CAMBRIDGE TEXTBOOKS IN LINGUISTICS

Number

Greville G. Corbett

NUMBER

GREVILLE G. CORBETT

UNIVERSITY OF SURREY

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

First published 2000

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times New Roman 9/13 pt *System* QuarkXPress™ [SE]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Corbett, Greville G.
Number/Greville G. Corbett.
p. cm. – (Cambridge textbooks in linguistics)
Includes bibliographical references and index.
ISBN 0 521 64016 4 (hardback) ISBN 0 521 64970 6 (paperback)
1. Grammar, Comparative and general – Number. I. Title. II. Series.
P240.8.C67 2000
415 – dc21 00-023604 CIP

ISBN 0 521 64016 4 hardback
ISBN 0 521 64970 6 paperback

1
Introduction

Number is the most underestimated of the grammatical categories. It is deceptively simple, and is much more interesting and varied than most linguists realize. This was recognized by Jespersen: 'Number might appear to be one of the simplest natural categories, as simple as "two and two are four." Yet on closer inspection it presents a great many difficulties, both logical and linguistic' (Jespersen 1924: 188). Lyons too pointed out its interest: 'The analysis of the category of number in particular languages may be a very complex matter' (Lyons 1968: 283). This book will illustrate the interest of number, and some first pointers are given in §1.1. We shall also see the challenges which Jespersen and Lyons allude to, one of the trickiest being the need to ensure that as we compare across languages we are really comparing like with like (§1.2). Hence the book is structured so as to work upwards from properties that are safe building blocks for comparison (§1.3). Finally in this introduction a few notes on presentation are needed (§1.4).

1.1 The special interest of number

Despite the significance of number, there are still surveys of linguistics where it receives a footnote's worth of attention. This is largely because there are some reasonable but incorrect assumptions about number, which are generally based on the consideration of a rather limited range of languages. In seeing where these assumptions are false we shall get an initial idea of how interesting the category really is.

> *First assumption:* number is just an opposition of singular versus plural

There are indeed languages with this basic opposition. But there are also many languages with richer systems, with a dual for two real world entities, some with a trial for three, others with a paucal for a small number. There are more exotic possibilities too, with the richest systems having five number values, as in Sursurunga. Moreover, some of the trickiest problems with number become much clearer when we look at the evidence from larger systems, that is those with more than the basic singular–plural distinction.

Second assumption: all relevant items (nouns, for instance) will mark number

We might expect that, say, all nouns would show number. That clearly is not the case, for instance English *honesty* does not mark plural. It seems natural to say that it is an abstract noun and that for certain abstract nouns number is not relevant. But this is a parochial fact about English; there are languages where the proportion of items for which number is relevant in this sense is quite small, and others where number marking is practically always available. The possible ranges of number marking are constrained in interesting ways.

Third assumption: items which do mark number will behave the same

Suppose that we carefully specify how many number values a particular language has and which types of noun mark number. Having avoided our first two false assumptions, we might assume that items would either fail to mark number or would show all the number values available. Once again, things are more interesting than that. In Maltese, for instance, just a few nouns have singular, dual, and plural, while the majority of nouns and the pronouns have only singular and plural. Or in Bayso, pronouns have two number forms while typical nouns have four.

Fourth assumption: number must be expressed

If number forms are available, then surely they must be used? This is an Anglocentric assumption and is quite false. We shall see instances where the marking of number is optional, and there are languages like Bayso where there are special forms which allow the use of a noun without any commitment to the number of entities involved. Linked to this assumption is the fact that number is usually thought of as prototypically inflectional. The inflection–derivation distinction is becoming a hot topic again in morphology and number is in fact highly problematic in this respect. This book will provide a good deal of relevant material; the presentation will be as neutral as possible in order to include the relevant data for a continuing debate in which the criteria are likely to change.

Fifth assumption: number is a nominal category

So far our examples have involved nouns and pronouns. But there are languages where number is a verbal category, marking the number of events rather than the number of individuals. We return to this distinction in the next section.

The point which is emerging is that English and other familiar Indo-European languages have quite *unusual* number systems; they occupy one corner of the typological space. It is clear that to understand the category of number we need to look at

a broad range of languages. Hale made a related point in a discussion of the problem of language endangerment:

> while the category of number is accessible, in an obvious sense, its surface realization across languages exhibits great diversity, and a great many individual languages fail to present the observable data which will permit us to get at the fundamental character of the oppositions involved and, thereby, to come closer to an understanding of the universal organization and inventories of the category of number. (Hale 1997: 75)

We shall see several instances of interesting systems which are essential for appreciating the full range of possibilities being found in languages which have few speakers and are clearly endangered. And the prospects for language loss are particularly serious for number. There are perhaps 6,000 languages spoken at present, of which around 250 are 'safe': they are likely to survive another hundred years at least. But these safe languages are not evenly distributed: over half of them belong to Indo-European or Niger-Kordofanian (Krauss 1992, 1993), while some families with many languages of special interest for number are hardly represented at all. It is therefore important to identify and investigate the most interesting systems while there is still time. Our 'linguistic tour' in the book will include over 250 languages. Several of these languages do not occur in the various typological samples and yet are vital for a full typology (Bayso is a good example). Hence this was a case where it was appropriate to examine as many languages as possible, rather than taking a defined sample. Many of the languages which were investigated will not be mentioned since they turned out to be similar in the relevant respects to others which are described here.

1.2 Comparing like with like

Since we shall look at a wide range of languages we must be careful to ensure that we are comparing like with like. For instance, how do we know that a language has number? Languages like English have the category of number, since we find correspondences like the following:

magazine	magazine-s
head	head-s
woman	women

There is a difference in meaning between *magazine* and *magazines* (obviously concerning the number of them), which corresponds to a difference in form. That same difference in meaning is found in *head/heads* and *woman/women.* The first member of each pair is said to be singular, and the second is plural. So when we say

that English has singular and plural we are referring to correspondences of meaning and form.

In many theoretical frameworks number, like comparable categories such as gender, case and person, is treated as a 'feature'. This feature is said to have certain 'values' (for number, these include singular and plural, and we have already come across others too).[1] These values of the number feature have meanings and forms associated with them. The main part of the meaning of the singular is that it refers to one real world entity, while the plural refers to more than one distinct real world entity. The formal expression of the plural in English is usually the addition of an ending, as in *magazines*, *heads*, while the singular is usually signalled (on nouns) by the absence of such a marker. But there are other ways of marking the plural too, as found in *women* and *geese*. It is the association of (a set of) meanings with (a set of) forms which allows us to talk of the singular and plural values of the feature number.

The plural may be realized in various different ways in a given language. Rather than listing all the forms on each occasion, linguists talk of 'plural forms'. Conversely, these plural forms may be used to express various related meanings, and here we may talk of 'plural meanings'. However, as a shorthand, people often talk of 'the plural' or 'the singular' when in fact just the meaning or just the form is intended. Normally the intention is clear but, particularly when comparing languages, it is important to be explicit about which we intend, for the following reason. We do not expect the form of the plural to be the same in English as, say, in Russian: even if the morphological means used are similar (mainly inflections in both languages); we anticipate that there will be phonological differences between them. Of course, we are correct (the items on the right are Russian translations of the English):

magazine	magazine-s	žurnal	žurnal-y
head	head-s	golov-a	golov-y
woman	women	ženščin-a	ženščin-y

The danger is that using the same term 'plural' for both forms and meanings may lead us to assume without question that though the forms differ the same meanings are expressed. In fact there are small but not insignificant differences between the English and Russian plural.

At this stage let us take an example where the differences are more obvious. English and Russian have singular and plural, while Sanskrit had singular, dual

[1] An alternative terminology has number as a 'category' and singular as a 'property' or 'feature' (Matthews 1991: 39–40). We retain 'category of number' as a wider term, to include all manifestations of number, including number words (for which see §5.1), as opposed to the category of gender, tense and so on.

and plural. Sanskrit used the dual for referring to just two real world entities, and the plural for more than two. Clearly the plural does not have the same meanings in English or Russian as in Sanskrit: it covers cases where two items are referred to in the former languages but not in Sanskrit (a general point made by Saussure 1916/1971: 161).

How then can we compare, say, the plural in different languages? The first answer must be 'with care', to ensure that we are indeed dealing with comparable things. Provided first that we can establish that in each language under consideration there is a regular correspondence of meanings and forms which allows us to demonstrate the existence of a number system, we can then compare the values in the two languages. Typically the value which includes in its meaning reference to the largest sets of referents will be called 'plural', whatever other meanings or restrictions it may have. It is therefore reasonable to compare the degree of overlap between the use of the plural in the different languages. (But we must be careful; for instance, in descriptions of Cushitic languages 'plural' is used to indicate a set of forms whose use does not always correspond to plural in most other languages, as we shall see in §6.1.1.) We shall find potentially confusing terminology for other number values too. The term 'collective' is used quite differently in different traditions. And there are subtler problems, for example where 'trial' is sometimes used of forms historically related to the numeral three but currently used for a small number ('few'). Thus although 'trial' is a possible term for the form in such languages we shall choose our terms favouring meaning and so would call this a 'paucal'.[2] The important thing in such cases is to be explicit about what is intended. As a general rule we shall give priority to meaning in our choice of terms.

The last question we need to tackle at this early stage in our investigation is: What type of category is number? The obvious answer, certainly for speakers of Indo-European languages, is that it is a nominal category, affecting primarily nouns and pronouns. In our examples above, the difference between *head* and *heads*, *golova* and *golovy* is the number of heads involved. Of course, number may be shown by verbs too in English (and Russian, and many other languages):

(1) my dog watches television
(2) my dogs watch television

Though number is marked on the verb here as well as on the noun, the essential difference between (1) and (2) is, of course, the number of dogs involved. This point can be seen particularly clearly in these examples:

[2] Though we favour terms based on semantics, this does not entail any claim that particular number values are always used according to meaning. Thus we label the form *cat* 'singular' because it is regularly used in expressions referring to a single cat; but there are also expressions, like *more than one cat*, where the singular is out of line with the semantics.

substantial amount of research in formal semantics on the nature of plurals; key references will be found in §2.5. That work is starting to connect with the wide range of number use in natural language: it is hoped that this book will be of use to semanticists for that purpose.

The orthography used in examples normally follows that of the source, to enable the reader to refer back easily, while for examples originally in a non-Roman script a standard transliteration is used. Examples are followed by glosses. These are intended to clarify the point at issue rather than being full glosses. When items are segmented in an example, this segmentation is mirrored in the gloss: *smile-s* smile-3.SG, in which the *s* is glossed as '3.SG'. Since the *s* cannot itself be segmented into constituent morphs representing third person and singular number separately, the glosses for these, abbreviations in this case, are joined by a stop. Abbreviations are listed on page xix.

2
Meaning distinctions

In this chapter we concentrate on the possible meaning distinctions in number systems. Often the situation in languages like English is taken as normal, whereas it represents only one of the possibilities. We will first consider whether number needs to be expressed; we shall see that for some languages the expression of number is in a sense optional, while in others it is a category which speakers cannot avoid. To investigate these systems we shall first consider the notion of 'general' number as a meaning distinction and base a partial typology upon it (§2.1). We then narrow our attention to the cases where number is expressed, and establish the main types of distinction within the category (§2.2). Thus §2.1 is devoted to the opposition of number and 'non-number', while §2.2 examines the possibilities within the number domain. In §2.3 we propose a typology, systematizing the material examined so far, and we go on to show that languages may simply not have a number system (§2.4); then we consider approaches to number within formal semantics (§2.5).

Our aim in this chapter is to find all the possible distinctions. At this stage we shall not be concerned about the type of nominal we look at, so long as we find those which show the greatest differentiation. Keeping any particular nominal 'still' as it were, we shall see how many different numbers it may have available, in the most favourable contexts. In the next chapter we consider the possibilities along the other dimension (holding a particular number distinction constant we shall examine which nominals can be involved in it). Then in chapter 4 we integrate the account of the possible number systems with the possible patterns of involvement of different nominals. In these chapters we concentrate on the semantic distinctions and we leave detailed consideration of the means used to express them for chapter 5. The more general typological point of this chapter is that as a first step we must cast our nets widely; a category as familiar as number proves to be remarkably varied once we examine a broad range of languages.

2.1 General number

In English we are usually forced to choose between singular and plural when we use a noun. However, there are languages for which number is less dominant,

languages in which the meaning of the noun can be expressed without reference to number. We shall call this 'general number', by which we mean that it is outside the number system. Various other terms have been used: Jespersen (1924: 198) writes of the lack of 'a common number form (i.e. a form that disregards the distinction between singular and plural)'; Hayward (1979) introduced the term 'unit reference', the German tradition is to use 'transnumeral', as in Biermann (1982). We follow Andrzejewski (1960) in using the term 'general'.

Given our definition of the meaning of general number, let us analyse its place in the number systems of various languages. It is found in the Cushitic language Bayso, which at the last count had a few hundred speakers on Gidicho Island in Lake Abaya (southern Ethiopia) and on the western shore of the lake. Bayso nouns have a form which represents the general meaning, that is, it is non-committal as to number (Corbett and Hayward 1987). *Lúban* 'lion' denotes a particular type of animal, but the use of this form does not commit the speaker to a number of lions: there could be one or more than that. Other forms are available for indicating reference specifically to one or to more than one lion, when required.

The situation in which a language would have both a form outside the number system and a minimal number contrast can be diagrammed as in figure 2.1. The meaning of the noun may be expressed independently of number, as occurs with the general meaning, or it may be expressed within the number system, which at its simplest means there will be a choice of singular or plural. In Bayso these meanings all have independent forms: as we have already noted, *lúban* 'lion(s)' is the general form. For reference to one lion, especially for reference to a specific lion, the singular *lubántiti* 'a/the particular lion' is used.[1] Bayso actually has one more possibility than the system in figure 2.1, since for reference to a small number of lions, two to

[1] Specificity plays a role with the other numbers too; for instance, in phrases consisting of noun plus numeral, number must be marked on the noun if there is a determiner or other modifying element in the phrase, but otherwise it need not be (Dick Hayward, personal communication). Compare:

(i) hiṇi deelel-jaa lama emeten
this.PL young.woman-PAUCAL two came.PL
'these two young women came'

(ii) deelel / deelel-jaa lama emeten
young.woman.GENERAL/ young.woman-PAUCAL two came.PL
'two young women came'

A subscript point (superscript in the case of 'ṗ') indicates glottalization, as in *hiṇi* 'these'. In the case of obstruents, glottalization is manifested as an ejective, but in the case of sonorants, it involves a preceding or following glottal stop. The labelling of the forms (as again with *hiṇi* 'these') is difficult, since controller and target numbers do not match in Bayso; see Corbett and Hayward (1987: 11–12) and §6.1.1. In (ii) use of the paucal for the noun is possible, but so is general number, while in (i) the paucal is required.

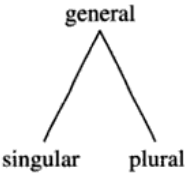

Figure 2.1 *System with separate general number*

about six, the paucal *lubanjaa* 'a few lions' is available; for more than that the plural *lubanjool* 'lions' can be used. We return to the relations of paucals to plurals in §2.2.4 and §2.3.2. Examples include (Dick Hayward, personal communication):

(1) lúban foofe
lion.GENERAL watched.1.SG
literally: 'I watched lion' (it could be one, or more than that)

(2) lubán-titi foofe
lion-SG watched.1.SG
'I watched a lion'

(3) luban-jaa foofe
lion-PAUCAL watched.1.SG
'I watched a few lions'

(4) luban-jool foofe
lion-PL watched.1.SG
'I watched (a lot of) lions'

For easy comparison with the systems which follow, let us use an artificial version of English based on Bayso. A language which distinguishes general, singular and plural allows the following contrasts:

(5) I saw dog (general: one or more)

(6) I saw dog-a (singular: exactly one)

(7) I saw dog-i (plural: more than one)

This three-way system is found in the Fouta Jalon dialect of Fula, which has over two million speakers in Guinea (and over 100,000 more in Sierra Leone and Senegal), but not in other dialects of Fula. Moreover, it is restricted to a part of the noun inventory. Fula nouns typically carry a marker which indicates number: *caa-ngol/can-ɗi* 'river/s', *gabb-ii/gabb-i* 'hippopotamus/es' (data are from Antonina Koval′ 1979 and personal communications, Barrie Evans 1994 and personal communications). In most instances, when a noun is used the speaker is required to

indicate singular or plural number, since one or other marker is required. But some nouns have a third form, which does not have such a marker:

general	*singular*	*plural*
toti 'toad(s)'	totii-ru 'toad'	totii-ji 'toads'
nyaari 'cat(s)'	nyaarii-ru 'cat'	nyaarii-ji 'cats'
gerto 'hen(s)'	gerto-gal 'hen'	gertoo-ɗe 'hens'[2]
boofo 'egg(s)'	woofoo-nde 'egg'	boofoo-ɗe 'eggs'
biini 'bottle(s)'	biinii-ri 'bottle'	biinii-ji 'bottles'

The forms which have no suffix express general meaning, that is they are used when number is irrelevant, for instance (Koval′ 1979: 11):

(8) ko biini tun waawi marde beere
PARTICLE bottle only can.PERFV preserve beer
'only a bottle/bottles can preserve beer'

Various nouns are able to show general number; as our list indicates, those denoting animals are well represented, and nouns denoting humans are included too. Barrie Evans (personal communication) gives interesting statistics on the availability of general forms; his database includes 180 items with a general form (11.5 per cent of all nominals in the database). They are spread across about half the genders of Fula; in the human gender, around 30 per cent of the nouns have general forms; thus 70 per cent do not, yet many nouns lower in terms of animacy do have general forms. In addition, there is the interesting restriction that the form without the suffix must have at least two syllables. Compare the following examples (Koval′ 1979: 12, 22):

(9) nyaari peɗay
cat(s) scratch
'a cat scratches/cats scratch'

Here the unsuffixed form, expressing general number meaning, is used (the singular, as noted above, would be *nyaarii-ru* and the plural *nyaarii-ji*). This is not possible in the next example:

(10) pucc-u latay
horse-SG kick
'a horse kicks/horses kick'

Here the unsuffixed form would consist of a single syllable, and so the singular form must be used to express general number. A further limitation on the use of

[2] The symbol *ɗ* indicates a preglottalized *d*.

Suppose that in a particular language it becomes established that number (plural) is to be marked where appropriate in particular circumstances, for instance, when humans are referred to. This means that all speakers treat number as mattering for human referents. A characteristic which favoured marking of number can come to favour it so much that number marking there is as good as obligatory. Then, if plural is not specified on a noun phrase with a human referent, the form with no number indication will imply singular. Number marking would have come to be obligatory for part of the system, but there could still be general number elsewhere (as we shall see in §4.5, it is common for there to be different systems in operation for different types of nominal). For instance, Smirnova (1981) shows that general/singular versus plural is a widespread system in Iranian languages, but not for nouns denoting humans; the latter have singular versus plural. And in Vai, a Northern Mande language (Welmers 1976: 45–6), there is a similar system. These reflect the tendency for number to be marked for nouns higher on the Animacy Hierarchy (see §3.1).[9]

Languages which have general/singular, for varying types of nominal, can be found widely distributed in the world. They include various West African languages in addition to Vai just mentioned (Manessy 1968), then also Sango (Samarin 1967: 134–7), Amharic (David Appleyard, personal communication) and Mangap-Mbula (Mangaaba-Mbula, Bugenhagen 1994). According to Aikhenvald (1994: 432) it is found in the majority of languages of South America. Elsewhere a particularly clear case is Even (sometimes called Lamut), a Tungusic language spoken by some 6,000 people scattered in north-east Siberia. Benzing (1955: 50) gives the following examples:

(14) zawod-la bəj gurgəwci-n
factory-LOC man work-3.SG
'in the factory, a man works/men work'

under which plural can be indicated (for instance, 'a noun is marked as plural only if it denotes humans . . .'). But this runs the risk of conflating two separate issues: we must specify the nouns for which marking is possible (for which see §3.2) and, as discussed above, the pragmatic conditions under which this possibility of number marking is actually taken advantage of. For an interesting case see the analysis of Tolai, a Melanesian language of New Britain (Mosel 1982: 129–40), where general number is compared to marking by number words (§5.1); both the type of noun and the context have a role.

[9] However, a general versus plural system may affect the pronouns too. Thus in Asheninca (an Arawakan language of the central Peruvian highlands; Reed and Payne 1986), the pronouns distinguish inclusive from exclusive for the first person (the inclusive implies more than one referent), and otherwise they have a second person form and a third person (distinguishing masculine and feminine). These pronouns may take the plural marker (*-payeeni*), which is the regular plural marker for nouns. However, this marker need not be used and is infrequent. Verbal suffixes which indicate number are also used relatively infrequently. Thus we seem to have a case of general number within the pronominal system. According to Reed and Payne (1986: 325) 'plural distinctions are not an integral part of the pronoun system of Asheninca'.

(15) tala asi gurgəwci-n
here woman work-3.SG
'here a woman works/women work'

Benzing's translations make clear that the general/singular form of the noun, together with singular agreement, can be used for reference to one individual or more than one.[10] A last example of this type is provided by the Austronesian language, Tagalog (David Gil, personal communication). In Tagalog, a form like *aso* can mean 'dog' or 'dogs'. Plural number can be expressed primarily by means of the interesting element *mga* [maŋa], probably best analysed as a clitic. It may occur before virtually any constituent, as in:

(16) mga aso
PLURAL dog
'dogs'

Thus the presence of *mga* indicates plurality, but its absence leaves the possibility of singular or general meaning.

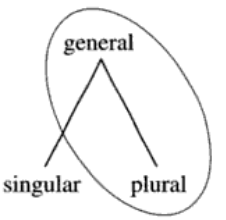

Figure 2.3 *System with general/plural versus singular*

The converse of the widespread system we have been examining would be one in which the general and plural meanings shared a form, as shown in figure 2.3. This system would allow the following contrasts in artificial English:

(17) I saw dog-a (singular: 'exactly one')

(18) I saw dog (general: 'one or more' OR plural: 'more than one')

[10] In the language isolate Nivkh, according to Panfilov (1958: 48, 52), when more than one entity is referred to, the subject and the verb may both be singular, or either one can be singular and the other plural, thus giving three possibilities. Lefebvre (1981: 76) gives examples from Quechua, in which there are four possibilities: subject and verb may each be marked as plural or have no plural marker. She discusses these in terms of the conditions under which plurality may be recovered by the hearer. There is an interesting research area here: we need to establish the languages where there are general forms and where number is simply not an issue, and to distinguish them from languages where number marking may be omitted, provided it can be recovered.

In this system, examples like (18) are vague; they cover the readings of examples (5) and (7) above; (18) could be used either for non-committal reference to a dog or dogs, or to indicate more than one.

This system does not exist in the pure form; that is, no language employs it as the normal case, forming the basis of its number system. This gap may be explained by appealing to markedness. It is generally accepted that the singular is the unmarked number as compared to the plural. When one member of an opposition is neutral with regard to the opposition, then this should be the unmarked form. Thus if one out of singular and plural can also be used as a number-neutral form (that is, having general meaning), then we would expect this to be the singular. (See §5.3.4 for zero expression, and §5.3.6 for further discussion of markedness.)

While the system of general/plural versus singular (figure 2.3) has not been found in pure form, it is nevertheless well established for sections of the noun inventories of particular languages (where the other system of general/singular versus plural is found for other parts of the noun inventory). According to Dick Hayward (personal communication) Cushitic languages normally have a form which is outside the number opposition. However, we frequently find that this general form is the same as the singular for some nouns, and the same as the plural for others, though fewer. This situation can be illustrated from the Cushitic language Arbore (Hayward 1984: 159–83). Many Arbore nouns have a general form (unit reference in Hayward's terms) which is 'semantically non-specific as to the "singular : plural" distinction' (1984: 161). We find pairs like the following:

general	*plural*
kér 'dog(s)'	ker-ó 'dogs'
garlá 'needle(s)'	garlá-n 'needles'

But we also find pairs in which the general form contrasts with a 'singulative'. ('Singulative' is a term relating to form; in meaning such forms are singular; 'singulative' is normally used when the singular form is derived from some other form, typically a collective or general form, and carries a number marker. It is not a significant term and we use it here only because we are quoting from sources which use it.)[11] Examples of general contrasting with singulative include:

singulative	*general*
tiis-in 'a maize cob'	tíise 'maize cob(s)'
lassa-n 'a loaf'	lássa 'bread'
nebel-in 'a cock ostrich'	nebel 'ostrich(es)'

[11] If one uses 'singulative' consistently for singular forms which correspond to a more basic plural form, then it would be logical to use the term 'plurative' for plural forms which correspond to a more basic singular, as in *kér* 'dog' ~ *ker-ó* 'dogs' above, as suggested by Dimmendaal (1983: 224).

There are also some instances in which the singulative contrasts with the plural:

singulative	*plural*
heero-nté 'a flood'	heeró-n 'floods'
farr-it 'a finger/toe'	farr-ó 'fingers/toes'

Thus nouns regularly have two number forms only. For most nouns, one of these forms can be used with general meaning; but this general meaning form may be paired with a singulative or with a plural. Other nouns have singulative versus plural.

The situation found in the Borana dialect of Oromo (another Cushitic language, previously called Galla) is comparable and equally interesting. When a noun has two forms these are normally singulative and general, or general and plural (singulative and plural is very rare). Andrzejewski (1960: 68) reports that:

> the vast majority of Nouns occur normally only in their *General Forms.* The *Plural* and *Singulative Forms* are seldom used and in fact it is possible to listen to conversations among the Borana for a whole day or even longer without coming across one *Plural* or *Singulative Form*. Nevertheless, there are Nouns whose *Plural* or *Singulative Forms* I have found in common use.

This quotation makes it quite clear that the difference in forms is not simply a morphological one; in this Oromo dialect, nouns have a general form, the one normally used, which gives no information as to number. A singulative or plural form may well be available when specificity as to number is required, but this need occurs infrequently. What then is the relative importance of singulatives as opposed to plurals in Oromo? Singulative forms are very rare and, with one exception are found only with nouns denoting persons (Andrzejewski 1960: 64n). The exception is the word for 'young bull'; cattle are of great significance in Oromo culture. Thus the general form appears to coincide with the singular more often than with the plural. That is to say, nouns following the model of figure 2.2 outnumber those following figure 2.3. This is what we would expect, if the singular is indeed unmarked with respect to the plural.

Most studies of Cushitic languages have, quite naturally, been concerned to establish the forms involved, and less has been done on the interesting question of the semantics of these systems. In languages where the preponderance in use of the general form is less great, where many nouns commonly use two number forms, it would be interesting to establish when the general form retains its general meaning and when, by regular contrast with either a singulative or a plural form, it becomes restricted in meaning to singular or plural. (There is further discussion of general number in §3.4, §4.5.3, §4.5.4 and §4.5.5.)

Before going on to the last system in our typology, it is perhaps worth making it quite clear how these systems differ from that of English. Of course, it is possible in English to use the singular, the unmarked number, for more general reference, as in: *the lion is a noble beast.* Here we are not referring to one lion, but to lions more generally, so this usage is sometimes called 'generic'. We can also say *lions are noble beasts.* Hence in this type of expression, number is not particularly important.[12] But in most contexts we are forced to choose singular or plural, and the choice is significant. Imagine that I can see three lions in the garden. If I then say *there's a lion in the garden*, this is true but misleading, since the use of the singular in English implies that there is exactly one lion in the garden.

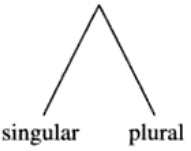

Figure 2.4 *System with singular versus plural*[13]

This leads us to the last possibility in our typology, and English is a good example of this. This is a system in which number must normally be expressed: there is no way of expressing general meaning (except by circumlocution), no forms outside the number system. We have the picture in figure 2.4. Let us consider our examples again, this time in normal English:

(19) I saw a dog (exactly one)

(20) I saw some dogs (more than one)

There is no form which would be appropriate for the readings of both (19) and (20). Indeed, this situation is taken by many people, including large numbers of linguists, to be completely normal and yet, as we have seen, there are many languages which employ rather different systems.

2.2 Number values

Having established the place of general number as outside the number system, we now turn to the distinctions which may be drawn within number. We shall look at that part of the system in a given language which is of greatest interest in terms of

[12] Nevertheless, *the lion / lions / a lion* are not interchangeable in such uses; see Rusiecki (1991) for an interesting discussion of the differences.

[13] Since English does not have general number, whether expressed by a unique form as in Bayso or by a form shared with another as in Japanese, the top node is unlabelled, leaving the opposition between singular and plural.

communication) as has Marrithiyel, another Daly family language (Ian Green 1989: 136–9). It occurs too in Anindilyakwa, the language of Groote Eylandt (the large island in the Gulf of Carpentaria, Northern Territory, Australia, about 1,000 speakers).[15] In traditional Anindilyakwa it would have been more correct to call the form a paucal since it could be used for three to five, and not just for three; however, younger speakers, who have been through school, tend now to use it as a strict trial for three only (Velma Leeding 1989: 225 and personal communication).[16]

We have seen languages with genuine trials, appropriate just when referring to three entities. There is a question as to whether there are also languages with quadrals (for reference to four entities). However, having raised the issue of paucals, we shall first continue the analysis of these, and only then return to the question of quadrals.

2.2.4 The paucal

The paucal is used to refer to a small number of distinct real world entities. It is similar to the English quantifier 'a few' in meaning, particularly in that there is no specific upper bound that can be put on its use. (Its lower bound, like that of the plural, will vary according to the system in which it is embedded.) As noted earlier, Bayso has a paucal, with singular and plural, giving the following system (in addition to general number):

singular paucal plural

The paucal is used in Bayso for reference to a small number of individuals, from two to about six. Bayso has this system in nouns, as we saw in §2.1, but not in its pronouns (§4.5.4).[17] The paucal is also found in Avar, but as a minor number there (§4.2.3).[18]

[15] The trial is also facultative in Ngan'gityemerri, Marrithiyel and Anindilyakwa. For other languages with trials we do not have enough information to know whether they are facultative or obligatory; it may be that trials are always facultative.

[16] A clear inflectional trial is reported in Lenakel and other Tanna languages, which are part of Oceanic (Lynch 1977); Lynch is specific about the trial 'marking three only' (1986: 262). For evidence on the use of the trial in languages of Victoria and in Arabana see Hercus (1966); it is suggested (1966: 337) that the forms there originally meant 'a group of people standing or sitting together or associated with each other in some way'.

[17] Walapai (Hualapai), a Yuman language of north-western Arizona appears to have singular–paucal–plural for nouns, pronouns and verbs (Redden 1966: 149–50, 159); Pilagá, a Guaykuruan language of Argentina, has this system for classifiers, but the full number system includes a dual (Vidal 1997); Kayapó, a Jê language of Brazil, may have singular–paucal–plural in its pronoun system (Wiesemann 1986: 361, 368) but this requires further investigation.

[18] The term 'restricted plural' may be found in place of 'paucal'. Sometimes the term 'paucal' is used for forms that are required with lower numerals. For instance, in constructions with

Systems with just a paucal in addition to singular and plural are rare. It is much more common to find it with a dual too, giving this system:

singular dual paucal plural

Here the meaning of the paucal changes to exclude two. This system is found, for instance, in Yimas, a Lower Sepik language with 250 speakers in the Sepik Basin of Papua New Guinea. The paucal is found in the pronoun and in the pronominal affixes on the verb. 'The paucal expresses a set of a few; more than two and usually less than seven, but the exact number varies quite widely according to context. Prototypically, however, it refers to a class of three to five individuals, and is always restricted to humans' (Foley 1991: 216). The restriction to humans is specific to Yimas, of course. The related language Murik formerly had this four-way number system for pronouns, nouns and agreeing adjectives (Foley 1986: 221–2). Another language with the system is Meryam Mir (Trans-Fly family) spoken in the eastern Torres Strait islands, but being squeezed out by Torres Strait Creole (Piper 1989). It too distinguishes singular, dual, paucal and plural (through complex morphology).

Dual and paucal are found in Fijian; for Boumaa Fijian, a dialect mutually intelligible with Standard Fijian, Dixon states that there is no fixed paucal–plural boundary, except that plural must be more than paucal. He points out a good example of its use in one of his texts:

> It is an announcement about village work, which every adult person must do each Tuesday. The message is called out, by Suliano, three times, each in a different part of the village; it should reach the ears of one-third of the villagers each time. Suliano uses the paucal second person pronoun in addressing his listeners – you (*dou*, 2pa) listen, our (*odatou*, 1incpa) people in this part of the village. Then he says: I'm calling out the tasks of you (*omunuu*, 2pl), the women, for today because this is our (*oda*, 1incpl) day for village work. (Dixon 1988: 52; 'inc' = inclusive, 'pa' = paucal, 'pl' = plural)

Here the paucal is used for about twenty people, one-third of the adult villagers, and the plural for them all (about sixty). Schütz (1985: 251) also discusses the

the numerals '2', '3' and '4' in Russian, when they are in a direct case form, a special form of the noun is required, almost always the same as the genitive singular, but unique at least in terms of stress for a few nouns, for example *dva časá* 'two hours, two o'clock' (the genitive singular is *časa*). However, this special form depends entirely on the presence of the numeral, it is not part of the number system. This is shown by that fact that it is not possible to say *časá* meaning 'a small number (2–4) hours'. Hence the use of 'paucal' is inappropriate here. The use of the genitive singular is taken up in §6.7.1.

paucal in Fijian, and says that it can be used for three and for twelve. Some consultants put the limit at fifteen, others put it higher. He points out that contrast is more important than the specific number, and mentions a text in which approximately thirty people are referred to sometimes with the paucal and sometimes with the plural. Andrew Pawley (personal communication) also says that its range varies considerably according to the situation.[19]

This system (with dual and paucal) is found widely in other Oceanic languages besides Fijian, for instance in Paamese, spoken in Vanuatu. The factors governing the choice of paucal and plural have been well described:

> The basic factor that is involved is the absolute size of the group being referred to. Intersecting with this parameter however is the question of relative size, i.e. whether the group being referred to is contrasted with some larger group within which it is subsumed. When the absolute number is low (say between three and about half a dozen), the paucal is generally used, whether or not there is any contrast with a larger group. (However, the plural will still very occasionally be used even with these low numbers when there is no such contrast.)
>
> When the absolute number is in the middle range (say, between about half a dozen and a dozen or so), the most significant parameter is that of relative number. For instance, one's own patrilineage will be referred to paucally when it is contrasted with the village as a whole, which will be plural. On the other hand, the patrilineage will be expressed in the plural when contrasted with the nuclear family, which will be in the paucal.
>
> As the absolute number increases over the middle range, relative number again becomes less significant, and the plural is generally used for all numbers over a dozen. (However, even with very large numbers, the paucal is occasionally used when the contrast in number is expressed. So, while the entire population of Paama will normally be expressed in the plural, even when contrasted with the country as a whole, it has been heard referred to paucally.) (Crowley 1982: 81)

Staying within Oceanic, the singular–dual–paucal–plural system also occurs in Manam, spoken on islands off the north coast of Papua New Guinea (Lichtenberk 1983: 108–9), and in Ambrym, which has around 400 speakers on the island of

[19] For the system in Wayan (a local language within the Fijian subgroup with 2000 speakers on two islands at the western margin of Fiji) see Pawley and Sayaba (1990: 152, 156).

Ambrym in the New Hebrides (Paton 1971: 12–13). This paucal is clearly a trial in origin, but is now used for small groups of persons. There are numerous instances of former trials becoming paucals, for instance in Kwaio, Sa'a, Langalanga and Lau, all Malaitan languages spoken in the Solomon Islands (Simons 1986: 33).[20]

It is found in Australian languages too, in Ungarinjin (Rumsey 1982) and in Murrinh-Patha, a Daly family language of north-west Australia. As I. Green (1993, chapter 6) points out, there have been two slightly different assessments of its paucal (which is found in the verb paradigm and in the free pronouns). Walsh (1976: 150) says the paucal is for 'no less than three individuals and up to about ten individuals', while Street (1987: 49) gives its range as three to approximately fifteen. This suggests again that its use varies from context to context.

Table 2.2 *Independent pronouns in Lihir*

	singular	dual	trial	paucal	plural
1 exclusive	yo	gel	getol	gehet	ge
1 inclusive	—	kito	kitol	kitahet	giet
2	wa	gol	gotol	gohet	go
3	e	dul	dietol	diehet	die

The paucal has been found in a more complex system too, with four other values, possibly in this configuration:

singular dual trial paucal plural

This system in question is found in Lihir, an Oceanic language spoken on a group of tiny islands off New Ireland (PNG). It is a member of the New Ireland Network, but does not belong to the same branch as Sursurunga and Tangga to be discussed below (Ross 1988: 258). The data are from Malcolm Ross (unpublished fieldnotes), from the dialect spoken on Lihir Island itself. As table 2.2 shows, five numbers are distinguished in each person, the only gap being the logically necessary one in that the first person singular cannot be inclusive. The same distinctions are found in the set of possessor suffixes (used on inalienables). The problem here is not the paucal, whose status is sure, but the 'trial', whose usage is not known. (If it is a paucal, giving the language a paucal and a greater paucal, then it would have a system like Sursurunga, discussed in §2.2.5 below.) Whether we have a trial or two paucals, Lihir is of considerable interest as a language with the maximum number of number values.

[20] The dual is less prone to this development. However, Blanc (1970: 45) notes that the former dual can be used for a small number in Arabic dialects, which suggests that a development of dual to paucal is possible.

Table 2.3 *Lihir number markers (Malcolm Ross, personal communication)*

Lihir number markers		Proto-Oceanic numerals		Lihir numerals	
-l	dual	*rua	two	lo	two
-tol	trial	*tolu	three	laktul	three
-het	paucal	*pati	four	burut	four

The origin of the forms can be traced back around 3,500 years to the Proto-Oceanic numerals, as given in table 2.3. Note that it is not only the paucal whose origin is obscured; comparison with the modern numerals shows that the origin of the dual and trial is no longer clear either. Such examples refute any suggestion that larger number systems might be no more than the use of numerals: in Lihir the number markers and numerals are evidently distinct.

We have seen that the paucal may be found in different systems, and the lower bound on the number of entities referred to varies accordingly.

2.2.5 The question of quadrals

We now consider whether there are languages with the following system:

singular dual trial quadral plural

Such languages would have a quadral, a set of forms specifically for the quantity four. If such languages exist, they are rare and all the claims come from within the Austronesian family. A well-documented suggested case is Sursurunga (Hutchisson 1986, and personal communications), which has some 4,000 speakers in southern New Ireland. It is one of the South New Ireland/West Solomonic languages, which form part of the New Ireland Network, that being a branch of Melanesian, within Oceanic, in turn part of Austronesian (Ross 1988: 258). The forms labelled quadral are restricted to the personal pronouns, but are found with all of them, the first person (inclusive and exclusive), the second and the third.

We retain the term 'quadral' in table 2.4 and in this section, while we give the reasons why it should be replaced for Sursurunga (as indeed should 'trial'); in later references to Sursurunga we replace it. Here is an example of a quadral form in use:

(23) gimhat káwán
1.EXCL.QUADRAL maternal.uncle:nephew/niece
'we four who are in an uncle-nephew/niece relationship'

Besides being used of four, the quadral has two other uses. First, plural pronouns are never used with terms for dyads (kinship pairs like uncle–nephew/niece in (23))

'When we (i.e. Sursurungas) want to build a cook house, we go up to the bush, that is we men and women.'

Here Hutchisson believed a quadral would not be used, since the group (Sursurungas in general) is too large. Similarly in the following example:

(27)

Má	máhán	a	kis	main	si	git	arwat
and	war	3.SG	exist	here	to	1.PL.INCL	enough

mai	a	hit	á	bet.
with	3.SG	seven	RELATER	year

'And the war was here among us (i.e. in this area) for seven years.' (From a story about World War II.)

Use of the quadral in this example would limit the area being referred to, say to a single village or a small group of villages, or it would limit the group, say to those alive during the war and affected by it directly. Hence the plural, as we would expect, is for numbers of entities larger than are covered by the quadral; however, there is no strict dividing line (certainly not at the number five).

If we use semantic labels, as we have done in the rest of the chapter, we should not call the forms trials and quadrals. Both have functions we have seen with paucals elsewhere. We may therefore represent the system in Sursurunga like this:

singular dual paucal greater paucal plural

The system is no less interesting since it has a well documented five-valued number category.

Another language with five values is Sursurunga's close relative Tangga (Capell 1971: 260–2; Beaumont 1976: 390; confirmed by Malcolm Ross, personal communication; note that Capell and Beaumont used the term 'quadruple'). Here we know that there are five forms, but we do not have such detailed information as we have for Sursurunga. Yet it seems clear that the forms which have the numeral 'four' as their source are not quadrals but rather paucals (Malcolm Ross, personal communication citing Maurer 1966; this is also Schmidt's view given in Capell 1971: 261). Unfortunately, as with Lihir, we have no information on whether Tangga has a genuine trial or whether it has two paucals.

The third language which has been claimed to have a quadral is more distantly related; it is Marshallese, a member of the Micronesian group within Austronesian, with some 20,000 speakers on the Marshall Islands. It has five number forms for the first, second and third person pronouns (Bender 1969: 8–9). We shall return to it when we discuss facultative numbers in §2.3.3. As in Sursurunga, the form which has been called the quadral has an additional use: with groups of more than four it is often used rhetorically to give an illusion of

intimacy (Bender 1969: 159). Again, then, it appears that this may not be strictly a quadral; we shall therefore treat it as a paucal. Byron Bender (personal communication) has no evidence for any comparable extension to the trial, so we shall treat Marshallese as having singular, dual, trial, paucal and plural.

These are the three best claims for quadrals. There are several false trails in the literature, that is, suggestions of other Austronesian languages with quadrals, which turn out in fact to have four number values not five. In such cases, the plural may have a form in which the numeral four can be reconstructed. We return to the development of such forms with plural meaning in §9.1.2; their existence as plurals suggests that there might have been instances of the quadral number since lost. Or it may be that once the numeral four becomes grammaticalized as a number value, it is inevitably used for groups larger than four. We have found no clear case of a quadral, by which we mean a grammatical form for referring to four distinct real world entities in the way that trials refer to three.

2.2.6 Greater numbers

Languages may have a secondary split into normal and 'greater' (sometimes termed 'lesser' and 'greater') within certain number values. The two which may be split are the paucal and the plural. There are relatively few known cases of split numbers and the account here is tentative.

Consider first the paucal. It is rare to find a split in the paucal, but that is exactly what we found in Sursurunga (§2.2.5). Either set of forms (those labelled 'trial' and those labelled 'quadral') would independently be reckoned a paucal on semantic grounds. We therefore treat them as a paucal and a greater paucal.

Splits within the plural are more common. Claiming such a split, into greater and normal plural, implies that both would independently count as plural. Since even the lesser is a plural (used where languages with just one plural would use it), we shall call it simply 'plural'.[23] The 'greater plural' typically implies an excessive number, sometimes called 'plural of abundance', or else all possible instances of the referent, sometimes called the 'global plural'. We shall use 'greater plural' to cover the different types (abundance, global). The evidence is limited, but it comes from a variety of languages and sources, sufficient to indicate that there is an interesting phenomenon that deserves study. More examples with careful descriptions of their meanings would be welcome. Again the definition is a semantic one. There are many instances of nouns taking more than one plural marker (these are 'double plurals', for which see §5.3.6). We are concerned here only with instances

[23] It may be objected that 'plural' is different if in opposition to a 'greater plural' as compared to when it is the only plural. But this is also true, as noted earlier, of 'plural' in a system with a dual and 'plural' in a system without. An advantage of avoiding 'lesser plural' is that this term is sometimes used for 'paucal'.

where the different plural forms have different meanings. (For instances where the ordinary plural is used with this effect see §7.3.2.)

Banyun

A potentially interesting case of a language with a greater plural is Banyun, a language of the West Atlantic branch of Niger-Kordofanian, spoken in Senegal and Guinea Bissau. There is a little information in Sauvageot (1967: 227–8). Nouns typically have singular and plural, distinguished by prefixes of the type shared by many Niger-Kordofanian languages:

(28)

bu-sumɔl	i-sumɔl[24]
SG-snake	PL-snake
'snake'	'snakes'

In addition there is a greater plural (which Sauvageot calls 'unlimited'), in this case *ba-sumɔl* 'snakes', which Sauvageot suggests is used when the number cannot be counted or the speaker feels it unnecessary. There are various prefixes available to signal the greater plural; they are not equivalent in that one of them, *ti-* as in *ti-sumɔl* 'snakes (unlimited)' implies more than *ba-* as in *ba-sumɔl.* Noun phrase modifiers such as adjectives agree, distinguishing the various singular, plural and greater plural classes.[25] (A similar distinction is reported in Senufo, see discussion of Sauvageot's paper, 1967: 236.)

Fula

Related to Banyun, since both are members of the West Atlantic branch of Niger-Kordofanian, is Fula, which is widely spoken across west and central Africa (the Fouta Jalon dialect was discussed in §2.1 above). Here some nouns have two plurals 'one to imply a normal number of items and the other to imply a very large number of items' (Evans 1994: 21.6), for example, as shown in table 2.5. In the

Table 2.5 *Plural forms in Fula*

singular	plural	greater plural	gloss
ngesa	gese	geseeli	field
wuro	gure	gureeli	herd

[24] Sauvageot calls this the 'limited plural'; this term is confusing because it too has been used as a synonym for 'paucal' (for instance, by Capell 1976: 15; see also the quotation from Laidig and Laidig 1990: 92 in §2.2.3). We shall therefore avoid the term 'limited plural'.

[25] The existence of different agreement markers distinguishes Banyun from several languages with a greater plural.

second example, *gure* 'herds (plural)' might be the herds of one man, his flock of sheep, his goats and his cows. This is a case, however, where there is insufficient evidence to know whether we really have two plurals or a paucal and a plural.

Arabic

In Arabic too, there are nouns with two plural forms, and in describing Syrian Arabic Cowell (1964: 369) gives helpful pointers to their status. Recall first from §2.1 that some types of Arabic noun have a general ('collective') form, for instance *dəbbān* for which, given real world considerations, the natural gloss is 'flies'. If it matters to specify one fly, then there is the singular *dəbbāne*; there is a corresponding dual *dəbbāntēn* and a plural *dəbbānāt*. It would appear that anything one might want to say about flies is provided for. But this is one of the instances where there is a fifth form *dababīn* 'many flies'. Cowell treats this as the plural of the collective. Such plurals may function, in his terms, as 'plurals of abundance' (for a formal approach to the semantics see Ojeda 1992a).

This is an instance of a recurring phenomenon, namely the formation of a plural whose predictable function is not required, and which takes on a different one. In Arabic, there is no obvious function for the plural of a collective when there is an ordinary plural available; where the 'extra' form exists, it may take other functions: one is the 'sort' reading as in English (§3.7.2), and another is the 'abundance' reading, so *dababīn* can mean 'various flies' or 'many flies'. When there is a plural of abundance (only certain of the nouns with collective forms have them) this may affect the meaning of the normal plural (sometimes then called the 'plural of paucity' so that the use of the latter implies that the entities referred to are few and are individually discriminated. However, this is not always the case.

Thus some Arabic nouns have two plurals; the relations between them vary. The existence of the greater plural may as it were 'push down' the ordinary plural into the position of a paucal (for the situation in Classical Arabic see Wright 1967: 234; the analysis is not uncontroversial, for discussion see Ratcliffe 1998: 79–81).

Hamer

In Hamer (or Hamar), it has been claimed by Lydall (1976, 1988) that a distinction can be drawn between a plural for a particular number ('particular plural', our 'plural') and plural for all instances ('global plural', a type of 'greater plural'). Hamer is a South Omotic language, which has about 15,000 speakers in the south-west corner of Ethiopia. Nouns have a general form, which, as in languages already discussed, stands outside the number system. Thus *k'úli* means 'goat' or 'goats'. This should be contrasted with the singular forms (Lydall 1988: 81–2), for which see table 2.6. The singular formed with the suffix *-tal-a* is for the male (of animates) and for the 'minor' singular of inanimates, used for something which is

Table 2.6 *Number forms in Hamer*

general form	singular	plural (particular plural)	greater plural (global plural)
k'úli 'goat(s)'	k'últa 'he-goat' k'úllo 'she-goat'	k'úlla 'the goats'	k'últono 'all goats'
goiti/goin 'path(s)'	goita 'path (little used)' goinno 'main path'	goinna 'the paths'	goitino 'all paths'

'small, minor, individual, infrequently used, or seldom found' (Lydall 1988: 79). The *-no* suffix, on the other hand, is used with certain nouns for a female (animate nouns) or, with inanimate nouns, for the 'major' singular, used for large and major things.[26]

The contrast between the number forms in the table, according to Lydall, is between a particular number of items, and all items.[27] Unfortunately Lydall gives little more information on the choice. In particular, it would be good to know how different this system is from those where there is an interaction of number with definiteness (for which see §9.2.4).

Kaytetye

Kaytetye is an Arandic language (part of Pama-Nyungan) spoken in Central Australia. Information is from Harold Koch (personal communication; see Koch 1990 for some of the morphology, and for textual examples see Koch and Koch 1993). The pronouns distinguish singular, dual and plural. Nouns need not mark number: marking is most likely for nouns denoting humans and least likely for those denoting inanimates. In addition to having the three-way distinction of the pronouns, nouns split the plural into a normal plural marked with the suffix *-amerne*, and a greater (global) plural ('all the X in the universe of discourse'), marked with the suffix *-eynenge*. Both plurals can serve as antecedents for the single set of plural pronouns. Kaytetye then is a clear instance of a language with a greater plural alongside the normal plural. (Kaytetye is also interesting in respect of facultative number, a topic we discuss in §2.3.3 below.)

[26] In the case of inanimate nouns, for single syllable nouns, and two-syllable nouns which end in a consonant, major singular and global plural will be identical in form (Lydall 1988: 80): *nu* 'fire' gives *nuno* 'large/main fire' (major singular) or 'fire considered as a whole' (global plural). In addition there is *nuta* 'small fire' and *nuna* 'the (particular) fires' (particular plural).

[27] As a curiosity, Larry Trask points out (personal communication) that there is a celebrated fictional example. J. R. R. Tolkien, in *The Lord of the Rings* invented a number of languages, including the elvish language Quenya. This language distinguishes a global plural from the ordinary plural: *el* 'star', *elen* 'stars', *elenath* '(all) the stars'.

2.2.7 Composed numbers

These are a rare phenomenon. Occasionally we find one number built as it were on another. Recall that we are making semantic distinctions here; using the form of one number as the base for building another is not unusual; it is also common to add number morphology 'to itself' as it were, typically adding plural morphology to an existing plural (giving a 'double plural', §5.3.6). But to base one number semantically on another is rare. Here is an instance from the Celtic language Breton, rightly reputed to have one of the most complex number systems (data from Ternes 1992: 417; compare Trépos 1957: 226–8, 265–6; Hemon 1975: 42; Ternes 1970: 191–2, 200–1; Denez and Urien 1980):

(29)	lagad	daou-lagad	daou-lagad-où
	eye.SG	DUAL-eye	DUAL-eye-PL
	'eye'	'(two) eyes'	'pairs of eyes'

First note that the dual is largely lost in Celtic,[29] but Breton has a new, quite recent dual. It is clearly from the numeral *daou* 'two', but it is a genuine dual; it is obligatory with those nouns which have it (nouns denoting parts of the body and items of clothing that occur in pairs), and when emphasis on two referents is required the numeral is used together with the dual (Ternes 1992: 416–17). To this dual, plural morphology can then be added, and the meaning is then a plurality of duals. The ordinary plural of *lagad* is *lagad-où* 'eyes'; note, however, that this is a restricted phenomenon; the same possibilities are not generally available to Breton nouns.

Breton can also 'compose' plural on plural (Ternes 1992: 415):

(30)	bugel	bugal-e	bugal-e-où
	child-SG	child-PL	child-PL-PL
	'child'	'children'	'groups of children'

In this example the first plural formation is highly irregular, and the second is a common one. The possibility of composing plural on plural is not freely available.[30]

[29] Though see §6.7.1. Breton also has nouns which take what looks like plural morphology but which are interpreted as duals, such as *maneg* 'glove', *manegoù* '(pair of) gloves' (Greg Stump, personal communication). For the plural, a second plural marker is added (an *-où* with the second marker *-ier* gives *-eier*), hence *manigeier* 'gloves'. This is not a composed form however: it does not mean 'several pairs of gloves', but merely 'gloves'. It can be used, for instance, for three gloves.

[30] A similar effect is found in Finnish, but in numeral phrases (data from Hurford forthcoming):

(i)	kolme-t	suka-t
	three-PL	sock-PL
	'three pairs of socks'	

Moreover the meaning of the form with two plural markers varies from noun to noun, and according to dialect (Stump 1989: 270–1).[31]

A very interesting and tricky example is Warekena (of Xie), an Arawakan language with some twenty older speakers in Brazil and around 200 in Venezuela. The data are from Alexandra Aikhenvald (1998: 300–4 and personal communications). Number is not required on nouns, but various plural markers are available (so this is a general/singular versus plural system). And once a referent is established as plural, plural marking is not repeated.

Let us look at the plural markers. First there are *-pe* and *-ne*, with *-ne* being used for nouns denoting animate non-humans and a few others, and *-pe* for the rest (human and non-human): *neɹima-pe* 'cousins' (ɹ is a lateral flap), *ʧinu-ne* 'dogs', *muɹupa-pe* 'canoes'. Then there is an 'emphatic plural' for 'very many' namely *-nawi*, as in *atapi-nawi* 'a great many trees'. It does not differentiate according to animacy and is used for larger numbers of entities than is *-pe*. And then, for particular emphasis on plurality, two plural markers can be used together: *-pe-nawi*. This gives:

(31)	abida-pe	abida-nawi	abida-pe-nawi
	pig-PL	pig-GREATER.PL	pig-PL-GREATER.PL
	'pigs'	'very many pigs'	'very many pigs indeed, too many to count'

Note that *-pe* rather than the expected *-ne* is used for this noun. We might think that *-pe* is really a paucal, but according to Alexandra Aikhenvald (personal

Normally the numeral stands in its singular form (the singular of *kolmet* 'three' would be *kolme*) and takes a noun in the singular. In the construction in (i), with the numeral and noun in the plural, the meaning is 'three groups of socks', and the natural group here is a pair.

[31] Comparable examples were recorded in Khamtanga (Chamir), a dialect cluster belonging to the Agaw group within Cushitic, by Reinisch (reported in Appleyard 1987: 252):

(i)	ieferā̀	iefír	iefírt
	child.SG	child.PL	child.PL.PL
	'child'	'children'	'crowds of children'

(ii)	lálā	lal	lálāle
	bee.SG	bee.PL	bee.PL.PL
	'bee'	'bees'	'swarms of bees'

Appleyard was working a century after Reinisch, and though he found similar forms, they were preserved as alternative plurals – his informants did not differentiate them semantically. It may be that dialect differences explain this; but it may also be that the distinction was lost in the intervening period. The point is that a plural of a plural is itself a plural; it would not be surprising if such a distinction, probably available for only a limited part of the noun lexicon, should be lost. Composed numbers are also reported in Arabic (Wright 1967: 190–1, 231–2, discussed in Ojeda 1992a).

communication) it can refer to quite large groups. If we treat *-pe* as a plural, then that makes *-nawi* a greater plural. What then of forms with both *-pe* and *-nawi*? The analysis is not as easy as with Breton, where the existence of the dual makes the distinction between a composed number (a plurality of pairs) and a vague greater plural clear. We should therefore be cautious about claiming that Warekena has a composed number (for a plurality of plurals), but it is a possible position. The alternative would be to claim that it, remarkably, has two greater plurals. In any case, the evidence that is beginning to emerge about Amazonian languages such as Warekena will be of special importance for understanding number systems.

2.3 Number systems (the Number Hierarchy and associated problems)

Having seen the ranges of number values, we would like to be able to predict the possible number systems which natural languages can have.[32] The main claim in the literature is the Number Hierarchy, which is often taken as unproblematic, yet which cannot account for some of the systems we have discussed. We shall first see why this is so (§2.3.1), then consider an alternative proposal (§2.3.2), and finally see how this alternative is also able to account for facultative numbers (§2.3.3).

2.3.1 The Number Hierarchy

A Number Hierarchy has been proposed, along the following lines:

singular > plural > dual > trial

Most of this can be derived from Greenberg's universal 34 (1963: 94):

> No language has a trial number unless it has a dual. No language has a dual unless it has a plural.

In addition to the positions derivable from Greenberg, the hierarchy is usually quoted with the singular in the first position. The logic of doing so is that just as the dual is more marked than the plural, so the plural is more marked than the singular. It might be argued that, in a sense, languages like Japanese (§2.1 above) have a plural without having a singular, and so the singular should not be included in the hierarchy. The view taken here is that such languages give the possibility of not specifying number (and using the form with no marker, the general form), but if number is specified, then there is a singular–plural opposition. The Number Hierarchy, then, is concerned only with the cases where number is specified.

[32] When checking grammars for counter-examples and confirming examples for such claims, it is important that the grammatical system as a whole should be considered; occasionally single lines are quoted from grammars, apparently suggesting quite exotic number systems, which prove to be rather more ordinary when the grammar and the examples are studied carefully.

A few points need to be clarified about the presentation of such hierarchies. The ordering of positions is crucial; the claim is that, for instance, the presence of a trial implies the presence of a dual. This differs from the way we presented the different systems above, where we ordered the number values simply from smaller to larger (for instance, singular, dual, plural). Though the ordering matters, the use of '>' or '<' is not important. We shall use '>' since there are more plurals than duals. This is true in three senses: in the sense of Greenberg's claim, that more languages have plurals than have duals; also in the sense that for languages which have both, speakers use the plural more frequently than the dual, and they also use the singular more frequently than the plural (see §9.3.2 for statistics); and third, for languages which have both values, there will be as many or more nouns with the plural as compared to the dual (§4.1). Those who prefer the use of '<' reflect the idea that the value to the right is the more marked.

There are two problems with the Number Hierarchy, which we will address in turn; the first is the problem of the systems with paucal number and the second is that it makes the wrong predictions when facultative number is involved.

2.3.2 Possible systems of number values: incorporating the paucal

The first problem with the Number Hierarchy is that it does not account for systems which include a paucal. Foley (1986: 133) and Croft (1990: 96–7) suggest this modified hierarchy (changed to our notation):

singular > plural > dual > paucal/trial

This would account for systems like the following:

singular	plural	(Russian)		
singular	dual	plural	(Upper Sorbian)	
singular	dual	trial	plural	(Larike)
singular	dual	paucal	plural	(Yimas)

However, it does not allow for systems which include the paucal in a different combination:

singular paucal plural (Bayso)

Lihir would also be a problem. We must allow for the paucal to be an option at more than one point, which makes it clear that no straightforward hierarchy will be adequate.

To make progress here we need to draw a distinction between 'determinate' and 'indeterminate' number values. These terms are to differentiate situations where, given the knowledge of the real world which the speaker has, we can determine that only one form is appropriate (determinate number) from those where we cannot

(indeterminate). Thus in a language with an obligatory dual, this would be an instance of determinate number, since to refer to two distinct entities only the dual is appropriate. The determinate numbers are basically the numerical ones: plural, dual, trial. Use of determinate number values is agreed across speakers (different speakers agree that, say, the dual is appropriate for referring to two referents),[33] it remains constant for the same speaker across different occasions, and it does not vary according to the referent (thus elephant-DUAL refers to two elephants just as ant-DUAL refers to two ants). The indeterminate number values are the paucal (and greater paucal) and the greater plural. These may vary across speakers (there is no clear dividing line between paucal and plural, recall the discussion of Murrinh-Patha in §2.2.4), for one speaker across occasions, and can vary according to the referent (elephant-PAUCAL may typically refer to fewer real world entities than ant-PAUCAL). While the determinate numbers can be defined in terms of numerals, the indeterminates correspond to other quantifiers: 'a few', 'many', 'all'.

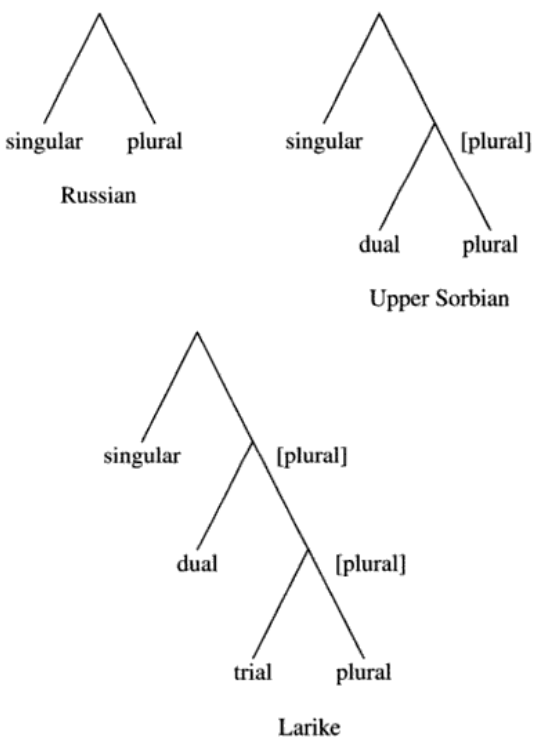

Figure 2.5 *Illustration of possible number systems*

[33] This statement must be qualified when facultative numbers are discussed.

according to whether the use of particular values is obligatory or 'facultative' (Greenberg 1966: 28). In §2.1 we looked at languages where the expression of number as a whole is not required. Here, however, we are concerned with particular values which may be facultative, even though the category of number is being expressed. For instance, in Ngan'gityemerri (see §2.2.3) there is singular, dual, trial and plural. The dual must be used to refer to two entities, the plural must be used for four and more. For three entities, the trial is used when the fact of there being three is salient (for example, at the first mention in discourse) but otherwise the plural is used for three. Recall that the trial is strictly for three, and is not a paucal (Reid 1990: 118–19 and personal communication). The trial is similarly facultative in Marrithiyel (I. Green 1989: 136, 138) and in Anindilyakwa (Velma Leeding, personal communication). Thus number must be expressed (in the sense that there are no general forms for avoiding it) but where it is expressed the use of the trial is facultative.

Consider now the systems with singular–dual–plural. The use of the dual may be obligatory, as in Sanskrit: 'The dual number is in regular use and of strict application, the plural practically never referring to two objects' (MacDonell 1927: 180); 'In Sanskrit, if there are two of something, whatever it is, the structure gives you no option but to use the Dual' (Diver 1987: 103). Or it may be facultative, as in the South Slavonic language Slovene. Here we do not find the same degree of choice as with the Ngan'gityemerri trial, but the important point is that the dual is not obligatory in the way that the plural is in Slovene:

> Normally, dual forms are used in pronouns and in verbal forms whenever two actual referents are involved, be they explicitly mentioned or only implicit. However, in non-pronominal noun phrases with, for example, body parts that come in pairs like 'eyes' and 'feet', dual forms tend to be used only when the quantifiers 'two' or 'both' are explicitly stated in the context, and are replaced by the plural when this quantifier is unstated, even if a pair of referents are obviously implicit. (Priestly 1993: 440–1)

Priestly gives the following example:

(32)

nóge	me	bolijo
foot.PL	1.SG.ACC	hurt.PL

'my feet hurt'

This is a fully appropriate utterance for a normal biped. It is assumed that two feet are referred to, and the dual is not required. We return to the difference between pronouns and nouns in §4.1. For present purposes, the important point is that nominals

express number obligatorily in Slovene; however, for referring to two entities, the use of the dual is not obligatory (see §9.3.2 for statistics on its use). The dual was obligatory in Classical Arabic, but is facultative in modern Arabic dialects (see Blanc 1970: 42–3 and references there; and see Cowell 1964: 367 for Syrian Arabic), as it was too in Classical Greek (see Diver 1987 for interesting statistical data). Just as the plural is different in English (no dual) and Sanskrit (with dual), so it is different in Sanskrit (with an obligatory dual) and Slovene (with a facultative dual). A plural in Slovene may be used for reference to just two real world entities.

Let us now consider how this relates to the Number Hierarchy, repeated here for convenience:

singular > plural > dual > paucal

If we have a system in which use of the plural is not required, then the less marked number, the singular, is used (as discussed in §2.1 above). If we have a system in which the dual is facultative, then in its place the less marked number, the plural, is used. It appears that the hierarchy is making useful predictions, based on markedness. Unfortunately this is only apparent here. Consider again Ngan'gityemerri: it has a trial which is facultative and so we would predict that the less marked dual could be used in its place. But of course this is not the case, the plural is used. This is what is expected if, as discussed earlier, the system is viewed as a set of binary choices (see figure 2.8).

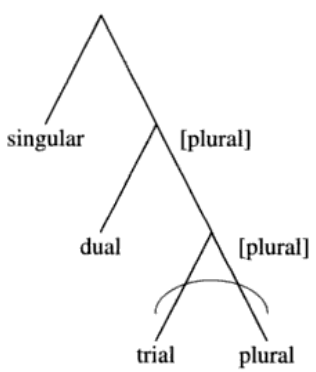

Figure 2.8 *The facultative trial of Ngan'gityemerri*

The point is that the last choice is facultative. If it is removed, as by the arc in figure 2.8, then Ngan'gityemerri has another possible system, singular–dual–plural, and the plural covers the area otherwise covered by trial and plural. In Slovene, the situation is as in figure 2.9.

If the dual–plural choice is not taken up, then the system reverts to a straightforward singular–plural system. It is tempting to suggest that facultative number can

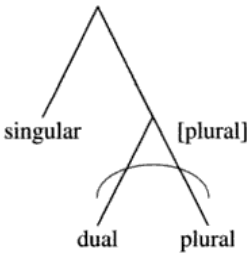

Figure 2.9 *The facultative dual of Slovene*

only affect the 'last choice' of number, as in the examples so far. However, there are three languages which show the situation is rather more interesting.

The first is Larike, which we considered earlier as an example of a language with a genuine trial. Unlike Ngan'gityemerri, it is not only the trial which is facultative, the dual is as well:

> The Larike plural forms may also be used when referring to quantities of two or three. Thus, in spite of the fact that duals and trials are used to specifically denote twos and threes, plural forms can still be used with the meaning of two or more. In these situations, the choice of whether to use plural versus dual or trial forms depends upon the speaker's desire to specify or focus upon the number of the referent nouns. Although the plural forms are probably most frequently used (even when referring to twos and threes), duals and trials are also quite common, and are often heard in routine conversations as well as in more formal language contexts. (Laidig and Laidig 1990: 93)

We represent this system in figure 2.10.

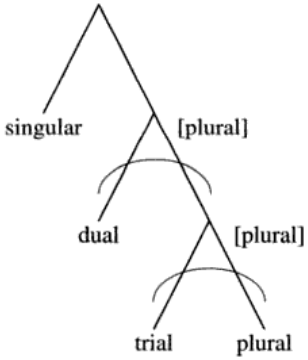

Figure 2.10 *The facultative numbers of Larike*

The second is Longgu, an Austronesian language (Cristobal-Malaitan subgroup) spoken on the north-eastern coast of Guadalcanal, Solomon Islands (data from Hill 1992). Longgu independent pronouns distinguish singular, dual, paucal and plural. The range of the paucal is not fully established, but Hill gives examples of its use referring to five people. There are also 'subject pronouns', which form the first element of a verb phrase and cross-reference the person and number of the subject noun phrase (Hill 1992: 92). These subject pronouns also distinguish four numbers. However, they do not necessarily match the number of the subject; they may match it or, in the third person, the plural may be used where the dual or paucal would be expected. This use of the plural occurs 'when the number of the subject has been established' (Hill 1992: 130) whether by a noun phrase or by a subject pronoun in a previous clause. For example:

(33) m-*arua* goni pilu na, *ara* goni-a
CONJ-3.DUAL build fence PERFV 3.PL build-3.SG.OBJ

pilu-i liva'a-na pilu ni boo-i m-*arua*
fence-SG like-3.SG fence LIG pig-SG CONJ-3.DUAL

na'i-a i ei
put-3.SG.OBJ LOC there

'and they both built a fence, they built a fence like a pig fence and they both put it there' (Hill 1992: 131)

The first subject pronoun *arua* is dual, and shows that there are two referents; the second clause partly repeats the first, and here the plural *ara* is used, rather than the dual. In the third clause, which involves a new event, the dual is again used. There is a similar usage with the paucal and, according to Deborah Hill (personal communication), replacement by the plural is more likely to occur with the paucal than with the dual. We represent this system in figure 2.11.

The more important use of number is its referential use with independent pronouns, where this facultative use does not apply; facultative use is restricted to subject pronouns, which we might argue are like agreement markers. However, we would expect the constraints we have developed to play a role here too, and indeed they do, in that the singular–plural choice is not affected, while the other two choices are. It is interesting too that the paucal is more likely to be replaced by the plural than is the dual.

The third language which requires us to weaken the constraint that only the last number choice can be facultative is Marshallese. Recall from §2.2.5 that the first, second and third person pronouns all have five number forms (Bender 1969: 8–9): singular, dual, trial, paucal (Bender's 'quadral') and plural. The dual, trial and

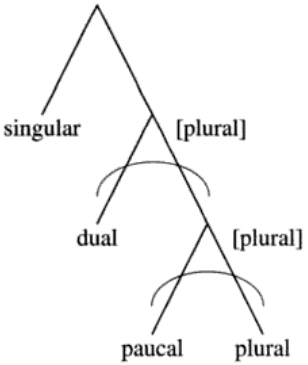

Figure 2.11 *The facultative numbers of Longgu*

paucal are all formed from the plural by the addition of regular suffixes and they are treated syntactically as singular (Bender 1969: 5). However, the important point is that according to Bender the use of these forms (dual, trial and paucal) is optional (1969: 5). Thus we have three facultative number choices. The system is given in figure 2.12.

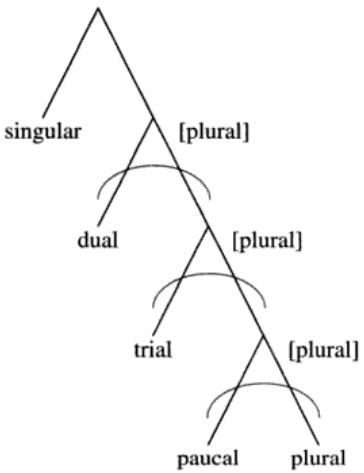

Figure 2.12 *The facultative numbers of Marshallese*

A pattern is emerging. We cannot restrict facultative number values to 'the last choice'; rather we must say that if there is facultative number it must involve 'the last choice'. It may involve other numbers, working up from the last choice. Thus it may affect the dual–plural choice in Larike, because it also affects the trial–plural choice; similarly it may affect the dual–plural choice in Longgu, because it also

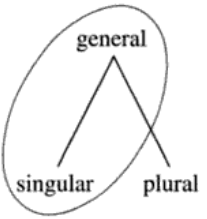

Figure 2.16 *System with general/singular versus plural*

more helpful representation of the system that we used earlier, given again as figure 2.16. This figure makes clearer a second type of problem with the other proposal: if we try to claim that the plural is facultative, then we get exactly the wrong pattern of overlapping: it is not consistent with the patterns found with the clear facultative numbers, in which it was precisely the plural that was used in place of facultative numbers (dual or trial). We should accept that the systems with general/singular forms are more radically different than figure 2.15 suggests: here it is the whole category of number which is optional (these systems are discussed further in §3.4 and §4.5.3). Therefore the position taken here is that such systems have general number (the choice is whether to express number or not); facultative number then occurs within systems where number is to be marked, but where a particular choice or choices need not be taken.

2.4 Languages without number

Linguists have claimed that all languages have number; that appears a reasonable claim, but when we look far enough we find counter-examples. Pirahã is the only remaining member of the Mura family and was spoken in 1997 by some 220 people along the Maici River (Amazonas, Brazil); it has been described by Everett (1986) on the basis of fourteen months of intensive contact with the Pirahã, and updated (1997) after five years of fieldwork. He states (1986: 217): 'there are no plural forms in Pirahã'. This holds even for pronouns, whose free forms are as in table 2.9 (1986: 280).

Table 2.9 *Personal pronouns in Pirahã (Everett 1986: 280)*

1st person	ti
2nd person	gíxai
3rd person	hiapióxio

He adds: 'There are no special plural forms for these pronouns.' This means that *hiapióxio* (third person) can be plural or singular, as this example shows (1986: 282):

Number is the most underestimated of the grammatical categories. It is deceptively simple yet the number system which philosophers, logicians and many linguists take as the norm – namely an obligatory distinction between singular and plural (as in 'cat' versus 'cats') – is only one of a wide range of possibilities to be found in languages around the world. Some languages, for instance, make more distinctions than English, having three, four or even five different values. Adopting a wide-ranging perspective, Greville Corbett draws on some 250 languages to analyse the possible systems of number. He reveals that the means for signalling number are remarkably diverse and are put to a surprising range of special additional uses. By surveying some of the riches of the world's linguistic resources this book makes a major contribution to the typology of categories and demonstrates that languages are much more varied than is generally recognised.

GREVILLE CORBETT is Professor of Linguistics at the University of Surrey. He is author of *Gender* (1991), also in the Cambridge Textbooks in Linguistics series, and co-editor (with Norman M. Fraser and Scott McGlashan) of *Heads in Grammatical Theory* (1993) and co-editor (with Bernard Comrie) of *The Slavonic Languages* (1993).

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org

ISBN 0-521-64970-6