

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.

Introduction

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 Anglo-centric 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

1.2 Comparing like with like

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/head-s* and *woman/women*. The first member of each pair is said to be singular, and the second is plural. So when we say

Introduction

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.

1.2 Comparing like with like

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.

Introduction

- (3) the sheep drinks from the stream
- (4) the sheep drink from the stream

Though the form of the noun does not change, and the marker of number is on the verb, it still indicates the number of sheep involved. (Example (4) cannot be used in English for the situation in which *one* sheep drinks several times.) In other words, we have nominal number which happens to be expressed on the verb (usually, in English, in addition to being expressed on the noun). Number in English is largely regular: words like *dog* ~ *dogs* greatly outnumber those like *sheep* ~ *sheep* and *criterion* ~ *criteria*. This suggests that number is an inflectional category in English: *dog* (singular) and *dogs* (plural) are forms of the same lexical item DOG.

There are many languages which, broadly speaking, are comparable to English in this respect. But there are also many languages in which number is fundamentally different: in particular it may be not a nominal category but a verbal one. Moreover it is often highly irregular and may not be an inflectional category. Let us consider briefly what verbal number is. The following examples are from Rapanui (the language of Easter Island, one of the Oceanic languages within Austronesian: data from Veronica Du Feu 1996: 191–2 and personal communication):

- (5) ruku
‘dive’
- (6) ruku ruku
‘go diving’

The form in (6) implies more than one dive, but not necessarily more than one diver. Verbal plurality is indicated (by reduplication here) since the event is in a sense plural. There are other possibilities for verbal number, just as nominal number can be more complex and varied than the English data suggest, as we shall see.

1.3 Structure of the book

We begin with nominal number, since it is the part of number where we find the greatest variety. In chapter 2 we look for as many meaning distinctions as we can identify in the world’s languages. We keep the nominal ‘still’ and see how many different number values it may have. Then in chapter 3 we hold the number value still, and see which nominals may be involved. The possible patterns of involvement in the number system are constrained by the Animacy Hierarchy, according to which, informally, the ‘more animate’ a nominal is the more likely it is to show number. We then allow both dimensions to vary together, that is to say, we attempt a typology of what number values are possible for what nominals (chapter 4). This integration of the two dimensions of the typology requires us to address the issues

1.4 Presentation

of minor numbers, associatives and distributives, among others. In chapter 5 we go on to the ways in which number is expressed, and in chapter 6 we discuss syntactic issues, mainly agreement but also including problems caused by numerals. Then we look at other ways in which the means for expressing number can be used, and see that there is a surprising range of uses, from honorific to evasive use (chapter 7). In chapter 8 we survey verbal number, covering meaning distinctions, the items involved in the verbal number system and the ways in which verbal number is expressed. In the concluding chapter we review what has been established about the category of number, draw together strands of the material particularly on the development of number systems (their rise and decline) and on the interaction of number with other categories; and then we look forward to new ideas for research into number.

1.4 Presentation

The book is designed for readers of several different types. For the student of linguistics, it is a guide to an area of obvious interest which has been neglected. And more importantly, it attempts to give a picture of the tremendous richness and diversity of the world's languages, by tackling a category where the familiar languages of Western Europe are overshadowed by the complexities of systems found elsewhere in the world. It is also intended to assist those researching particular languages or groups of languages, whether for a major research project or an undergraduate essay. Seeing familiar material analysed in a typological context can give a new perspective. This is particularly important for those areas where the terminology has become misleading, suggesting differences and similarities which do not hold. It is hoped especially that the book will prove valuable to field-workers by giving them both helpful leads for analysis and the awareness of the types of data which will enable us to understand the category of number more fully. The task in this area is urgent since, as mentioned above, many of the crucial languages are endangered. The picture presented in the book has been built up out of many small pieces, and this should be made evident; hence there are many references in the text, though where possible the detail is given in notes. There are extensive references for those who wish to go further. Since the work is organized thematically, special care has been taken so that those seeking data on specific languages can find the relevant references through the index.

The relevance of the book to typologists is evident: it is another example of the approach to typology which examines categories rather than constructions. Furthermore each chapter can be taken as illustrating a particular typological point, and so the book may be used as a hands-on introduction to typology. For morphologists, it should provide grist to the mill for those concerned with the relations of inflectional and derivational morphology (as noted in §1.1). There is also a

Introduction

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,

Meaning distinctions

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 ‘p’) 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.