:

81. *Tetapi, akhirnya ia yakin bahwa yang dihadapinya*
but finally 3SG sure that REL US.face.3
adalah Kwalik
COP Kwalik
'But, in the end he is sure that it is Kwalik who will face him.' (McKay)

82. *dia takut kalau-kalau Retno tiba-tiba ingin ikut*
3SG afraid if.DUP Retno suddenly wish follow
bersamanya ke Jakarta
together.3 to Jakarta
'He was afraid that Retno might suddenly want to go with him to Jakarta.' (PYD: 181)

A variety of complementizers are possible in this construction. Besides the two seen in the preceding examples, the possibilities include *jika* 'if', *untuk* 'for', and various interrogative expressions such as *bagaimana* 'how' and *mengapa* 'why'.

The second possibility is the same as the first, but with the complementizer omitted:

83. *Yang sudah tidak peduli orang lain percaya*
REL PERF NEG care.about person other believe
atau tidak dengan ucapannya
or not with statement.3
'Who did not care whether others believed his statements or not.' (SDM: 231)

84. *Aku tidak suka kau bergaul dengan Nicky terlalu dalam*
1SG NEG like 2SG socialise with Nicky too inside
'I don't like you to socialise too closely with Nicky.' (PYD: 82)

See Kana (1986: 241-252) for discussion of whether this type of structure should be analysed as 'raising to object'.

Finally, and most commonly, the subject of the complement clause is co-referential with the subject of the matrix and is omitted, a control structure:

85. *Aku masih takut jatuh cinta*
1SG still afraid fall love
'I am still afraid to fall in love.' (PYD: 12)

86. *Mata Sandy seakan tak percaya melihat mereka*
eye Sandy as.if NEG believe AS.see 3PL
"It was as though Sandy's eyes could not believe that they saw them.' (PYD: 202)

The predicate *suka* 'like' occurs very commonly in this construction:

87. *Karena Tika memang gadis yang baik dan suka menolong*
because Tika certainly girl REL good and like AS.help
'Because Tika was certainly a good girl and liked to help.' (PYD: 61)

88. *Kamu lebih suka bermain-main mencari kesenangan sendiri daripada bekerja*
2SG more like play AS.seek happiness self from work
'You prefer to play around seeking your own happiness rather than working.' (SDM: 61)

This word can also be used as an auxiliary in the same way that *mau* and *ingin* can (see section 4.1.1):

89. *Nasi itu suka dia makan*
rice that like 3SG eat
'That rice, he likes to eat it.' (Waruno Mahdi, p.c.)

This usage is plausibly a grammaticalization of the control structure described above, and the process has progressed to the extent that the selectional restrictions which *suka* exerts over its subject, which must normally be human or minimally conscious, is relaxed:

90. *Bis kota itu suka terlambat*
bus city that like delayed
'That bus is often / usually behind schedule.' (Waruno Mahdi, p.c.)

4.1.2.8 NOMINAL USES

Emotion and cognition words can occur in typically nominal environments. The most common of these is as apparent object of the verb (*me)rasa(kan)* 'feel':

91. *supaya tidak merasa bosan*
in.order.that NEG AS.feel bored(om)
'in order not to feel bored(om) (SDPGS: 58)

It is possible though that *merasa* selects an adjective as its complement, and that the emotion word here is not a noun. This interpretation is favoured by the fact that such complements are negated by *tidak*, which negates verbs and adjectives, rather than *bukan*, the usual negator for nouns:

92. *Ia merasa tidak puas*
3SG AS.feel NEG satisfied
'He felt unsatisfied.' (SDPGS: 229)

However, emotion and cognition words can appear as the complement of other verbs in more clearly nominal contexts:

93. *Pertemuan [...] itu tampaknya mengundang simpati*
meeting that apparently AS.invite sympathy
dunia terhadap kebijaksanaan Vietnam
world towards policy Vietnam
'The meeting has apparently brought the sympathy of the world to the policy of Vietnam.' (McKay)

94. *Namun ia sempat menyatakan prihatin atas*
however 3SG still.be.able AS.evident.CAUS concern on
sikap anti-Asia itu
attitude anti-Asia that
'However, he still has the opportunity to express concern about the anti-Asia attitude.' (McKay)

In example 93, the emotion word *simpati* has a possessor (*dunia),* and it is also possible for an emotion word to be preceded by a quantifier, another typically nominal environment*:*

95. *semua benci pada Caetano Calazar*
all hatred to Caetano Calazar
'all the hatred of Caetano Calazar' (McKay)

Word	Gloss	Noun	ke- -an	Gloss
bangga	proud		+	s.t. one is proud of
benci	hate	+	+	i. hatred, ii. hostility
bimbang	anxious		+	i. doubt, ii. in doubt
bosan	bored	+	+	boredom
cemas	worried		+	anxiety
cemburu	jealous		+	jealousy
cinta	love		+	i. love, ii. beloved
curiga	suspect	+	+	suspicion
ganggu	upset			
gemar	fond of		+	i. fondness, ii. hobby
gemas	annoyed		+	i. anger, ii. passion
gila	insane	+	+	madness
heran	surprised	+	+	miracle
ingat	remember	+		
jemu	fed up		+	boredom
jengkel	annoyed	+	+	annoyance
kangen	long for		+	what is missed
kasih	love	+	+	pity
kasihan	pity	+		
kenal	know s.o.			
kuatir	afraid	+	+	anxiety
lupa	forget	+	+	forgetfulness
malu	ashamed		+	shame
marah	angry	+	+	anger
mimpi	dream	+		
mirip	resemble		+	resemblance
peduli	care about	+		
percaya	believe	+	+	i. trust, ii. one who can be relied on
prihatin	concerned		+	concern
puas	satisfied	+	+	satisfaction
sadar	aware		+	i. awareness, ii. realization
salut	respect			
sayang	pity, love	+	+	i. sorrow, ii. love, iii. darling
senang	like	+	+	i. happiness, favourite
setia	loyal		+	loyalty
simpati	sympathetic	+		
suka	like	+	+	i. pleasure, ii. hobby
takut	afraid	+	+	fear
tega	have the heart to	+		
yakin	sure			

Table 4.3 – Indonesian Emotion and Cognition Words

Use as nouns and nominalizations

Many emotion and cognition words can also serve as the base for a nominal derived with the circumfix *ke- -an:*

96. *Matanya lebar ketakutan*
eye.3 wide NOM.afraid.NOM
'Her eyes were wide with fear.'

Table 4.3 shows which of the words listed in Table 4.1 are judged acceptable as nouns by informants on the basis that they can co-occur with the demonstrative *itu*, and which derive *ke- -an* nouns[7]. It is surprising that there is no obvious blocking effect where the root can be used as a noun. Several factors may play a role. Firstly, nominal uses of the roots are not common and they may be innovative. Several of the examples given above are taken from journalism and could be direct translations of English. Secondly, the morphologically derived nouns belong to a higher register than the zero derivations - there is a general tendency for morphological simplification in spoken Indonesian. Thirdly, some of the *ke- -an* derivatives do not have only the predictable meaning and this possibly allows the two forms to co-exist[8].

4.1.2.9 ATTRIBUTIVE USES

Indonesian codes attributes in noun phrases in two ways: either as simple modifiers, or as predicative modifiers in relative clauses. In some cases, either possibility is grammatical:

97. *rumah besar*
house big
'the big house'

98. *rumah yang besar*
house REL big
'the big house'

The structure in example 98 includes a relative clause, therefore any clausal predicate can occur in such structures and we would expect emotion and cognition words to be included. This is indeed the case:

99. *orang-orang warga negara Turki yang simpati*
person.DUP citizen state Turkey REL sympathetic
terhadap kelompok separatis Chechnya itu
towards group separatist Chechnya that
'the people who are sympathetic towards the Chechnya separatist movement' (McKay)

---

[7] The circumfix *ke- -an* has other functions, as discussed in section 3.1.1; irrelevant derivations are ignored.

[8] The secondary meanings of the *ke- -an* nominals often include a causative meaning although no causative morphology appears. Thus, *kesenangan* 'favourite' can be derived from *senangkan* 'make happy, make like'. This possibility appears elsewhere in the morphological system - see footnote 24 and section 3.3.1.

In general, verbs cannot be used as simple modifiers, but many emotion and cognition words can be used in such a structure. The most common collocation is with the root *rasa* used as a noun (compare the discussion of this word as a verb in the previous section):

100. *Menurut Blainey, kaum penganggur ini membangkitkan*
AS.follow Blainey class unemployed this AS.get.up.CAUS
rasa cemburu warga pembayar pajak negeri itu
feeling jealousy citizen payer tax state that
'According to Blainey, this class of unemployed gives rise to a feeling of jealousy in the tax-paying citizens of the country.' (McKay)

101. *Miss Nevill, yang telah mendapatkan kembali rasa*
Miss Nevill REL PAST AS.able.APPL return feeling
percaya dirinya
believe self.3
'Miss Nevill, who was able to recover a feeling of self-belief' (SDPGS :113)

Although these collocations are common, they should not be considered as compounds which provide the source of the verbal construction discussed in the previous section. If they were compounds, verbal suffixes would be expected to appear outside the second element, as is the case with the compound *beritahu* 'inform, (LIT. give news)'. This compound root can take the causative suffix *–kan* outside the entire unit:

102. *Ia sengaja memberitahukan kode kepada musuh*
3SG deliberate AS.inform.CAUS code to enemy
'He purposely revealed the code to the enemy.' (E&S : 74)

In contrast, the *rasa* + emotion or cognition word collocations do not behave in this way:

103. *Saya merasakan takut*
1SG AS.feel.APPL afraid
'I feel fear.'

104. **Saya merasatakutkan.*

Therefore there are two distinct ways in which the root *rasa* collocates with emotion and cognition words, as a verb with an attributive complement and as a noun with an attributive modifier.

Emotion and cognition words do occur as modifiers of other head nouns, but these usages are less common:

105. *Kedua, petani adalah pendukung setia Partai*
second farmer COP supporter loyal party
Demokratik Liberal yang dikalahkan oleh
democratic liberal REL US.defeated.CAUS by
Hosokawa
Hosokawa
'Secondly, the farmers are loyal supporters of the Liberal Democratic Party which was defeated by Hosokawa.' (McKay)

106. *kemudian Barnes bertanya dengan nada heran*
then Barnes ask with tone surprised
'then Barnes asked in a surprised tone' (SDPGS : 318)

Almost the same meaning as example 104 would be expressed if the word *nada* was omitted. The use of attributive words with the preposition *dengan* 'with' is very common in Indonesian. An example of this structure is the following:

107. *Japp memandangnya dengan heran*
Japp AS.look.at.3 with surprised
'Japp look at him, surprised.' (SDPGS : 269)

As will be discussed below, this construction is important in resolving the question of whether Indonesian has a lexical class of adjectives.

The examples of this and the preceding section have shown that emotion and cognition words (or most of those considered here) can be used as nouns and in attributive constructions without morphological derivation. On the other hand the construction discussed in section 4.1.2.2 gives the appearance of using such words as verbs. Therefore the question of whether these words can usefully be assigned to a particular lexical category must be addressed, and it will be the topic of the next section.

4.1.3 Lexical categorisation

Indonesian allows all major categories to act as the predicate of a clause. Many examples with verbal predicates have been seen in chapter 2, but in addition it is possible for nouns and prepositions to act as main predicate:

108. *Pak Bodomo guru bagus*
Mr Bodomo teacher excellent
'Mr Bodomo is an excellent teacher.'

109. *Ibu saya di Surabaya*
mother 1SG LOC Surabaya
'My mother is at Surabaya.'

It is also possible for words which the English speaker thinks of as adjectives to act as predicates:

110. *Rumah kita besar sekali*
house 1PL.INCL big very
'Our house is very big.'

Therefore, it is not self-evident that emotion and cognition predicates should be assigned to the lexical category verb. Indeed, they do not consistently exhibit the morphological possibilities characteristic of verbs as discussed in chapter 2 (excepting those words which can occur as underived transitive verbs - see section 4.1.2.4), and in many respects their behaviour is reminiscent of that of English adjectives. For example, English constructions of the type *I was angry with him* are common, and control constructions are possible with emotion adjectives: *I was happy to leave the city.* The possibility that the Indonesian words should be treated as adjectives must therefore be considered.

Before addressing this issue, however, two other questions must be resolved. Firstly, the same emotion and cognition words which function as predicates also appear in contexts which require nouns. As nouns can be the main predicate of a clause in Indonesian, the possibility that the words in question are basically nouns must also be considered. Section 4.1.3.1 considers and rejects this possibility. Secondly, the question must be asked: does Indonesian have the lexical category adjective at all? It is well known that languages without this category (or with a very restricted membership) exist (Dixon 1977) and a number of scholars (Cumming 1991, Thompson 1988) have argued that Indonesian is such a language, and that the words which express property concepts in Indonesian are a subclass of intransitive verbs. Section 4.1.3.2 examines the arguments and evidence relevant to this general question, and concludes that the distinction between adjective and verb does exist in Indonesian, although the two categories are more similar than they are in Indo-European languages. Section 4.1.3.3 then turns to discuss what the appropriate classification of emotion and cognition predicates might be.

4.1.3.1 NOMINAL USES OF EMOTION AND COGNITION WORDS

There are three arguments against treating the construction in which an emotion or cognition word takes a bare DP complement as structures with nominal predicates. Two arguments are syntactic and one is semantic. The first syntactic argument hinges on the DP following the emotion word. If a true nominal predicate has a dependent nominal following it, this can almost always be treated as a possessor, and therefore an attached pronoun can be substituted for the dependent:

111. *Pak Bodomo guru anak saya*
Mr Bodomo teacher child 1sg
'Mr Bodomo is the teacher of my child'

112. *Pak Bodomo gurunya*
Mr Bodomo teacher.3
'Mr Bodomo is her teacher.'

However, the bare DP complement of an emotion word can never be substituted by a reduced pronoun (examples 39-42 of chapter 2 repeated):

113. *Saya* *sayang* *pada* *Siti*
1SG pity to Siti

114. *Saya* *sayang* *Siti*
1SG pity Siti

115. **Saya* *sayangnya*
1SG pity.3

116. *Saya* *menyayanginya*
1SG AS.pity.APPL.3
'I pity Siti / her.'

The impossibility of this type of structure was discussed in section 2.2.2.1 in relation to the properties of verbs; here it also provides evidence against an analysis of examples like 114 as involving nominal predicates.

The second syntactic argument against this analysis is based on the type of negation used in clauses like example 114. As noted in section 4.1.2.7, the normal negator of nouns is *bukan*, and this is true for nominal predicates:

117. *Pak Bodomo* *bukan* *guru*
Mr Bodomo NEG teacher
'Mr Bodomo is not a teacher.'

118. **Pak Bodomo* *tidak* *guru*

A clause such as example 114 would always normally be negated with *tidak*[9]:

119. *Saya* *tidak* *sayang* *Siti*
1SG NEG pity Siti
'I don't feel sorry for Siti.'

It is not possible to make a categorical statement here, as almost any Indonesian clause can be negated with *bukan* if particular emphasis on the negation is intended:

120. *Saya* *bukan* *sayang* *Siti*
1SG NEG pity Siti
'I don't feel sorry for Siti at all.'

However, in the unmarked case negation provides a reliable test which distinguishes nominal predications from other types of predicates, and this test shows that example 114 and similar clauses are not nominal predications.

[9] In more colloquial registers, other negators are such as *nggak* are also available. They have the same distribution as *tidak* and I ignore them in the following discussion.

Finally, the semantic interpretation of nominal predications is not consistent with the meaning of clauses with emotion words and bare DP complements. Nominal predications in Indonesian are always equational, the two elements are understood as being identical:

121. *Dokter itu wanita*
doctor that woman
'The doctor is a woman.' (MacDonald & Darjowidjojo 1967 : 213)

122. *Hari ini hari gajian*
day this day pay
'Today is pay day.' (MacDonald & Darjowidjojo 1967 : 213)

But such an interpretation is meaningless for a clause such as example 114: *I* am not identical to *pity for Siti*. The equational interpretation is equally impossible in the case of a clause with a cognition predicate:

123. *kita lupa keyword yang sudah ditentukan*
1.PL.INCL forget keyword REL PERF US.fixed.CAUS
'we forgot the Keyword which had just been established.' (McKay)

We is not the same as *forget the keyword ...*; again the equational interpretation makes no sense.

The arguments advanced in this section, from both syntax and semantics, show that although many emotion and cognition words can be used as nouns, as discussed in section 4.1.2.8, without morphological change, clauses in which they appear followed by a bare DP complement cannot be treated as nominal predications.

4.1.3.2 ADJECTIVES AND VERBS IN INDONESIAN

Thompson (1988) gives explicit arguments that in Acehnese "property concepts do not form a separate lexical category from Verbs" (p169). She then includes Indonesian in a list of languages "for which similar arguments can be made" (p170), although this is immediately followed by the following qualification:

> That is, in each of these languages, whether or not a class of Adjectives can be identified on language-internal distributional grounds, Property Concept Words share many features with verbs. (p170)

It is thus not clear exactly what claim Thompson is making with respect to Indonesian. I will initially assume that her claim is to be read in the strongest fashion, that is, that Indonesian is like Acehnese and has no separate lexical category adjective.

Thompson offers four arguments for the lack of adjective as a lexical category in Acehnese, following Durie's analysis (1985: 101-103). The first argument is based on the possibilities for cross-referencing clitics and this is not applicable to Indonesian. The second argument is that Property Concept words can occur with aspect markers such as the inchoative *ka*. This

argument can also be made for Indonesian. Clauses in which a Property Concept word is the main predicate can have verbal auxiliaries[10]:

124. *Risikonya akan berat sekali*
risk.3 FUT heavy very
'The risk will be very great.' (IRG: 235)

125. *Minuman itu harus dingin*
drink that must cold
'The drinks must be cold.' (IRG: 235)

Thompson's third argument is that verbs in Acehnese can be used freely to code attributes, that is there is no syntactic difference between the DP-internal modifiers in examples 126 and 127:

126. Acehnese
aneuk muda nyan
child young that
'that young child'

127. *ureueng pula padè nyan*
person plant rice that
'that person planting rice'

This argument is not true for Indonesian, at least not in the same way that it is true for Acehnese. Where a verb is used attributively within a noun phrase in Indonesian, it always follows the relativizer *yang*:

128. *orang yang menanam padi itu*
person REL AS.plant rice that
'that person planting rice'

Words attributing properties to a noun can appear in the same structure:

129. *anak yang mudah itu*
child REL young that
'that young child'

But in this case, the possibility also exists to place the Property Concept word in direct construction with the head noun:

130. *anak mudah itu*
child young that
'that young child'

But this is not possible with a verb; the equivalent structure is predicative[11]:

[10] Teeuw (1962) claims that Indonesian adjectives cannot co-occur with modals such as *harus*. His study is based on writings of a single author published between 1920 and 1950 (see also Teeuw 1977 for details of methodology). Examples 124 and 125 suggests that what is accepted as standard has changed over the intervening period.

131. *?orang menanam padi itu*
person AS.plant rice that
'The person is planting that rice.'

Note that the structure here is radically different from that of example 128, where the demonstrative *itu* qualified the whole DP. In example 131, it only qualifies the object noun *padi*. Attributive constructions in Indonesian clearly differentiate between two classes of words, verbs and Property Concept words. Thompson's final argument from Acehnese is that many intransitive verbs which do not take an actor argument ('non-controlled verbs' in Durie's terminology) can be modified by the intensifier *that* 'very'. Indonesian does allow intensifiers with verbs (see example 68), therefore this argument is relevant:

132. *Dia mengerti sekali*
3SG AS.undertsand very
'She understood very well.'

Clearly, the arguments which show that verbs and adjectives are not distinct lexical classes in Acehnese do not transfer convincingly to Indonesian. The claim that Thompson can hope to sustain is that made in the passage quoted above: that verbs and adjectives are distinct lexical classes, but are also very similar in some respects. I will now argue that this is the correct position[12].

Teeuw (1962) gives detailed morphological and syntactic criteria to justify a class of adjectives in Indonesian. The most important morphological criteria are the following:

1) addition of the prefix *se-* derives a comparative form. Thus *besar* 'big' gives *sebesar* as in *sebesar telur ayam* 'as big as a chicken's egg'. This prefix is impossible with verbal roots: *beli* 'buy' *> *sebeli*.

2) addition of the prefix *ter-* derives a superlative form. Thus *terbesar* 'biggest'. There are many verbs formed with a homophonous prefix. The meaning relation between the original verb and the derived form is that the degree of control of the actor is lessened in the derived form: *bangun* 'wake up' > *terbangun* 'be wakened suddenly, come to one's senses'. The importance of this prefix is also pointed out by Mahdi (1998).

3) addition of the 3rd person bound pronoun *-nya* derives a noun:

[11] The following example is marginal because the object is more definite than the subject. Such structures are avoided in Indonesian.

[12] It should be noted that on the basis of the Durie/Thompson arguments for Acehnese the case of emotion and cognition predicates in that language is clearcut. Thompson's Property Concept words in Acehnese are non-controlled verbs, that is, they can only occur with optional undergoer enclitics and never with obligatory actor proclitics. But as seen in examples 11 and 12 above, there are emotion predicates in Acehnese which can take actor proclitics sometimes. They therefore do not form part of the same lexical class of verbs as Property Concept Words in Acehnese.

133. *Tingginya orang itu seratus tujuh puluh*
high.3 person that one.hundred seven ten
sentimeter
centimetre
'That person's height is one hundred and seventy centimetres.' (IRG: 306)

The bound pronoun *-nya* can only be attached to verbs with the prefixes *meN-* and *di-* (see extensive discussion in section 2.2.2.1), and it always codes a non-subject argument of the verb.

A crucial syntactic criterion separating adjectives and verbs which Teeuw discusses is the possibility for adjectives to form adverbial phrases with the preposition *dengan* 'with':

134. *Jumlah penduduk di kota Jakarta menanjak*
sum dweller LOC city Jakarta AS.ascend
dengan cepat sekali
with fast very
'Jakarta's population is rising very quickly.' (IRG: 208)

This construction is not possible with verbs:

135. **Jumlah penduduk di kota Jakarta menanjak*
sum dweller LOC city Jakarta AS.ascend
dengan mempercepat
with AS.*per*.fast

Nouns are possible following any preposition, including *dengan*, but other criteria establish a clear difference between adjectives and nouns. Firstly, the demonstratives *ini* and *itu* never modify an adjective or an adjective phrase. As adjectives follow the head in noun phrases, the two word classes are often adjacent, but without the head noun, or following a headless relative clause, their co-occurrence is ungrammatical:

136. *rumah besar ini*
house big this
'this big house'

137. *yang besar ini*
REL big this
'this big one'

138. **besar ini*
big this

Similar restrictions apply to the co-occurrence of adjectives and numerals[13]. Teeuw (1962) thus establishes a distinction between verbs and the other two

[13] An additional criterion for Teeuw is that verbs can be preceded by modal auxiliaries, whilst adjectives cannot. See examples 124 and 125, and footnote 10 for contrary evidence from contemporary Indonesian.

major lexical classes, and a distinction between nouns and adjectives, thereby showing that adjectives are a lexical class in Indonesian.

Where Thompson (1988) treats the similarities between verbs and property concept words by extending the class of verbs, Teeuw takes the opposite approach and argues that at least some verbs cannot be excluded from the class of adjectives. There is a class of intransitive verbs, traditionally referred to as basic verbs or *grondwoordelijke werkwoorden* in Dutch, which do not participate in the morphological alternations available to most verbs in the language. The class includes words such as *datang* 'come', *turun* 'descend' and *tidur* 'sleep'. An early twentieth century source[14] cited by Teeuw (1962: 416) even uses the term 'adjectives of action' for these words. The major consideration which influenced these scholars was that these verbs do not have forms with *meN-* or *di-*, nor do they occur in the Pro-V construction. As was argued in detail in chapter 2 of this work, however, these clause types are those available for verbs with two direct arguments. The morphological alternations involved signal the linking of actor or undergoer to the subject syntactic function, and this is not applicable to intransitive verbs except in a split-intransitive system. Such a system was discussed in relation to Acehnese data above, but clearly Indonesian does not have such a system. Therefore there is no reason to expect intransitive verbs to participate in the relevant morphological patterns. The other tests identified by Teeuw must therefore carry the full burden of telling us whether there are two word classes or only one, and for several tests, the evidence is unequivocal. There are no *ter-* prefixed derivatives of basic verbs having a superlative meaning, but some of them have such derivatives with a decontrolled meaning, for example *tidur* 'sleep' > *tertidur* 'fall asleep'. Adverbial phrases with *dengan* are impossible for basic verbs unless there is a homophonous noun. So the phrase **dengan turun* 'with descend' is not possible. Finally, Teeuw has a single example from his corpus of a basic verb deriving a form prefixed by *se-*:

139. *Akan* *tetapi* *di* *pasar* *belum* *seturun* *itu* *benar*
FUT but LOC market not.yet *se*.descend that correct
'But in the market it had not really gone down as much as that' (Teeuw 1962 : 418)

In general, such derivations do not seem to be possible. This evidence distinguishes the two classes clearly. Nevertheless, the point made by Thompson is once again shown to be valid: intransitive verbs and adjectives in Indonesian have a good deal in common.

Having established that there is a lexical class of adjective in Indonesian, I now turn to the question of what category emotion and cognition words should be assigned to, whether it is useful to claim that they belong to a single category as lexical roots. Whatever the answer to this

[14] C.A. Van Ophuisjen (1915) *Maleische spraakkunst Leiden,* publisher not known.

question, the further question can be asked whether this analysis can accommodate the clause type discussed in section 4.1.2.3 above.

4.1.3.3 THE LEXICAL CATEGORISATION OF EMOTION AND COGNITION WORDS

The range of possible uses of emotion and cognition words discussed in section 4.1.2 immediately suggests that it will not be possible to analyse all of these uses as structures involving a single lexical category. In at least some cases, zero-derivation of different categories will be necessary. The best analysis will be the one which minimises the necessity for such derivations. Given the arguments advanced in section 4.1.3.1 and the fact that not all the words under consideration can be used felicitously as nouns, I will assume that treating the words as basically nouns is not a good analysis. The choice then is between analysing them as verbs or as adjectives.

The criteria discussed in the previous section do not give an unequivocal answer, but they do favour the adjectival analysis for most of the words in question. As discussed in section 4.1.2.9, emotion and cognition words can be used attributively without the relativizer *yang* (examples 100, 101, 105 and 106) and they can form manner adverbials with the preposition *dengan* (example 107). The morphological behaviour of these words is also in general more adjectival than verbal, but the productivity of the adjectival morphology is limited. Of the words listed in Table 4.1, only the following have fully acceptable comparative forms with the prefix *se-*:

140.

berani	brave
cemas	worried
gila	mad
marah	angry
mirip	similar
peduli	concerned
prihatin	concerned
setia	loyal
takut	afraid
tega	heartless

Speakers find clauses which use those derivatives which do exist to be very mannered, and would not use them under normal circumstances. Thus the following is grammatical:

141.

Dia	*tidak*	*setakut*	*saya*	*dengan*	*anjing*
3SG	NEG	*se*.afraid	1SG	with	dog

'(S)he is not as afraid of dogs as me.'

but the meaning is more naturally expressed as follows:

142.

Dia	*tidak*	*takut*	*dengan*	*anjing*	*seperti*	*saya*
3SG	NEG	afraid	with	dog	as	1SG

'(S)he is not as afraid of dogs as me.'

One word which fits the criteria for the group under discussion only appears as a predicate with the prefix *se-* attached. The root *rupa* means 'form, kind', and *serupa* means 'to be similar'.

143. *Ia kembali dengan sebuah buku yang serupa*
3SG return with class book REL similar
dengan kepunyaan Mr. Morley
with possession Mr Morley
'He returned with a book which resembled Mr Morley's.' (SDPGS : 50)

I assume that this is a case of a homophonous prefix with a different function.

The superlative with *ter-* also is only sporadically possible, although rather more productive than prefixation with *se-*. The following words have derivatives of this type:

144.

bangga > terbannga	most proud
berani > terberani	most brave
cemas > tercemas	most worried
cinta > tercinta	most loved
gila > tergila	most crazy
jengkel > terjengkel	most disturbed
kasih > terkasih	most loved
khawatir > terkhawatir	most worried
marah > termarah	most angry
mirip > termirip	most similar
peduli > terpeduli	most caring
prihatin > terprihatin	most concerned
puas > terpuas	most satisfied
sayang > tersayang	most loved
senang > tersenang	happiest
setia > tersetia	most loyal
takut > tertakut	most afraid
tega > tertega	most heartless

Two points can be noted about these derivatives. Firstly, all the words which have *se-* derivatives also have *ter-* superlatives. These words therefore would seem to the prototypically adjectival members of the group. Secondly, some of the derivations are not regular. All three of the words which mean something like 'love', *cinta, kasih* and *sayang*, have derivatives in which the thematic role of the subject of the predication has switched, and indeed the glosses given in 144 are not quite accurate for these words. E&S give as a gloss for *tercinta* 'beloved', and it seems likely that the other two derivatives in this group are formed by analogy with this word and should have a similar translation. Two other words have similar derivatives: *kenal* 'know' > *terkenal* 'well-known, famous', and *percaya* 'believe, trust' > *terpercaya* 'trustworthy'. In all these cases, the prefix *ter-* adds a meaning of intensity in these derivations, which is compatible with its general function with adjectives, but it also changes the relations encoded by the predicate.

This leads to a further problem: as noted above the prefix *ter-* also derives verbs for which the participant's control of the situation is decreased, to the extent that a passive is often the best English translation. A number of emotion and cognition words have derivatives of this type:

145.	*heran > terheran*	astonished
	ingat > teringat	be reminded
	lupa > terlupa	forget accidentally
	pikir > terpikir	come to mind
	sadar > tersadar	realize, recover consciousness

An example of one of these derivatives is the following:

146.	*Untuk*	*pertama*	*kalinya*	*Fandi*	*teringat*	*lagi*
	for	first	time.3	Fandi	*ter*.remember	again
	pada	*Tuhan*				
	to	Lord				

'For the first time, Fandi was reminded again of God.' (SDM :272)

An additional complication is that the syntactic valence of this derived form is apparently also not fixed. As well as taking a prepositional phrase complement, it can also appear with a bare DP:

147.	*Tiba-tiba*	*saja*	*dia*	*teringat*	*suaminya*
	suddenly	EMPH	3SG	*ter*.remember	husband.3

'All of a sudden, she was reminded of her husband.' (SDM : 193)

There are two possible explanations of this structure. Firstly, it may be an example of an oblique argument adjoined to the verb and licensed in that way rather than with a preposition. The other two cases in which I have argued that this occurs both involve a morphologically complex verb, which is a point of similarity with the current example. This analysis is discussed further in section 5.3.5 Secondly, it should be recalled that *ingat* is one of the words which can serve as a verbal root without an applicative suffix, and this fact may explain the last example[15].

One particularly interesting example is the word *senang*. This word occurs quite commonly in clauses with a single argument and is usually translated as 'happy' in such contexts:

148.	*Indonesia*	*akan*	*senang*	*sekali*	*bila*	*PM*	*datang*
	Indonesia	FUT	happy	very	if	PM	arrive
	sebelum	*penyelenggaraan*		*pertemuan*	*itu*		
	before	implementation		meeting	that		

'Indonesia will be very happy if the PM arrives before the start of the meeting.' (McKay)

The same word also occurs in clauses with a second argument coded as a prepositional phrase, and in such contexts it is usually translated as 'like':

[15] *teringat* has other strange behaviour, described by IRG (116).

149. *Aku senang rumah itu dan aku senang akan ketenangannya*
1SG like house that and 1SG like about peace.3
'I like the house and I like its peacefulness.' (E&S : 499)

Such translations suggest that there is a *senang* which is an adjective, and one which is a verb. But it is not entirely clear that the translation does not impose this analysis, rather than any facts of Indonesian. It is not hard to find examples which suggest strongly that the translation dichotomy is wrong:

150. *Saya senang sekali atas hasil consolidasnya.*
1SG like/happy very about result consolidation.3SG
?'I like very much the result of the merger.'
'I am very happy about the result of the merger.' (McKay)

One can imagine a context where this would be translated as 'I like very much the success of the merger', but 'I am happy about the success of the merger' seems a good deal more natural. *Senang* also occurs with a complement clause, the subject position of which is controlled by the subject of *senang*. In many examples of such constructions, either translation is quite possible, with 'like' giving an habitual reading, and 'happy' suggesting a single occurrence:

151. *Dia memang senang berdiam berlama-lama di bawah pancuran*
she actually happy/like stay for a long time LOC under shower
'She was actually happy to stay under the shower for a long while' OR
'She actually liked to stay under the shower for a long time' (SDM:8)

Note that the difference in the translation here is basically one of aspect only, suggesting that the two concepts 'like' and 'happy' are close. There really seems no reason to assume two distinct words in Indonesian[16].

The evidence discussed in this section suggests that it may not be realistic to treat emotion and cognition words as though they all belonged to the same basic lexical category. A small group allow comparative and superlative formations like true adjectives, others can function as verbal roots without morphological derivation. The most realistic view is probably that there is a cline between adjectives and verbs in Indonesian, with some emotion and cognition words closer to one end of the continuum and some closer to the other. If this is the case, then the only useful approach would be to consider each word in detail and make a decision as to lexical category on a case by case basis, and this will not be attempted here. Rather, I turn to the question of whether the clause type in which an emotion and cognition word is followed by a bare noun phrase is a verbal clause or an adjectival clause.

[16] The case of *suka* makes it clear that these semantics are not arbitrary. Although it is not used as an intransitive predicate, *suka* is listed in the dictionary as a noun meaning 'joy, pleasure, happiness' as well as in its verbal sense, 'like'.

Previous studies (Kana 1986, Stevens 1970) took this for granted, as do the standard reference grammars (McDonald and Darjowidjojo 1967, IRG) in so far as they treat this type of clause at all. However, the discussion to this point indicates that the question is not straightforward, particularly as several of the most adjectival words, those with both comparative and superlative derivatives, can appear in this type of clause (*gila, marah, mirip, takut).* The crucial point is whether there is any evidence elsewhere in Indonesian that predicative adjectives can take a DP argument other than their subject. As far as I can discover, the only instance of this is in the comparative construction discussed previously (see example 141). In this case, the DP has a specific semantic role, object of comparison, and is presumably licensed by the morphology rather than by the adjectival base. There is evidence for this interpretation in the behaviour of the small group of words mentioned in this paragraph. For example, the word *takut* can take either a prepositional phrase or a noun phrase as complement*:*

152. *Saya takut dengan anjing*
1SG afraid with dog

153. *Saya takut anjing*
1SG afraid dog
'I am afraid of dogs.'

When it takes the prefix *se-*, the object of comparison is a noun phrase, but the stimulus, the cause of the emotion, can only be a prepositional phrase*:*

154. *Dia tidak setakut saya dengan anjing*
3SG NEG *se*.afraid 1SG with dog
'(S)he is not as afraid of dogs as me.'

155. **Dia tidak setakut saya anjing*

This data reinforces the impression that the clause type with a prepositional phrase is adjectival and that with a noun phrase is verbal.

An additional argument comes from the fact that the emotion and cognition words in the bare DP clause type share properties with a group of other words which are unequivocally verbal. Stevens (1970) discusses this evidence, and some of it will be important in the discussion of section 4.2. I have previously noted that a number of Stevens' 'pseudo-transitive' verbs are used as auxiliaries (example 19 and related discussion), and one of the emotion and cognition words shares this property. *suka* 'like' can be used in this way, as can be seen from the possibility of it appearing to the left of the actor in a Pro-V clause:

156. *Nasi itu suka saya makan*
rice that like 1SG eat
'I will eat that rice.' (Waruno Mahdi p.c.)

That the verb is semantically bleached here is shown by the fact that the selectional restriction on its subject has been relaxed. When used as a verbal or adjectival predicate, *suka* requires a human subject, or an animate portrayed as capable of experiencing the feeling of liking. But in this use, it is

possible to have an inanimate subject. The other words with this property lack the possibility of appearing with a prepositional phrase complement (but see section 4.1.1 for discussion of *ingin*), but, with this exception, *suka* behaves exactly like them. Arguing by analogy, *suka* must be a verb at least in the clause type with a bare DP complement.

One final point should be considered here. In a recent paper, Wierzbicka (1995) has argued that there are systematic differences in meaning between verbs and adjectives. She notes the fact that many emotion verbs in Russian have English adjectives as their closest translation equivalents, and claims that it is possible to detect a consistent difference in the degree of agency attributed to the experiencer across the two types of expression. A similar point is made in a study of adjective / verb minimal pairs by Pustet (2000). If such observations are true, it would be surprising if the prepositional complement clause type and the bare DP complement clause type should be headed by respectively an adjective and a verb and still have apparently identical meaning (see section 4.1.2.3 for a relevant example). This might be taken as an argument for treating the prepositional complement clause type as verbal also, although I would be reluctant to make this move for reasons which will be discussed in chapter 5. The argument from semantics is less compelling if we note that the studies referred to above deal with translation equivalents from two languages and pairs of words within a single language. But the data under discussion consists of words which may appear in two lexical categories without any morphological derivation. In such cases, it is legitimate to ask how much the semantic structure of the two items might differ. The null hypothesis in this case must be that where the meaning of a word is compatible with the semantic constraints associated with constructions into which two lexical categories can be inserted, then the word can be inserted into either with no change, or minimal change, to its semantic structure. The case of *senang*, discussed above, is very relevant. Although outside the scope of this study, it should not be too difficult to provide a semantics for this word which is compatible with both the intransitive uses and the transitive ones. That the result will not correspond closely to either *like* or *happy* in English is irrelevant.

To summarise this discussion, there is evidence that many, even most, emotion and cognition words have adjectival characteristics in Indonesian and the clause type which I take as criterial for membership of the word class is an adjectival construction. But many of the words also appear in the bare DP clause type, and the evidence just discussed suggests strongly that this is a verbal construction. The next section will take this as an assumption in its discussion, and chapter 5 will present arguments that the properties of that clause type only make sense within the full picture of Indonesian clausal syntax if that assumption is maintained.

4.2 Emotion verbs and grammatical functions

This section deals with the group of Indonesian verbs identified in the previous section and examines their status in regard to the syntactic transitivity system of that language. These verbs do not appear in the range of construction types analysed in chapter 2 and which I have assumed to be

typical of transitive verbs in Indonesian. The question this section seeks to answer is what grammatical function is assigned to the stimulus argument in a clause headed by an emotion or cognition verb and including two bare DPs (see examples in section 4.1.2.3 above).

I attempt to answer this question using the theoretical resources of Lexical Functional Grammar (LFG). This framework assumes a constrained list of grammatical functions (GFs): subject (SUBJ), object (OBJ), oblique (OBL_T) and second object (OBJ_T). The last two GFs are subscripted with a theta because in the case of OBJ_Ts their thematic role is assumed to be restricted within any particular language (Bresnan and Kanerva 1989), and in the case of OBL_Ts their thematic role is indexed by the adposition or case marker which licenses them. These GFs are not configurationally defined. They are elements of f(unctional)-structure, a structure which is distinct from but parallel to c(onstituent)-structure (see section 1.3.1 and Bresnan 2001a: 50-56). Syntactic properties other than configurational ones are therefore of primary interest in this investigation.

The group of predicates of interest here includes the following words:

157.

berang	angry, irate	*mimpi*	dream
bosan	bored	*mirip*	resemble
gila	crazy, obsessed	*peduli*	care about, pay attention
kangen	miss, long for	*sayang*	love, pity
kasih	love	*senang*	happy, like
kasihan	love, pity	*suka*	like
kuatir	fear	*takut*	fear
lupa	forget	*yakin*	sure
marah	angry		

This list is not exhaustive; rather it includes just those words for which I have clear examples from texts. Also, several words are left out of this list which might have a claim to be included. These words (*benci* 'hate', *ingat* 'remember' and *percaya* 'believe') all occur in the construction under consideration. But they also all have related transitive verbs which are not derived with an applicative suffix (*membenci*, *mengingat* and *mempercaya*). Therefore it is not possible to say whether clauses with these predicates and a DP stimulus are examples of the relevant construction, or whether they are rather examples of the unprefixed transitive verb construction discussed in chapter 2. Given this unclarity, it is preferable to omit these predicates from the discussion.

Section 4.2.1 examines the evidence that the emotion and cognition verb clauses are distinct from normal transitive clauses in Indonesian, and shows that the stimulus is neither a SUBJ nor an OBJ. Sections 4.2.2 and 4.2.3 examine evidence which might show that the stimulus is an OBL_T or an OBJ_T respectively, concluding that the second possibility is to be preferred.

4.2.1 Transitive Clauses - SUBJ and OBJ

The overwhelming majority of transitive verbs in Indonesian can appear in four types of clause:

158. *Dia membaca buku itu*
3SG AS.read book that
'S/he read the book.'

159. *Buku itu dibaca (oleh) Ali*
book that US.read (by) Ali
'The book was read by Ali.'

160. *Buku itu saya baca*
book that 1SG read
'The book, I read.'

161. *Dia baca buku itu*
3SG read book that
'S/he read the book.'

Various restrictions apply to the constructions exemplified in 158 to 161; these have been discussed in detail in chapter 2. Clauses with emotion predicates look the same as the type exemplified in 161, but the other clause types are not possible:

162. *Aku bosan hidup terus-menerus di Surabaya*
1SG bored life continue-AS.continue LOC Surabaya
'I am fed up with living all the time in Surabaya.' (PYD: 180)

163. **Aku membosan hidup terus-menerus di Surabaya.*

The type of clause illustrated in example 159 is not possible with a first person actor, but substitution of a third person actor does not make this construction possible with this predicate:

164. **Hidup terus-menerus di Surabaya dibosannya.*

The type exemplified in 160 is not possible either, although the direct manipulation of example 162 yields a sentence which does have a reading:

165. *Hidup terus-menerus di Surabaya aku bosan.*
'Living all the time in Surabaya, I am fed up with it.'

This clause requires an intonation distinct from that associated with clauses of the type seen in example 160 with a pause after the initial DP which is a topicalized constituent, and the addition of an auxiliary verb to the sentence makes the distinction absolutely clear. In a construction of the type in example 160, the actor pronoun remains adjacent to the main verb even when an auxiliary or modal is included, but in emotion predicate clauses, the actor pronoun appears to the left of an auxiliary:

166. *Buku itu akan saya baca.*
book that FUT 1SG read
'I will read that book.'

167. *Hidup terus-menerus di Surabaya aku akan bosan.*

'This monotonous life in Surabaya, I will be fed up with it.'

168. **Hidup terus-menerus di Surabaya akan aku bosan.*

This evidence shows that these predicates do not participate in the system of clauses associated with verbs assumed to be transitive in Indonesian. Therefore it is not possible to describe the emotion and cognition verbs as syntactically transitive.

This status alone may be enough to show that the stimulus argument of an emotion predicate is not an OBJ. If OBJ is defined as the non- SUBJ argument of a transitive verb, then stimuli are not objects because the predicate that they are associated with is not syntactically transitive[17]. A more stringent test requires OBJs be able to become the SUBJ of a related clause, typically a passive. Chapter 2 has provided detailed arguments as to whether the relationships between the possible clause types for transitive verbs in Indonesian can be compared to the active-passive relation in a language like English. Leaving aside that issue, it is clear that stimuli also fail this test as seen in example 164. Many of the emotion predicates derive true transitive verbs with an applicative suffix, and a comparison of their properties with the basic predicates is instructive. The applicativized verbs appear with both *meN-* and *di-* prefixes:

169. *Dia merasa perempuan itu sedang* ***mengasihaninya***
3SG AS.feel woman that PROG AS.pity.APPL.3
'She feels that that woman is pitying her.' (SDM: 86)

170. *Dan dia tidak suka* ***dikasihani***
and 3SG NEG like US.pity.APPL
'And she does not like to be pitied.' (SDM: 86)

Further examples are given above in section 4.1.2.4. The contrast in syntactic behaviour between the derived transitive verb which does have an OBJ and the emotion predicate is clear.

The possibility that the stimulus is the SUBJ of its clause can also be eliminated for various reasons. Firstly, it would be surprising that an argument alternated between being an OBL_T (in the prepositional construction) and a SUBJ with no verbal morphology. Secondly, it can be seen from example 169 that the experiencer is the SUBJ of the corresponding applicative verb with *meN-* prefix, the clause type which corresponds to an active clause. Applicativization does not affect SUBJs, therefore the experiencer should be the SUBJ of the basic predicate. A clause can only have one SUBJ, therefore the stimulus cannot be one. Further, Indonesian has a third person pronoun, *ia*, which is restricted (more or less) to appear in SUBJ position. This pronoun can occur as experiencer of an emotion predicate clause:

[17] I assume that, as SUBJ and OBJ are syntactic concepts, syntactic transitivity is the relevant notion for this argument.

171. *Ia* *kuatir* *suratnya* *tidak* *sampai*
3SG fear letter.3 NEG arrive
'She was afraid her letter would not arrive.' (ES: 313)

Once again, the fact that a clause can only have one SUBJ forces the conclusion that the stimulus cannot be SUBJ. Finally, the examples in section 4.1.2.7 with complement clauses with a controlled subject position showed that the experiencer was the sole nominal argument of the emotion predicate, and controller of the reference of the missing argument. It must therefore be SUBJ in such examples.

This section has demonstrated that emotion predicates are not transitive verbs in Indonesian. Therefore, a bare DP stimulus argument following such a predicate cannot be an OBJ. And the (remote) possibility that such arguments are SUBJs has also been eliminated.

4.2.2 OBL_T

Having established that a DP stimulus of an emotion predicate cannot be assigned the GFs SUBJ or OBJ, I now turn to examine the arguments as to whether such elements might be classified as OBL_T. I have discussed above (section 4.1.2.3) the apparent lack of semantic differentiation between the clause type with a prepositional complement and that with a bare DP complement. Another example (see also examples 58 and 59) which makes this clear is the following, in which the two types are co-ordinated:

172. *Aku* *senang* *rumah* *ini,* *dan* *aku* *senang*
1SG like house this and 1SG like
akan *ketenangannya*
about NOM.quiet.NOM.3
'I like this house, and I like it's peacefulness.' (E&S: 499)

Such a difference in syntactic coding is normally correlated with a semantic difference. Therefore, it might be possible to argue from the lack of semantic differentiation that the stimulus clauses with a bare DP complement is still an OBL_T, but one with non-standard coding. There are two pieces of evidence which suggest that this would be a wrong conclusion.

But before presenting this evidence, there is once more a definitional issue. Indonesian is not a case-marking language, and therefore the obvious distinction between core and non-core arguments is the presence of prepositions. This is not a definitive criterion however. The literature includes examples both of DPs with the coding properties of obliques being direct arguments, and of DPs with the coding properties of direct arguments being obliques. Examples of the first type are animate, referential objects in Spanish, which are preceded by the preposition *a*:

173. *Busco* *a* *mi* *amigo*
seek.1SG *a* 1SG.POSS friend
'I'm looking for my friend.' (Hopper & Thompson 1980: ex11b)

and various subjects in Icelandic with 'quirky' case-marking:

174. *Henni hefur alltaf þótt Ólafur leiðinlegur*
she.DAT has always thought Olaf.NOM boring.NOM.SG
'She has always considered Olaf boring.'
(Zaenen, Maling & Thráinsson 1985: ex13))

An example of the second type is the use of accusative case in Icelandic to mark some temporal adjuncts (Smith 1996: 40-42):

175. *Ég var þar tvo daga*
1SG.NOM was there two.ACC days.ACC
'I was there for two days.' (Smith 1996: 41, ex49)

Also relevant here is the syncretism of ergative case, which I assume marks terms, with some oblique case, typically instrumental or genitive, in many languages with ergative-absolutive morphology (Blake 1977: 60). The oblique arguments adjoined to verbs discussed in chapters 2 and 3 are also examples of this type. Such examples show that coding properties alone are not sufficient to establish the syntactic status of an argument. Behavioural tests are necessary also, and I now turn to two such tests which show that the stimulus of the first conjunct of example 172 and similar clauses does not have the properties of an oblique argument.

The first way in which DP stimuli behave like direct arguments is that a quantifier can be floated from them. This is not possible with a PP stimulus, an oblique argument, or a PP which is an adjunct:

176. *Anak-anak itu suka gula-gula itu semuanya.*
child.DUP that like sugar.DUP that all
'All the children like the sweets.' OR
'The children like all the sweets.'

177. *Anak-anak itu suka dengan gula-gula itu semuanya*
child.DUP that like with sugar.DUP that all
'All the children like the sweets.' NOT
'The children like all the sweets.'

178. *Orang-orang Sasak datang dengan anak-anaknya*
man.DUP Sasak come with child.DUP.3
semuanya
all
'All the Sasak people came with their children.' NOT
*'The Sasak people came with all their children.'

In each case, the quantifier can be read as having floated from the SUBJ, and evidence to be presented below will show that other direct arguments also have this property. But only with the DP complement of an emotion verb can the quantifier be read as having floated from the other DP in the clause; when that DP is within a PP the quantifier cannot be construed with it[18]. Section

[18] One of my consultants allows both readings of example 177, but the second reading seems to be marginal and is strongly rejected by the other consultants.

2.3.1.3 has presented evidence that the adjoined oblique actor in a *di*-V-DP clause cannot float a quantifier, as has section 3.2.2.1 for the non-subject argument of *ke- -an* verbs .

The second piece of evidence comes from extraction facts. Indonesian subjects can always be extracted leaving a gap and this generalisation is agreed on by all sources. The facts regarding other arguments of transitive verbs are in dispute; they have already been mentioned in section 2.2.2.2 and will be discussed in detail in chapter 5. For emotion predicates, I am only aware of one discussion of extraction in the literature, Stevens (1970), and this source claims that extraction of DP stimuli is grammatical[19]. This judgment is shared by the native speakers I have consulted. Gapped extraction from PPs is completely impossible however[20], neither preposition-stranding nor pied-piping is available in Indonesian. The status of the PP involved, argument or oblique, is not relevant:

179. *gula-gula yang anak-anak suka itu*
sugar.DUP REL child.DUP like that
'the sweets that the children like'

180. **gula-gula yang anak-anak suka dengan itu*
sugar.DUP REL child.DUP like with that

181. **gula-gula dengan yang anak-anak suka itu*
sugar.DUP with REL child.DUP like that

182. **anak-anak yang orang Sasak datang dengan itu*
FOR: 'the children that the Sasak people came with'

Section 2.2.2.2 has presented evidence that the adjoined oblique in a *di*-V-DP clause cannot be relativized, and section 3.2.2.1 has made the same point in respect of the non-subject argument of adversative verbs. Clearly, the behavioural properties of DP stimuli are different from those associated with the grammatical function OBL_T in Indonesian, regardless of whether the OBL_T is coded as a PP or adjoined to the verb. Therefore it is not possible to maintain an analysis which treats these constituents as obliques, and it seems that the only possible analysis within LFG's inventory of grammatical functions is to treat the stimuli as OBJ_Ts. I turn to an examination of the evidence for this position in the following section.

The data presented in section 4.1.2.2 suggesting that in the prepositional construction the stimulus is assigned a semantic role directly by the predicate, with the preposition making no semantic contribution may also be relevant here. This behaviour is not typical of true OBL_Ts, which display semantic interdependence with the preposition or case-marker which licenses them. Therefore this evidence suggests that rather than the bare DPs being

[19] Kana (1986) discusses emotion predicates in some detail (see also below), but does not consider extraction.

[20] Extraction with a resumptive pronoun is possible: see IRG :289-291 and section 5.1.1.

non-standard OBL_Ts, it is actually the arguments coded as OBL_Ts which are non-standard. The head of the DP is the semantic head of the whole constituent. For this reason, when developing an explicit analysis of the emotion and cognition verbs in section 4.3, I treat these prepositions as case-markers rather than as heads.

4.2.3 OBJ_T

As mentioned above, OBJ_T is the grammatical function of the second object of ditransitive verbs for LFG. Indonesian has a small number of underived verbs which are ditransitive, and a large number derived with the applicative suffixes *-i* and *-kan*. I will discuss here three properties of the second objects of such verbs: access to SUBJ in a related clause, quantifier float and extraction and compare them with the properties of DP stimuli.

The second object of a ditransitive clause with a *meN-* prefixed verb cannot become subject of a clause with a *di-* prefixed verb:

183. *Ali mengirimi wanita itu surat itu*
Ali AS.send.APPL wanita that letter that
'Ali sent the woman the letter.'

184. **Surat itu dikirimi wanita itu*
letter that US.send.APPL woman that
('The letter was sent the woman.) (Chung 1976a: ex.63a)

In the case of derived ditransitive verbs such as that in example 182, the same thematic argument would be OBJ in a related clause with an underived, *meN-* prefixed verb and has access to SUBJ function in a clause with a *di-*prefixed unsuffixed verb:

185. *Surat itu dikirim kepada wanita itu*
letter that US.send to woman that
'The letter was sent to the woman.'

The motivation for applicative is to make the third argument available to syntactic processes which require core status; there is no motivation to maintain the status of other arguments. In respect of access to SUBJ, the behaviour of the OBJ_Ts and the DP stimuli is similar. DP stimuli cannot become SUBJ of a related clause either unless the verb is morphologically altered. They can become SUBJ of a related clause headed by a *di-*prefixed derived verb (example 170) and the behaviour of the two argument types diverges here. Second objects cannot be SUBJ or OBJ with derived verbs while stimuli can. However, the relevant verbal derivation, applicative, promotes an oblique argument to core status, therefore the source of the applicativized emotion verbs must be the construction with a PP stimulus and the comparison between the two types is not in fact direct on this point[21].

[21] I ignore here the issue of applicative derivations in Indonesian which take a transitive verb and derive another transitive verb with a different meaning e.g. *memergok* 'catch someone by surprise', *memergoki* 'catch someone redhanded'.

In the previous section, examples were given which showed that a quantifier could float from a SUBJ. This is also possible from OBJs and from OBJ_Ts:

186. *Saya mukul anak-anak itu kemarin semuanya*
1SG hit child.DUP that yesterday all
'I hit all the children yesterday.'

187. *Saya memberinya hadiah itu semuanya*
1SG AS.give gift that all
'I gave her all the presents.'

The descriptive generalisation is that quantifiers can float from any term argument in Indonesian. And as previously shown, a quantifier can also float from an DP stimulus (example 176, repeated here):

188. *Anak-anak itu suka gula-gula itu semuanya.*
child.DUP that like sugar.DUP that all
'All the children like the sweets.' OR
'The children like all the sweets.'

In this case, the stimuli behave like term arguments; they are not SUBJ or OBJ, so OBJ_T is the remaining possibility.

Many descriptions of Indonesian state that extraction (relative clause formation and question formation) is only possible with SUBJs. The facts are a good deal more complicated than this as soon as anything other than the prescriptive standard is investigated; a more complete account is given by Voskuil (1996: Ch 8), see also sections 2.2.2.2 and 5.1.The discussion of chapter 2 suggests that the correct generalisation is that SUBJ and OBJ can be extracted in the absence of a verb prefix. Thus an OBJ can be extracted from a clause of the type seen in example 161 (auxiliary added to disambiguate the construction)[22]:

189. *buku yang dia akan baca itu*
book REL 3SG FUT read that
'the book that she will read'

Chapter 2 did not consider the extractability of second objects, and judgments vary for these arguments. As might be expected, such arguments cannot be extracted when the verb carries a prefix, and no native speaker I have consulted will allow extraction when the verb is derived with an applicative suffix. In the case of clauses with underived, unprefixed ditransitive verbs, however, at least some speakers will permit extraction of the second object[23]:

[22] Michael Ewing (p.c.) informs me that he has observed numerous examples of such constructions in recorded conversation of educated Indonesians.

[23] This is not true for the verb *ajar* 'teach' for my primary consultant. I have no explanation of this fact.

190. *Saya akan beri dia buku itu*
1SG FUT give 3SG book that
'I will give him the book.'

191. *?buku yang saya akan beri dia itu*
book REL 1SG FUT give 3SG that
'the book that I will give him'

As demonstrated above (example 179), this type of extraction is also possible with a DP stimulus:

192. *orang yang saya suka itu*
person REL 1SG like that
'the person that I like'

On this test, the DP stimulus possesses the syntactic properties of a direct argument more clearly than OBJ_Ts. I have demonstrated that DP stimuli cannot be plausibly assigned any of the other three GFs recognised by LFG, and that their syntactic properties are very similar to those of OBJ_Ts. Within the theory as currently formulated, this would seem to be the only possible analysis. Section 3.3.2.3 has already raised the possibility of LMT predicting such an array of GFs, although in that case the data showed that it could not be correct. But in the case of emotion and cognition verbs, the data do support such an analysis. The only alternative is to add a GF to the inventory, and the final section of this chapter will examine which of these two possibilities is preferable from a theoretical point of view. First, I discuss the formal issues which arise in implementing such an analysis.

4.3 Formalising the analysis

4.3.1 An LFG analysis

There are two issues to be addressed in analysing the Indonesian emotion and cognition verbs using the apparatus of Lexical Mapping Theory. The obvious question is what feature should be lexically specified for the second argument in order to ensure that it is mapped to the OBJ_θ GF. There are two possibilities, and these will be discussed in section 4.3.1.2. Before that, I deal with a less obvious question: what is the status of the first argument, the experiencer, of these verbs?

4.3.1.1 TRANSITIVE OR TWO-PLACE UNACCUSATIVE

The principles in LMT which assign features to semantic roles give particular emphasis to patientlike roles (Bresnan 2001a: 309). The first such role in the a-structure is assigned the feature [*-r*], with other similar roles assigned the value [*+o*][24]. These principles have two effects. Firstly, they ensure that the arguments of ditransitive verbs are linked correctly. The first object will have the feature [*-r*] and the lowest role it can link to is thus OBJ.

[24] I follow the principles given by Bresnan (2001a: 309) for English. As far as I can tell, these principles also work for Indonesian.

The second object, in a language such as English or Indonesian which does not allow symmetrical objects (Bresnan and Moshi 1990), will have the feature [*+o*], and the lowest role with which it is compatible is OBJ_{θ}. The second effect is to make a distinction between unaccusative and unergative intransitive verbs. Because unaccusatives have a patientlike argument, they are assigned the feature [*-r*], while the argument of unergatives receive the feature [*-o*] by a default principle. While this distinction has no significance in the mapping to GFs (the argument of all intransitive verbs maps to SUBJ), it does allow other effects to be explained neatly (see Bresnan 2001a: 312-314 for discussion of unaccusatives and resultatives).

This analysis of intransitive verbs allows the possibility that two-place unaccusative verbs might exist. This possibility has been discussed in detail within the P&P framework in Belletti and Rizzi (1988) in an analysis of Italian psych verbs to which I will return. Note that in the current framework, such verbs would have exactly the properties I am suggesting the Indonesian emotion and cognition verbs possess. The first argument in the a-structure, being patientlike will be assigned the feature [*-r*]. The second argument is also patientlike, and given that the language does not allow symmetric objects, the only possible feature assignment is [*+o*]. In the absence of a logical subject with the feature [*-o*], a [*-r*] argument can map to SUBJ and therefore the first argument will do so. The second argument then maps to the lowest role compatible with its feature specification, which is OBJ_{θ}. The attraction of this analysis is that once the type of the verb is specified, everything else falls out from general principles of the theory. If this analysis were correct, we would predict that the emotion and cognition predicates which can have a bare DP second argument would all be unaccusatives. It is then a question of some importance for this discussion whether some or all emotion and cognition verbs in Indonesian can be considered as unaccusatives.

The suggestion is not preposterous on semantic grounds. The existence of verbs in many languages which code the experiencer as object, such as *frighten* in English, shows that it is possible to construe an experiencer as affected by the event denoted by the verb. Some accounts emphasise that in such cases there are reasons from the semantics of the arguments (Pesetsky 1995) or the aspectual structure of the event (Grimshaw 1990) to take the experiencer as the less prominent argument. But Dowty (1991) discusses the respective proto-Agent and proto-Patient entailments of the two arguments of verbs denoting psychological processes, and concludes that both have something in common with each proto-role. Being affected by the event is the entailment which Dowty sees as most clearly identifying the experiencer with the proto-Patient role. In addition, the aforementioned study by Belletti and Rizzi (1988) sees no inconsistency in treating at least some psych verbs as having two internal arguments.

Kana (1986: chapter 3) identifies two morphological tests which she claims separate unergative and unaccusative verbs in Indonesian. Firstly, she notes that the suffix *-kan* can be used to add a second argument to some intransitive verbs. In some cases, the additional argument is an object:

193. *Pegawai itu bekerja di kantor gubernur*
employee that *ber*.work LOC office governor
'That employee works in the governor's office.'

194. *Sri mengerjakan pekerjaannya dengan baik*
Sri AS.work.APPL work.3 with good
'Sri carried out her work well.' (Kana 1986: 56, ex20b,a)

In other cases, it is a subject:

195. *Dia tidur di atas tikar*
3 sleep LOC top straw.mat
'She sleeps on a straw mat.'

196. *Ibu menidurkan anaknya*
mother AS.sleep.CAUS child.3
'Mother put her child to sleep.' (Kana 1986: 57, ex.26b,a)

As my glosses above indicate, the traditional analysis is that the suffix has two functions, as an applicative and as a causative. Kana claims that the suffix is a transitivizer whose effect depends on the valence of the base verb. Where the base is unergative, the suffix adds an object (i.e. applicative), and where the base is unaccusative it adds a subject (i.e. causative)[25]. Similar correlations have been noted for multi-functional affixes in other languages (Austin 1997a).

Secondly, Kana claims that out of a range of nominalising morphologies available in Indonesian one is restricted to being used with verbs which have a patient (an initial 2 for Kana) in their lexical structure. The circumfix[26] *peN- -an* derives nouns from verbs which have the meaning of 'the action of Ving', and they can take the undergoer of the original verb as an argument of the noun:

197. *bunuh* *pembunuhan* *serdadu*
kill NOM.kill.NOM soldier
'murder of the soldier'

198. *sampai* *penyampaian* *surat*
arrive NOM.arrive.NOM letter
'delivery of a letter'

Kana does not provide detailed negative evidence establishing that the verbs she identifies with this test are unaccusative, but when such evidence is examined it suggests that this test is not reliable, or that large numbers of Indonesian verbs have multiple lexical entries. Several of the verbs which Kana uses to exemplify the function of *-kan* with unergatives turn out to have *peN- -an* nominals derived from them:

[25] For a different account, which also includes the suffix *-i*, see Voskuil 1996.

[26] The capital *N* here has the same meaning as it does in the actor subject prefix *meN-*: the nasal undergoes various assimilations to following consonants. See chapter 1 for details.

199. *Orang tua saya akan bercerai*
person old 1SG FUT *ber*.separate
'My parents are going to get divorced.'

200. *Anwar menceraikan isterinya*
Anwar AS.separate.APPL wife.3
'Anwar divorced his wife.' (Kana 1986: 60, ex34b,a)

201. *penceraian* - divorce, separation

202. *Saya ingin berbicara dengan saudara*
1SG wish *ber*.speak with cousin
'I want to speak with you.'

203. *Mereka membicarakan soal itu*
3PL AS.speak.APPL matter that
'They discussed the problem.' (Kana 1986: 60, ex37b,a)

204. *pembicaraan* - discussion

Given such data, and the fact that *peN- -an* nominalisation is not widespread for emotion and cognition verbs, I will not use this test in what follows[27]. This leaves the evidence of *-kan* suffixation as the only test for the status of these verbs.

For the purposes of the following discussion, it is necessary to assume either that uses of emotion and cognition verbs with a single argument are verbal rather than adjectival contrary to the conclusions of section 4.1.4.2, or that adjectival predicates in Indonesian take a subject and that this argument has a semantic role which is not agentlike, that is, that they are unaccusative predicates. As discussed in section 4.1.2.4, some of the Indonesian words under consideration do form causatives with *-kan*. Very often the use of such derived verbs is quasi-adjectival, with only a single argument:

205. *Orang Inggris adalah orang yang sangat membosankan*
person England be person REL very AS.boring.CAUS
'English people are very boring.' (SDPGS: 92)

But examples do occur in which they do take two arguments and it is clear that the object of the causative verb is the experiencer:

[27] Two of the verbs which are potential unaccusatives, *kenal* and *yakin* do have *peN- -an* nominalizations. Interestingly, in each case the nominal would seem to be based on the causative verb derived from the base even though the *-kan* suffix does not appear in the nominal. Thus, *kenal* 'know' has a causative *kenalkan* 'introduce' and *pengenalan* means 'introduction, and *yakin* 'sure' has a causative *yakinkan* 'convince' and *peyakinan* means 'act of convincing'. If this possibility is systematic, Kana's claim as to the nominalisation possibilities of unaccusatives might be dependent on the parallel possibility of causativization. See also footnote 6 and section 3.3.1 for the morphological issue.

206. *Ia menyukakan anaknya pada anjing*
3SG AS.like.CAUS child.3 to dog
'She taught her child to like dogs.' (E&S: 530)

207. *Bermimpi jahat tentu menakutkannya*
dream evil certainly AS.afraid.CAUS.3
'He surely was frightened by a nightmare.' (E&S:544)

As discussed in section 4.1.2.4, there are also emotion and cognition verbs which use the suffix *-kan* with applicative effect. However, closer examination suggests that there may be a degree of correlation between being able to take a bare DP second argument and forming causative derived verbs with *-kan*.

There are five verbs which are apparent counter-examples to this correlation. One of them, *ingat* 'remember', has a corresponding unsuffixed transitive verb. It is therefore possible that the causative derivation is based on this verb. Another potential counter-example is the verb *takut* 'fear', and as discussed in section 4.1.2.4, this verb can form applicatives and causatives with both suffixes, *-i* and *-kan*. It would be hard to base any useful argument on the behaviour of this verb. Of the remaining three verbs, for one, *peduli* 'care, attend', I have a single example in which it takes a bare DP second argument which is also not reliable evidence. The remaining two verbs, *lupa* 'forget' and *mimpi* 'dream' both provide more serious counter-evidence. *lupa* appears with a bare DP following in several examples:

208. *Kamu lupa rumahku?*
2SG forget house.1SG
'Have you forgotten my house?' (SDM: 96)

and it is quite common with the *-kan* suffix (21 instances in my database), always with the experiencer as subject. The verb *mimpi* is less common, but the facts seem to be similar. These two verbs then are real problems for the unaccusative analysis.

One possibility is to take *lupa* and *mimpi* (and others which I have not come across, no doubt) to be a separate class of verbs for which the experiencer is treated as agentlike. To do this would in turn raise the question of whether there was any semantic basis for this classification, or whether it was a case of lexical idiosyncrasy. The division between unergative and unaccusative verbs is notoriously idiosyncratic (Rosen 1984), and this might be a possible account. It would still be necessary, however, to also give an account of the group of verbs which take a bare DP second argument and form applicatives with *-i*. Kana (1986) does not claim that this suffix can only attach to intransitives of one type. Her claim is rather that *-i* signals that a locative has been advanced to object status.

This is an additional complication for the two-place unaccusative account. Three classes of verbs must be recognised: transitives with a non-locative second argument (at least *lupa* and *mimpi*), two-place unaccusatives (the verbs with *-kan* causatives) and verbs with a locative as second argument (those with *-i* applicatives). The status of the first argument of this last group

of verbs would be unresolved; either they are not unaccusative because they do not take *-kan* as an applicative suffix, or locative advancement takes priority in the morphological rules. There is however good evidence to reject the analysis of the stimulus argument as an underlying locative. As documented in detail in section 4.1.2.2, the true locative preposition *di* never occurs with emotion and cognition predicates, and the other prepositions which code static location in Indonesian are also almost unrepresented. Together with the fact that for almost every one of the predicates listed in Table 4.2 both *pada* and *dengan* are possible, this means that an extremely liberal interpretation of the notion locative would be needed to carry the analysis through.

An alternative, and to me, preferable analysis is to assume that the emotion and cognition verbs form a unified class, and that they are genuinely transitive, that is, that they have an external argument. This means accepting that there is some idiosyncrasy in the uses of the suffix *-kan*, an unremarkable proposition when the behaviour of the verb *takut* is recalled, and that the correlation between emotion and cognition verbs and *-kan* applicatives is suggestive but not very significant.

4.3.1.2 LEXICAL FEATURES OF THE STIMULUS

Following the arguments of the previous section, I will treat Indonesian emotion and cognition verbs as two-place verbs with an external argument. This means that the first argument in the a-structure will not be taken as patientlike and associated with the value [*-r*]. But the second argument cannot be associated with that value either, or it would map to the OBJ GF, which is not the desired result. In order for the second argument to map to the OBJ_θ function, it must be associated with either the value [*+r*] or the value [*+o*]. The argument must map to the most marked compatible function: if the value [*+r*] is associated with it the options are OBJ_θ and OBL_T, if the value [*+o*] is associated the options are OBJ or OBJ_θ. In either case, OBJ_θ is more marked and will be the function mapped to. The question is whether there is any reason to prefer specifying one feature in the lexicon rather than the other.

I will argue here that there is a strong reason for preferring to specify the feature [*+r*], and this argument is based on one of the claims of LMT, that there are natural classes of argument functions. In section 4.1.2.2, I argued that even when emotion and cognition predicates are followed by a prepositional phrase, they assign a semantic role to the argument inside the PP. The essential difference between that construction and the construction with a bare DP complement lies in the ability or inability of the predicate to syntactically license a second argument. Depending on one's view of the evidence discussed in section 4.2, this is dependent either on the difference between an adjective and a transitive verb, or an intransitive verb and a transitive one. Regardless of which analysis one prefers, it is the case that the complement functions in the two constructions form a natural class in LMT. OBJ_θ and OBL_T appear in the same column of the table, example 63 of chapter 1, that headed by the feature [*+r*]. Therefore, if this feature is used in the lexical entry of emotion and cognition verbs, we can capture the generalisation that in the two different constructions, these predicates have

the same semantic argument array. The differences in surface coding are due to other properties of the predicates.

I therefore propose the following lexical entries for the two possibilities for the predicate *senang* 'like':

209. *senang*, V_{tr} 'like' < x, y >
|
[+*r*]

senang, V_{itr}/A 'like' < x, : y >
|
[+*r*]

In either case, the first argument will be assigned the feature value [-*o*] by the default principle. In the first case, the second argument maps to the lowest compatible GF, OBJ_θ:

210.
$$\begin{bmatrix} \text{PRED} & \text{'senang < SUBJ, OBJ-}\theta\text{ >'} \\ \text{SUBJ} & \begin{bmatrix}\text{PRED} & \text{'Ali'}\end{bmatrix} \\ \text{OBJ-}\theta & \begin{bmatrix}\text{PRED} & \text{'rumah'} \\ \text{DEF} & + \end{bmatrix} \end{bmatrix}$$

Ali senang rumah itu.

while in the second case it is not a direct argument and therefore can only be realised as an OBL_T. Following the discussion of section 4.1.2.2, I do not treat the preposition as the predicator in the prepositional phrase, but only as a case marker:

211.
$$\begin{bmatrix} \text{PRED} & \text{'senang < SUBJ, OBL-}\theta\text{ >'} \\ \text{SUBJ} & \begin{bmatrix}\text{PRED} & \text{'Ali'}\end{bmatrix} \\ \text{OBL-}\theta & \begin{bmatrix}\text{PRED} & \text{'rumah'} \\ \text{DEF} & + \\ \text{C-MARKER} & \text{'dengan'}\end{bmatrix} \end{bmatrix}$$

Ali senang dengan rumah itu.

But this realisation is still compatible with the feature value specified in the lexicon. This generalisation across the two possibilities would not be possible if the feature specified lexically were [+*o*], which is incompatible with the GF OBL_T.

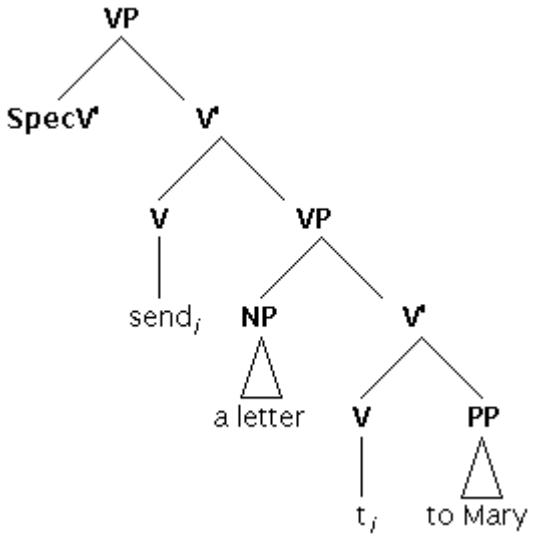
4.3.2 P&P analyses - VP shells and internal arguments

The previous sections have argued that the bare DP complement of an emotion or cognition verb has syntactic properties which resemble those of a second object, an OBJ_θ in LFG terms. This grammatical function is accepted as basic in LFG, but other syntactic theories do not accord it the same status. In particular, P&P syntax, since Larson (1988), analyses second objects as the

result of DP-movement in a particular VP configuration. The aim of this section is to show that such an analysis of the Indonesian data under consideration is implausible.

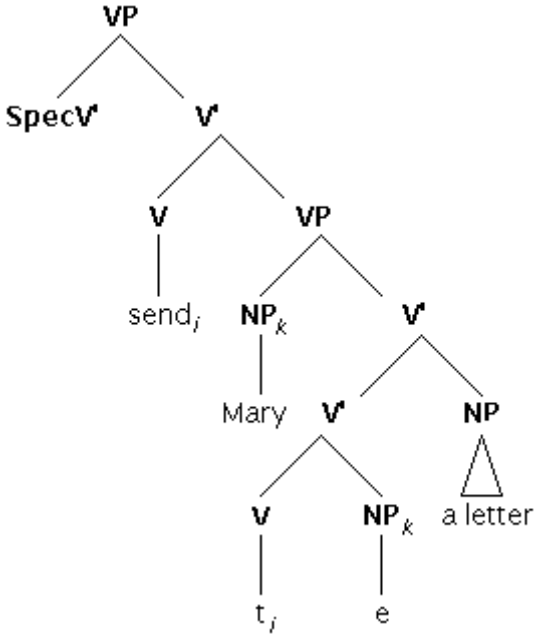
In Larson's analysis of the double object construction (1988, afterwards ODOC), ditransitive verbs project nested VP configurations. The basic structure for a clause containing such a verb is the following:

212.



The double object configuration is the result of an operation in the lower VP which ODOC describes as similar to passive in that the specifier or subject position becomes athematic. By Burzio's generalisation (Burzio 1986), the verb therefore is not able to assign case to its complement; that argument moves to the specifier position and the other argument (the theme) is realised as a V' adjunct:

213.



In all cases, the verb raises from the lower VP to the higher VP (before moving to a higher functional head if necessary).

The first problem in applying this type of analysis to the Indonesian data is to decide what the lexical characteristics of the verb should be, that is, what d-structure does it project. If a layered VP is only projected by a verb with two internal arguments, then we must conclude that Indonesian emotion verbs are of this type. Further, we must conclude that in their basic configuration, these verbs project the experiencer argument to a lower structural position than the stimulus. This must be the case as it is the stimulus which will end up as the second object and it must therefore originate as the specifier of the lower VP. Then a passive-like operation must obligatorily apply in the lower VP, and, following this, the experiencer must move to subject position. These instances of DP movement are not the problematic part of the analysis. An unaccusative verb does not assign a thematic role to a subject position, and as noted by Burzio (1986), does not assign case to its object. It is plausible to extend these generalisations to the question of two-place unaccusatives by assuming that such verbs cannot assign case from either of the positions they occupy within VP. Thus, in a structure such as 212, the conditions are met for Larson's PASSIVE operation: no thematic role is assigned to the specifier of the lower VP. This results in a structure like 213, and on the assumption suggested here, the verb in its higher position cannot case-mark the DP in the lower specifier position. Therefore this DP has no option but to move upward again, finally to the case position, specifier of IP. The second object can receive case from the verb if Belletti and Rizzi's (1988) revision of Burzio's generalisation is adopted. Their version of this principle states that a verb which does not assign a thematic role to its subject, cannot assign structural case to an object. Thus, the second object can be assigned inherent case by the verb without affecting the other elements of the analysis[28]. There is a possible analysis that gives the desired result: that the stimulus argument is a second object derived by movement in the lower VP of a layered structure.

However, such an analysis can be ruled out on both empirical and theoretical grounds. Firstly, the discussion of section 4.3.1.1 has shown that either the empirical evidence as to the unaccusative status of these verbs is unreliable, or that there are two classes of verbs. In the first case, as concluded above, there is no reason to accept that any of the Indonesian verbs are two-place unaccusatives. In the second case, the Larsonian analysis would be unable to account for the fact that there are Indonesian emotion and cognition verbs which are not unaccusative on the test proposed (*-kan* suffixation), but which nevertheless have second arguments which have the same properties as those which are unaccusative on the same test.

Theoretically, there is a consensus across all versions of the thematic hierarchy that experiencers are high on the thematic hierarchy, certainly higher than stimuli (or themes)[29]. Where surface syntactic relations do not reflect this ranking, for example with a verb such as English *frighten*, various

[28] ODOC claims that the lowest V' bar is reanalysed as a V for this, but the details are not relevant here.

[29] Dowty's 1991 analysis, mentioned previously, is perhaps the closest to being an exception to this generalisation.

sorts of special pleading have been considered necessary. As previously mentioned, Pesetsky (1995) claims that in such cases the stimulus is not a theme, it does not have the same semantic role as the second argument of *fear*. He argues for a more fine-grained analysis of semantic roles, specifically that the role *causer* is important for inverse psych verbs. The question of what is the correct level of detail at which to make generalisations about semantic roles is an important one, but one which is beyond the scope of this study[30]. Another approach is that of Grimshaw (1990), who accepts the thematic hierarchy, but dissociates thematic and aspectual information. In the case of verbs such as *frighten*, Grimshaw argues that information from the event structure conflicts with the thematic hierarchy, that is, the stimulus has a causal role and is more prominent in the event that the experiencer, and that this information is treated as more important in argument structure. Note however that Indonesian has only the one type of emotion and cognition verb - there is no syntactic evidence requiring us to propose more fine-grained distinctions here, either in thematic or aspectual structure.

The claim that some verbs have two internal arguments and no external argument, that they are two-place unaccusatives, has been made by Belletti and Rizzi (1988) for two of the three classes of psych-verbs in Italian. However, the essential feature of their analysis is that the d-structure for such verbs reflects the thematic prominence of the arguments although the surface relations are reversed (that is, the stimulus is the surface subject)[31]. For the remaining class of verbs (the *temere* class), the experiencer is subject and the theme is object. The null hypothesis in this case must be that that if surface thematic relations are as we expect, then so are relations at any other level. Belletti and Rizzi do not question this assumption and argue for an analysis of the *temere* class which respects it, that is, that the experiencer is an external argument. There is an otherwise plausible analysis in Belletti and Rizzi's scheme which captures the Indonesian facts - that the verbs are two-place unaccusatives which assign case inherently to the theme argument, explaining its second-object properties and also why it is the experiencer which must move - but this is ruled out by the considerations just discussed.

4.3.3 Summary

This section has argued that an LFG analysis is possible which captures the fact that the second argument of Indonesian emotion and cognition verbs has properties very similar to those of second objects. It has also been argued that two influential approaches in the P&P framework, those of Larson (1988) and Belletti and Rizzi (1988) are not capable of providing such satisfactory analyses. The following section turns to the question of whether it is a theoretically or empirically adequate conclusion to accept that the arguments

---

[30] Dowty 1991 and Foley and Van Valin 1984 present arguments for only recognising two roles as relevant to syntax. At the other extreme, McRae, Ferretti and Amyote (1997) present psycho-linguistic evidence suggesting that no such generalisations are used in comprehension.

[31] Belletti and Rizzi's analysis is thus a P&P reworking of the Relational Grammar inversion analysis of such constructions (Blake 1990: 46-50)

in question are second objects, or whether taking a different view of Indonesian clausal syntax as a whole can offer a better solution.

4.4 Theoretical and empirical problems

The previous section was mainly concerned to establish that the technical apparatus of LFG was adequate to present the analysis of Indonesian emotion and cognition verbs which was proposed in section 4.2, and to make the analysis explicit. I now turn to the question of whether this analysis sits comfortably within the broader LFG framework, more specifically, whether the use made of the grammatical function OBJ_θ is consistent with the theory. I argue in section 4.4.1 that the analysis is consistent with the theory but is also typologically surprising, and that this is contrary to the spirit of LFG. The crucial assumption of the analysis is that the system analysed in chapter 2 is the paradigm for transitive verbs in Indonesian. Section 4.4.2 asks whether emotion and cognition verbs provide empirical evidence for questioning that assumption and provides an initial suggestion of how an alternative analysis of the system might look.

4.4.1 The status of OBJ_θ in LFG

Classic LFG (Kaplan and Bresnan 1982) admitted OBJECT-2 as a primitive grammatical function. The GF was necessary for the analysis of, for example, English double object constructions such as:

214. Mary baked James a cake.

But in a theory where the inventory of GFs did not depend on any other theoretical construct, the status of this GF could be marginal; it could be a necessary evil with no theoretical consequences. However, with the introduction of LMT (Bresnan and Kanerva 1989), the GF OBJ_θ has a rather different theoretical status. The apparatus of LMT predicts that four grammatical functions, and only four, must exist. The four are placed in a markedness hierarchy, but the theoretical apparatus accords them all equal legitimacy. The restricted role played by OBJ_θ is then somewhat surprising, as is the fact that not all languages have this GF. This point can be made clear by considering the asymmetry between the two markedness relations established by LMT. SUBJ is the unmarked GF with both OBJ and OBL_T more marked. Typologically, this plausibly translates as the claim that SUBJs appear more freely in a language than the other two GFs, but not as a claim that there are languages which have SUBJs but which lack all other GFs. Any such claim would be incoherent to the extent that SUBJ is a relational notion: it is the grammatical function of the most prominent or most syntactically privileged DP in the clause and this clearly implies that it must be possible for there to be other DPs in the clause with some other grammatical function[32]. But the relation between OBJ and OBL_T on the one hand and OBJ_θ on the other is

[32] This argument applies in cases such as Mel'cuk's analysis of Lezgian (1983) as a language without transitive verbs. The non-subject DPs still have to have some grammatical function, in this case oblique. A more complex case is provided by the analysis of Tagalog as having only equational sentences - this is discussed in detail in chapter 5.

different: in this case the increase in markedness can mean that the most marked function is optional. This does not follow from the apparatus of LMT, where the theoretical difference in markedness is the same in the two cases.

It might be suggested that the evidence presented here justifies proposing a more fine-grained analysis of non-subject arguments in Indonesian than is allowed by LFG. The syntactic properties of DP stimuli and OBJ_θ s are not identical, and the considerable overlap might be due only to the fact that both are types of direct argument. There might be more than three types of non-subject argument in the language. However, the LMT offers a strong theory-internal argument against adopting such an analysis. Two binary-valued features allow exactly the four GFs assumed by the theory. and introducing a new GF would disrupt this aspect of the theory. To do so must mean that the two features used thus far do not exhaust the relevant information, and another feature would be required. This in turn would mean that rather than adding one GF, at least two would have to be added, on the assumption that the new feature would be dependent on one of those already assumed. If the new feature was independent of the existing ones, then four new GFs would be added to the system. To my knowledge, there is no empirical support for reducing the constraints on the theory in this way. This consideration suggests that LFG should either seek an explanation of the asymmetry discussed in the previous paragraph, or should seek analyses, such as that proposed here, which accord a greater importance to the OBJ_θ grammatical function.

Against this is the aim of LFG to be what might be called a 'typologically responsible' framework. That is, the framework aims to deal with the full typological range of languages without doing violence to the facts. And the typological facts are that second objects are not used in all languages and that where they do exist their occurrence is limited. There is empirical evidence from a number of languages, some referred to in previous sections, that subject and indirect object, or a nominative and dative case array, is a diathesis that occurs in many languages and that this is associated with predicates of low transitivity (see Blume 1998 for a partial survey). The data under consideration here might be thought to fall under this category, but the match is in fact poor. Indonesian is very similar to English in its treatment of three-place events (modulo the presence of morphological applicative). Recipients and beneficiaries, the prototypical semantic roles which receive dative case in case-marking languages, are coded either as obliques or as direct objects; they are never second objects. Some of the examples of this clause type in other languages are restricted to verbs denoting psychological processes, and these might be considered particularly relevant. But as discussed in previous sections, such verbs have the experiencer in dative case. There is then a double mismatch with the Indonesian data: dative is not the relevant semantic relation and experiencer is not the semantic role which is treated in non-standard fashion. Therefore on typological grounds the proposed analysis of Indonesian emotion and cognition verbs is extremely unusual. The fact that the apparatus of LFG can accommodate the analysis is not a sufficient reason for accepting it without examining alternatives.

4.4.2 Empirical issues

A basic assumption that has guided the discussion in this chapter since section 4.2 is that a two-place verb which does not enter into the system analysed in chapter 2 is not a transitive verb in Indonesian. The analysis of the second argument of emotion and cognition verbs as a second object is a direct result of this assumption. However, in addition to the concerns discussed in the preceding section there are empirical reasons to be uncomfortable with the analysis and to question the assumption.

Firstly, the syntactic properties of the arguments in question are not identical to those of second objects. Relative clause constructions with the stimulus argument as head are accepted by all speakers as fully grammatical, as noted by Stevens in 1970 (see example 179 above). On the other hand, relative clauses with second objects as head are impossible for many speakers, are marginal for some speakers, but subject to lexical idiosyncrasy even for them (see example 191). Secondly, the type of clause in which emotion and cognition verbs appear is strikingly similar to the type of clause referred to as a bare verb clause in chapter 2. In this clause type, an unaffixed verb appears and the actor is always subject. In particular, extraction of the second argument is fully grammatical from this type of clause:

215. *Inilah* *buku* *yang* *Badu* *sudah* *baca*
this.EMPH book REL Badu PERF read
'This is the book which Badu has read already.'
(Voskuil 1996: 189, ex4)

On the assumptions adopted to this point, the second argument is an OBJ in one case and an OBJ_θ in the other case, despite the great similarity in their syntactic properties.

I argued in section 2.3.2 that the impossibility of extracting non-subject arguments was characteristic of the type of clausal organization found in Philippine languages, and that some Indonesian clause types showed features of that type of organization. I will now suggest that this line of thinking can be pushed rather further, and that Indonesian should be analysed as having two types of clausal organization available side-by-side. One type has a verb phrase as a major constituent of the clause, whilst the other has a predicate phrase as a major constituent, which may in turn be headed by a verbal root. Instead of taking the ability to appear in the full range of clause types discussed in chapter 2 as criterial for transitive verbs, it is better to take only those clauses with an unprefixed (where prefix includes morphologically attached pronoun) verb and two arguments as transitive verbal constructions. The other clause types in which some verbs appear are then analysed as subject-predicate clauses, in the same way that the clauses with nominal, adjectival and prepositional heads are traditionally analysed. The function of the verbal prefixes can then be analysed as derivational: they allow verbs to head predicate phrases. Emotion and cognition verbs are regular transitive verbs on this account; their oddity is that they do not allow derivation of predicate heads without prior affixation with other morphology.

The division between these two modes of clausal organization correlates with a difference in how arguments are licensed. In addition to changing the status of verb roots, the prefixes *meN-* and *di-* must also specify the linking of an argument to the SUBJ function, as argued in chapter 2. This means that LMT does not play a role in the licensing of SUBJs in these clause types. On the other hand, the linking properties of both bare verb clauses and emotion and cognition verb clauses are exactly as would be predicted by LMT: the more thematically prominent argument is linked to the SUBJ function consistently, and, if the new proposal is correct, the non-subject argument is an OBJ as expected. The difference in the properties of non-subject arguments between the two clausal structures, verb phrase and predicate phrase, may also be a result of some difference in how the arguments are licensed. In chapter 5, additional evidence for the dual structure hypothesis is examined, and some problems in stating the hypothesis in the LFG framework are discussed.

5 Two types of clause structure

This chapter examines in more detail the hypothesis introduced at the end of the preceding chapter: that Indonesian has two types of clause structure which exist side by side, but which have different properties. I will take it that the crucial contrasts to be accounted for are the possibilities for extraction of the non-subject argument and the distribution of attached pronouns, both of which depend on the type of verb in the clause. That the two phenomena are closely linked is made clear by the examination of the secondary strategy for relativization in Indonesian, which uses *-nya* as a resumptive pronoun. This is discussed in section 5.1 and arguments are presented to show that the relativization possibilities of various syntactic positions cannot be derived from the binding theory properties of the pronoun *-nya*. The distribution of *-nya* is examined in section 5.2, and I argue that the generalisation which describes that distribution makes crucial reference to the notion of possible predicative elements, with prefixed verbs falling into this class. Section 5.3 presents evidence that this is a plausible approach to the prefixed verbs based on both language-internal evidence and comparative and historical considerations. Section 5.4 argues that the intuitions which are the basis for treating prefixed verbs as different in category to unprefixed verbs cannot be captured by an LFG analysis in any straightforward way, and that the resources of the theory are not adequate to deal with the data examined here. It is not my intention to reformulate the theory, so I offer only very brief suggestions as to how the problem might be overcome. Section 5.5 summarises the conclusions of this chapter, and of the work as a whole. An appendix considers the question of whether Indonesian is a counter-example to the Accessibility Hierarchy of Keenan & Comrie (1977).

5.1 Strategies of relativization

In previous chapters, many examples have been given which included relative clauses. All of these have exemplified the major relativization strategy of Indonesian, gapping of the coreferential nominal in the embedded clause:

1. *gadis*$_i$ *yang* *Ø*$_i$ *dilihat* *Umar* *itu*
   girl REL US.see Umar that
   'the girl that was seen by Umar'

I have argued that the possibility of a non-subject argument being the gap in such relative clauses is criterial for identifying what I will call the verb phrase type of clause in Indonesian, and this test shows that the bare verb actor subject clause type and the emotion verb clause type with a DP non-subject belong to this type. There is another strategy for forming relative clauses in Indonesian, however, and its distribution reveals additional information about clause structure.

5.1.1 Resumptive pronoun strategy

Relative clauses can also be formed with a resumptive pronoun taking the place of the head nominal in the embedded clause. The resumptive

pronoun is always the attached pronoun *-nya*, and this strategy can therefore be used for any DP which occupies a place in the relative clause where *-nya* is possible. This means that possessors, non-subject arguments of *meN*-prefixed verbs, and the objects of some prepositions can head relative clauses:

2. *orang*$_i$ *yang* *mobilnya*$_i$ *dicuri*
person REL car.3 US.steal
'the person whose car was stolen' (IRG: 288)

3. *[sebuah lagu]*$_i$ *yang barangkali saudara akan*
one.CLASS song REL perhaps brother FUT
menyukainya$_i$
AS.like.APPL.3
'a song which perhaps you will like' (IRG: 289)

4. *rumah*$_i$ *di yang belakangnya*$_i$ *ada pohon mangga*
house LOC REL behind.3 exist tree mango
'the house behind which there is a mango tree' (IRG: 290)

I have argued in previous chapters that the gapped strategy, although the most common means for forming relative clauses, can only be used in two cases: where the head noun is also subject of the embedded clause, and where the head noun is the object of an unprefixed verb in the embedded clause. These are also positions in which an attached pronoun is impossible. I have demonstrated this for the case of the unprefixed verbs elsewhere and now I will present data showing that subjects are never attached pronouns in Indonesian.

5.1.2 Subject and reduced pronouns

There is no typological objection to the possibility that subjects should be reduced pronouns. Many scholars assume that the subject is the most topical entity in a clause (in the sense that it is the typically the most continuous in discourse, not in the sense that it is an example of the LFG discourse function TOPIC) and topical elements can often be expressed in a reduced fashion, with zero anaphora, cliticization or pronominal expression often preferred (Givon 1983). Nor is it true that other languages from the Austronesian family disallow subject reduced pronouns. For example, in Tagalog, pronouns generally cliticize to a position 'after the first daughter of the smallest maximal projection which contains them' (Kroeger 1993: 121), and this is true for both nominative (subject) clitics and others[1]:

5. *Dito siya magtatayo ng=bahay*
here 3SG.NOM FUT.AV.build GEN=house
'Here he will build a house.' (Kroeger 1993: 122)

[1] Glossing as in the source: AV - active voice, OV - objective voice. The = sign indicates that the case particles form a phonological unit with the following word, the form *kay* in example 6 is a suppletive form used with proper names.

6. *[Para kay=Pedro] ko binili ang=lauran*
for DAT=Pedro 1SG.GEN PERF.buy.OV NOM=toy
'For Pedro I bought the toy.' (Kroeger 1993: 122)

The correct analysis of the nominative argument in Tagalog is the subject of considerable debate (Cena 1995, Kroeger 1993, Schachter 1976, 1995, Sells 1998, among others). There is a consensus however that the argument glossed here as NOM(inative) is the most syntactically privileged, and it is clear that with respect to cliticization, it is treated in the same fashion as the non-subject argument of a transitive clause.

Other Austronesian languages for which the grammatical function subject may not be applicable still show this sort of symmetry with respect to the two arguments of transitive verbs. Acehnese (Northern Sumatra), which Durie (1985) argues has no subject function at all, uses clitics to code both actor and undergoer of a transitive clause. The actor clitic is obligatory and the undergoer clitic is optional[2], in each case the clitic is capable of being the sole exponent of the argument in the clause:

7. *ji= kap =keuh*
3 bite 2
'It'll bite you!' (Durie 1985: 47, ex4-3)

Similar facts can be observed in the Menó-Mené variety of Sasak (South-Central Lombok). It is not clear whether this variety can usefully be analysed in terms of subject and object functions (Austin 1997b, Kroon 1998), but both arguments of a transitive verb can be represented by clitic pronouns. As in Acehnese, positional coding distinguishes the two possibilities: actor clitics attach to auxiliaries and undergoer clitics attach to verbs:

8. *Yaq-n bèng-k hadiah-nò*
FUT-3 give-1SG present-that
'He will give me that present.'

Again, a clitic pronoun can be the only exponent of the semantic role in the clause.

Indonesian completely lacks the option of having the subject coded as a reduced pronoun. The reduced pronouns seen in some Pro-V clauses have already been treated in detail in chapter 2, and they are not subjects (see also section 1.2.2). The other set of reduced pronouns which exist are *-ku, -mu* and *-nya*. These attach to the right edge of a preceding word. As subjects are very often the first element in the clause, linear order might be assumed to prevent this type of expression for subjects. But even when some element precedes the subject, it is not possible for that element to host an enclitic[3]:

[2] This is true of the dialect from North Aceh which Durie describes. Southern dialects omit the undergoer clitics altogether (Bukhari Daud, p.c.)

[3] This is possible in Sasak, for example.

9. **Ali memikirkan bahwanya sakit*
Ali AS.think.APPL that.3 sick
(FOR: 'Ali thinks that he is sick.')

The possibility of a reduced pronoun attaching to the right of an auxiliary also does not exist in Indonesian:

10. **Akannya pulang*
FUT.3 go.home
(FOR: 'He will go home.')

The only type of structure which seems to contradict the claim made here is the 'raising-to-object' type:

11. *Siti menganggap dia bodoh*
Siti AS.consider 3SG foolish
'Siti considers him foolish.'

In this type of structure, the subject of the embedded clause can be encliticized to the matrix verb:

12. *Siti menganggapnya bodoh*
Siti AS.consider.3 foolish
'Siti considers him foolish.'

But there is good evidence in such cases that the subject of the embedded clause is the non-subject argument of the matrix verb. It can be made subject in a passive *di*-V clause:

13. *Dia dianggap bodoh oleh Siti*
3SG US.consider foolish by Siti
'He is considered foolish by Siti.'

The generalisation can therefore be maintained: subjects in Indonesian are never reduced pronouns. The positions which can be gapped in relative clause formation are in complementary distribution with the positions where a resumptive pronoun can occur (there are also positions where neither is possible).

5.1.3 Complementary distribution of relativization strategies

The data presented above show that there is a complementary distribution between the positions in which reduced pronouns can occur and the positions from which gapped relativization is possible. This fact might lead one to search for an explanation for the restrictions on gapped relativization in terms of the binding properties of either the resumptive pronoun or of the crucial positions. This section will present evidence that no such explanation is possible, as well as briefly discussing one analysis along those lines.

5.1.3.1 TOPIC-COMMENT STRUCTURES

The resumptive pronoun strategy for relativization is generally agreed to be related to the possibility of Topic-Comment structures of the following type (Alsagoff 1992: chapter 7, IRG: 288-291):

14. *[Orang itu]$_i$, saya melihat -nya$_i$/dia$_i$*
person that 1SG AS.see 3SG
'That person, I saw them.'

Fundamental assumptions of all versions of binding theory tell us that the topicalized DP in such structures must be outside the binding domain that contains the resumptive pronoun. In terms of Chomsky's formulation (1981), this is a Principle B effect: a pronoun cannot be coreferential with a c-commanding nominal in the relevant domain. The topicalized element and the pronoun co-refer and, equating leftward position with height in the tree, the pronoun is c-commanded by the topicalized element. Therefore, they cannot be in the same binding domain. The same argument can be made in the LFG version of binding theory, where the functional nucleus is taken as the relevant domain (Bresnan 2001a: chapter 10). Where a possible antecedent of either pronoun occurs within the relevant domain, coreference with the pronoun is not possible:

15. *[Orang itu]$_i$, Ali$_j$ melihat -nya$_{i/*j}$/dia$_{i/*j}$*
person that 1SG AS.see 3SG
'That person, Ali saw them.'

When the free pronouns in examples 14 and 15 are replaced by *-nya*, precisely similar arguments apply in that case also.

But in topic-comment structures, the distribution of *dia* is wider than that of *-nya*. The free pronoun can appear in subject position which is impossible for *-nya*:

16. *[Anak itu]$_i$, dia$_i$ mencubit doktor itu*
16' **Anak itu, nya mencubit doktor itu*
child that 3SG AS.pinch doctor that
'The child, she pinched the doctor.'
(Adapted from Alsagoff 1992: 228, ex22a)

According to Alsagoff (1992: 228), the resumptive pronoun can also appear in second object position. Such examples are considered ungrammatical by my consultants; whether this difference is due to variation between Indonesian and Malaysian, or whether it should be attributed to some other reason, is not clear. But at least some speakers of Malaysian allow examples such as the following, and again *dia* has a wider distribution than *-nya*:

17. *Mariam$_i$, Ali menawarkan doktor itu dia$_i$*
17' **Mariam Ali menawarkan doktor itu nya*
Mariam Ali AS.offer doctor that 3SG
'Mariam, Ali offered her to the doctor.' (Alsagoff 1992: 229, ex25a)

The binding properties of *dia* have not changed between these various examples: coreference with the topic is perfectly acceptable in all of these cases. But examples 14 and 15 showed that the binding conditions on *dia* and *-nya* must be the same. Therefore, the impossibility of *-nya* in 16' and 17' cannot be due to any binding theory violation, and must be due to some restriction on possible hosts. This might be phonological, or syntactic.

Additional evidence comes from the fact that only some prepositions allow an enclitic pronoun as their object. There is no obvious semantic generalisation about which prepositions have this property and which do not, nor is there an obvious phonological generalisation. The two monosyllabic prepositions *di* and *ke* do not allow *-nya* as their object, but monosyllabic nouns are fully grammatical with *-nya* as possessor, therefore it seems unlikely that this could be a relevant factor. I will therefore treat the property as a lexical idiosyncrasy[4]. Regardless of whether the preposition can host a clitic, *dia* can always follow it as the resumptive pronoun in a topic-comment structure:

18.	*Lisa$_i$,*	*semua*	*orang*	*ingin*	*bertemu*	*dengan*	*dia$_i$*
	Lisa	all	person	wish	meet	with	3SG
	'Lisa, everyone wants to meet her.'						

19.	*Lisa$_i$,*	*semua*	*orang*	*ingin*	*berbicara*	*tentang*	*dia$_i$*
	Lisa	all	person	wish	talk	about	3SG
	'Lisa, everyone wants to talk about her.'						

The enclitic *-nya* is only possible when the preposition has the lexical property of allowing it:

18'	*Lisa$_i$,*	*semua*	*orang*	*ingin*	*bertemu*	*dengannya$_i$*
	Lisa	all	person	wish	meet	with
	'Lisa, everyone wants to meet her.'					

19'	*Lisa$_i$,*	*semua*	*orang*	*ingin*	*berbicara*	*tentangnya$_i$*
	Lisa	all	person	wish	talk	about

Again there is a restriction on possible hosts for *-nya*, but in this case it is due to some lexical property of prepositions. The binding properties of the two forms of the pronoun are identical, therefore they cannot be the explanation for the impossibility of *-nya* in one clause but not in the other.

In relative clauses, the resumptive pronoun can only be *-nya*. But the properties of such structures as far as binding theory is concerned are effectively identical to the topic-comment structure. The arguments given in the preceding paragraph apply in exactly the same fashion: only *-nya* can appear in relative clauses, but this cannot be due to the binding properties of such structures.

[4] I have not investigated the question in detail, and I certainly do not rule out the possibility that further research might reveal some interesting pattern. This issue is discussed further in section 5.2.

5.1.3.2 VOSKUIL 1996

Voskuil (1996: chapter 8, afterwards V) offers an account of the impossibility of A-bar movement from the position of the non-subject argument of prefixed verbs in Indonesian in the P&P framework. The first crucial assumption that V makes is that verb prefixes license *pro*, a zero pronoun, as the object of the verb they are attached to. Both prefixed verbs and bare verbs allow zero anaphora of their non-subject arguments, and V claims that a different empty category is involved in the two cases. Following a prefixed verb, *pro* will be left, but this is not possible for a bare verb. In that case, V assumes that a zero operator is moved to a topic position and that this binds a trace in the position of the non-subject argument:

20. *Lupus langsung menghampiri* pro
Lupus immediately AS.approach
'Lupus immediately approached (him).' (V: 193, ex29)

21. Op$_i$ *mak sudah beli* t$_i$
1SG PERF buy
'I already bought (some).' (V: 193, ex30)

Some of the evidence supporting this assumption has been discussed in section 2.2.2.2 and I will not repeat it here.

V's second crucial assumption is the following principle:

22. An empty category is interpreted as pro if morphological conditions for pro are met, irrespective of derivational history.

This principle is motivated with arguments from Italian and Irish. Irish has anti-agreement effects: an overt subject and agreement on the verb cannot co-occur. When a subject is extracted, agreement cannot occur on the verb of the clause from which the extraction has occurred[5]. V reasons that in such cases, the presence of agreement would license *pro* and therefore force interpretation of the empty category as *pro*, which cannot be the foot of a chain. The principle 22 therefore predicts that a gap in the position of the non-subject argument of a *meN*-prefixed verb cannot be interpreted as a trace, it cannot be the result of movement. This leaves open the possibility that it can be interpreted as *pro*, but coreferential with a DP outside the binding domain of *pro*. Nothing in V's account, or in the general principles which he assumes, rules out the possibility of *pro* being a null resumptive pronoun in a structure such as that seen in example 3 (repeated here):

3. *[sebuah lagu]*$_i$ *yang barangkali saudara akan*
one.CLASS song REL perhaps brother FUT
menyukainya$_i$
AS.like.APPL.3
'a song which perhaps you will like' (IRG: 289)

[5] The Irish data are discussed in detail and given an LFG analysis in Andrews (1990).

3'. *[sebuah lagu]$_i$ *yang* *barangkali* *saudara* *akan*
one.CLASS song REL perhaps brother FUT
menyukai pro$_i$
AS.like.APPL.3
'a song which perhaps you will like' (IRG: 289)

pro has the feature matrix [-anaphor, +pronominal], that is, it is a pronoun and is predicted to have the same distribution as an overt pronoun (Chomsky & Lasnik 1995: 36). The structure shown in 3' should therefore be possible, but it is not.

I will give only a brief sketch of V's analysis of this problem, because the analysis is based on a significant empirical error. V claims that examples like 3 are ungrammatical, contradicting the sources which I have consulted (Alsagoff 1992: chapter 7, IRG: 281) as well as native speaker judgments. The question V attempts to answer is why 3 is not possible, the answer to this will rule out 3' and complete the account of the ungrammaticality of the extraction of non-subject arguments. V's account treats resumptive pronouns as variables and claims that the binding relation between a variable and some other element must be due to movement, except when other principles prohibit movement from the position of the variable. In the case of structures like 3, no principle bars movement from the non-subject position, the equivalent in English is well-formed. Therefore, the variable must be derived from movement, and 3 is not well-formed[6]. The empirical error in V's account unfortunately distracts his attention from the real question: why 3 is grammatical, but any similar structure without the resumptive pronoun is not.

5.1.3.3 SUMMARY

The preceding sections have shown that the distribution of the attached pronoun *-nya* and the distribution of the two strategies for forming relative clauses in Indonesian are intimately linked. I have also argued that no explanation for this distribution can be found in the binding theoretic properties of the pronouns involved. The question that therefore needs to be answered is: what factors do constrain the distribution of *-nya*, and hence of relativization strategies? I take up this question in the following section.

The attentive reader may have noted in the discussion of the secondary relativization strategy that the position of a second object is not a position in which *-nya* can occur. It is also not a position from which gapped extraction is possible, with the exception noted in section 4.2.3 (example 191). But possessors and objects of (some) prepositions can be relativized with a resumptive pronoun. These facts suggest that Indonesian may provide a counter-example to one of the universals proposed by Keenan & Comrie (1977), that a relativization strategy should apply to a continuous segment of the hierarchy of syntactic functions. This issue is taken up in the Appendix to this chapter.

---

[6] V's account is less stipulative than might appear from this summary. But in view of the empirical problem, giving the details seems pointless.

5.2 The distribution of *-nya*

The distribution of *-nya* (and the other two enclitic pronouns) has been discussed briefly in section 5.1.3.1. Here I analyse this question more carefully, and establish a generalisation as to the positions in which enclitic pronouns can occur. Three positions are possible: after prefixed verbs, after some prepositions, and as possessor in a DP. I will discuss them in this order, which reflects the increasing complexity of the issues involved. The case of prefixed verbs has been discussed in detail in chapter 2. All that is necessary to the present discussion is to recall that an enclitic can encode the non-subject argument of a prefixed verb, and that its syntactic status is different in the case of the two prefixes. The pronoun is a true clitic in the case of *meN-* prefixed verbs, and I assume that it is generated as sister of the head in c-structure. With a *di-*prefixed verb, the pronoun is an affix attached in the lexicon. This fact already suggests that a generalisation which covers all cases will have to be stated in terms of f-structures.

Only a subset of prepositions can have enclitics as their complements. This group includes the following prepositions:

23.		
	pada	at, to
	kepada[7]	to
	daripada	than, from
	bagi	for
	buat	for
	untuk	for
	bersama	together with
	dengan	with
	terhadap	towards
	oleh	by
	dari	than, from

A text example of a preposition hosting an enclitic is the following:

24.	*Akan*	*kubuatkan*	*susu*	*untukmu*
	FUT	1SG.make	milk	for.2

'I will prepare some milk for you.' (SDM: 152)

The group of prepositions listed in 23 does not include the prototypical locative preposition, *di*[8]. However, enclitics do occur in some locative expressions, for example:

25.	*di*	*atasnya*
	LOC	top.3

'on top of him/her/it/them'

Expressions of this type are best analysed as preposition plus locative noun sequences, rather than as compound prepositions; bracketed as in 26 not 27:

[7] *kepada* is typically used when the complement is a person.

[8] This preposition also does not occur with emotion and cognition verbs (see section 4.1.2.2).

26. [PP [P di] [DP atasnya]]
27. [PP [P di atas] [DP -nya]]

The enclitic in such cases is then an example of the possessive type, and this analysis can be maintained even when *di* is omitted in informal Indonesian, e.g. *atasnya* 'on top of it'. It is not the case, however, that there is a neat divide between static locative prepositions and all others with respect to enclitics. For example, *tentang* 'about, concerning' does not host clitics. The prepositions used to code recipients and beneficiaries are all present in 27. As discussed in section 2.2.1.3, Indonesian pronouns have only human reference for conservative speakers, and historically, this generalisation has covered *-nya* also. It is possible, then, that the prepositions which allow enclitics are those which are more likely to appear with human complements, and those which do not allow enclitics are those which have prototypically non-human complements, such as *di*. Aristar (1997) argues analogously that, in case-marking languages, the non-grammatical cases include in their definition some specification of their typical argument. I will assume without argument that prepositions in Indonesian are lexically specified to take a DP complement, and that this is the sister of the P head in c-structure.

Enclitics are hosted very commonly by nouns, in which case they code possession of the head noun:

28. *Dan kamu akan menyesalinya seumur hidupmu!*
and 2SG FUT AS.regret.3 se.age life.2
'And you will regret it all your life!' (SDM: 125)

The structure of DP in Indonesian is significantly different from that in English as far as possessives are concerned. Common analyses for English have the possessive as the head of D (see e.g. Abney 1987), or more conservatively, as the specifier of DP. The possessive in Indonesian is not a specifier or the head of D. Unlike English, but like for example Italian, possessives and determiners can co-occur, assuming the argument given in section 1.2.3 that quantifiers are exponents of D in Indonesian:

29. *Kelima anaknya pandai*
ORD.five child.3 clever
'All five of her children are clever.' (IRG: 133)

30. *Dua pertiga orang Indonesia tinggal di pulau Jawa*
two FRAC.three person Indonesia live LOC island Java
'Two thirds of Indonesians live in Java' (IRG: 133)

Possessives are not specifiers in NP either; they follow the head but are not always final in the constituent. They can be followed by demonstratives:

31. *satelit kita itu*
satellite 1PL.INCL that
'that satellite of ours' (McKay)

32. *negeri saya ini*
country 1SG this
'this country of mine' (McKay)

by relative clauses:

33. *kepala Ardi yang tadi ditahan oleh sandaran*
head Ardi REL before *di*.support by support
kepala kursi
head chair
'Ardi's head that was previously supported by the head rest of the chair' (SDM: 202)

and by some other adjuncts.

Possessors are not necessarily the modifier closest to the head noun. Adjectives and modifying nouns can intervene:

34. *guru bahasa saya*
teacher language 1SG
'my language teacher'

35. *sikap tegas bapak*
attitude steady father
'father's firm attitude'

and the possessor can be realised as a bound pronoun in this position:

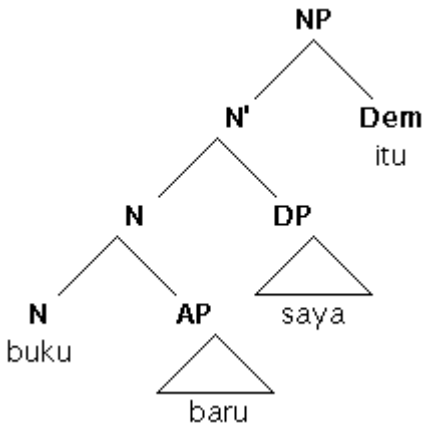
36. *guru bahasaku*
teacher language.1SG
'my language teacher'

When a modifying noun intervenes between the head noun and the possessor, there is a potential ambiguity and this is not dependent on whether or not a bound pronoun is used. Thus both example 34 and example 36 have a (marginal) reading 'teacher of my language'.

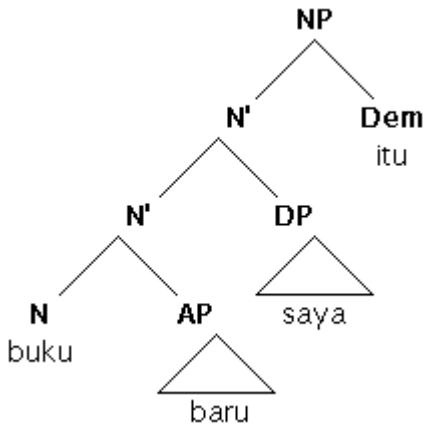
The facts described above make it clear that in Indonesian DPs, possessors are part of the NP complement of the functional head. The fact that constituents such as relative clauses, unambiguously part of the projection of N, also follow possessors shows that possessors are not specifiers within NP, they must be either complements or adjuncts under an N' projection. Given that they are not obligatorily closest to the head noun, and can be preceded by constituents which are certainly not complements (adjectives), two structures seem possible: either modifying nouns and adjectives always form compound nouns in Indonesian, or N' is a freely recursive category in Indonesian. These two structures are illustrated in the following trees[9]:

[9] I assume that demonstratives are specifiers of NP, but not determiners. Nothing depends on this assumption.

37.a



b



38. *buku baru saya itu*
book new 1SG that
'that new book of mine'

The first alternative is an extremely stipulative solution; allowing it would considerably weaken X-bar theory as applied to Indonesian. Therefore, the second structure is preferable.

This result means that the three cases in which enclitics are allowed cannot be brought under a single generalisation at c-structure, even if we ignore *di*-prefixed verbs. For *meN*-prefixed verbs and for prepositions, the clitic is the immediate complement of the head, but this is not the case for nouns. The structure argued for above places possessors in a position typical of adjuncts, under an X' but not under the lowest X'. If a unifying generalisation exists, it must therefore be sought in f-structure.

At first sight, this approach does not seem much more promising. For prefixed verbs (both types now), the attached pronoun encodes the non-subject argument, the second argument. For prepositions, the enclitic encodes the complement, possibly the only lexically specified argument. And for nouns, the enclitic encodes an optional argument. However, this description is overly reliant on familiar concepts from English grammar. Taking closer account of the facts of Indonesian grammar gives a more revealing picture. Section 4.1.3 presented data showing that all major categories in Indonesian can be the main predicate of a clause, and this fact is the clue to the generalisation which describes the distribution of enclitics: an enclitic can only be the non-subject argument of a (potential) predicate. This statement covers the case of *di*-prefixed verbs also, as they have the actor as non-subject argument in f-structure. The line of argument being developed here requires a corollary of this generalisation: unprefixed verbs cannot be predicates in the same sense, as they have non-subject arguments which can never be enclitics. I now consider the question of how this contrast might be understood as the outcome of other factors in the grammar of Indonesian.

5.3 Potentially predicative elements and arguments

The previous section has argued that the correct characterisation of the positioning of attached pronouns in Indonesian makes reference to the non-

subject argument of potential predicates. I will refer to such elements as being potentially predicative. If these potential predicates have a non-subject argument, the implication is that they also have a subject argument. But it is rather implausible that, for example, all Indonesian nouns have a subject argument as a lexical property. This section will argue that this objection can be met when we consider more closely the properties of potentially predicative elements, and that there is a significant similarity between prefixed verbs and other potential predicates in the relation between predicate and subject.

5.3.1 Semantic roles

Verbs assign idiosyncratic semantic roles to their arguments. Thus the subject of the English verb *fall* is understood to move (involuntarily) from a higher point to a lower point during the falling event, while the subject of *rise* moves in the opposite direction and may do so voluntarily. Similarly the non-subject arguments of two verbs such as *cut* and *weigh* are understood very differently. Linguists use thematic roles as convenient generalisations over these individual roles, but the semantics that verbs assign to arguments are more complex. This is not true in the same fashion for non-verbal predicates. Such predicates may assign a variety of roles to an argument internal to the predicate, sometimes but not always indicated by preposition choice in English: consider the difference between *a student of linguistics* and *an opponent of linguistics*. But the roles assigned to the subject of predication are much more limited.

Lyons (1977: 469) lists four non-verbal sentence schemata covering equative, ascriptive, locative and possessive predications, and other classifications in the literature are similar (e.g. Hengeveld 1992). This classification, or a very similar one, can be justified on syntactic grounds in various languages and the semantic distinctions are intuitively clear. However, it is possible to make a semantic generalisation over these four types: all of them can be seen as ascribing a property to the subject of predication. Besides simple properties, there is the property of being identical to something else, the property of being in a certain location and the property of being possessed by someone. In this sense, the subject of all these predication types has only a single semantic role, that of bearer of a property, and the predicate itself has only one semantic function, to specify the property[10]. Notice also that such structures do not allow the possibility of switching the direction of predication, except in the special case of identity. While it is common for languages to have syntactic and even occasionally lexical resources which allow an event to be represented from different viewpoints, this is not generally possible for non-verbal predications except by some type of paraphrase. This is, of course, linked to the fact that languages do not have extensive resources for this type of predication. Either, as in Indonesian, there is no copula in most such predications, or there are limited

[10] I am aware that I am sidestepping a number of ancient and contentious philosophical debates at this point. The reader is referred to Seuren 1998: 121-133 and Moro 1997: 248-261 for some discussion.

lexical resources. Hengeveld (1992) reports languages with several copulas and what he terms semi-copulas, but these are used for different types of predication, not for altering the directionality of the relationship. Thus, *I gave Helen some money*, and *Helen received some money from me* are representations of the same event using different lexical predicates, but to change the viewpoint with a predication such as *I was in London*, the only option is rather stilted paraphrase *London was my location*. The essential point here is that the semantic role of the subject of such predications is fixed, or at least limited to a very small set of possibilities.

5.3.2 Grammatical functions

The fact that all major categories in Indonesian can head a predication raises questions about the nature of the grammatical function SUBJ in the language. Do all categories include in their lexical entries the information that SUBJ is a GF that they can assign? There is some relevant empirical evidence, but it is not conclusive. There is a strong tendency for adjectives which modify nouns to be introduced by *yang*. That is, many adjectival modifiers are formally relative clauses, with the modified noun as the gapped SUBJ:

39. *Kemudian dipukulnya bola itu sekuat-kuatnya.*
after US.hit.3 ball that *se*. force.DUP.3
Servis yang sempurna
service REL perfect
'Then she struck the ball forcefully. A perfect service.' (SDM: 5)

In some cases, for example when more than one adjective is used, the relative clause structure is obligatory:

40. *Dari Amsterdam Ardi membawa Marisa ke Paris.*
from Amsterdam Ardi AS.take Marisa to Paris
Kota cantik yang hangat (...) yang ramah dan eksotis
city beautiful REL warm REL friendly and exotic
'From Amsterdam, Ardi took Marisa to Paris. The beautiful, warm city ... friendly and exotic.' (SDM: 50)

Thus adjectives have at least a tendency to appear in predicative structures. This is certainly related to the blurring of the distinction between adjectives and intransitive verbs in Indonesian, discussed in section 4.1.3.2.

Prepositional phrases often modify nouns also, and relative clause structures like the adjectival ones are possible:

41. *anak yang seperti Tini*
child REL like Tini
'children who are like Tini' (IRG: 287)

but this is less common than with adjectives. Most such modifiers of nouns are simple PPs:

42. *masih adakah tempat untuk dendam di hatimu?*
still exist.EMPH place for revenge LOC liver[11].2
'Is there still a place for revenge in your heart?' (SDM: 55)

43. *Marisa menatap hasil pemeriksaan laboratorium*
Marisa AS.observe result examination laboratory
untuk urine-nya itu dengan kesal
for urine.3 that with cross
'Marisa stared at the result of her urine test crossly.' (SDM: 128)

Treating this type of structure as syntactically consisting of a subject and a predicate would be very problematic. There are clear reasons against such an analysis both from the point of view of constituent structure (the PP is a modifier, not in a specifier-head relation with the noun it modifies) and from the point of view of functional structure (the modified noun would have either have two functions assigned it, or a controlled position would have to be postulated). The last example, example 43, also raises another problem in the analysis of PPs as having subjects: they can be modifiers of events as well as of entities as the PP *dengan kesal* is in this example.

Nouns only have a subject in predicative structures. The kind of arguments that have been made in respect of English to treat the possessor position as a subject position (M.Anderson 1983, Grimshaw 1990, Stowell 1983 among others) are not relevant to Indonesian. As discussed in chapter 1 and in section 5.2 above, possessors in Indonesian are arguments internal to the projection of the head noun, and not specifiers of any projection. There is no reason to assume that nouns have SUBJ as an assignable function in their lexical entries. On the basis of the evidence discussed above, there is also little reason to assume that prepositions have subjects either, and consistency would then suggest that adjectives ought to be treated in the same way. I take it that this conclusion is related to the position argued for in the previous section: the lack of a SUBJ function is related to the lack of a contentful semantic role for that SUBJ to be linked to.

I argued in section 5.2 above that the possible positions of attached pronouns should be understood in terms of elements that are potentially predicative, and that this should include prefixed verbs. This would imply that clauses containing prefixed verbs also have a subject-predicate structure and that such verbs do not have a SUBJ in their lexical forms. This is an intuitively implausible idea: how can it be that a verb does not always have a SUBJ in its argument list? Nevertheless, I will suggest that there are reasons for adopting such an analysis.

The most important reason for adopting the suggestion is the very striking parallel between prefixed verbs and the other predicative categories with respect to the assignment of a semantic role to the subject of the predication. I argued above that the semantic role assigned to the subject of non-verbal predications is predictable on general grounds, and there is

[11] Indonesian treats the liver as the seat of the emotions.

therefore no work to be done by any sort of mapping theory. The assignment of a semantic role to the subject of prefixed verbs is accomplished in a similar fashion if the account of the verb prefixes presented in chapter 2 is accepted. That proposal has the consequence that mapping theory is irrelevant for such verbs: the mapping to subject is directly specified by the prefix. The account of chapter 2 will need some revision in light of the proposal presented here, but this crucial insight will be maintained.

Under the current proposal, the prefixes *meN-* and *di-* are derivational. Their effect is to alter the attributes of the verb root so that its maximal projection will be a predicate rather than a verb phrase. In chapter 1, I claimed that all morphological processes in Indonesian were derivational, therefore this analysis has the advantage of including the verb prefixes in this generalisation. The insight that the SUBJ function of prefixed verbs is not assigned by the operation of LMT can be incorporated in two ways. One possibility is that the verb prefixes have the effect of allowing the maximal projection of the verb to be interpreted as an attribute of the SUBJ. The predicate would have a semantic role Attributee to be associated with the SUBJ as suggested for nonverbal predicates previously. This would mean that the literal reading of a clause such as *Saya melihat gadis itu* would be along the lines of 'I have the attribute of being the seer of the girl'. This possibility would give a consistent account of all the predicate types, and it also fits into a comparative and historical perspective on Austronesian syntax which will be discussed in section 5.4.3. But this analysis is very close to the nominalisation analysis discussed in section 2.3.3.1, and there are good reasons to reject that analysis for Indonesian. There are actor and undergoer nominalizations available, distinct from the prefixed verb forms (although the morphology is related in the case of the actor forms):

44.

Root	**Gloss**	**A-nominal**	**Gloss**	**U-nominal**	**Gloss**
tatar	upgrade	*penatar*	trainer	*petatar*	trainee
suruh	order	*penyuruh*	one who gives orders	*pesuruh*	messenger, one who takes orders

And although the non-subject argument of a prefixed verb has the same coding properties as the possessor of a noun (i.e. enclitics are possible), their behavioural properties are different. The preferable alternative is that the verb prefixes remove an argument from a-structure and thus from the operations of LMT. The semantic role associated with the argument is still available in the lexical entry of the verb and can be semantically identified with the SUBJ in a subject-predicate structure. The lexical entry of the prefixes *meN-* and *di-* would be the following:

45. *meN-*, prefix, __V < x, y ...>, [SPEC ARG x]
46. *di-*, prefix, __V < x, y...> [SPEC ARG y]

and the morphological rule which creates prefixed verbs might be:

47. Verb prefixing

Prefix + V < (a), [SPEC ARG], (b) > $\Rightarrow$ V [+pred] < (a),(b) >, θ_j

|

θ_j

For each prefix, an argument is specified as the argument targeted by the operation of the morphological rule. The rule removes that argument from the a-structure but leaves the semantic role associated with it available for semantic interpretation. Similar changes would be required for the morphological rule for Pro-V clauses (chapter 2, example 152); given the speculative nature of the current discussion, I do not specify this change here.

Aside from the issue of semantic roles, the analysis of prefixed verbs as potentially predicative elements makes a prediction about the syntactic properties of such words which is supported by evidence. The prediction is that such words will be able to head nonpredicative constituents without a SUBJ. The data which supports this prediction is discussed in the following sections.

5.3.3 Nonpredicative uses of prefixed verbs

The three categories which can head non-verbal predicates, nouns (projecting to DP), adjectives and prepositions, also can all head non-predicative constituents. The fact that these usages are common, even basic, and do not require the assignment of a SUBJ function (except possibly in the case of adjectives) was taken previously as an argument in favour of the idea that potentially predicative elements need not have a SUBJ in the argument list of their lexical entries. If prefixed verbs are potentially predicative elements also and need not have a SUBJ as a lexical property, then it can be predicted that they should also be able to appear in positions other than that of predicate. This is indeed the case, and in the contexts which might be analysed in this way, there is a contrast in acceptability between prefixed verbs and unprefixed verbs which is in line with the account offered here.

Voskuil (1996: 206-208) identifies constructions involving what he calls three different verb modes: infinitive, present participle and gerund. This classification, on the basis of the data Voskuil presents, is due to the nature of the translations of the different possibilities rather than to any clear differences in the Indonesian[12]. IRG (302-308) treats the topic under two heads: predicate nominalisation and simple nominal clauses, which seem to correspond respectively to a subclass of Voskuil's gerunds, and everything else. Sneddon's predicate nominalizations have different characteristics from the other types, and I discuss them last.

Simple nominal clauses consist of the maximal projection of a prefixed verb or an intransitive verb:

[12] Voskuil also cites another work (Lappia, M. & J.E.Voskuil (1992) Nominal infinitivals. In S.Barbiers & C. Levelt (ed) *Proceedings of the Third Leiden Conference of Junior Linguists*) as giving a more detailed treatment. This work was not available to me.

48. *Mencari pekerjaan di kota tidak begitu mudah*
AS.find work LOC city NEG like.that easy
'Finding work in the city is not very easy.' (IRG: 307)

49. *Dipukul Yanto sungguh tidak enak!*
US.hit Yanto surely NEG nice
'Being beaten by Yanto is surely not nice.' (Voskuil 1996: 207, ex17)

Such nominalizations do not have the same distribution as DPs, although the overlap is considerable. In the above examples, the nominalised clause is the subject of an adjectival predicate. They can also be the subject of a verb:

50. *Membaca buku menambahkan pengetahuan umum*
AS.read book AS.increase knowledge common
'Reading books increases one's general knowledge.'
(Voskuil 1996: 207, ex16)

or the complement of a preposition:

51. *Dengan mengucapkan terima-kasih, dia meninggalkan*
with AS.say thank.you 3SG AS.leave
ruang itu
room that
'Saying thank you, he left the room.' (Voskuil 1996: 206, ex7)

Such clauses do not seem to be possible as non-subject arguments of a verb. Generally, substituting a bare verb for the prefixed form results in unacceptability:

52. **Cari pekerjaan di kota tidak begitu mudah*

53. **Baca buku menambahkan pengetahuan umum*

54. **Dengan ucapkan terima-kasih, dia meninggalkan ruang itu*

Voskuil (1996: 206) claims that, in some contexts and for some verbs, a bare verb form is marginally acceptable, and my consultants allow the possibility of such structures in very informal speech[13]. Overall, bare verb forms are marginal in these structures, and the difference in grammaticality is definitely in the predicted direction.

Familiar analyses of such structures in English appeal to the existence of a null pronoun in subject position (P&P) or to anaphoric control of the SUBJ function (LFG). Such accounts depend on the fact that the verb forms used have distinct properties, traditionally summarised by the term *non-finite*. But there is no distinction in verb form which can be appealed to in Indonesian in the same way: exactly the same prefixed verb forms appear in the 'non-finite' contexts as appear as the predicate of free clauses. Any analysis which

[13] Example 54 is judged unacceptable under any circumstances, for reasons to do with register.

postulates a finite/non-finite distinction in Indonesian introduces a theoretical assumption not based on the facts of the language. And such an analysis still has to explain the ungrammaticality of bare verb forms in these contexts: if a subject is projected either in syntax (null pronoun) or in f-structure (anaphoric control), why can it satisfy the properties of one verb form but not the other? The other contrasts between the properties of prefixed and unprefixed verbs which have been discussed to this point referred to properties of non-subject arguments. No data examined previously has suggested that the subjects of the two types of verb differ in any way, therefore this contrast is surprising. Additionally, one of the facts about English that has to be accounted for is that in some non-finite contexts (before *-ing* verbs) a subject is possible. However, in Indonesian, a subject is never possible in these constructions. These facts support the view that the grammatical function SUBJ is not present at all in these cases.

The analysis of prefixed verbs developed here can provide a straightforward analysis of these constructions. A prefixed verb is a potentially predicative category, and it need not have a SUBJ in its lexical entry, it can therefore appear without a SUBJ. Such verbs nevertheless carry semantic information about a potential subject, and that information is used in interpreting these constructions, even though it is not represented syntactically. Note that this account implies that verb roots have a logical subject argument specified in such a way as to ensure that it is mapped to SUBJ, but that prefixation removes this and leaves the possibility (actually high probability) of a SUBJ being added back to the argument list by lexical redundancy rule. This is admittedly messy. An alternative analysis would use a category-changing rule to make a prefixed VP into a DP, with an additional rule required to state that the SUBJ of the verb need not (cannot) appear in such a case. This solution is also messy, and the first alternative at least has the advantage that all the processes assumed to affect potentially predicative words are confined to the lexicon.

The difference between the structures just discussed and what Sneddon calls *predicate nominalizations* is that the subject is expressed in this latter type. Only adjectival clauses, clauses with intransitive verbs, and clauses with *di*-prefixed verbs can be nominalised in this way, in which *-nya* is added to the predicate head to make it a noun and the subject follows it as a possessor:

55.	*Lalu lintas*	*menjadi*	*macet*	*karena*	*ditutupnya*
	traffic	AS.become	stuck	because	US.close.3
	beberapa	*jalan*			
	several	road			

'Traffic became jammed because of the closure of several roads.' (IRG: 302)

This type of nominalisation raises interesting questions which are not relevant to the present discussion, but they will be discussed briefly in section 5.4.3.

5.3.4 Overt copulas

Indonesian has two elements which can be analysed as copulas: *adalah* and *ialah* (IRG: 237-238). Both words include the foregrounding particle (IRG:

261-263) or focus marker (Alsagoff 1992: 114-117) *-lah*, the other morpheme is the existential verb *ada* in one case and a third person pronoun in the other. These words can be optionally inserted between the subject and the predicate in a non-verbal predication (*ialah* can only follow a third person subject), and therefore the account developed here would predict that prefixed verbs could follow these words also. This is possible, and therefore the data give some support to my hypothesis. However, in the case of nominal, adjectival and prepositional predicates, the copulas do not alter the meaning:

56. *Soal ini (adalah) penting sekali*
problem this COP important very
'This problem is very important.'

The appearance of the copula in such clauses is generally to make the structure unambiguous, particularly when a nominal clause has a subject with dependent nouns. The copula then clarifies the break between the two constituents.

A copula is only possible before a prefixed verb where that verb is to be understood as a nominalisation:

57. *Tugas saya adalah menjual koran*
task 1SG COP AS.sell newspaper
'My job is selling newspapers.' (IRG: 308)

The nature of such structures makes it almost impossible to have a minimal pair of examples. However, comparison of example 57 and the following gives some idea of the meaning difference which comes with the copula:

58. *Saya menjual koran*
1SG AS.sell newspapers
'I sell newspapers.'

If prefixed verbs and the other possible predicative elements are as similar as I am claiming, then this semantic change accompanying the introduction of the copula would not be predicted, and this would be evidence against my hypothesis.

However, an alternative view of *ialah* and *adalah* is possible, one which removes the conflict between my hypothesis and the data just discussed. As mentioned previously, both the words in question include a pragmatic particle in their form. One is clearly a grammaticalization of a left-dislocation construction: X_i, *ia*$_i$*-lah* Y, and the other quite possibly derives from a similar type of construction, as *ada* is an existential and presentative verb, very commonly initial in a clause:

59. *Ada koran di meja*
exist newspaper LOC table
'There is a newspaper on the table.'

It is therefore just as plausible to say that these two words are pragmatic elements used to clarify the structure of the clause, as it is to claim that they

are copulas. Then, the contrast in meaning of clauses with prefixed verbs depending on the presence of *adalah* or *ialah* would be predictable. The previous section has shown that prefixed verbs and their maximal projection can be nominalizations, and this would mean that a subject-predicate structure with such a phrase as predicate would be ambiguous between two readings, one with a nominal predicate and one with a verbal predicate. It is not hard then to imagine that *adalah* and *ialah* would come to be used to disambiguate, with their presence always indicating that a nominalised clause followed. This interpretation is supported by some facts about negation. Firstly, *adalah* and *ialah* are definitely not verbs as they cannot be preceded by the negator *tidak*. Any negation in a clause containing one of these words is a part of the predicate constituent following *adalah* or *ialah*:

60. *Pernyataan ketua koperasi itu (*tidak) adalah tidak benar*
statement chairman co-operative that PRT NEG correct
'The statement by the chairman of the co-operative is not correct.' (IRG: 238)

Secondly, when a nominalised clause functions as predicate, it can only be negated by the nominal negator *bukan*. When this word is present, there is no possible ambiguity and *adalah* and *ialah* 'rarely occur' (IRG: 308):

61. *Kegemaran saya bukan memelihara binatang*
hobby 1SG NEG AS.take.care.of animal
'My hobby is not looking after animals' (IRG: 308)
(Looking after animals is not my hobby)

This can be compared with:

62. *Kegemaran saya adalah tidak memelihara binatang*
hobby 1SG PRT NEG AS.take.care.of animal
'My hobby is not looking after animals.'
(Not looking after animals is my hobby)

These two pieces of evidence support the analysis of *adalah* and *ialah* as pragmatic particles which are used to clarify potentially ambiguous structures. On this interpretation, the meaning change that occurs when one of them precedes a prefixed verb is not counter-evidence to my hypothesis. Rather, the phenomenon is predicted.

5.3.5 Intransitive verbs: the prefixes *ber-* and *ter-*

The proposal being advanced here is that Indonesian has two types of clausal structure in which verbs can appear, and that a crucial diagnostic as to which type is being used is morphology: the prefixes *meN-* and *di-* derive potential predicates from verbal bases. Some one argument verbs in Indonesian are monomorphemic (e.g. *pergi* 'go', *datang* 'come'), but many are morphologically complex (e.g. *ber-temu* 'meet', *ter-tidur* 'fall asleep', and the verbs discussed in chapter 3). Section 3.3.3 already argued that the possibility of licensing an oblique argument by adjunction to the verb existed for one

type of intransitive verb, and the analysis presented here predicts that this possibility should be more widespread. The behaviour of some verbs prefixed with *ber-* and *ter-* supports this prediction.

The prefix *ber-* derives intransitive verbs (IRG: 61-64, Verhaar 1984a), such as the following:

63. *Bara cemburu berkobar lagi di hati Marisa*
fire jealous *ber*.flare.up again LOC liver Marisa
'The fire of jealousy flared up once more in Marisa's heart.' (SDM: 100)

64. *Aku tidak mau bertengkar*
1SG NEG want *ber*.dispute
'I don't want to quarrel.' (SDM: 71)

But many verbs with this prefix can occur with a DP complement:

65. *Arneta seorang janda berumur tiga puluh empat tahun*
Arneta one.CLASS divorcee *ber*.age three ten four year
'Arneta was a divorcee aged thirty four.' (SDM: 63)

66. *Banyak siswa belajar bahasa Perancis*
many student *ber*.study language France
'Many students study French.' (IRG: 266)

As would be predicted if such DPs are adjoined obliques, they cannot be separated from the verb:

67. **Arneta seorang janda berumur katanya tiga puluh empat tahun*
Arneta one.CLASS divorcee *ber*.age word.3 three ten four year
(FOR: 'Arneta was a divorcee claiming to be aged thirty four.')

68. **Banyak siswa belajar dulu bahasa Perancis*
many student *ber*.study before language France
(FOR: 'Many students used to study French.')

And in some cases, an alternation between the prepositional and the adjoined realisations can be observed:

69. *Semua orang ingin bertemu dengan Lisa*
all person wish *ber*.meet with Lisa
'Everyone wants to meet Lisa.'

70. *Saya ingin bertemu guru saya*
1SG wish *ber*.meet teacher 1SG
'I want to see my teacher.' (E&S: 565)

I assume that in cases where such an alternation is not possible, this is due to the lack of a suitable preposition, as was argued in section 3.3.3.

Similar data is discussed by Postal (1977), who suggests that *ber-* may be a marker of an antipassive clause type. Leaving aside the issue of whether this analysis is satisfactory, Postal also cites a personal communication from Dreyfuss to the effect that the post-verbal argument is restricted to generic and non-specific nominals, and that therefore this type of clause is an instance of noun incorporation[14]. Example 65, with a quantified nominal following the verb, and example 70, with a possessed nominal, suggest that this account is not correct (see also argumentation is section 2.3.1.2).

The prefix *ter-* also generally derives intransitive verbs (IRG: 112ff) which often translate English passives, and carry an implication that the event was out of the control of the participant (Cartier 1978)[15]. A post-verbal nominal is possible with some *ter-* verbs (see also example 147 of chapter 4):

71. *Dia sedang terlibat affair dengan seorang*
3SG PROG *ter*.involved affair with one.CLASS
pegusaha wanita yang umurnya sepuluh tahun
worker woman REL age.3 one.ten year
lebih tua
more old
'He was having an affair with a female worker who was ten years older.' (SDM: 59)

and an alternation between PP and DP occurs in some cases:

72. *Obat itu terminum (oleh) anak saya*
medicine that *ter*.drink by child 1SG
'The medicine was accidentally drunk by my child.' (IRG: 115)

The preposition is obligatory when a pronoun follows, a constraint which also applied to *ke- -an* verbs (see section 3.2.2.1), and the now familiar adjacency requirement also applies:

73. **Obat itu terminum dulu anak saya*

This data shows that the split between two types of clausal organization extends to the intransitive verbs of Indonesian, and confirms that there is an essential relation between the subject-predicate clause type and morphologically complex verbs. The possibility of an oblique adjoined to the verb should be limited to the case of morphologically complex (i.e. potentially predicative) verbs.

[14] Postal (1977: n44) suggests that the two analyses are not inconsistent; indeed he claims that antipassive in his sense is a prerequisite for incorporation.

[15] Following the discussion of section 3.3.1, it is interesting to note that where *ter-* is prefixed to a verb root with a suffix, the suffix is almost always dropped and is always unacceptable for some speakers (IRG: 115-116).

5.3.6 Historical and comparative evidence

The discussion of the preceding two sections of prefixed verbs heading nominalizations is related to some material presented in section 2.3.1.1. That section discussed the hypothesis put forward by Voskuil (1996) that the actor in a clause with a *di*-prefixed verb is syntactically equivalent to a possessor in a nominal projection, it has genitive case. The position of such actors and the fact that *-nya* can appear in this position are the main pieces of evidence in support of this position. Other material in chapter 2 made it clear that the various possibilities for the expression of the actor in clauses with *di*-prefixed verbs could not be treated together, which already suggests that Voskuil's analysis is not correct. The data presented in section 5.3.3 provides an additional argument against the proposal. If Voskuil is correct, we would expect that the enclitic could be used to represent the actor of a *di*-verb in a nominalisation. If such actors are in genitive case anyway, this would seem to be a natural prediction. But it is not correct. In the one construction where *-nya* must appear in a nominalisation, predicate nominalisation, it does not express the actor. Example 55, repeated here, does not have the meaning that some identifiable person or persons acted to close the roads; rather the nominalised clause has the reading of an agentless passive.

55. *Lalu lintas* *menjadi* *macet* *karena* *ditutupnya*
traffic AS.become stuck because us.close.3
beberapa *jalan*
several road
'Traffic became jammed because of the closure of several roads.' (IRG: 302)

The function of *-nya* in such constructions is clearly non-referential as can be seen by comparing an intransitive clause with its nominalisation:

74. *Kerusuhan* *terjadi* *di* *kampus*
riot occur LOC campus
'Riots occurred on the campus.'

75. *Mereka* *prihatin* *atas* *terjadinya* *kerusuhan* *di*
3PL concerned on occur.3 riot LOC
kampus
campus
'They were concerned about the occurrence of riots on the campus.' (IRG: 303)

The pronoun does not represent a participant in the event here, and it does not do so in example 55 either. If anything can be analysed as a possessor in example 55, it is the subject of the original clause, the undergoer[16]. This is directly against the prediction which Voskuil's analysis would make.

[16] One function of *-nya* is to act as a ligature between the head noun and a possessor, especially where the structure might be ambiguous otherwise. For example, *ibu Suparjo* is ambiguous between 'Mrs Suparjo' and 'Suparjo's mother', while *ibunya Suparjo* has only the second meaning.

These considerations, as well as the fact that the projections of prefixed verbs are ambiguous between a nominal and a predicative reading, rule out any attempt to analyse Indonesian as having non-verbal predication as its predominant clause type. Nevertheless, it is valuable to consider the respects in which the current proposal resembles such analyses, in order to bring out the characteristic Austronesian traits which Indonesian possesses. This type of analysis has been proposed for all clauses in Tagalog and other Philippine languages (Dahl 1973: 117-118, Naylor 1995, Schachter & Otanes 1972: 62), and for non-actor subject clauses in Tagalog (Capell 1964, Lopez 1941), Atayal (Egerod 1966: 346), Toba Batak (van der Tuuk 1971) and Rukai (Li 1973: 202-211). The basic idea is that all the relevant clauses are equational, with a subject and a nominalised predicate. This can be illustrated using Tagalog examples. The following paradigm was given as examples 84-87 of chapter 2. It is repeated here with translations amended to indicate the equational analysis (the most natural English wording uses clefts which introduce an unwanted pragmatic effect unfortunately):

76. *Bumili* *ang* *lalake* *ng* *isda* *ng* *pera* *sa*
AF-buy NOM man GEN fish GEN money OBL
tindahan
store
'The man is the one who bought fish with money in the store.'

77. *Binili* *ng* *lalake* *ang* *isda* *ng* *pera* *sa*
UF-buy GEN man NOM fish GEN money OBL
tindahan
store
'The fish is what the man bought with money in the store.'

78. *Binilhan* *ng* *lalake* *ng* *isda* *ng* *pera* *ang*
LF-buy GEN man GEN fish GEN money NOM
tindahan
store
'The store is where the man bought fish with money.'

79. *Ipinambili* *ng* *lalake* *ng* *isda* *ang* *pera* *sa*
IF-buy GEN man GEN fish NOM money OBL
tindahan
store
'The money is what the man bought fish with in the store.'

The morphology attached to the verb roots is assumed to have a particular character. It is derivational, and it orientates the word so that it may be used to refer to one of the participants involved in the event denoted by the base (Himmelmann 1991). Similar processes can be seen on a limited scale in English. The noun *control* is not oriented to any participant, and both arguments of the verb *control* can appear as modifiers in the noun's projection: *Hoover's control of the FBI*. But if the orientating nominalizer *-er* is used, the resulting noun refers to one of the participants and that argument cannot appear in the noun's projection: *the/*Hoover's controller of the FBI*.

Whilst I do not claim that prefixed verbs in general head nominalised constituents, the nature of the verb prefixes clearly has something in common with this view of Tagalog morphology. The prefixes are derivational, in that they prevent the root from projecting a regular verb phrase, and they are orientating, in that they specify the semantics of the argument of which they can be predicated. While I reject the view that prefixed verb clauses in Indonesian are equational, I leave open the possibility that they are attributive in nature. On this view, the prefix indicates that the subject of the predication has a certain role or attribute in regard to the event which is denoted.

The sample of languages mentioned above as those for which equational sentence analyses have been proposed is biased towards the Formosan and Philippine languages which are generally assumed to have separated early in the Austronesian dispersal (Tryon 1995). This suggests that the clausal patterns which have suggested such analyses to various scholars were a feature of early stages of the history of the Austronesian language family. Starosta, Pawley & Reid (1982) argue that even if the Philippine-type systems we know today do not nominalise all verbs, the origin of such systems is in subject-predicate clauses in which the predicate was a nominalisation with an orientating affix. There is thus a considerable body of scholarly opinion which justifies treating the prefixed verb constructions of Indonesian as conservative in the context of the Austronesian language family.

There is also good evidence that the bare verb constructions are innovative in the context of related languages. It is a characteristic of many Western Malayo-Polynesian languages that verbs can occur in two forms, a bare form and a form with a nasal segment prefixed to the root. These two forms are ordered differently with respect to the arguments of a transitive clause, the accessibility of the two arguments to syntactic processes varies with the verb form, and there may be a pragmatic difference in meaning between the two. These points are illustrated with examples from Balinese and Sasak (Ngenó-ngené variety):

80. Balinese
Nyoman lempag tiang
Nyoman hit 1SG
'I hit Nyoman' (Artawa & Blake 1997: ex2)

81. *Tiang nglempag Nyoman*
1SG N.hit Nyoman
'I hit Nyoman' (Artawa & Blake 1997: ex8)

On the basis of various tests, Artawa and Blake show that the preverbal argument is the subject in each case, that is, the bare verb form is associated with undergoer subject.

82. Sasak
Aku balé beli
1SG house buy
'It is a house that I buy.' (Austin 2000: ex3)

83. *Aku* *mbeli* *balé*
1SG N.buy house
'As for me, I buy a house.' (Austin 2000: ex4)

The meaning difference here suggests that the bare verb form is more orientated towards the undergoer, and at least some syntactic processes show the same effect: only the patient can be questioned and can head a relative clause if the verb form is bare. Although it is not made evident in the glossing of the Tagalog examples above, it is also true in that language that bare verb forms are orientated towards the undergoer (Blake 1988, Himmelmann 1991)[17]. Starosta, Pawley & Reid (1982) claim that Proto-Austronesian had an ergative orientation, at least in part, and the sort of evidence just surveyed from daughter languages is compatible with such a claim. We would expect undergoer subject verb forms to be less-marked in ergative languages, and such a pattern might be retained from a proto-language[18].

In this context, the Indonesian bare verb clause types are unusual. The two types discussed previously in this study, bare verb actor subject clauses and emotion verb clauses, both resemble typical clauses of a language organised along nominative-accusative lines rather than the examples from other Austronesian languages examined above. These considerations, along with the tendency among grammarians to stigmatise these clause types, support the idea that they represent an innovative pattern in Indonesian. Section 2.2.1.4 discussed the historical origin of the prefix *di-*, and the evidence suggests that it is not a very old feature of the Malay subgroup. It is tempting to think that the introduction of this prefix and the introduction of actor subject bare verb clauses might be linked. If previously, the bare verb form occurred in undergoer subject clauses, the introduction of an alternative pattern might well provide the sort of pressure which might lead to innovative morphological marking on the older clause type. I stress that this is pure speculation: I have collected no data on the relative antiquity of the two phenomena in Malay[19].

5.3.7 Summary

The data and arguments presented here show that the idea that prefixed verbs in Indonesian have a good deal in common with non-verbal predicates in the language is plausible aside for reasons other than the contrasts which I seek to explain. However, all the issues discussed above relate to the status of the subject and the relation between it and the predicate. I am trying to account for contrasts in the behaviour of non-subject

[17] Aspect morphology is not glossed separately in the Tagalog examples for ease of reading. When this information is included, it becomes clear that the voice morpheme in example 66 is zero.

[18] I do not intend to imply any claim that any of the languages discussed in the preceding paragraph are organised on ergative-absolutive lines. The issue has been debated for some years without any definite conclusions being reached. In addition to the works cited, see Maclachlan & Nakamura 1997, Manning 1996a, Payne 1982.

[19] Voskuil (1996: 203-204) gives examples of bare verb actor subject clauses from classical Malay, but does not provide full details of the sources.

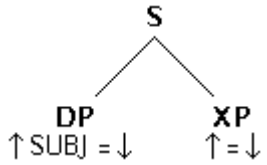
arguments. While it is a useful first step to be able to show that the constituents within which the non-subject arguments occur differ in status, this in itself is not an account of the differences in behaviour. I now turn to a discussion of the difficulty of expressing the intuition about the nature of prefixed verbs in LFG, and the closely related difficulty of giving a theoretically satisfying account of the different properties of the non-subject arguments of the two clause types.

5.4 Problems for an LFG account

5.4.1 Where does subj come from?

Bresnan (2001a: 112) argues that the category S may be expanded as NP (or DP) and XP in a subject-predicate structure, that is, that S may be top node of a configurational structure, despite the exocentricity of the category. The phrase-structure schema in question, with functional annotation, is the following:

84.



The origin of the SUBJ function in this schema raises some problems. The functional annotation makes it clear that the GF must be part of the lexical information associated with the head of XP: the ↑ = ↓ equation makes this clear. But the schema allows any category of head here. Is it plausible or desirable that, for example, all nouns have an item in their lexical argument list which maps to a SUBJ function in f-structure? Various solutions to this problem seem possible.

Firstly, there is the possibility of a null copula analysis. This sort of account would assume that there is a lexical item in Indonesian which takes a SUBJ and a predicative phrase as its arguments, and which has no phonological content:

85. (PREDICATE) V < SUBJ, XP [+pred] >

This alternative is allowed by the framework of LFG as it stands, although the introduction of null elements is against the spirit of the theory as I understand it.

Another possibility is that the potentially predicative categories in Indonesian are systematically ambiguous. The LFG architecture does not really permit grammatical functions to be optional in lexical entries; optionality is a shorthand for multiple entries, with each one instantiating one option. We could assume that nouns, prepositions and adjectives in Indonesian all have two lexical entries, linked by redundancy rules:

86. N → N < SUBJ >

One entry has no SUBJ function in its argument list, whilst the other does have this attribute. However, in line with the prior discussion, I assume that the grammatical function would be introduced directly by lexical specification. There is no contentful semantic role to be listed in the lexicon of these words, and therefore the GF cannot be the result of the operations of LMT. At most, the lexical forms with the SUBJ function have a semantic role, call it Attributee, available for the semantic interpretation of the SUBJ, although even this may not be a necessity. Note that these assumptions mean that every time one of those heads is introduced into a structure, two parallel f-structures must be created, one with a SUBJ function and one without. In every case, one of these structures will be ill-formed: where the item was not predicative, the f-structure with a SUBJ will be ruled out as incoherent (the GF would not be required by any predicate), and where the item was predicative, the f-structure without a SUBJ will be ruled out as incomplete (a GF required by a predicate would not be present). The approach is therefore computationally intractable, as the number of f-structures to be considered must increase exponentially. This is a considerable drawback.

There is another disadvantage to this approach: it cannot capture the intuition behind the proposal introduced in section 5.3. I suggested there that the essential difference between prefixed verbs and unprefixed verbs was that they appeared in different types of structure. The significant difference between the two was, I proposed, that unprefixed verbs had a SUBJ function in their argument list, but that prefixed verbs did not. This approach allows both types to project the same sort of constituent in syntax (a VP) using the same phrase structure rules, while ensuring that only one of these constituents could result in a well-formed f-structure under the standard X-bar theory assumed. The prefixed verb would have no SUBJ function, the expansion of IP has the functional annotation (↑ SUBJ) = ↓ on its specifier node, by general principles, and the structure-function correspondence must therefore fail for such verbs. But if the SUBJ function is added to potential predicates by lexical redundancy rules, then this would no longer be the case, and nothing would prevent a prefixed verb from appearing in a standard SUBJ-VP structure. Some diacritic feature would need to be added to the prefixed verbs to ensure that this did not happen, and this in turn would require a separate set of phrase structure rules to account for the projections of prefixed verbs. Also, if a lexical redundancy rule similar to 86 applies to verbs, this must mean that the verb root has an argument which will map to SUBJ in its lexical entry (bare verbs have a SUBJ), the prefix must remove this when it makes the verb a potential predicate, and then the redundancy rule restores a SUBJ in a new lexical entry. This is messy and implausible.

Bresnan (2001a: 276) proposes a solution to this problem as part of a discussion of predicative complements in English. While prepositions and nouns are only predicates of main clauses when supported by the copula, they can appear alone as predicates in embedded clauses with controlled subject functions:

87. Jogging keeps Susan [in a bad mood].

Bresnan proposes a lexical predication template 'which augments the lexical form for a preposition or nominal to one which has a subject of predication' (276):

88. 'in <(↑OBJ)>' ⇒ 'be-in <(↑SUBJ) (↑OBJ)>'

This proposal differs from the redundancy rule approach discussed above in that in addition to altering the argument list of the lexical item, the template also alters the listed item. In the context of the discussion in which the proposal is made, it is not clear how this is to be understood. As a main clause predicate, a preposition such as *in* requires the copula, but in an XCOMP, the copula is impossible:

89. *Jogging keeps Susan [is/be in a bad mood].

Nonfinite verb forms can appear in such environments, therefore the impossibility of any form of the copula is surprising if it is indeed a part of the lexical item. If that is the intention, the template requires also a rule deleting *be* in non-finite environments. In any case, this approach will not transfer to Indonesian, unless it is taken to imply the zero copula analysis already discussed. In that case, it is simpler to assume that the zero copula licenses the SUBJ function, rather than including both steps. If on the other hand, the change in the lexical item is to be understood as semantic, then this proposal is identical to the redundancy rule discussed previously, as far as syntax is concerned. The case which Bresnan discusses as involving a clear semantic change, that of locative prepositions, also does not transfer to Indonesian where the purely locative preposition *di* can be used as a predicate:

90. *Mereka di Jakarta sekarang*
3PL LOC Jakarta now
'They are in Jakarta now.' (IRG: 237)

The notion of a lexical predication template is therefore not helpful.

There is additional evidence in Indonesian that a solution which assumes that the SUBJ is introduced by a lexical process, such as a redundancy rule, must be incorrect[20]. Indonesian uses headless relative clauses freely. Such constituents have the same distribution as other DPs, they can fill argument positions:

91. *Yang terbayang di depan matanya cuma Marisa*
REL visible LOC front eye.3 only Marisa
'What was visible before his eyes was Marisa only.' (SDM: 103)

[20] I am grateful to Peter Austin for pointing out the importance of the argument advanced in this paragraph.

92. *Andi memilih yang besar*
Andy AS.choose REL big
'Andy chose the big one.' (IRG: 301).

And, like other nominal constituents, headless relative clauses can be predicates:

93. *Mobil saya yang biru*
car 1SG REL blue
'My car is the blue one.' (IRG: 301)

This possibility is problematic for any account which assumes that the SUBJ function is added in the lexicon. There is no lexical head which projects the nominal constituent and which could have the SUBJ function associated with it by a lexical process. Three alternative analyses must be considered. Firstly, a null nominal head might be proposed for Indonesian, and the SUBJ could be associated with it in the relevant cases. This proposal is open to the same objection as the null copula analysis discussed previously. In addition, the null head would always and only co-occur with *yang*, as there is no independent motivation for null nominal heads in the language. Therefore, this analysis would be stipulative. The second possibility is that *yang* itself can acquire a SUBJ by lexical rule. This proposal is also stipulative because it would be necessary either to claim that *yang* is the only member of some lexical category, or that all other members of that category, perhaps complementizers, are barred from the redundancy rule which adds SUBJ to the lexical entry. It would not, however, be necessary to stipulate that *yang* could not appear as a predicate on its own. Nouns and adjectives have no obligatory arguments specified in their lexical entries, therefore they can be single-word predicates. Prepositions in Indonesian have obligatory complements and therefore can only be part of a phrasal predicate. A similar approach could be taken for *yang*: it has an obligatory clausal complement. This proposal would also add to the computational problem discussed before. If *yang* has two possible lexical entries, and every noun does also, any nominal constituent including a relative clause would have four possible f-structures associated with it. The final alternative is that *yang* always introduces a predicate, and that the head noun which is modified by the relative clause is always its SUBJ. Such an analysis is semantically awkward, in that the head noun always has some function in the relative clause, and syntactically improbable, in that it would imply a claim that any DP including a relative clause was a category-shifted clause. None of these three analyses is plausible, therefore the use of headless relative clauses as predicates is strong evidence against any proposal which allows potential predicates to acquire a SUBJ function by some lexical process.

A solution to the problem under consideration would be to have the phrase structure rule which allows DP, XP as an expansion of S to be associated with some f-structure mechanism which introduces the SUBJ function. Such a proposal would capture the intuition that the non-verbal categories are predicative in only this syntactic environment. However, this solution conflicts with a basic principle assumed in LMT, the subject condition:

> Every lexical form must have a subject. (Bresnan & Kanerva 1989: 28)
>
> or:
>
> Every predicator must have a subject. (Bresnan 2001a: 311)

This principle apparently rules out any analysis in which the SUBJ function is introduced constructionally.

However, the subject condition is qualified in both the sources cited above. Bresnan & Kanerva (1989) note that there are languages, such as Acehnese (Durie 1985), for which no clear case can be made for analysing a subject grammatical function. Additionally, Bresnan (2001a) specifies that the condition applies only to the a-structures of predicators taking core arguments, that is SUBJ, OBJ and OBJ_θ. Even on this basis, the condition does not apply in all cases. For example, Bresnan (2001a: 275-280) analyses locative PPs in English as oblique arguments. The complement of the preposition is analysed as an OBJ, a core argument function, and the subject condition might be expected to apply to the prepositional predicator. But it is explicitly stated (p277) that such structures lack a subject of predication. These considerations suggest that the statements of the subject condition quoted above are not adequate, and that this condition should be considerably more restricted in its scope[21]. Therefore I will assume that, in restricted circumstances, grammatical functions can be introduced by non-lexical means.

The concrete details of my proposal would be as follows[22]. The phrase structure rule allowing subject-predicate structures also introduces a PRED value 'is-predicated-by' taking a SUBJ and a PREDICATE as its arguments. This item is not associated with any c-structure position:

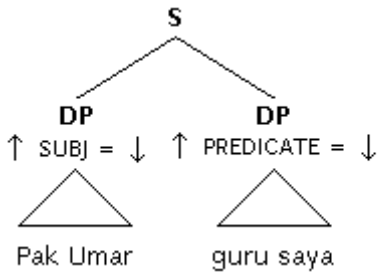
94. S → DP XP
[PRED 'is-predicated-by' < SUBJ PREDICATE >]
where predicate = XP [PREDICATIVE +]

In the case of a simple nominal predicate, this rule allows structures of the following type:

[21] Joan Bresnan (p.c.) suggests that "the subject condition must be restricted to a subset of verbal predicators".

[22] I am grateful to Joan Bresnan for suggesting improvements to the formulation of this proposal.

95.



96.

$$\begin{bmatrix} \text{PRED} & \text{'is-predicated-by } \langle \text{SUBJ, PREDICATE} \rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'NAMED-Mr Umar'} \\ \text{PREDICATIVE} & + \end{bmatrix} \\ \text{PREDICATE} & \begin{bmatrix} \text{PRED} & \text{'teacher'} \\ \text{PREDICATIVE} & + \\ \text{POSS} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

In the case of lexical items other than verbs, the feature [PREDICATIVE +] is assigned to all lexical heads by a redundancy rule. The feature is relevant for subject-predicate structures and for the distribution of enclitic pronouns and its presence has no other effects. Verbs only receive this feature as a result of the morphological operation of prefixation, and therefore bare verbs cannot fill the place of the XP in rule 94.

5.4.2 Status of the non-subject argument

There are two issues to consider in relation to the status of the non-subject argument of a *meN*-prefixed verb: what grammatical function is assigned to such arguments, if any; and how is that function assigned? I will propose two possible answers to the first question, both of these will lead naturally to the discussion of the second issue.

One possible answer to the question of what GF is assigned to the non-subject argument of a *meN*-prefixed verb is to say that it is an OBJ, and that any differences between such arguments and the non-subject arguments of bare verbs are due to the fact that one is inside a predicative constituent and the other is not. Such an analysis implies a stipulation that arguments cannot be extracted from within predicative constituents. This stipulation is empirically accurate, but is not derivable from other properties of the construction discussed here. The advantage of this approach is its intuitive appeal: the non-subject arguments of prefixed verbs have a good deal in common with the non-subject arguments of bare verbs. They are certainly terms, and non-subject terms are generally OBJs. The disadvantage is that this account requires the stipulation as to the impossibility of extracting from predicative constituents. This is clearly *ad hoc*: it is not possible to extract from other potentially predicative constituents when they appear in either predicative or non-predicative environments, nor is it possible to extract from nominalizations of clauses:

97. **buku yang membaca menambahkan pengetahuan*
book REL AS.read AS.increase knowledge
umum
common
(FOR: 'books of which the reading increases general knowledge')

Thus, the stipulation would only be inserted to ensure the desired result in the one case where extraction would be predicted to be possible.

The alternative approach is to assume that the non-subject argument of a prefixed verb is not an OBJ. Then the constraint on forming relative clauses can be stated in terms of a natural class of arguments defined by LMT: only [*-r*] arguments (SUBJ and OBJ) can head gapped relative clauses. This is preferable to the stipulation discussed in the previous paragraph, but leaves the questions about the status of the non-subject argument still unanswered. Evidence has previously been given that this argument is a term: it can float a quantifier, and in the special case of the ergative undergoer subject clauses, it can antecede a reflexive which is also a term. This means that the only possibilities other than OBJ are OBJ_θ, or some GF not currently recognised in LFG. The reasons discussed in section 4.4.1 for not introducing another GF are compelling, but so were the reasons discussed there for not allowing OBJ_θ as the first non-subject argument of a verb. These reasons are reinforced in this case by the fact that prefixed verbs with two objects occur in Indonesian. If it is strange to analyse the single non-subject argument of a verb as an OBJ_θ, then it is even stranger to analyse both non-subject arguments of a ditransitive verb as having that GF. It is interesting to note that the same point has been reached in chapter 4 and here, arguing from opposite assumptions: if the prefixed verbs are true transitive verbs, then the non-subject arguments of the bare verbs must have a GF other than OBJ with OBJ_θ the only available option; and if the prefixed verbs are not true transitive verbs, then the same must be true of their non-subject arguments. Again, this suggests that the LFG framework is too restricted to accurately describe the data in this case.

One possible solution is to appeal to the distinction between terms (direct arguments) and non-terms (obliques and adjuncts). The version of a(rgument)-structure proposed by Manning (1996a,b) Arka and Manning (to appear) and Wechsler and Arka (1998) incorporates information as to the term or non-term status of arguments. Section 3.3.2.3 has presented evidence that this version of a-structure gives a better account of some Indonesian data, and if we accept this version of a-structure, then the notion *term* is a part of the theory. Bresnan (2001a: 96) recognises a core / non-core distinction, but it is not clear what theoretical consequences this distinction has. The notion is independently required in any case. Many languages distinguish terms from non-terms in their grammar, by restricting certain syntactic processes to apply only to one class. Where terms includes OBJ_θ, the class of terms cannot be stated as a natural class using the feature specifications of LMT. SUBJ and OBJ share the feature [*-r*], OBJ and OBJ_θ share the feature [*+o*], but SUBJ and OBJ_θ do not share any feature. Some notion such as term is necessary in these cases in order to state syntactic generalisations. We might then allow, in special circumstances, an argument to be licensed in an f-structure purely by virtue of

being a term. Such arguments would not derive their GF from the operation of LMT, and would therefore have no feature specification. If prefixed verbs have non-subject arguments of this type, then the constraint on gapped relativization suggested previously would give the desired results. However, this is almost as much a stipulation as the alternative solution to the problem discussed previously, with the proviso that the term / non-term distinction is a fact of Indonesian grammar: terms can float quantifiers and non-terms cannot.

I will advance only one more speculation on this topic. The previous paragraph already suggested that it might be possible that LMT does not operate on the SUBJ argument of prefixed verbs, and by implication, of other potential predicates. It might be possible then that LMT does not apply to potential predicates at all. This idea has some plausibility if we consider the other types of head which can be predicates. Nouns can have a possessor argument, but this is a property of the category N, and is not mediated by LMT. Adjectives (at least in Indonesian) do not license a direct argument (see section 4.1.3) and an oblique is the only possibility for a non-subject argument in that case. Prepositions are perhaps more problematic: an argument introduced by a preposition is an oblique, but what GF is assigned to the complement of the preposition: is it also an OBL_{θ}, or is it an OBJ as traditional terminology might suggest? Complement of prepositions cannot be athematic, therefore I assume that OBL_{θ} is the correct answer. But in either case, it is plausible that the GF is not assigned via LMT. However, it is not at all clear that similar arguments can be made in regard to prefixed verbs. Apart from anything else, there are ditransitive verbs to be accounted for, and although this approach may give an answer to the question of how non-subject arguments of predicates receive a GF, it cannot answer the question of what GF the non-subject argument of a prefixed verb actually has. I am forced to conclude that this question is not answerable in a satisfying way in LFG as currently formulated.

5.5 Conclusions

I have argued in this chapter that the contrasting properties of non-subject arguments of prefixed and bare verbs suggest a divide between two types of clause structure in Indonesian. There is evidence to support the idea that the type of structure that prefixed verbs appear in has much in common with non-verbal predications, and this idea also fits well into an historical and comparative view of Austronesian syntax. The type of clause in which bare verbs appear has more in common with the subject-VP structure familiar to linguists trained in English-speaking countries, and this is an innovative pattern in the context of the Austronesian family. I have also argued that there is no straightforward way to express this idea using the available theoretical vocabulary of LFG. To the extent that the intuition underlying my proposal regarding prefixed verbs is correct, this conclusion exposes inadequacies in the theory.

In terms of the objectives of this study as a whole, these conclusions mean that it has not been possible to give a complete account of non-subject arguments in Indonesian. However, the investigation has produced interesting results:

- morphological realisation of arguments can allow the possibility of ergative clauses in a language whose basic alignment is not ergative-absolutive;
- Indonesian allows the OBL_θ grammatical function to be licensed by adjunction to a morphologically complex verb where the argument is assigned a semantic role by the verb. In one case, this strategy was the only one available as no preposition exists in the language which is semantically compatible with the relevant semantic role, patient;
- Indonesian also requires the OBL_θ grammatical function in one clause type to be realised with a preposition although the semantic role assigned to the argument comes direct from the predicate. The preposition does not make a semantic contribution;
- there is an example of predicate composition in the language where the composed argument structure is the result of purely semantic considerations;
- there is evidence that the language has (at least) two types of clausal organization side-by-side.
- the mapping of arguments to GFs is satisfactorily accounted for by LMT in only one of the two types of clause, the innovative one. LMT is of limited use in analysing the conservative clause structure, and this finding is in line with a number of recent studies of Austronesian languages (Arka 1998, Foley 1998, Sells 1998, Wechsler & Arka 1998).

Indonesian has strategies for the licensing of arguments outside of the scheme introduced at the beginning of this work. Two of these seem to be of particular importance, and these two are linked in an important fashion. Firstly, terms can be expressed morphologically in Indonesian and secondly oblique arguments can be licensed by adjunction to the verb which assigns them a semantic role. In both cases, adjacency to the head which governs the argument and the morphological composition of that head are essential features of the licensing mechanism. Both of these processes are possible only in clauses which use the conservative type of structure identified in chapter 5. This suggests that better understanding of that type of clausal structure and of the ways in which arguments are licensed in it will depend in part on investigation of the historical development of Indonesian (or Malay) syntax. Two other questions for future research arise from the results reported here:

- to what extent does the split in clausal organization extend to intransitive verbs? The discussion of chapter 3 and of section 5.3.5 has brought forward some suggestive data, but a more detailed analysis of the problem is needed.

- what is the significance of the particular properties attributed to functional categories in Indonesian in this work, and do they correlate with other typological characteristics?

Appendix: Indonesian and the accessibility hierarchy

Keenan & Comrie (1977; afterwards K&C) propose some universal restrictions on the formation of relative clauses. One of the languages in their sample is Malay, although the nature of the data used is not entirely clear. K&C acknowledge the assistance of Chung with Indonesian data, they cite McDonald & Dardjowidjojo (1967), which is a reference grammar of Indonesian, but they also state in a footnote (n2) that their two informants were from Malaya. Whatever the source and status of their data, K&C's claims do not tally with my understanding of Indonesian grammar, and the actual facts may constitute a counter-example to their second universal

> Any R(elative)C(lause)-forming strategy must apply to a continuous segment of the A(ccessibility)H(ierarchy).

K&C report that Malay has two relativization strategies, one of these is, in their terms, [-case] which is the gapped strategy. The second is [+case] in K&C's terms, because a preposition can precede the head noun. The example which they give of this strategy is distinct from either of the strategies described above:

98. *perempuan kepada siapa Ali beri ubi kentang itu*
woman to who Ali give potato that
'the woman to whom Ali gave the potato' (K&C: ex15)

K&C acknowledge that this example was elicited under pressure, and it is judged strange by native speakers. The summary table given by K&C shows this strategy to be possible for indirect objects, obliques and genitives. The primary strategy, gapping, is shown as possible for subjects and some objects. However, the example which is given to show object relativization is ungrammatical according to most sources and according to native speakers:

99. *?*Ali bunuh ayam yang Aminah sedang memakan*
Ali kill chicken REL Aminah PROG AS.eat
'Ali killed the chicken that Aminah is eating.' (K&C: ex13)

Unlike the example used by Chung (1976a) in making a similar point, there is no question here as to the type of clause which is embedded: it is clearly a *meN*-V clause (see discussion in section 1.1.2).

Presenting the facts of interest in K&C's study as seen in the data given in this study reveals a problem. If we equate the non-subject argument of all transitive structures with K&C's Direct Object, then it is true that the primary strategy is available for all subjects and some objects, the same result that K&C report, if for rather different reasons. And if we take the resumptive pronoun strategy as the secondary strategy, ignoring the possibility shown in K&C's example quoted above (example 96), this strategy is available for some objects, some obliques and genitives. A comparison of the two sets of results looks like this[23]:

[23] I omit the column for object of comparison, which is not relevant.

100.

	Relativizable Positions				
K&C	Subj	DObj	IObj	Obl	Gen
Strategy 1	+	+/-	-	-	-
Strategy 2	-	-	+	+	+
Musgrave					
Strategy 1	+	+/-	-	-	-
Strategy 2	-	+/-	?	+/-	+

It is not immediately clear how to include K&C's indirect object in the present framework. Although the Accessibility Hierarchy appears to generalise over syntactic positions, K&C treat indirect object as a semantic entity. The example from Malay given previously (example 96) apparently is the basis for the claim that indirect objects are relativizable in Malay. But the recipient in example 96, and in the simple clause related to the embedded clause in that example, is formally an oblique: it is introduced by a preposition. In addition, K&C do not seem to recognise second objects as relevant to their investigation, and it is hard to see how they would integrate that grammatical function into their framework. Ditransitive verbs in asymmetrical object languages (Bresnan & Moshi 1990) have the recipient (K&C's indirect object) as their object, with the theme argument as second object. If we reconstruct the accessibility hierarchy as a syntactic phenomenon, and use the meta-language of LFG for this, the following is the result for Indonesian:

101.

	Relativizable Positions				
	SUBJ	OBJ	OBJ_θ	OBL_θ	POSS
Strategy 1	+	+/-	-	-	-
Strategy 2	-	+/-	-	+	+

If this is a licit reinterpretation of K&C's hierarchy, then Indonesian is a counter-example to one of the universals that they propose: the secondary strategy does not apply to a continuous portion of the hierarchy. For LFG, this result is not surprising. As discussed in section 1.3.3, LMT predicts that OBJ_θ will be the most marked GF. OBJ and OBL_θ would be adjacent by the predictions of the theory, according more closely with the Indonesian facts. The relative ranking of OBJ_θ and POSS would still be an issue, but I will assume that a plausible argument can be made that possessors are more common in the languages of the world than second objects. Therefore, on theory-internal grounds, an LFG account gives a hierarchy which more accurately reflects the Indonesian facts as well as avoiding the problem in K&C's account of mixing syntactic and semantic categories in formulating their hierarchy.

Sources of Examples

E&S - John M. Echols & Hassan Shadily (1994) *Kamus Indonesia-Inggris* Jakarta: Penerbit PT Gramedia (3rd edition revised and edited by John U. Wolff and James T. Collins assisted by Hassan Shadily)

PYD - Fredy S (1991) *Perasaan yang ditinggalkan* Jakarta: Gultom Agency

SDM - Mira W (1995) *Sekelam dendam Marisa* Jakarta: Penerbit PT Gramedia

SDPGS - Agatha Christie, tr. Alex Tri Kantjono W (1986) *Satu, Dua, Pasang Gesper Sepatunya* Jakarta: Penerbit PT Gramedia

References

Abney, Stephen P. (1987) *The English Noun Phrase in its Sentential Aspect*. PhD Dissertation, MIT.

Adams, K. & A. Manaster-Ramer (1988) Some questions of Topic/Focus choice in Tagalog *Oceanic Linguistics* 27: 79-101

Adelaar, K. Alexander (1992a) The relevance of Salako for Proto-Malayic and for Old Malay epigraphy. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148:381-408

Adelaar, K. Alexander (1992b) *Proto Malayic: The reconstruction of its phonology and part of its lexicon and morphology* (Pacific Linguistics C-11) Canberra: Research School of Pacific and Asian Studies

Aissen, Judith (1999) Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 19:673-711

Alsagoff, Lubna (1991) The Notions of Subject and Topic in Malay. *Proceedings of the Berkeley Linguistics Society* 16: 21-75

Alsagoff, Lubna (1992) *Topic in Malay: the other subject* PhD Dissertation, Stanford University

Alsina, Alex (1997) A Theory of Complex Predicates: Evidence from causatives in Bantu and Romance. In Alex Alsina, Joan Bresnan & Peter Sells (ed) *Complex Predicates* Stanford CA: CSLI Publications

Anderson, Edmund A. (1983) The meaning of variation in Indonesian. *Linguistic Studies of Indonesian and Other Languages in Indonesia (NUSA 15)* pp1-26

Anderson, Mona (1983) Prenominal genitive NPs. *The Linguistic Review* 3: 1–24

Andrew, Avery D. (1984) Lexical insertion and the elsewhere principle in LFG. Unpublished manuscript, ANU.

Andrews, Avery D. (1990) Unification and morphological blocking. *Natural Language and Linguistic Theory* 8: 507–557

Andrews, Avery D. & Christopher D. Manning (1999) *Complex Predicates and Information Spreading in* LFG Stanford CA: CSLI Publications

Aristar, Anthony R. (1997) Marking and hierarchy types and the grammaticalization of case-markers *Studies in Language* 21: 313-368

Arka, I Wayan (1993) *Morpholexical aspects of the* -kan *causative in Indonesian* MPhil thesis, University of Sydney

Arka, I Wayan (1998) *From morphosyntax to pragmatics in Balinese* PhD thesis, University of Sydney

Arka, I Wayan & Christopher D. Manning (to appear) Voice and grammatical relations in Indonesian: A new perspective. In Peter K. Austin and Simon Musgrave (ed) *Voice and Grammatical Functions in Austronesian*

Arka, I Wayan & Jane Simpson (1998) Control and complex arguments in Balinese. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG98 Conference* CSLI Publications. Available at: http://csli-publications.stanford.edu/LFG/3/arka-simpson.html

Artawa, Ketut (1999) *Ergativity and Balinese Syntax* (3 parts) NUSA 42

Artawa, Ketut & Barry J. Blake (1997) Patient primacy in Balinese. *Studies in Language* 21: 483–508

Artawa, Ketut, Put Artini & Barry J. Blake (1997) Balinese Grammar and Discourse. Unpublished MS, Udayana University, STKIP, Singaraja and La Trobe University.

Austin, Peter (1997a) Causatives and applicatives in Australian Aboriginal Languages, in Kazuto Matsumura and Tooru Hayasi (eds) *Dative and related phenomena*, pp165-225. Tokyo: Hitsuji Shobo.

Austin, Peter (1997b) Ergativity, clitics and grammatical relations in Sasak. In M.Butt & T.Holloway (ed) *Proceedings of the First Lexical Functional Grammar Conference, Rank Xerox Research Centre, Grenoble* Palo Alto CA: Stanford University

Austin, Peter K. (2000) Verbs, voice and valence in Sasak. In Peter K. Austin (ed) *Working Papers in Sasak, vol.2* pp5-24, Melbourne: Lombok and Sumbawa Research Project, The University of Melbourne

Austin, Peter K. (to appear) Lexical-Functional Grammar. In *International Encyclopedia of Social and Behavioural Sciences* (Vol.3, part 9 article 20) London: Elsevier Scientific

Baker, Mark (1988) *Incorporation* Chicago: University of Chicago Press

Belletti, Adriana & Luigi Rizzi (1988) Psych verbs and θ-theory. *Natural Language and Linguistic Theory* 6:291-352

Benjamins, Geoffrey (1997) Affixes, Austronesian and Iconicity in Malay. *Department of Sociology Working Papers No.132* Singapore: National University of Singapore

Blake, Barry J. (1977) *Case-marking in Australian Languages* Canberra: Australian Institute of Aboriginal Studies

Blake, Barry J. (1988) Tagalog and the Manila-Mt Isa axis *LaTrobe Working Papers in Linguistics* 1: 77-90

Blake, Barry J. (1990) *Relational Grammar* London/New York: Routledge

Blake, Barry J. (1994) *Case* Cambrdige UK: Cambridge University Press

Blake, Frank R. (1906) Expression of case by the verb in Tagalog. *Journal of the American Oriental Society* 27: 183–189

Blevins, J.P. (1995) Syncretism and paradigmatic opposition *Linguistics and Philosohy* 18: 113-152

Bloomfield, Leonard (1917) *Tagalog Texts with Grammatical Analysis* Urbana: The University of Illinois (University of Illinois Studies in Language and Literature Vol. III, part 2-4)

Blume, Kerstin (1998) A contrastive analysis of interaction verbs with dative complements *Linguistics* 36: 253-280

Bresnan, Joan W. (1973) Syntax of the comparative clause construction in English *Linguistic Inquiry* 4:275-343

Bresnan, Joan W. (1982) Control and complementation. *Linguistic Inquiry* 13:343-434

Bresnan, Joan W. (2001a) *Lexical-Functional Syntax* Malden MA/Oxford UK: Blackwells Publishers

Bresnan, Joan W. (2001b) Explaining morphosyntactic competition. In Mark Baltin & Chris Collins (eds) *Handbook of Contemporary Syntactic Theory* pp11-44, Oxford UK: Blackwells Publishers

Bresnan, Joan W. & Jonni Kanerva (1989) Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20: 1-50

Bresnan, Joan W. & Sam A. Mchombo (1987) Topic, pronoun and agreement in Chichewa. *Language* 63: 741-782

Bresnan, Joan W. & Lioba Moshi (1990) Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21: 147-185

Burzio, Luigi (1986) *Italian Syntax: a Government and Binding approach* Dordrecht: Reidel

Capell, Arthur (1964) Verbal systems in Philippine languages *Philippine Journal of Sciences* 93: 231-248

Cardinaletti, Anna & Michal Starke (1996) Deficient pronouns: A view from Germanic. In Höskuldur Thráinsson, Samuel David Epstein & Steve Peter (eds) *Studies in Comparative Germanic Syntax, Vol. 2* pp21-65. Dordrecht: Kluwer (*Studies in Natural Language and Linguistic Theory 38*)

Carter, Richard (1976) Some linking regularities. In Beth Levin & Carol Tenny (ed) *On Linking: Papers by Richard Carter* Cambridge MA: Center for Cognitive Science, MIT (Lexicon Project Working Papers No.25)

Cartier, Alice (1978) On *ke*-verb sentences in Indonesian. In S.A.Wurm & Lois Carrington (ed) *Second International Conference on Austronesian Linguistics Proceedings, Fascicle 1* Canberra: Pacific Linguistics (Series C-61)

Cartier, Alice (1979) De-voiced Transitive Verb Sentences in Indonesian. In Frans Plank (ed.) *Ergativity: Towards a theory of grammatical relations* pp161-183 London: Academic Press

Cena, Resty (1995) Surviving without relations. MS, Edmonton, Alberta.

Chafe, Wallace L. (1970) *Meaning and the Structure of Language* Chicago: University of Chicago Press

Chomsky, Noam (1957) *Syntactic Structures* The Hague: Mouton

Chomsky, Noam (1981) *Lectures on Government and Binding* Dordrecht: Foris Publications

Chomsky, Noam (1991) Some notes on economy of derivation and representation. In R. Friedin (ed) *Principles and Parameters in Comparative Syntax* Cambridge MA: MIT Press (also in Chomsky *The Minimalist Program* MIT Press [1995])

Chomsky, Noam & Howard Lasnik (1995) The Theory of Principles and Parameters. In N.Chomsky *The Minimalist Program* pp13-127, Cambridge MA: MIT Press

Chung, Sandra (1976a) An Object-Creating Rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41-87

Chung, Sandra (1976b) On the subject of two passives in Indonesian. In Charles Li (ed) *Subject and Topic* pp57-98, New York: Academic Press

Chung, Sandra (1978a) Stem sentences in Indonesian. S.Wurm & L.Carrington (ed.) *Second International Conference on Austronesian Linguistics: Proceedings Fascicle 1* (*Pacific Linguistics* C-61) pp335–365 Canberra: Research School of Pacific and Asian Studies

Chung, Sandra (1978b) *Case Marking and Grammatical Relations in Polynesian* Austin: University of Texas Press

Clynes, Adrian (1995) *Topics in the Phonology and Morphosyntax of Balinese* PhD dissertation, Australian National University

Cole, Peter & Gabriella Hermon (2000) Word order and binding in Toba Batak. Paper presented at AFLA 7, Amsterdam.

Cumming, Susanna (1991) *Functional Change: The Case of Malay Constituent Order* Berlin: Mouton de Gruyter

Dahl, Otto Christian (1973) *Proto-Austronesian* Lund: Studentlitteratur (Scandanavian Institute of Asian Studies Monograph 15)

Dardjowidjojo, Soenjono (1978) The semantic structure of the adversative ke--an verbs in Indonesian. In S.Udin (ed) *Spectrum: Essays presented to Sutan Takdir Alisjahbana on his seventieth birthday* pp107-124, Jakarta: Dian Rakyat

Davies, William & Carol Rosen (1988) Unions as multi-predicate clauses *Language* 64: 52-88

De Casparis, J. G. (1956) *Selected inscriptions from the 7th to the 9th century A.D.* Bandung : Masa Baru (Prasasti Indonesia ; no. 2)

De Guzman, Videa P. (1988) Ergative analysis for Philippine languages: an analysis. In R. McGinn (ed) *Studies in Austronesian Linguistics* Athens OH: Center for International Studies (Ohio University Monographs in International Studies, Southeast Asia Series 76)

De Guzman, Videa P. (1997) Some remarks on the Grammatical Function of the 'Non-Subject' Agent in Tagalog. Paper presented at the 8th International Conference on Austronesian Linguistics, Academia Sinica, Taipei.

Di Sciullo, Anne-Marie & Edwin Williams (1987) *On the Definition of Word* Cambridge MA: MIT Press

Dixon, R.M.W. (1977) Where have all the adjectives gone? *Studies in Languge* 1:1-80

Dixon, R.M.W. (1979) Ergativity *Language* 55: 59-138

Dixon, R.M.W. (1994) *Ergativity* Cambridge UK: Cambridge University Press

Donohue, Mark (1999) Syntactic roles vs. semantic roles: external possession in Tukang Besi. In Doris L. Payne & Immanuel Barshi (ed) *External Possession* pp373-401, Amsterdam/Philadelphia: John Benjamins

Dowty, David R. (1991) Thematic proto-roles and argument selection. *Language* 67: 547-619

Dryer, Matthew (1986) Primary objects and secondary objects *Language* 62: 808-845

Durie, Mark (1985) *A Grammar of Acehnese* Foris: Dordrecht

Echols, John M. and Hassan Shadily (1961/1989) *Kamus Indonesia-Inggris* Jakarta: Penerbit PT Gramedia (3rd edition revised by John U. Wolff and James T. Collins in association with Hassan Shadily)

Egerod, Søren (1966) Word order and word classes in Atayal *Language* 42: 346-369

Falk, Yehuda N. (to appear) *Lexical-Functional Grammar: An introduction to parallel, constraint-based syntax* Stanford CA: CSLI Publications

Fillmore, Charles J. (1968) The case for case. In E.Bach & R.T.Harms (ed) *Universals in Linguistic Theory* pp1-88, London: Holt, Rinehart & Winston

Foley, William A. (1991) *The Yimas Language of New Guinea* Stanford CA: Stanford University Press

Foley, William A. (1998) Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd Lexical-Functional Grammar Conference, Brisbane. Available from http://www.sultry.arts.usyd.edu.au/LFG98/austro/workshop.htm.

Foley, William A. and Robert D. Van Valin Jnr (1984) *Functional syntax and universal grammar* Cambridge UK: Cambridge University Press

Givón, Talmy (1983) Topic continuity in discourse: An Introduction. In T.Givón (ed) *Topic Continuity in Discourse: A Quantitative Cross-Language Study* pp5-41, Amsterdam: John Benjamins

Gonda, J. (1949) Prolegomena tot een theorie der woordsoorten in Indonesische talen. *Bijdragen tot de Taal-, Land- en Volkenkunde* 105: 275-331

Grimes, Barbara F. (ed. 1996) *Ethnologue* (13th edition) Dallas: Summer Institute of Linguistics

Grimshaw, Jane (1990) *Argument Structure* Cambridge MA: MIT Press

Grimshaw, Jane (1997) Projection, heads and optimality *Linguistic Inquiry* 28:373-422

Gruber, J.S. (1965) *Studies in Lexical Relations* PhD Thesis, MIT

Guilfoyle, Eithne, Henrietta Hung & Lisa Travis (1992) SPEC of IP and SPEC of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10: 375–414

Haegeman, Liliane (1991) *Introduction to Government & Binding Theory* Oxford: Blackwell Publishers

Hall, Barbara (1965) *Subject and object in modern English* PhD dissertation, MIT

Harris, Z.S. (1965) Transformational theory *Language* 41: 363-401

Hengeveld, Kees (1992) *Non-verbal Predication* Berlin/New York: Mouton de Gruyter

Hilman & Boim (1992) *Lupus. Ih, Syereem!* Jakarta: Gramedia Pustaka Utama

Himmelmann, Nikolaus P. (1991) *The Philippine Challenge to Universal Grammar* Köln: Institut für Sprachwissenschaft (Arbeitspapier No.15, Neue Folge)

Himmelmann, Nikolaus P. (1996) Person marking and grammatical relations in Sulawesi. In H. Steinhauer (ed) *Papers in Austronesian Linguistics No. 3* pp115-136, Canberra: ANU (*Pacific Linguistics* A-84).

Hiorth, Finngeir (1976) Active/passive and non-substantival *ke- -an* forms in Bahasa Indonesia *Review of Indonesian and Malayan Affairs* 8:62-88

Hopper, Paul (1979) Some discourse origins of ergativity, *Hawaii Working Papers in Linguistics* II:137–153.

Hopper, Paul (1988) How ergative is Malay? In R.McGinn (ed) *Studies in Austronesian Linguistics* pp441-454, Athens OH: Ohio University Press

Hopper, Paul & Sandra Thompson (1980) Transitivity in grammar and discourse *Language* 56: 251-299

Iskandar, N.S. (1946) *Djangir Bali* Jakarta: Balai Pustaka

Iskandar, N.S. (1952) *Udjian Masa* Jakarta / Groningen: J.B.Wolters

Jackendoff, Ray (1987) The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18:369-411

Jackendoff, Ray (1990) *Semantic Structures* Cambridge MA: MIT Press

Jaeggli, Osvaldo A. (1986) Passive *Linguistic Inquiry* 17: 587-622

Jelinek, Eloise (1984) Empty categories, case, and configurationality *Natural Language and Linguistic Theory* 2: 39-76

Junus, Umar (1971) *ke- -an* construction in Indonesian *Linguistics* 17: 24-44

Kachru, Yamuna (1990) Experiencer and Other Oblique Subjects in Hindi. In M.K.Verma and K.P.Mohanan (ed)*Experiencer subjects in South Asian languages* pp59-75, Stanford CA: CSLI Publications

Kana, Marit (1986) *Grammatical Relations in Bahasa Indonesia* Cornell University: PhD dissertation

Kaplan, Ronald M. & Joan W. Bresnan (1982) Lexical-Functional Grammar: A formal system for grammatical representation. In Joan w. Bresnan (ed) *The Mental Representation of Grammatical Relations* pp173-281, Cambridge MA: MIT Press

Kaswanti Purwo, Bambang (1984) The categorial system in contemporary Indoensian: pronouns. In John M. Verhaar (ed) *Towards a description of contemporary Indonesaian: preliminary studies Part 2 (NUSA 19)* pp55-74

Kaswanti Purwo, Bambang (1988) Voice in Indonesian: A discourse study. In M.Shibatani (ed) *Passive and Voice* pp195-241, Amsterdam: John Benjamins

Keenan, Edward L. (1976) Towards a universal definition of 'subject'. In Charles Li (ed) *Subject and Topic* pp303-333, New York: Academic Press

Keenan, Edward L. (2000) Morphology is structure: a Malagasy test case. In I. Paul, V. Phillips & L. Travis (ed) *Formal Issues in Austronesian Linguistics* pp27-48, Dordrecht: Kluwer

Keenan, Edward L. & Bernard Comrie (1977) Noun Phrase accessibility and Universal Grammar *Linguistic Inquiry* 8: 63-99

Kiparsky, Paul (1982) From Cyclical Phonology to Lexical Phonology. In Harry van der Hulst and Norval Smith (ed) *The Structure of Phonological Representations (Part 1)* pp131-175, Dordrecht: Foris Publications

Kroeger, Paul R. (1993) *Phrase Structure and Grammatical Relations in Tagalog* Stanford CA: CSLI Publications

Kroeger, Paul R. (1998) Nouns and verbs in Tagalog: A reply to Foley. Paper presented at the 3rd Lexical-Functional Grammar Conference, Brisbane. Available from http://www.sultry.arts.usyd.edu.au/LFG98/austro/workshop.htm.

Kroon, Yosep (1998) The *isiq* construction and its grammatical relations. In P.K. Austin (ed) *Working Papers in Sasak Vol.1* pp105-118, Melbourne: Lombok and Sumbawa Research Project, The University of Melbourne

Larson, Richard K. (1988) On the Double Object construction. *Linguistic Inquiry* 19: 335-391

Legendre, Géraldine, Raymond Wilson and Paul Smolensky (1993) An Optimality-Theoretic typology of case and grammatical voice systems. *Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society* pp464-478

Li, Charles (ed. 1976) *Subject and Topic* New York: Academic Press

Li, Paul Jen-kuei (1973) *Rukai Structure* Taipei: Academia Sinica (Academia Sinica special publications 64)

Lopez, Cecilio (1941) *A manual of the Philippine national language* (3rd ed) Manila: Bureau of Printing

Lyons, John (1977) *Semantics* Cambridge UK: Cambridge University Press

Maclachlan, Anna & Masanori Nakamura (1997) Case-checking and specificity in Tagalog *The Linguistic Review* 14: 307-333

Mahdi, Waruno (1988) *Morphophonologische Besonderheiten und Historische Phonologie des Malagasy* Berlin: Dietrich Reimer Verlag

Mahdi, Waruno (1998) Posting to Bahasa list, 13 February 1998. Available from: http://w3.rz-berlin.mpg.de/~wm/PAP/BHS-L/BHS-adjectives.html

Manning, Christopher D. (1996a) *Ergativity: Argument Strcuture and Grammatical Relations* Stanford, CA: CSLI Publications

Manning, Christopher D. (1996b) Argument structure as the locus for binding theory. Presented at LFG Colloquium, Grenoble

Manning, Christopher D. & Ivan A. Sag (1999) Dissociations between Argument Structure and Grammatical Relations. In Gert Webelhuth, Jean-Pierre Koenig and Andreas Kathol (eds.), *Lexical And Constructional Aspects of Linguistic Explanation*, CSLI Publications, pp. 63-78

Martin, J.R. (1996) Transitivity in Tagalog: A Functional Interpretation of Case. In Berry, Butler, Fawcett and Huang (ed.) *Form and Meaning* pp229-296, Norwood NJ: Ablex Publishing Corporation

McCrae, Ken, Todd R. Ferretti & Liane Amyote (1997) Thematic roles as verb-specific concepts *Language and Cognitive Processes* 12: 137-176

McCune, Keith (1979) Passive Function and the Indonesian Passive. *Oceanic Linguistics* 18:119-169

McDonald, Ross D. & Soenjono Darjowidjojo (1967) *Indonesian Reference Grammar* Washington DC: Georgetown University Press

Mel'cuk, Igor (1983) Grammatical subject and the problem of the ergative constructions in Lezgian *Folio Slavica* 5: 246-293

Miller, P. (1992) Postlexical cliticization v. affixation: co-ordination criteria. *Papers from the 30th Chicago Linguistic Society* 30: 382-396.

Mithun, Marianne (1984) The Evolution of Noun Incorporation. *Language* 60:847-894

Moeliono, Anton M. (1994) Contact-induced language change in present-day Indonesian. In Tom Dutton & Darrell T.Tryon (ed) *Language Contact and Change in the Austronesian World* pp377-388, Berlin/New York: Mouton de Gruyter

Moeliono, Anton M. & S. Dardjowidjojo (ed. 1988) *Tata Bahasa Baku Bahasa Indonesia* (Standard Grammar of Bahasa Indonesia) Jakarta: Balai Pustaka & Departemen Pendidikan dan Kebudayaan

Moro, Andrea (1997) *The raising of predicates* Cambridge UK: Cambridge University Press

Musgrave, Simon (2000) A Note on Animacy Hierarchy Effects in Sasak and Sumbawan. In P. Austin (ed) *Working Papers in Sasak* Volume 2 pp49-56

Myhill, John (1988) Nominal agent incorporation in Indonesian. *Journal of Linguistics* 24:111-136

Nakamura, Masanori (1998) Reference Set, Minimal Link Condition, and Parametrization. In P.Barbosa, D.Fox, P.Hagstrom, M.McGinnis and D.Pesetsky (eds) *Is the best good enough? Optimality and Competition in Syntax* Cambridge MA: MIT Press

Naylor, Paz B. (1980) Linking, Relation Marking, and Tagalog Syntax. In P.B.Naylor (ed) *Austronesian Syntax* (Papers from the 2nd Eastern Conference on Austronesian Languages, Ann Arbor 1976) Ann Arbor: Center for South and Southeast Asian Studies (Michigan Papers on South and Southeast Asia, No.15)

Naylor, Paz B. (1995) Subject, Topic and Tagalog Syntax. In David C. Bennett, Theodora Bynon and G. George Hewitt (eds) *Subject, Voice and Ergativity: Selected Essays* pp161-201, London: School of Oriental and African Studies

Payne, Doris L. & Immanuel Barshi (1999) External possession: what, where, how and why. In Doris L. Payne & Immanuel Barshi (ed) *External Possession* pp3-29, Amsterdam/Philadelphia: John Benjamins

Payne, Thomas E. (1982) Role and Reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog *Studies in Language* 6: 75-100

Perlmutter, David & Paul Postal (1983) Some basic laws of clause structure. In David Perlmutter (ed) *Studies in Relational Grammar I* pp81-128, Chicago: University of Chicago Press

Pesetsky, David (1995) *Zero Syntax* Cambridge MA: MIT Press

Plank, Frans (ed. 1984) *Objects: towards a theory of grammatical relations* New York: Academic Press

Poedjosoedarmo, Gloria (1986) Subject selection and subject shifting in Indonesian. In Dardjowidjojo (ed) *Miscellaneous Studies of Indonesian and Other Languages of Indonesia, Part VIII (NUSA 26)* pp1-17

Pollock, Jean-Yves (1989) Verb movement, Universal Grammar and the structure of IP *Linguistic Inquiry* 20: 365-424

Poser, William J. (1992) Blocking of phrasal constructions by lexical items. In Ivan A. Sag and Anna Szabolsci (ed) *Lexical Matters* pp111-130, Stanford CA: CSLI Publications

Postal, Paul M. (1969) On So-Called Pronouns in English, In D.Reibel & S.Schane (ed.) *Modern Studies in English* Eaglewood Cliffs NJ: Prentice-Hall pp201–224.

Postal, Paul M. (1977) Antipassive in French *Lingvisticae Investigationaes* 1: 333-374

Prince, Alan & Paul Smolensky (1993) *Optimality Theory: Constraint interaction in generative grammar* RuCCs Technical Report #2. Rutgers University Center for Cognitive Science

Pustet, Regina (2000) How arbitrary is lexical categorization? Verbs v. adjectives *Linguistic Typology* 4:175-212

Rosen, Carol G. (1984) The interface between semantic roles and initial grammatical relations. In David M. Perlmutter & Carol G. Rosen (ed) *Studies in Relational Grammar 2* pp38-77. Chicago: The University of Chicago Press

Sadler, Louisa (1998) Syntactic clitics, the structure-function mapping and morphological blocking. MS, University of Essex

Schachter, Paul (1976) The subject in Philippine languages: Actor, Tpoic, Actor-Topic, or none of the above?. In Charles Li (ed.) *Subject and Topic* pp493–518. New York: Academic Press

Schachter, Paul (1995) The subject in Tagalog: still none of the above. *UCLA Occasional Papers in Linguistics No. 15* Los Angeles: Department of Linguistics, UCLA

Schachter, Paul & Fe T. Otanes (1972) *Tagalog Reference Grammar* Berkeley: University of California Press

Seiter, William J. (1978) On the syntactic character of middle objects in Polynesian. In S.A.Wurm and Lois Carrington (ed) *Second International Conference on Austronesian Linguistics: Proceedings, Fascicle 2* pp 1289-1306, Canberra: Australian National University (Pacific Linguistics C-61)

Sells, Peter (1998) The functions of voice markers in the Philippine languages. In S.G.Lapointe, D.K Brentari & P.M. Farrell (ed) *Morphology and Its Relation to Phonology and Syntax* pp111-137, Stanford CA: CSLI Publications

Sells, Peter (to appear) Form and function in the typology of grammatical voice systems. In J.Grimshaw, G. Legendre & S.Vikner (ed) *Optimality-theoretic Syntax* Cambridge MA: MIT Press

Seuren, Pieter A.M. (1998) *Western Linguistics: An Historical Introduction* Oxford: Blackwell Publishers

Shibatani, Masayoshi (1985) Passives and related constructions: A prototype analysis. *Language* 61:821-848

Shibatani, Masayoshi (1994) An integrational approach to Possessor Raising, Ethical Datives, and Adversative Passives *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 20: 461-486

Silverstein, Michael (1976) Hierarchy of features and ergativity. In R.M.W. Dixon (ed) *Grammatical Categories in Australian Languages* pp112-171, Canberra: Australian Institute of Aboriginal Studies

Simatupang, Iwan (1993) *Merah merahnya* Jakarta: Haji Masagung

Smith, Henry (1996) *Restrictiveness in Case Theory* Cambridge UK: Cambridge University Press

Sneddon, James Neil (1996) *Indonesian Reference Grammar* St Leonards NSW: Allen & Unwin

Sneddon, James Neil (2000) Another look at *ke- -an* verbs. In Bambang Kaswanti Purwo (ed) *Kajian Serba Linguistik untuk Anton Moeliono: Pereksa Bahasa* pp511- 529, Jakarta/Atma Jaya: Universitas Katolik Indonesia and PT BPK Gunung Media

Soemarmo (1970) *Subject-Predicate, Focus-Presupposition and Topic-Comment in Bahasa Indonesia and Javanese* PhD dissertation, University of California, Los Angeles

Sproat, Richard (1985) Welsh syntax and VSO structure *Natural Language and Linguistic Theory* 2:173-216

Starosta, Stanley, Andrew K. Pawley & Lawrence A. Reid (1982) The evolution of focus in Austronesian. In Amran Halim, Lois Carrington & S.A.Wurm (ed) *Papers from the Third International Conference on Austronesian Linguistics* vol. 2 (*Pacific Linguistics* C-75*)* pp145-170, Canberra: Research School of Asian and Pacific Studies

Steinhauer, Hein (1994) The Indonesian language situation and linguistics *Bijdragen tot de Taal-, Land- en Volkenkunde* 150: 755-784

Stevens, Alan M. (1970) Pseudo-transitive verbs in Indonesian. *Indonesia* 9:67-72

Stowell, Tim (1983) Subjects across categories *The Linguistic Review* 2: 285-312

Sukamto, Katarina (1999) The accessibility of inferrables in Indonesian. Austronesian Informal Seminar Series, The University of Melbourne, 30 April 1999.

Tampubolon, Daulat P. (1983) *Verbal Affixation in Indonesian, a semantic exploration* (*Pacific Linguistics* D-48) Canberra: Research School of Pacific and Asian Studies

Tanner, N. (1967) Speech and society among the Indonesian élite: a case study of a multilingual community *Anthropological Linguistics* 9: 15-40

Teeuw, A. (1962) Some problems in the study of word-classes in Bahasa Indonesia *Lingua* 9: 409-421

Teeuw, A. (1977) The morphological system of the Indonesian adjective. In Amran Halim (ed) *Miscellaneous Studies in Indonesian and Languages in Indonesia Part II (NUSA 3)* pp1-18

Tenny, Carol (1994) *Aspectual Roles and the Syntax-Semantics Interface* Dordrecht: Kluwer

Thomas, Michael R. (1978) Indonesian's unmarked verbs. In John M. Verhaar (ed) *Miscellaneous Studies in Indonesian and Languages in Indonesia Part V (NUSA 6)*, pp 7-10

Thompson, Sandra (1988) A discourse approach to the cross-linguistic category 'adjective'. In John A. Hawkins (ed) *Explaining Language Universals* pp167-185, Oxford: Blackwells

Tryon, Darrell (1995) Proto-Austronesian and the major Austronesian sub-groups. In Peter Bellwood, James J. Fox and Darrell Tryon (ed) *The Austronesians: Historical and Comparative Perspectives* Canberra: The Australian Nationalk University, pp17-38

van den Berg, René (1996) The demise of focus and the spread of conjugated verbs. In H. Steinhauer (ed) *Papers in Austronesian Linguistics No. 3* pp89-114, Canberra: ANU (*Pacific Linguistics* A-84).

van der Tuuk, H.N. (1971) *A Grammar of Toba Batak* The Hague: Martinus Nijhoff (Translation by J. Scott-Kemball of *Tobaasche Spraakkunst* Amsterdam 1864 and 1867)

Van Valin, Robert D. Jr. & Randy LaPolla (1997) *Syntax: Structure, meaning and function* Cambridge UK: Cambridge University Press

Verhaar, John W.M. (1984a) Affixation in Contemporary Indonesian. In Bambang Kaswanti Purwo (ed.) *Towards a Description of Contemporary Indonesian: Preliminary Studies Part 1 (NUSA* 18) pp1–26

Verhaar, John W.M. (1984b) The Categorial System in Contemporary Indonesian: Verbs. In Bambang Kaswanti Purwo (ed.) *Towards a Description of Contemporary Indonesian: Preliminary Studies Part 1 (NUSA* 18) pp27–63

Voskuil, Jan (1996) *Comparative Morphology: Verb Taxonomy in Indonesian, Tagalog and Dutch* The Hague: Holland Academic Graphics (HIL Dissertations No 21).

Wechsler, Stephen, and I Wayan Arka (1998) Syntactic Ergativity in Balinese: an Argument Structure Based Theory. *Natural Language and Linguistic Theory* 16:387-441

Wierzbicka, Anna (1995) Adjectives vs. verbs: the iconicity of part-of-speech membership. In M. E. Landsberg (ed) *Syntactic iconicity and linguistic freezes*, pp223-245. Berlin/New York: Mouton de Gruyter.

Winstedt, R.O. (1913) *Malay Grammar* Oxford: Clarendon Press

Zaenen, Annie, Joan Maling & Hoskildur Thráinsson (1985) Case and grammatical function: the Icelandic passive. *Natural Language and Linguistics Theory* 3: 441-483

Zoetmulder, P.J. (1982) *Old Javanese - English Dictionary* s'Gravenhage: M.Nijhof

Author/s:
Musgrave, Simon

Title:
Non-subject arguments in Indonesian

Date:
2001-10

Citation:
Musgrave, S. (2001). Non-subject arguments in Indonesian. PhD thesis, Department of Linguistics and Applied Linguistics, University of Melbourne.

Publication Status:
Unpublished

Persistent Link:
http://hdl.handle.net/11343/38779