Verb Classification in Australian Languages

Empirical Approaches to Language Typology

25

Editors

Georg Bossong
Bernard Comrie

Mouton de Gruyter
Berlin · New York

Verb Classification in Australian Languages

by

William B. McGregor

Mouton de Gruyter
Berlin · New York 2002

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

Library of Congress – Cataloging-in-Publication-Data

McGregor, William.
Verb classification in Australian languages / by William B. McGregor.
p. cm. – (Empirical approaches to language typology ; 25)
Includes bibliographical references and index.
ISBN 3 11 017141 4 (cloth : alk. paper)
1. Australian languages – Verb phrase. 2. Australian languages – Semantics. 3. Grammar, Comparative and general – Syntax. I. Title. II. Series. III. Series: Empirical approaches to language typology ; 25
PL7003 .M34 2002
400'.15–dc21 2001059196

Die Deutsche Bibliothek – Cataloging-in-Publication-Data

MacGregor, William :
Verb classification in Australian languages / by William B. McGregor. – Berlin ; New York : Mouton de Gruyter, 2002
(Empirical approaches to language typology ; 25)
ISBN 3-11-017141-4

Preface

The roots of this book go back a long way, perhaps to the early 1980s when I began working on Gooniyandi, and was confronted with the task of describing its unusual verbal construction. It seemed clear that a set of morphemes—that apparently derived historically from verbs—served as verbal classifiers, in basically the same way as nominal classifiers in standard systems of nominal classification. This observation sowed the seeds for a fascination in the verbal constructions in the languages of northern Australia, leading me to wonder about the extent to which verbal classification was exhibited in them, and to an interest in the typology and semantics of the phenomenon. It also led me to puzzle about how the phenomenon might be analysed grammatically, and whether there are psycholinguistic implications—does verb classification imply something about the perception of events or their representation in the minds of speakers?

By the mid 1990s a sufficient number of descriptions of northern Australian languages had appeared that were rich enough in detail to permit a preliminary typological investigation of the phenomenon. A couple of linguists were also beginning doctoral research that targeted this aspect of the verb in particular languages, Edith Nicolas (Bardi and Bunuba), and Eva Schultze-Berndt (Jaminjung). This encouraged me to begin to think seriously about verb classification in the wider perspective, and to commence writing the present work.

My first attempts in a typological investigation were in a presentation to the Association for Linguistic Typology Conference II, held in September 1997 at the University of Oregon. That conference also brought me an awareness of similar phenomena in other languages, specifically instrumental prefixes to verbs in some Amerindian languages. At about the same time I was fortunate enough to be supervising Xiaokang Zhou's doctoral research on Mandarin Chinese, which alerted me to the existence of a rather different type of verbal classification. The horizons were further widened in 1999 at the Association for Linguistic Typology Conference III in Amsterdam, where I con-

ducted a workshop on verb classification. This workshop attracted presentations on a variety of languages, Australian, Sinitic, South American, and Chukotko-Kamchatkan.

Many of the ideas elaborated in this book were previously presented in departmental seminars at University of Melbourne, Amsterdam University, and Max-Planck Institute for Psycholinguistics, as well as at workshops and conferences, including (in addition to the above mentioned pair): Linguistics Association of Great Britain Spring Meeting (Lancaster University, 1998), Third International Workshop on Australian Aboriginal languages (Max Planck Institute for Psycholinguistics, 1998), and Ideophones Conference (University of Cologne, 1999). I am grateful to the audiences of all of these occasions for their questions, constructive comments, and criticisms.

In the Spring Semester 2000 Jan Rijkhoff and myself ran a course entitled *Nominal and verb classification* at the Department of Linguistics, Aarhus Universitet. I learnt a good deal about nominal classification and the noun phrase from Jan, and benefited considerably from presenting my ideas about verb classification in this course.

I owe a considerable debt of gratitude to the many Australianist linguists who provided me data, answers to questions, access to unpublished manuscripts, and/or comments on draft sections of the book dealing with languages of their expertise, including Barry Alpher, Peter Austin, Barry Blake, Peter Carroll, Howard Coate, Nick Evans, Ian Green, Komei Hosokawa, Joyce Hudson, Emily Knight, Frances Kofod, Patrick McConvell, Graham McKay, David Nash, Edith Nicolas, Rachel Nordlinger, Rob Pensalfini, Nick Reid, Alan Rumsey, Thomas Saunders, Eva Schultz-Berndt, Janet Sharp, Margaret Sharpe, Bronwyn Stokes, Tsunoda Taskau, David Wilkins, Melanie Wilkinson, and Stephen Wilson. I benefited greatly from extended discussions on verb classification with Eva Schultz-Berndt, as well as from her excellent doctoral thesis on verbal constructions in Jaminjung (Schultze-Berndt 2000)—though we steadfastly maintain divergent views on the topic. Acknowledgement is also made to Hilary Chappell, Kari Fraurud, Steve Levinson, Stephen Matthews, David Wilkins, and Xiaokang Zhou for discussion of various issues relating to the topic of verbal and nominal classification. I am especially thankful to Jean-

Christophe Verstraete for reading and commenting on the entire manuscript. Needless to say, I alone am responsible for any errors of fact or interpretation.

My own fieldwork, which represents the ultimate data-base for this investigation, was funded by the Australian Institute of Aboriginal and Torres Strait Islander Studies, the National Aboriginal Languages Program, and the Australian Research Council (Grants A58930745 and A9324000). Many of my ideas about verb classification were formulated during my tenure of an Australian Research Council Research Fellowship (A59332055) held at the University of Melbourne, 1993-1997, and subsequently refined and revised under a Research Fellowship at the Catholic University of Leuven, during 1998, and a Guestship at the Max Planck Institute for Psycholinguistics, Nijmegen in 1999–2000. I am grateful to all of these institutions for their financial support, and to the Max Planck Institute for Psycholinguistics for intellectual stimulation.

Ultimately the book would not have been possible were it not for the input from various Aboriginal people who I and other Australianist linguists worked with. I am personally grateful to †Jack Bohemia, †Carmel Charles, Reni Chestnut, †Dave Lamey, Suzie Lamey, Maudie Lennard, †Freddy Marker, †Bill Munro, David Street, Mervin Street, and †Magdalene Williams for their contributions to my language learning since 1980.

The editor of this series, Georg Bossong, provided useful advice and comments on the penultimate draft; my editor at Mouton de Gruyter, Ursula Kleinhenz, has consistently and cheerfully answered my incessant formatting questions.

W.B.McG.

Århus
December 2001

Contents

List of Figures

List of maps

List of Tables

Abbreviations and Conventions

ABL	ablative
ABS	absolutive
AC	active
ACC	accusative
ACP	accomplishment
ADV	adversative
AKT	Aktionsart
ALL	allative
ALSO	also, including
A-mark	argument-mark
ANIM	animate
AO	agent-oriented verb
ARG	argument
ASP	unmarked for aspect
AT	spatial function 'be located at'
aug	augmented
AUX	auxiliary
CARD	cardinal (pronoun)
CHAR	characteristic nominaliser
CL	classifier
COMIT	comitative
COMP	complementiser
CONJ	conjunction
CONT	continuous aspect
COTEMP	contemporary
CVC	compound verb construction
DAT	dative
DEF	definite
DEM	demonstrative
DS	different subject
du	dual
DUBIT	dubitive
EMP	emphatic
EN	epenthetic nasal
ERG	ergative
ETC	etcetera
exc	exclusive
EXIST	existent
EXT	extendible
FACT	factive mood
FEM	feminine
FOC	focus
FUNC	function
FUT	future
GEN	genitive
Ident	identificational
IMP	imperative
IMPFV	imperfective
IN	incorporated noun
inc	inclusive
INCH	inchoative
INF	infinitive
INST	instrumental
INTRANS	intransitive
IO	indirect object
IRR	irrealis
IT	iterative
IV	inflecting verb
IVR	inflecting verb root
LM	landmark
LOC	locative

LOQ	delocutive
LP	lexical prefix
MAS	masculine
MD	middle
min	minimal
Mp	manipulative prefix
N	nominal
NI	noun incorporation
NO	number
NOM	nominative
NP	nominal phrase
N/P	neuter plural
npc	non-past completive
npp	non-past progressive
OBL	oblique
ORIG	originative
PA	past
PART	participle
pau	paucal
pc	past completive
PER	perlative
PF	perfective
PHAB	past habitual
pl	plural
PO	patient-oriented verb
POT	potential
PRES	present
PRIV	privative
PRO	pronominal
PROG	progressive
PROX	proximal
PUNCT	punctual
PURP	purposive
PV	pivot form/nominal
REF	reflexive
REM	remote
REP	repetition, repeated
RPC	remote past continuous
SEQ	sequential
SER	serial verb construction
SFOC	sentence focus
SG	singular
SoA	state of affairs
SS	same subject
STAT	stative
SUBJ	subjunctive
SVC	simple verb construction
TNS	tense
TOP	topic
TR	transitive
TJ	trajector
USIT	usitative
UV	uninflecting verb
V	verb
VC	verbal construction
VOC	vocative
VP	verb phrase
1	first person
2	second person
3	third person
ø	zero morpheme
~	alternating allomorphs
-	morpheme boundary
+, ‖	tight morpheme boundary in Gooniyandi verbs
'	morpheme boundary in Bunuba verbs; also indicates a glottal

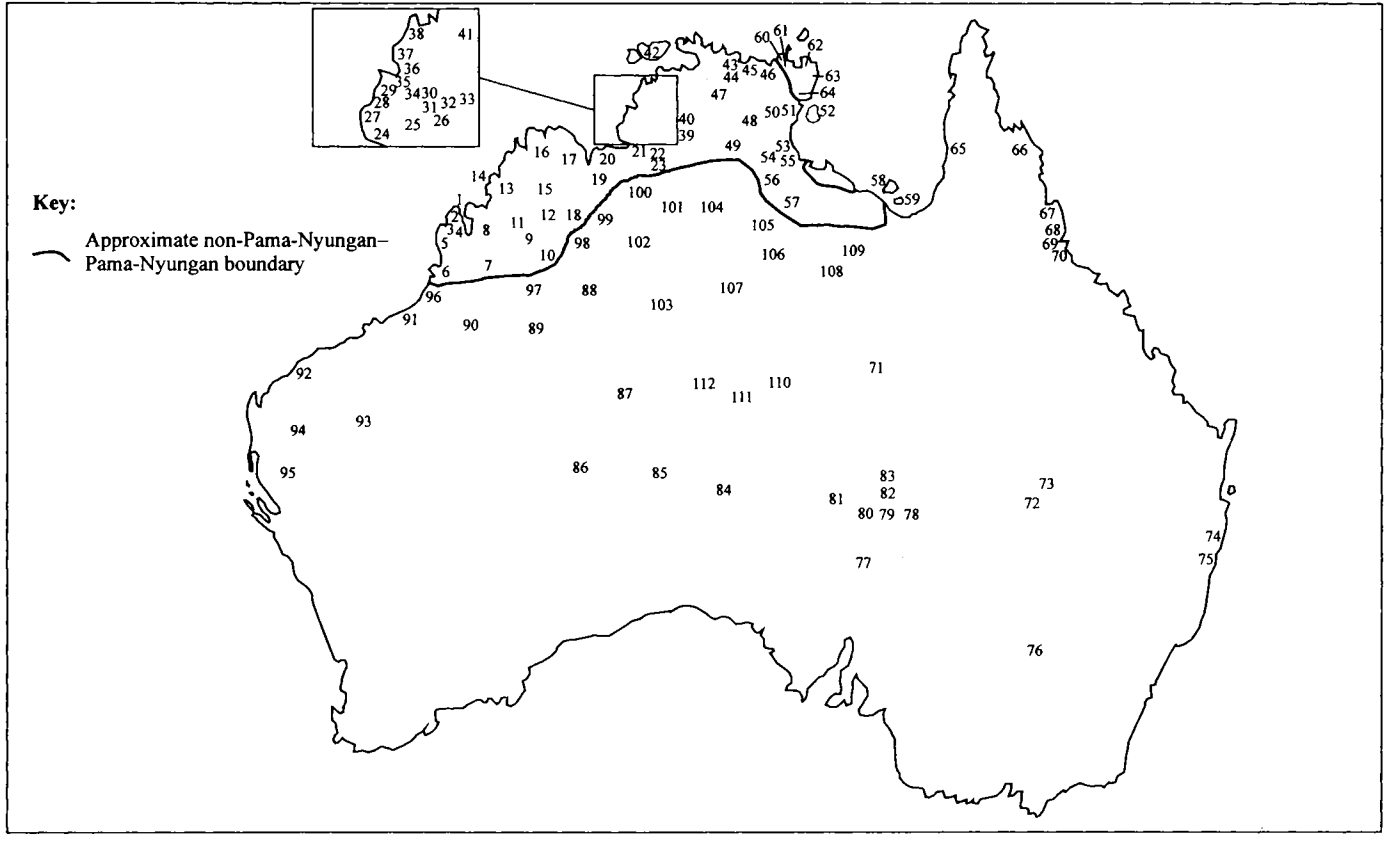
	stop, and sometimes distinguishes /n/-/g/ from /ŋ/ and /ŋg/	;	glosses pause within intonation unit
|	collocates with	→	changes to (via phonological rule); acts on
/	conflated with		
<	derives from		
()	optional elements	↔	alternates with
,	end of intonation unit	*	ungrammatical or unacceptable form or reconstructed proto-form
:	vowel length in cited words; also separates words in complex		

The root forms of inflecting verbs are generally cited in small capitals; this is done consistently in languages that distinguish two classes of verbs, inflecting and non-inflecting, the latter being cited in italics. Verbal classifying morphemes in Gooniyandi are also cited in small capitals, preceded by the + sign.

Labels for grammatical roles are given with initial capital letters.

As far as possible accepted orthographies have employed in transcribing words in Australian Aboriginal languages, rather than adopting a single orthographic system. However, examples taken from Nekes and Worms (1953) are represented as in the source, in IPA font, in the first line; this is followed in the second line by a retranscription in the accepted or recommended orthography for the language. Appendix 2 provides basic information on the phonologies of Australian languages, and describes the main orthographic conventions employed in writing the languages.

Capitalised DH, NH, and NY indicate proto-phonemes or archiphonemes that account for the following alternations, respectively: *th* ~ *j*, *nh* ~ *ny*, and *ny* ~ *ny* or *n*.



Map 1. Approximate locations of traditional Aboriginal languages referred to

Key to language names

Chapter 1
Introduction

1.1. What is verb classification?

Noun classification is a well known and popular topic, that holds considerable fascination to linguists. Recent literature exhibits a variety of approaches and perspectives, including typological, descriptive, semantic, grammaticalisation, discourse, psycholinguistic, sociolinguistic, and philosophical—some examples are Dixon (1968); Allan (1977); Leeding (1989); Nichols (1989); Seiler (1989); Corbett (1991); Sands (1995); Downing (1996); van Berkum (1996); Aikhenvald (2000); and articles in Craig (1986b); Harvey and Reid (1997); Unterbeck, Rissanen, Nevalainen, and Saari (1999); and Senft (2000). We now know that noun classification is widespread throughout the languages of the world, and comes in a variety of forms and types. It is becoming increasingly apparent that most systems are to some extent semantically motivated (e.g. Zubin and Köpcke 1986; Corbett 1991; Corbett and Fraser 2000; Dahl 1999; and Fraurud 1999). Thus, one of the fascinations of noun classification is that it might provide insight into the "mind" of speakers of a language: into how they perceive, understand, conceptualise, and categorise the world (e.g. Allan 1977: 308; Lakoff 1986).

What all systems of noun classification have in common is that they are language-internal systems that overtly typologise nominal words, phrases, and/or their conceptual referents into classes or categories. That is to say, the language possesses a system of lexical or grammatical devices that group nominals, and/or their referents, into categories.

There is nothing special about nouns that make them the only part-of-speech that can be classified grammatically. Though not widely known, comparable systems exist for the other major part-of-speech, verbs. A not inconsiderable number of languages show systems in which verbs—and/or the events they refer to—are overtly cate-

gorised into types by grammatical means. The term *verb classification* is employed in this book specifically in reference to such overt systems of grammatical categorisation of verbs. The restriction to overt systems is motivated primarily by the need to delimit a manageable and coherent domain of investigation, and provides a principled and linguistically significant means for doing so: it is grammatically founded, with potential cognitive repercussions. A full account would, of course, include both overt and covert systems.

This sense of *verb classification* is not well established in linguistics, and is found in neither theoretical linguistics nor linguistic typology; just a few descriptive traditions invoke it. Investigations of northern Australian languages have for some years employed the term (or something like it) in this way; see, for instance, Capell (1979); Silverstein (1986); Green (1989); McGregor (1990a); Reid (1990); Nicolas (1998, 2000); Schultze-Berndt (2000); and Knight (2001).[1] It is also used in similar ways, albeit fairly marginally, in the Sinitic (e.g. Killingley 1983; Matthews and Yip 1994, 1999) and Austronesian literature (Nojima 1996).

To be sure, labels such as *verb classification*, *verb classifiers*, *classificatory verbs*, *verb classes* and the like are quite widely used in linguistics. But they are used rather differently, generally in reference to covert systems of classification. Thus, many writers use terms such as these to refer to lexical subcategorisation, the classification of verbs into categories such as transitive vs. intransitive, and telic vs. atelic—for example, Fillmore (1968); Taylor (1971); Longacre (1976); Clark (1978); Chen (1987); Levin (1993); Van Valin (1993); Radford (1997). The main exceptions, where overt systems of verb classification are referred to, are conjugation classes—usually considered to be relevant only to minutiae of verbal morphology, and of marginal interest in the overall scheme of grammar.

In a few traditions the terms refer to overt systems of classification that are marked on or by verbs, but actually classify nouns in certain grammatical roles, usually subject or object. This is what is usually meant by the term *classificatory verb* in Amerindianist linguistics, for instance in Athapaskan studies, where it refers to a small set of verbs of handling that classify objects by shape (Hoijer 1945; Carter 1976;

Young and Morgan 1980; Silver and Miller 1997: 32–34). Blankenship (1997) uses the same term for a similar phenomenon in Cherokee. Likewise, in Papuan linguistics the term often refers to existential verbs that classify nouns according to their existential status, the positional state they prototypically occupy—e.g. Lang (1975); Seiler (1985); Merlan, Roberts, and Rumsey (1997); Levinson (1999).

This chapter sets the scene for the remainder of the book. Section 1.2 begins by outlining the range of variation within systems of overt grammatical classification, identifying the major parameters of variation. But in a way, this is putting the cart before the horse: it is not yet clear what grammatical classification actually is—how one would recognise a system of grammatical classification. To this end, in §1.3 I first propose conceptual models, intended to elucidate significant features of classification systems that will be encountered in later chapters. I then propose a formal characterisation of overt grammatical classification according to distributional criteria.

Of course, it would be nonsensical to claim that the proposed formal characterisation is right, or the only one possible. Two observations count in its favour. First, it correlates with everyday understandings of the notion of classification and so is not outlandish. Second, it includes most grammatical phenomena that are generally regarded as systems of nominal and verbal classification, and excludes many non-contenders. It includes other things as well, though. Most, if not all of these, I submit, can and should be excluded on functional-semiotic grounds, rather than by fiat. My motivation for taking this approach is that formal-distributional systems provide the substance for functional-semiotic systems; they constitute necessary (though not sufficient) foundations on which functional-semiotic systems may be constructed.

The remainder of the chapter focuses on verb classification in Australian languages, providing an overview of the phenomenon that highlights the main points developed in subsequent chapters. Section 1.4 identifies the main formal features of the systems, and maps the distribution of the major types across the continent. Section 1.5 discusses the semantic bases for verb classification systems in Australia, and makes some general comments on the nature of the semantics of classification systems. Finally, §1.6 outlines the aims and organisation of the book.

1.2. Towards a typology of classification

Three parameters relevant to the typology of nominal classification systems also apply to systems of verb classification. We discuss these in the following subsections.

1.2.1. Superclassification and subclassification

The first parameter concerns the "direction" in which classification takes place: whether all (classified or classifiable) items are assigned to one of a number of universal types for the language, or the items are assigned to local types, peculiar to each lexeme. I refer to these as, respectively, *superclassifying* and *subclassifying* systems, following McGregor (1997b: 180). In nominal superclassifying systems one takes a nominal (N) and assigns it (or its referents) to one of a relatively small and fixed number of types or classes; in nominal subclassifying systems, by contrast, the N (or the set of its conceptual referents) is divided into a number of types, generally different for each N. Yidiny (Queensland) shows a superclassifying system: about a score of types are distinguished, to which any nominal can be assigned according to criteria such as inherent characteristics and uses (Dixon 1977). On the other hand, English has a system of nominal subclassification, in constructions such as *steam train*, *electric train*, *diesel train*, *passenger train*, *goods train* etc.; *tank engine*, *diesel engine*, *petrol engine*, *two-stroke engine* etc.; *little finger*, *middle finger*, *index finger*, etc.. Here subtypes of the N are distinguished—the N is not assigned to a generic type (such as vehicle, long thin object, or whatever).[2] A similar system is found in Gooniyandi (Western Australia) (McGregor 1990a: 260–264), and various other Australian languages (e.g. Martuthunira (Western Australia) (Dench 1995), and Kayardild (Queensland) (Evans 1995a); see also Harvey 1992).

McGregor (1997a) proposes that certain types of noun incorporation—specifically, Mithun's Type I *lexical compounding* and a subset of Type II *manipulation of case* (Mithun 1984)—involve verb subclassification (see also Craig 1994: 566). Warray (Northern Territory) has

this type of noun incorporation. Consider *-mirral-lagi-* (sun-put) 'put in the sun (to dry)', and *-wik-lagi-* (water-put) 'soak' (Harvey 1995: 144), which refer to different culturally relevant types of putting; putting actions are categorised into subtypes according to the resulting location of the moved entity. Clearly the action of putting is not itself assigned to a particular category or type of event. Certain noun-verb compounds in English (e.g. *hand-pick*, *pistol-whip*, *horse-whip*, *test-drive*, etc.) also represent a type of verbal subclassification: they specify subtypes of the event denoted by the verb.

Gooniyandi, by contrast, shows a system of verb superclassification in which every verb is assigned to one or more of a dozen types (McGregor 1990a: 557–572, and Chapter 2 below). This system works in a very different way to the Warray subclassifying system. Whereas in Warray an incorporated noun serves to distinguish a particular subtype of the action referred to by the verb, in Gooniyandi the verb classifier indicates to which category the event belongs. Thus, the Gooniyandi verb *jaa* 'soak' is classified as an action of putting or throwing (McGregor 1990a: 564).[3] This works in the opposite direction to Warray *-wik-lagi-* (water-put), where *-wik* specifies the type of putting, namely in water—i.e. soaking. While 'put' is involved in both cases, it is involved in an abstract, schematic way in Gooniyandi, in identifying categories of events that show abstract features suggestive of the gloss 'put'. However, not all of these events are describable as instances of putting (*yood-*). In Warray, the action of putting is involved in a concrete referential way, and *-lagi-* 'put' would apply to (and be implied by) any verbal construction involving this root and an incorporated N. In other words, 'put' is implied by the noun incorporation construction in Warray, but not by the Gooniyandi verbal construction.

1.2.2. Classes and categories

Nominal superclassification systems are usually divided into *noun class* or *gender* systems and *noun classifier* systems. For reasons that will become apparent in the next subsection, I refer to systems of the latter type as *category* systems, rather than classifier systems. The units

in these systems are accordingly referred to as classes and categories. Scholars disagree on how the two systems are best distinguished. But one widely accepted scheme is Dixon's proposal that a cluster of co-varying morpho-syntactic and other criteria are involved:

> ... noun classes constitute a closed grammatical system, with a finite (usually fairly small) and determinable number of choices involved: each noun belongs to one class and few (or none) will correspond to more than one class. Noun class is shown by morphological processes which can apply to the noun itself and must apply to some other constituent; they sometimes combine information about noun class and number, case or definiteness. The morphological process most often involves adding an affix or clitic to constituents of a number of specified grammatical types. Classifiers [i.e. category markers — WMcG] comprise a largish (often semi-open) set, whose members may not be exhaustively listable; each classifier is either a free form or else a root, to which a numeral affix or clitic may be added. Not every noun may take a classifier; many nouns occur with one of a number of different classifiers, sometimes with a difference in meaning and sometimes not. Classifiers, but not noun classes, may be used in different ways in different speech styles, within a language. (Dixon 1982: 217–218)

The contrast as set up here is peculiar to nouns. Can it be generalised to verbs, or other parts of speech? The answer is clearly yes (see also Silverstein 1986: 510), by relaxing the requirement that the morphological marking in class systems be located on something other than the classified entity itself (see also Corbett 1991). The distinguishing features are thus as shown in Table 1.

This characterisation is not unproblematic.[4] Cases exist where it is difficult to categorise a system as noun class or noun category due to misalignments among the parameters (see also Craig 1986a: 4). For instance, Ngan'gityemerri shows a system of about sixteen noun classes marked by a mixture of free lexical forms, proclitics and enclitics (Reid 1997). Nearby Murrinh-Patha also has a noun classification system with features of both, and seems to be an intermediate case (Walsh 1997: 278). So also does the Gooniyandi system of verb classification (see Chapter 2). Nevertheless, the features tend to co-vary.

Various types of nominal category systems can be distinguished, including (to use well established labels): numeral classifiers (typical of

Table 1. Major distinguishing characteristics of class vs. category (i.e. "classifier") systems, according to Dixon (1982, 1986)

Class systems	Category systems
Bound morphosyntactic mark, not necessarily centred on lexeme	Lexical marker, always within the phrase
Smallish number	Largish number
Closed system	Open (relatively) system
Exhaustive—applies to all words of a certain part-of-speech	Not necessarily exhaustive
Largely disjoint	More significant intersection
No (or little) registerial variation	Possible differences in categorisation according to register
Obligatory use	Discourse sensitive use

languages of South-East Asia), noun classifiers (many Australian languages), possessive classifiers (Micronesian languages), and verb classifiers (Athapaskan—see §1.1 above).

1.2.3. Classifiers and classifying constructions

I have eschewed the widely-used term *classifier* in deference to the fact that in many cases it is the grammatical construction that categorisation is associated with, not a particular class of linguistic unit (see also Wilkins 2000). Units that serve a categorising function are not necessarily restricted to this function, and may have other uses in addition, in other grammatical environments. This means that we can have classification without dedicated classifiers. The term *classifier* is reserved in this book exclusively for dedicated classifiers; the terms *classification construction* and *categorisation construction* apply to any construction in which some linguistic unit serves to classify or categorise a lexeme or referent. Classifiers thus occur in classification constructions, but classification constructions do not require classifiers. It is possible for a language to show classification constructions involving dedicated classifiers exclusively, or no classifiers whatever; alternatively, a mix-

ture of dedicated classifiers and other items might be involved.[5]

Subcategories are normally marked constructionally, less commonly by dedicated classifiers (e.g. cranberry morphs). Supercategories are always associated with particular constructions, and may or may not involve dedicated classifiers; classes are generally marked by dedicated classifiers. Occasionally, systems intermediate between class and category systems are marked by a combination of dedicated classifiers and other units, both occurring in classifying constructions. This seems to be the case in Ngan'gityemerri (Reid 1997).

1.2.4. Concluding remarks

To sum up the preceding discussion, and clarify the terminology, Figure 1 shows the main types of grammatical classification systems.

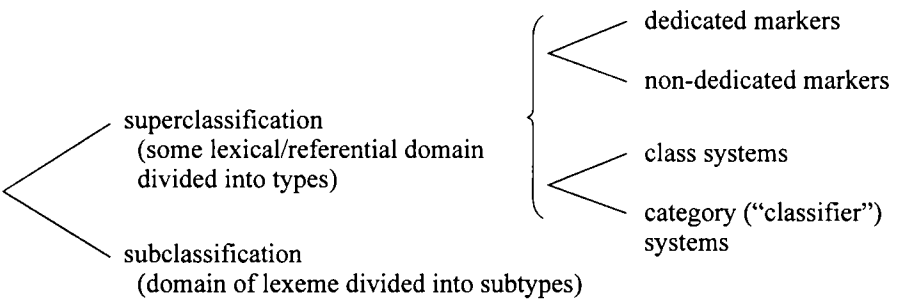


Figure 1. Gross typology of grammatical classifying systems

Problems with the typology have already been mentioned, especially the distinction between classes and categories, which is not always clear-cut. So also are there grey areas between subclassification and superclassification. This seems to be the case for some generic-specific classification systems in various Australian languages. The typology of classification systems according to the parameters set out in the previous three subsections is at best an ideal one, which systems approximate by degrees. Types should not be reified, and I make no attempt to define them. Figure 2 below illustrates graphically typical as-

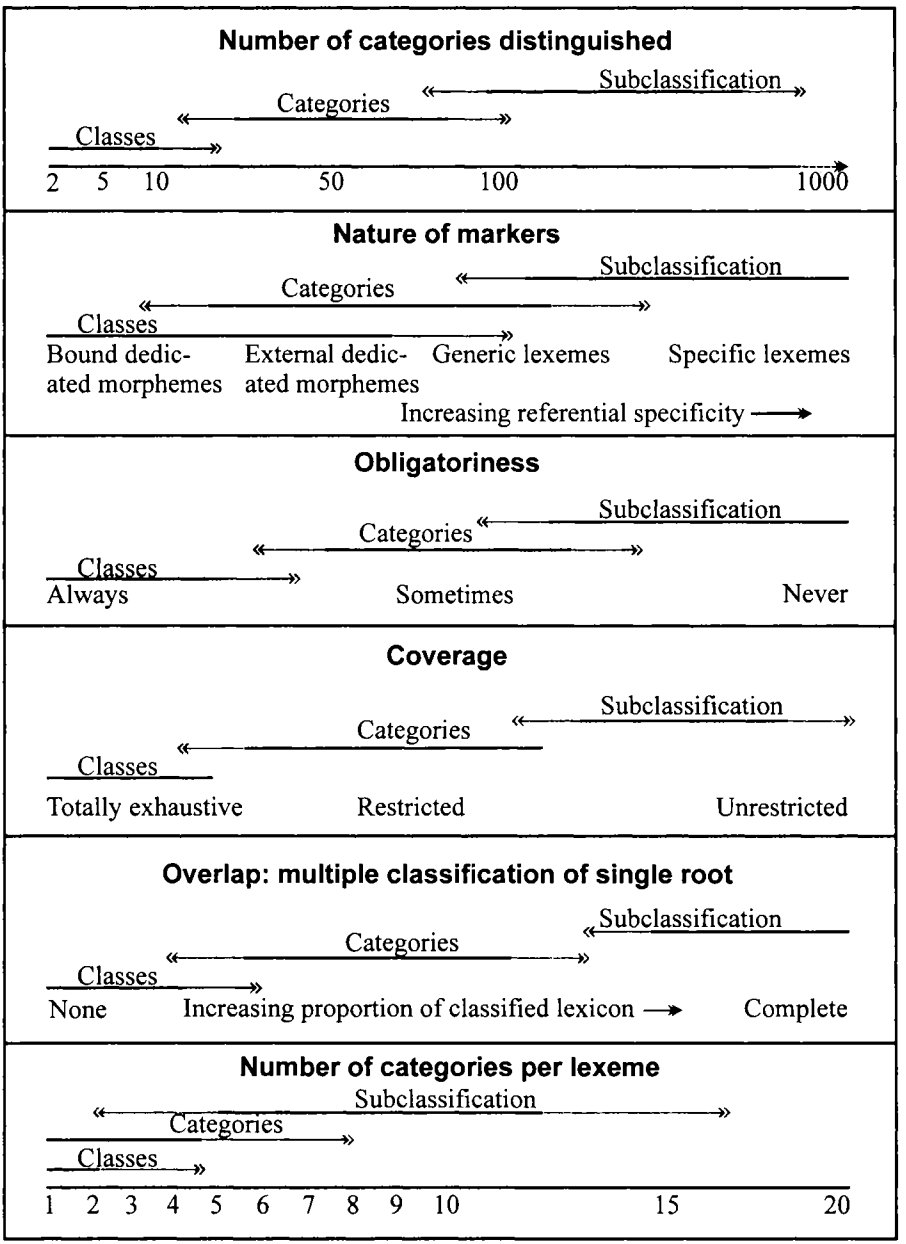


Figure 2. Comparison of types of classification

sociations of systems and parameters, as well as the range of variation within each. Parameters on the left tend to covary, and define classes; those in the centre tend to covary and define categories; those to the right covary and define subclassification. They are however independent, and covariation is a tendency.

The typology provides a useful conceptual handle on the range of cross-linguistic variation, and I use it for convenience. What is required for each language is a detailed description of the system, not merely placement on the ternary typology.

1.3. Understanding grammatical superclassification

1.3.1. A conceptual model

In this section I propose a conceptual model for overt grammatical superclassification, as a first tack on the problem of understanding what sort of phenomenon it is. Subclassification and covert grammatical classification differ in many respects, and are excluded from the discussion.

But first I should explain why I do not consider taxonomies such as are associated with the biological sciences as prototypes for grammatical classification systems. True, when we speak of classes, taxonomic groups are generally what we have in mind. For instance, in various human societies people are assigned to classes, such that each person belongs uniquely to a class; multiple class membership is precluded. This is not, however, a defining feature, as some class systems do admit membership reassignment. In Western societies, for instance, one can be reassigned over time, although at any point one will be a member of a unique class. Even the latter condition is sometimes relaxed, as in some Australian Aboriginal societies—e.g. Wangkajunga people of Wangkatjungka community, near Fitzroy Crossing (Kimberley, Western Australia)—in which membership of subsections (eight classes into which the social universe is divided) is not entirely disjoint; persons can belong to more than one subsection at a given point in time.

Linnean taxonomies might be taken as prototypes for class systems.

But there are systems in which partitioning into non-intersecting sets is neither a requirement nor an ideal. As well as social classes, Westerners operate with social roles; these do not partition the members of a society, and indeed most people belong to more than one such category. One might be an unemployed plumber, an academic linguist, a working mother, a motel owning factory worker, and so forth. Regardless, these roles serve as an effective and convenient ways of categorising people.

The model I propose for linguistic classification is not based on persons in social life, but on libraries. Books in a library are usually divided into labelled subsets according to subject matter, each book being assigned to a labelled shelf, facilitating its location. The books themselves might also have a label marking their subject matter, as an extra precaution, e.g. to facilitate replacement after use. Non-fiction books are usually assigned to shelves, or sections of shelves, labelled according to subject matter, thus facilitating location of those dealing with a particular subject. Corresponding systems of grammatical classification tend to be found in noun class systems, where surrounding material is marked—perhaps facilitating location of a noun in the mental lexicon. Verb class systems, by contrast, often show the mark only on the lexeme itself. This is like a system used in the fiction section of some libraries, where each book is marked by a coloured spot indicating whether it is a mystery, romance, thriller, science fiction, or whatever. Here the members of a class are scattered throughout the shelves, and class specification facilitates identification of members of a class within a system that is organised differently—and would be equally effective in a library where the books were randomly assigned to shelves. Both systems are useful, but in different ways, corresponding to different ways readers tend to use and access books of the two super-types.[6]

There is no reason why a lexical item could not belong to more than one set, any more than a book can not be marked with different dots, or placed—in multiple copies—on different shelves.[7] But if so, in any usage one will need to be selected, and the choice will be motivated,[8] typically according to some feature of the lexeme or referent. For instance, the nominal *dimana* 'horse' in Kija (Western Australia) belongs to both masculine and feminine classes; it is masculine when the referent

is a stallion, feminine when the referent is a mare (see §1.3.3). Similarly, in Nyulnyul, the verb -MARR 'burn, cook' belongs to both conjugation classes; the selection depends on clausal transitivity.

Library systems such as described in the previous two paragraphs model systems of grammatical classes. Grammatical category systems, on the other hand, might be thought of in terms of a simplified model of a card-cataloguing system. A lexical item of some part-of-speech is written on one or more cards and assigned to labelled trays that are thematically organised. A given lexeme might be multiply represented and classified, just as a given book is multiply carded according to the topics it treats. But the conceptual referent is in each case different. To the reader who accesses *A Functional Grammar of Gooniyandi* from the 'Australian Aboriginal languages' tray, the work is a different one to what it is to the reader who accesses it through, say, 'mother-in-law languages', 'descriptive grammars', or 'ergativity'. According to the mode of access, the work can be expected to be used in distinct ways, even though it designates the same slice of physical reality. Likewise, the Gooniyandi speaker who as it were accesses a lexical card—say for *mila-* 'see'—in the 'extendible bivalent action' tray is doing so because they wish to retrieve or invoke a particular sense of the verb, namely its use as way of referring to the ordinary activity of someone seeing something. Had they categorised it as 'effect accomplishment' they would be using it as a means of referring to the activity of someone glancing at something.

A library with a card categorisation system might contain uncatalogued books (e.g. dictionaries and encyclopaedias). But if it has a class system marked externally to the books, every item in that part of the system must belong to a class—all must go to some shelf (if only a residual one), which will generally be labelled.[9] If, however, the class is marked on the books themselves, some books might not be marked. This has parallels in systems of linguistic classification. Category systems do not always apply to every member of a part-of-speech, whereas where classes are marked externally to a word, every member must be assigned to a category, indexed on a nearby word; where marked on the word itself, marking may be partial. (The marking system on the books and the external system need not be identical, again

paralleling the situation for "head" and "agreement" classes—e.g. Evans 1997a.)

This model does not furnish criteria for distinguishing classes from categories. Rather, it provides an intimation of what might be involved in the two phenomena. Two means of classifying lexical items have been suggested: by virtue of their belonging to sets (lexical classes), or by assigning (filing) representations of the items, to sets. Note that according to this model, in both class and category systems lexical types correspond to book tokens; in category systems, the model treats lexical tokens as card tokens in a filing system. This is admittedly a linguist's conceptualisation; there is no implication that it corresponds with anything in the minds of speakers—though it might.[10]

The library model also provides a clue as to why mutual exclusion tends to be associated with classes, overlap with categories—and why these features are not criterial. To divide a set of material objects into physical subsets, requires that each be placed in a unique subset. Two instances will be required if an item is to be placed in two subsets. In a library, two copies of a book will be necessary—and for economic reasons, this will be the exception rather than the rule. But to categorise, it is not the objects themselves that are assigned to subsets, but their representations. It is much more economic to divide representations into physical sets than the actual objects.

Putting things in a slightly different way, the relation 'belongs to' or 'is an element of' is germane to classes: books (respectively lexemes) 'belong to' classes. But this is not the case for categories: books (lexemes) do not 'belong to' categories. Here the pertinent relation is 'is assigned to': books-in/for-use are 'assigned to' categories through their catalogue representations, which indicate how they are to be used in particular cases. What this means for the linguistic system is that it is the instantiated use of the lexeme, the discourse token, that is assigned to a category. This distinction is obscure because in language the representation of a lexeme is indistinguishable from the lexeme itself—it stands for itself.

In a category system, this process of assignment constrains the range of referents of a particular lexeme-instance; only certain aspects of its meaning potential will be activated. As a result, category systems

can generally be used to serve either or both functions: (a) *reference modification*—facilitating modification of the reference of a lexeme (Bolinger 1967: 20); and/or (b) *referent construal*—indicating how the referent is to be construed in the particular instance, via the knowledge domains invoked.

Returning to the library model, a cataloguing system also serves to classify the universe of subject matter of books, which could be said to lie "behind" the system. In the same way a category system in language effectively serves as a system of *referent classification*. The world of referents of words of that part-of-speech—or a substantial part thereof—is typologised. Clearly there is no reason why a class system cannot do the same, though the fact that classes are generally fewer in number than categories means that the referent classification effected by class systems will usually be less systematic.

The library analogy permits an understanding of how a system of classification can be considered to apply both to linguistic items—the words of a part-of-speech—and entities in the world. Thus I disagree with Lucy (2000: 328–331), who proposes that in general nominal classification systems should be understood as systems not for classifying linguistic items, but for classifying the external world of experience, as suggested also by Green (1997) and Schultze-Berndt (2000), among others. Lucy (2000: 331) proposes that it is only in class systems that linguistic items are classified. The basis for this claim appears to be that it is only in class systems that classification can be based exclusively on inherent features of the lexical items themselves, rather than the referent. But having a basis in the features of the referent rather than the lexeme does not preclude classification of linguistic items, as the catalogue model demonstrates. That inherent lexical features are largely irrelevant to category assignment in category systems is not germane: the difference is that in the one system it is prototypically a lexical type that is classified, whereas in the other it is a lexical token.

I agree with Lucy (2000) and Dahl (1999) that it is important to distinguish between lexical and referent classification.[11] The reason, however, is not that one precludes the other, but rather that we can have either without the other—or both together. A typical class system exem-

plifies a lexical classification system, and a typical category system exemplifies a referent classification system. Plain referent classification systems, that are not accompanied by lexical classification systems are not difficult to find. An obvious example is the third person singular pronominal set in English: the three terms effect a referent classification that is not accompanied by lexical classification. (We return to this issue in the next subsection.)

In this book I frequently view verb category systems from the perspective of the event classification they effect. By *event* I mean the referent—conceptual and/or real—of an instance of use of a verb together with the words and morphemes that enter into immediate syntagmatic relations with it. Together, these items form a unit I will refer to as a *verb phrase* (VP), a phrase consisting of a verb together with auxiliaries, qualifying adverbs, and so forth—but excluding object NPs (McGregor 1997b: 122–125). A VP is thus effectively a contextualised instance of use of a verbal lexeme, a token; it refers just to that which unfolds over time, excluding associated entities, props, and circumstances. All of these things together will be referred to as the *situation*, which is the referent of an entire clause. Verbal lexemes—the types—do not, strictly speaking, refer to anything; they denote, rather than refer (Lyons 1977: 208). The distinction between denotation and reference is not made systematically in terminological choice in this book, mainly to avoid an overly pedantic style. Likewise, the term *event* designates both the referent of a VP and the denotation of a verbal lexeme. However, these distinctions are important, and, where necessary, are made by descriptive expressions rather than terminologically.

It is important to be clear on the types of relationship that I am suggesting are involved in the various sorts of classification we have dealt with in this section. To reiterate, for classes it is 'belongs to', or 'is a member of', and operates on the domain of lexical types; for categories, it is 'is assigned to', and operates on the domain of lexical tokens; and for referent classification it is 'kind of' or 'type of', which applies to tokens of events in the world. Thus, if in Gooniyandi the lexical verb *mila-* 'see' is assigned—in a particular instance of use—to a certain category by the classifier +A, then the referent event is specified as being of a certain type: namely, one that is potentially extendible in time,

and externally directed. The 'belongs to' relation is irrelevant.

The library analogy might perhaps suggest that class and category systems are discrete types, thus conflicting with the linguistic reality it models. This is not so. One can imagine variants of a library class system that approximate category systems—for instance, a well-endowed library with multiple copies of many books, each copy being assigned to a different set. Such a system will become increasingly expensive as it approximates a genuine system of categories. Nor is it difficult to imagine functional systems that combine formal features of both types.

Summing up, I have suggested that the Linnean model of linguistic classification is misleading in that it presumes that classification systems tend to the ideal of partitioning the target domain into disjoint sets. Even class systems do not necessarily aspire to this as an ideal. Some sort of categorisation—not necessarily taxonomic—of a referent domain may be achieved by a system of linguistic classification. The significance of such systems of categorisation is not denied, and it may be important in language acquisition. But grammatical classification systems always have some functions beyond themselves, beyond the fact of assigning lexical items to categories: they are not there for mere decoration. They have uses—for example, lexical retrieval and reference management. In this regard they are like library systems, and unlike biological taxonomies, that merely describe segments of reality. The uses temper the taxonomic tendency with tendencies to lexical economy, referential perspicuity, and so forth. Thus linguistic category systems aim for some balance between requirements of learnability and usability, somewhat in the manner of a library cataloguing system.

1.3.2. Distributional criteria for grammatical superclassification

The library catalogue model provides an interpretation of grammatical classification, a way to understand the difference between classes and categories, and how they shade into one another. The assumption that the system be semantic is not essential. A system might be based on the size or colour of the book, or the first letter of the title, or author. And two or more different systems might be used in the one library—e.g. a

subject system for non-fiction, an alphabetic-by-author system for fiction. In the same way, a nominal or verbal classification system might employ semantic characteristics for some lexemes, phonological for others. For a system to be useful, however, it should not be a complete mish-mash, and there should be some principled way of deciding when to employ one sub-system, and when another.

The library catalogue model, however, fails to model the mechanism of grammatical classification, and gives no recognition criteria. The question I am most often asked about verb classification is: how do you know when you are dealing with it?—how can you be certain that a phenomenon actually is an instance of verb classification? Strangely, the question is rarely asked of nominal classification, even though there is far from complete agreement amongst investigators (Craig 1994: 565). Literature searches have failed to reveal usable criteria, let alone arguments that a certain system in language L is a nominal classifying system, not something else. What we typically find instead is discussion of distinctions among types of nominal classifying system, presuming that we have delimited the range of phenomena in the first place. This is the case, for instance, for influential accounts such as Dixon (1982, 1986); Craig (1986a); and Silverstein (1986).[12] In this section we attempt to remedy this situation.

Perhaps the most crucial requirement for a classification system, irrespective of what is classified, is difference: whatever set is classified, its members must show different behaviours. If no difference is discernible within a particular set of entities according to some parameter, it is not reasonable to speak of classification. The articles *the* and *a ~ an* in English are, quite rightly, not considered as nominal class markers or classifiers, although we can easily define sets of nominals according to the ability to occur with them. These sets satisfy most of the criteria for classes laid out in Dixon (1982: 217–218), at least as well as many genuine class systems do. But any nominal can occur with either article, so no difference is detectable, and it makes no sense to speak of classification.

Also essential to categorisation is commonality. No system that failed to group things somehow would count as a categorisation system; if every member of the set were differently assigned, the result

would be as useless as a classification system as if no distinctions were drawn. What is required is something in between the two extremes.

Overt grammatical categorisation requires manifestation in language form; this is a necessary though not sufficient condition. For suppose that a language has two linguistic items X_1 and X_2, such that expressions involving X_1 always designate single entities, those involving X_2, sets of larger cardinality. Clearly X_1 and X_2 distinguish different categories, namely sets of one thing vs. larger sets. If there is no corresponding difference among the designating linguistic expressions in terms of distribution with respect to X_1 and X_2—any expression can occur with either, depending only on the nature of the referent reality—it would be inappropriate to speak of grammatical classification. And this corresponds with the usual treatment by linguists. Such a situation arises in English—almost—and no linguist speaks of a classification system defined by the suffixes -ø and -*s*, even though they clearly do categorise referents.[13] Instead, they speak of grammatical categories singular and plural, not a system of nominal classification.

Languages contain many such items, that need to be excluded from the set of potential category markers. Of course, not all cases are as clear as the number system just discussed. A few linguists, for example, speak of grammatical genders in English, distinguished by the three third person singular pronominals. The problem is that these pronominals do not enter into structural relations with nominals, but non-structure-bound referential relations; referent groupings are on the whole effected according to semantic gender, not nominal category.

This indicates the need for distributional requirements for overt systems of grammatical superclassification. Categories must be discernible by virtue of differential collocations of the members of some lexical set and the members of another set, in well defined environments. The distributional requirements are tentatively specified as follows:

(i) the domain $\mathcal{D}$ to be classified contains all—or a substantial part of—a set of linguistic items, $\mathcal{P}$, usually an open lexical class;
(ii) items in $\mathcal{D}$ collocate—in well defined grammatical environments—with members of another set of linguistic items $\mathcal{M}$, such that:
(a) all members of $\mathcal{M}$ can occur in collocation with some member

of $\mathcal{D}$; (b) all items with this potential are included in $\mathcal{M}$; and (c) the Ms are of the same "order" of reality;

(iii) the cardinality of $\mathcal{M}$ must be greater than one, but significantly less than the cardinality of $\mathcal{D}$; and

(iv) members of $\mathcal{M}$ must show differences in patterns of collocation with items in $\mathcal{D}$; specifically, there must exist at least one pair (M_i, M_j) for which the set of Ds they collocate with, $\{D_{i,1}, D_{i,2}, D_{i,3}, \ldots\}$ and $\{D_{j,1}, D_{j,2}, D_{j,3}, \ldots\}$, are significantly different.

(i) specifies that what is classified should represent the bulk of an open part-of-speech $\mathcal{P}$. The wording admits exclusion of some members of $\mathcal{P}$, and inclusion of members of different word classes. The openness requirement on $\mathcal{P}$ ensures that the categorised domain is relatively large and open. It is possible to speak of categorisation of smaller domains, though the point of doing so diminishes as cardinality decreases.

(ii) and (iii) guarantee that the categories are formally indexed in some environments, and that they number more than one, but significantly fewer than $\mathcal{D}$.[14] No stipulation is made on the nature of the markers in $\mathcal{M}$: (ii) is sufficiently general to admit morphemes, lexemes, or combinations of the two; it is not even specified that they must be morphemes: they could as well be meaningless elements of linguistic form. However, they must be items of the same "order" of linguistic-semiotic reality (stipulation (ii)(c))—mixtures of phonemes and morphemes are thus excluded, as are allomorphs and morphemes. Usually members of $\mathcal{M}$ will be of the same formal type, that occur in the same structural position in the relevant grammatical environment. The Ngan'gityemerri system of nominal classification—in which, it will be recalled, the markers include both free lexical items, proclitics and enclitics, that do not occur in a single structural place—shows that this should not be written in as a requirement (Reid 1997).

On the other hand, especially (but not only) in noun class systems different systems of markers can be found—some attached to determiners and/or adjectives in the same NP, others to the noun itself, and still others to verbs. Often the different systems make the same distinctions, and define the same classes. But sometimes they do not, and it is

necessary to distinguish between head and agreement classes, as in Mayali (Evans 1997a). If the same classes are defined, $\mathcal{M}$ could consist of marker of different formal types; otherwise, it is probably necessary to distinguish different $\mathcal{M}$ sets associated with different classification systems. Criteria (i)–(iv) are not sufficiently refined to allow the Ngan'gityemerri and Mayali situations; it is not clear how they should be revised.[15]

Condition (ii) does not preclude Ms from occurring in environments where Ds are absent; nor is there any requirement that they (or a word they form a distributional part of) cannot simultaneously serve other grammatical functions. They need not be dedicated classifiers. If and when the Ms occur outside of the relevant constructions, they no longer serve a classifying function. (Once a classifier does not imply always a classifier.) (i)–(iv) are therefore not requirements for disembodied classification systems, but for systems rooted in grammatical constructions.

(iv) ensures categorical diversity: if every M defined the same set of Ds, no classification of $\mathcal{D}$ would be effected. This condition also ensures that the members of $\mathcal{D}$ show sufficient diversity in their patterns of collocation with Ms: for any D_i there will be at least one D_j that behaves differently from it with respect to the Ms it can collocate with.

Two points should be highlighted in regard to (iv). First, it stipulates that the collocation sets for at least one pair of Ms should be not just different, but significantly different; what counts as significant is unspecified. If sets $\{D_{i,1}, D_{i,2}, D_{i,3}, \ldots\}$, $\{D_{j,1}, D_{j,2}, D_{j,3}, \ldots\}$ differ in just a few members for every pair (M_i, M_j), we could reasonably conclude that the collocational differences result from inadequacies in the data, or the effect of some minor conflict or inconsistency between certain Ms and certain Ds. The imprecision means that there may be cases where opinions differ as to whether or not classification is involved. For instance, Gooniyandi has three stance verbs, *wara-* 'stand', *warang-* 'sit', and *bagi-* 'lie', the use of which in clauses of existence and attribution depends on whether the referent is abstract, and if not, stance (if animate), and salient dimension (if inanimate). When we contrast the sets of Ns collocating with these verbs, opinions could differ as to whether there are sufficient differences amongst them to war-

rant treating them as existential classifiers (see §1.1 above).

Second, (iv) requires only that at least two Ms differ significantly in their collocate sets. That some Ms show identical collocate sets must be permitted. For instance, in Bantu languages the Ms would include, among other things, nominal prefixes and agreement morphemes attached to adjectives. Many of these morphemes also specify number, and singular-plural pairs such as nominal prefixes *mo-* and *ba-*, or ø- and *bo-*, in Sesotho (Demuth 2000: 274) would, it seems, define identical noun collocate sets.

Granted the above qualifications and requisite emendations to (i)–(iv), I contend that they are necessary requirements for overt systems of grammatical categorisation. They are also sufficient: if satisfied, it is reasonable to speak of a system of overt grammatical classification in which items in $\mathcal{D}$ are assigned to distinct categories by the members of $\mathcal{M}$. The basis on which the assignment is effected remains unspecified.

Conditions (i)–(iv) do not guarantee categories that are either interesting or significant. I seriously doubt whether explicit criteria could be devised that succeed in delimiting that which is interesting or significant from that which is not. Beyond (i)–(iv) it boils down to considerations of usefulness: what is worth distinguishing?

The present approach takes distributional considerations as fundamental. Any system of nominal or verbal classification must have a foundation in linguistic substance, crucially, in distributional substance. If the distributional criteria are satisfied, whatever analysis is imposed on the phenomenon in question, the categories still have to be accounted for. This can be done without invoking functional or semantic considerations. Saying this is not to suggest that such factors can be ignored, or that they are unimportant; they are essential to "significant" classification systems. But they can be teased apart—to some extent at least—from the distributional facts. It seems procedurally useful to do so, just as it is procedurally useful to distinguish etic substance from emic form. And just as not all items of grammatical substance count as morphemes (e.g. formatives), so too do not all items of this sort of distributional substance count as significant classification systems.

Classification is pervasive in language—indeed, in life—and is certainly not restricted to the range of phenomena included under nominal

or verbal classification. For present purposes it is useful to distinguish two types:

- *grammatical classification*: systems of overt or covert classification of lexemes of a certain part-of-speech, that is manifest in behaviour in determinate grammatical environments; and
- *epistemological classification*: systems of linguistic units that categorise a domain of (conceptual) referents.

Epistemological classification—an extension of Mark Durie's (1985: 151–168) "epistemological classifiers"—need not be grammatical classification: for instance, *wh-* words in English are not grammatical classifiers, though they do serve as epistemological classifiers. Most grammatical classification systems (at least the interesting ones) are also epistemological systems.

A final comment is in order. Corresponding to a system of overt classification by members of $\mathcal{M}$ may be a covert system of classification, defined by the potential of members of $\mathcal{D}$ to occur with Ms. Some Ds may show similar potentials for collocation with Ms, while others have different potentials. Categories so defined do not correspond directly with Ms, and it is only by putting together knowledge of all Ms that occur with given Ds that this classification can be effected. In Nyulnyul, for instance, it is possible to distinguish three covert classes of inflecting verbs corresponding to the two overt conjugation classes marked by prefixes *na-* and ø-. These are *na-*, ø-, and *na-*/ø-, where the members of the first collocate exclusively with *na-*, of the second, only with ø-, and of the third, with both. Whether or not for a system $\mathcal{M}$ of overt classifiers there is a corresponding useful covert categorisation system is an empirical question, the answer to which is rarely as straightforward as in the Nyulnyul example.

1.3.3. An example: the Kija noun class system

To render the proposals of the previous section more concrete, we briefly discuss the nominal classifying system of Kija, described in

Kofod (1996: 19–20, nd). Criteria (i)–(iv) are satisfied, as follows:

(i) The domain $\mathcal{D}$ is the entire set of nouns. (ii) Adjectives, possessive pronouns, and interrogatives take different suffixes depending on the noun they modify: *-ny* ~ *-ji*; *-l* ~ *-el*; and *-m* (the second form in each case occurs following a consonant). These three items constitute M, whose members collocate, in the NP, with the members of $\mathcal{D}$, as shown by the following examples:

(1) *timana-ny ta-ny jirrawu-ny* Kija
horse-MAS that-MAS one-MAS
'that one stallion' (Kofod nd)

(2) *timana-l ta-l jirrawu-l* Kija
horse-FEM that-FEM one-FEM
'that one mare' (Kofod nd)

(3) *timana-m ta-m melakawu-m* Kija
horse-N/P that-N/P many-N/P
'those many horses' (Kofod nd)

As these examples reveal, every word of an NP, including the referring noun, is usually marked by an M. For the noun, however, is not invariably present and not all final consonants *ny*, *l*, or *m* are Ms—e.g. the final *l* of *miyal* 'meat' belongs to the root itself, since other words in the NP would have *-ny* attached.

(iii) The cardinality of $\mathcal{M}$ is three—more than one, but significantly less than the cardinality of $\mathcal{D}$. Finally, the members of $\mathcal{M}$ show different patterns of collocation with members of $\mathcal{D}$. In brief, with *-ny* ~ *-ji* we find nouns denoting many human and large higher order animates, various other animates, and many inanimates; similarly with *-l* ~ *-el* we find nouns denoting many human and large higher order animates, and various other animates and inanimates. Although both Ms define sets of human and higher order animate nouns that overlap considerably, the nouns denoting lower order animates and inanimates are virtually disjoint. The third M, *-m*, actually collocates with all nouns that collocate with the other two Ms, as well as with a few other nouns. Putting

these facts together, it would seem that the sets are significantly different, and (iv) is satisfied.

Is this an interesting or significant classification system? The answer is a definite "Yes", since the collocations have an obvious semantic basis. Nouns referring to human beings normally collocate with *-ny ~ -ji*, *-l ~ -el*, or *-m* according to whether the referent is masculine, feminine or plural. The same goes for nouns referring to large animals, at least when sex is obvious. Nouns denoting most other animates collocate with one of either *-ny ~ -ji* or *-l ~ -el* regardless of the referent's sex, provided it is singular. Sex is sometimes relevant for these nouns, not because of biological gender, but rather by virtue of mythologically associations. For example, in the Dreamtime the turtle was a woman, and the word for 'turtle' usually collocates with *-l ~ -el*: *tarntal*, *wayiwurrul*, *pilitpal*, *palarnel* (dialectal variants); the crocodile and bat were men, and the nouns *lalangkarrany* 'crocodile' and *pinyjirrminy* 'bat' collocate with *-ny ~ -ji*. When more than one animate is referred to, the nominal collocates instead with *-m*.

For inanimates, choice of M is a complex matter, though generally a noun referring to a single thing will collocate exclusively with either *-ny ~ -ji* or *-l ~ -el*, while nouns referring to more than one thing will collocate with *-m*. However, some inanimate nouns collocate with *-m*, even though only one thing is referred to: e.g. *marnem ~ thunpam* 'fire', *kurrngam ~ kurlum* 'water' and *mayim* 'vegetable food', and most body part nouns. (In some circumstances even these nouns can collocate with *-ny ~ -ji* or *-l ~ -el*, especially when a particular individual is being referred to.)

Clearly it makes sense to identify three nominal classes according to the M that occurs in the NP—the system shows more characteristics of a class than a category system (see Table 1). It is reasonable to label the classes masculine, feminine, and neuter/plural, as per Kofod (1996, nd). What is striking about the system, however, is that the relevant relation is not so much class membership—nouns are in the majority of instances not accorded classes as lexical features—but class assignment, on the basis of properties of the referent.

1.4. Verb superclassification Australian style

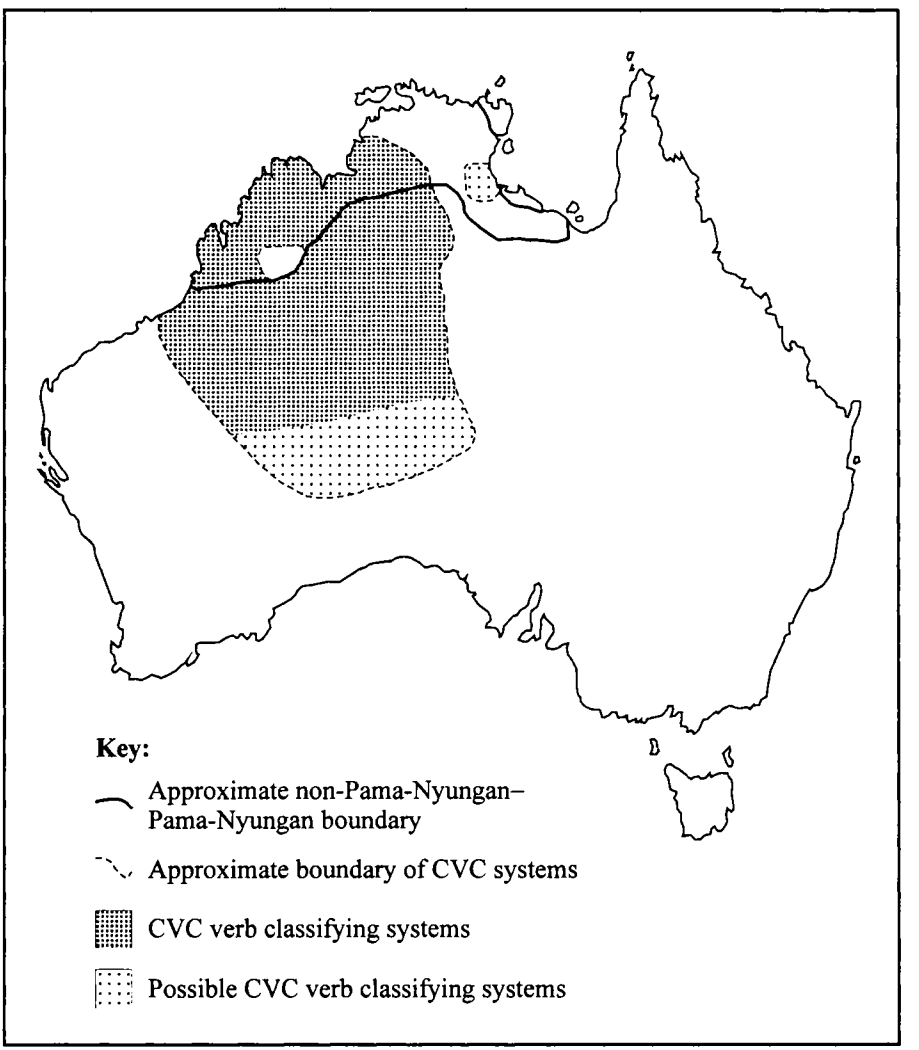
The primary concern of this book is with systems of verb superclassification in Australian Aboriginal languages. Most attention is devoted to verb category systems. These are spread over a large region of northern and north-western Australia, crossing many genetic boundaries, including the Pama-Nyungan–non-Pama-Nyungan divide; at least one isolated system is found in the south-east of the continent. Verb class systems, or conjugation classes, are also dealt with, albeit in a more cursory fashion. Maps 2 and 3 provide approximate indication of the geographical distribution of the two phenomena.

Verb category systems in Australia are closely associated with a particular construction type, a type of *compound verb construction* (e.g. Capell 1979). This prototypically involves two verbal elements forming a compound: an *inflecting verb* (IV) that is inflected for verbal categories such as tense, mood, and aspect, as well as (in many languages) person and number of the subject and/or object; and an *uninflecting verb* (UV)[16] that takes no inflections, and is largely morphologically invariant. In (4), for instance, the IV -MARA 'take' takes all the inflections, while the UV *jarug* 'push back' occurs in bare root form.

(4) *jarug* *andu-ø-ma-nga-lu* Ngarinyin
push:back 3plACC-3sgNOM-take-PA-PROX
'He pushed them back this way.' (Rumsey 1982a: 112)

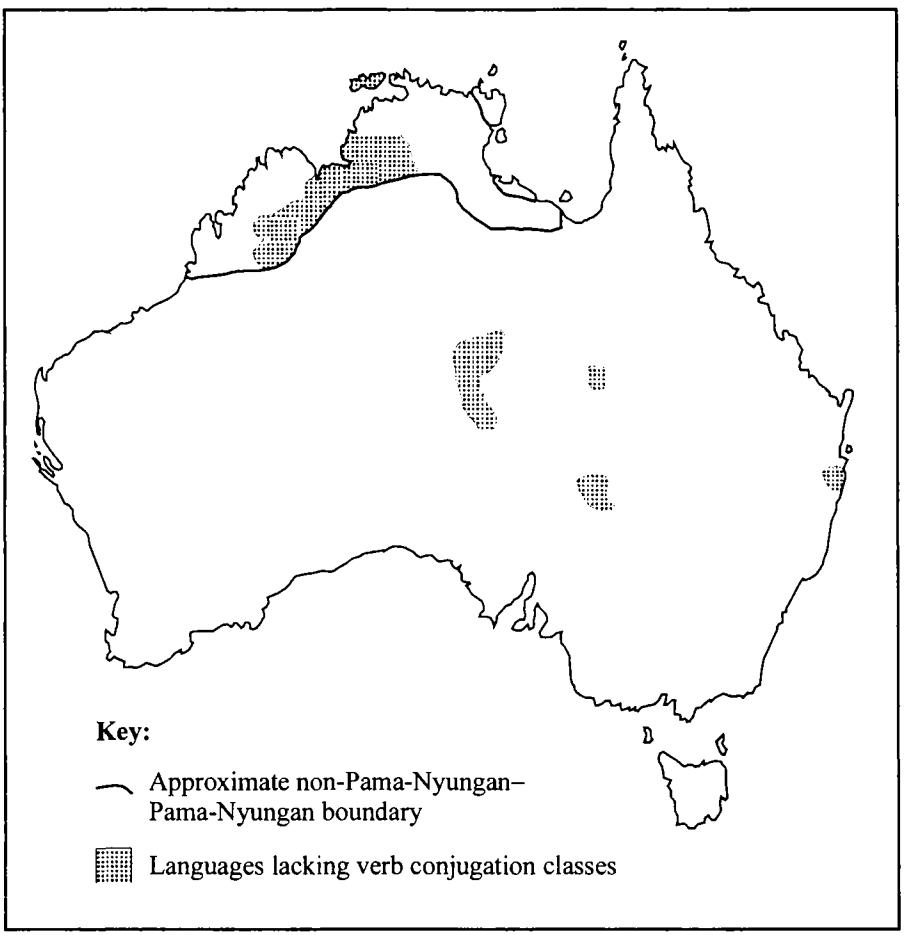
This type of compound verb construction—consisting of an IV and a UV and serving as a vehicle for a verb category system—will be designated by the acronym CVC; other types will be referred to by the full appellation. In terms of the characterisation of §1.3.2, UVs constitute $\mathcal{P}$, IVs that occur in the CVC, $\mathcal{M}$. In fact, lexical units of other parts-of-speech (e.g. nominals and adverbials—see §3.1 below) can normally occur in the position occupied by UVs, so $\mathcal{D}$ is a superset of $\mathcal{P}$. Thus the IV discharges the classifying function in this construction: in (4) -MARA 'take' classifies *jarug* 'push back' as an event of caused motion (see §3.3 below).

In languages with CVCs, IVs and UVs are distinct parts-of-speech, usually with few shared members. The class of IVs is generally fairly small and closed, while UVs form an open class. In some respects—including morphological properties, the grammatical categories they distinguish, and their environments of occurrence—UVs appear less "verby" than IVs, and show features in common with nominals



Map 2. Approximate distribution of CVC syperclassifying systems in Australia

(Schultze-Berndt 2001a; see also §3.1). The appropriateness of the label *verb classification* for superclassifying CVCs might thus be questioned. While I acknowledge that the term is potentially misleading, within the VP the categorised units "function as" verbs. More accurately, they serve the same grammatical function in CVCs as IVs do in *simple verb constructions* (SVCs), constructions that involve just an IV (see §3.1). Specifically, as proposed in §6.2 below, the UV and the IV serve—in the CVC and SVC respectively—the grammatical role of Process: they designate an event.[17] Thus, whilst imperfect, the term



Map 3. Approximate distribution of verb conjugation classes

verb classification is not unreasonable. In an attempt at terminological precision, I offer the following definition:

- In compound verb constructions *verb classification* refers to a type of superclassification satisfying (i)–(iv) of §1.3.2; the CVC is a classifying construction, not a classifier system. What is classified is the lexical token that occurs in the Process role in that construction.[18]

The correlation between verb category systems and compound verb constructions is imperfect. Compound verb constructions are more widespread, and need not be the locus of category systems; this seems to be the situation in Gunwinjguan languages of Arnhem Land (§3.11). Conversely, a few languages do not have CVCs but do distinguish verb categories. It can be argued that some such systems derive historically from CVC systems that have undergone grammaticisation resulting in the delexicalisation of former inflecting verbs. One language of western New South Wales has a verb superclassification system associated with a rather different type of compound verb construction (§7.2.2).

Verb class systems are even more widespread, being found across almost the entire Pama-Nyungan region, as well as much of the non-Pama-Nyungan region. Conjugation classes are usually indexed by distinct sets of tense, mood, and/or aspect inflections, within which a conjugation marker is sometimes discernible, though not necessarily as a morpheme. It is proposed in Chapter 8 that a number of these verb class systems have derived historically from verb category systems—in fact, from CVCs. Conjugation markers in these systems represent historical remnants of inflected forms of IVs.

In contrast to CVC category systems, verb class systems are strongly associated with items of a single part-of-speech; a unit of any other part-of-speech must normally be reclassified by a derivational morpheme before it can occur in the verbal construction, and be assigned to a class. The question of the appropriateness of the term *verb class system* does not arise. As we will see, verb class and category systems are in a gradual relation (as per Figure 2). This, plus the historical relation mentioned in the previous paragraph, constitutes additional motivation for using the term *verb classification* in reference to the

CVC. To treat the two phenomena as incommensurate misrepresents linguistic reality.

1.5. Semantic basis of verb superclassification

1.5.1. Fundamental parameters

A relatively small set of semantic features recurs in nominal superclassifying systems in language after language, including animacy, gender, shape, and use. Can we identify a comparable small set of recurrent semantic features in overt verb superclassifying systems?

The answer seems to be in the affirmative. Three semantic features of event "configuration" are particularly relevant: *vectorial configuration* (an abstract representation of action configuration), *Aktionsart* (lexical/event aspect), and *valency* (lexical/event transitivity).[19] Vectorial configuration is the verbal analogue of shape in nominal classification systems; Aktionsart is the analogue of what Rijkhoff (in press) refers to as Seinsart, "the way a nominal property is represented in the spatial dimension in terms of Shape and Homogenity"; valency has no obvious analogue in semantic features of nominal classification systems.

As to the other semantic features characteristic of nominal classification, animacy and gender are obviously irrelevant to verb classification. Although events could be classified according to use (e.g. purpose or goal), this feature does not seem to be employed in Australia. No system I know of groups together instances of hitting and kicking that have the purpose of say disabling a person, as against unintentional or non-serious hitting and kicking. The contrast intentional vs. unintentional is never implemented.

Below I provide some discussion of the three macro-features.

[1] Vectorial configuration refers to an abstract relational pattern characterising an event in terms of vectors reminiscent of the trajector-landmark configurations of Cognitive Grammar (e.g. Langacker 1987). Some categories are specified in terms of a vectorial configura-

tion, so that any event satisfying it—that displays the relevant action-shape—may be assigned to that category. In some cases it seems conceptually useful to diagram the configuration in the manner of Cognitive Grammar representations; for instance, the Nyulnyul category marked by the IV -KAL 'wander' can be described by the vectorial configuration shown in Figure 3 (McGregor in preparation: chapter 11). Here the vector represents the "action-path", the fact that it is not straight indicating that the action is not specifically directed towards a particular goal or endpoint, but meanders. This configuration might be met in various ways, depending on how abstractly the path is interpreted. It can be given a concrete interpretation whereby the trajector traces a crooked path of motion; alternatively it could be interpreted in terms of the "path" the activity itself takes, i.e. the action is uncontrolled or non-goal-directed. Thus the event of slipping or sliding is assigned to this category, as is playing.

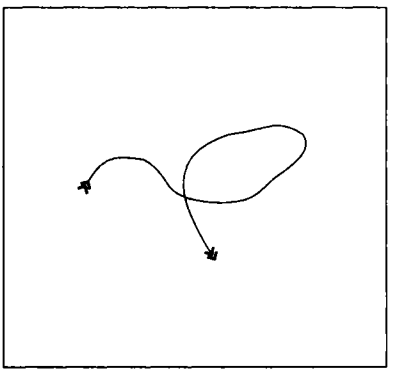


Figure 3. Vectorial configuration for -KAL 'wander' category in Nyulnyul

In other cases diagrammatic representation does not seem so useful, and a description in words appears more revealing. For instance, in Ngarinyin the category marked by the IV -Y_1IBU 'throw' specifies events involving action on an entity causing it to move in a horizontal direction (Rumsey 1982a: 118). Although this clearly specifies a vectorial feature, it seems easier to state it verbally than diagrammatically. (Which is not to say that a diagram, or perhaps a sequence of diagrams,