5
The expression of number

It is now time to look at the ways in which number is expressed. This rather basic issue is a surprisingly novel one. Of course, many grammars describe number marking in individual languages but little has been done towards a typology. We shall therefore give an initial typology, one which aims to list the possibilities (claiming that the listed types can exist and no more). The obvious candidates for number expression are all found: special words (§5.1), syntax (§5.2), morphology (§5.3) and lexical means (§5.4). We shall then examine three types of system which are distinctive and which belong here in a discussion of means because they do not give rise to new semantic distinctions. (This is why they were not treated in chapter 2.) These are inverse systems (§5.5), minimal-augmented systems (§5.6) and 'constructed' numbers (§5.7). Finally we take up the discussion of the reduced expression of number, considering the form of items which are not (or not fully) within the number system (§5.8).

The main typological point is the importance of comparing like with like. In previous chapters we examined number values and their ranges of availability and compared each of these across languages. Now we turn to the means of expression, and must continue to be clear about when our claims relate to meaning and when to the means of expression. Thus we have values such as dual or paucal (which are expressed in various ways), and we have systems of expression (such as inverse number); we cannot therefore treat inverse as a value to put alongside dual and paucal. Naturally, in this chapter we concentrate on the variety of means of expression, and on the patterns which are claimed to underlie this variety.

5.1 Number words

Some languages have special ‘number words’, just for the purpose of indicating number. Thus in Tagalog, virtually any constituent can be pluralized by the word *mga* [maŋa], perhaps best characterized as a clitic (David Gil, personal communication):

"..."
Another clear case is provided by Miskitu, a Misumalpan language of the Caribbean Coast of Nicaragua and Honduras (T. Green 1992):

(5) aras
   ‘horse’

(6) aras kum
    horse SG
    ‘a horse’

(7) aras kumkum
    horse several
    ‘several horses’

(8) aras nani
    horse PL
    ‘horses’

(9) yang kauhw-ri
    1 fall-1.PAST.INDEF
    ‘I fell’

(10) yang nani kauhw-ri
     1 PL fall-1.PAST.INDEF
     ‘We (exclusive) fell’

Note particularly that the pronoun in (10) takes the plural word in a regular way (recall §3.5). There is interesting work to be done on the syntactic category of

---

1 Other examples of this are Canela-Krahô, a Jê language of Brazil (Popjes and Popjes 1986: 175–6, 185–6, examples (309) and (368)), and Golin, a Chimbu language of Papua New Guinea.
number words, which varies considerably (Dryer 1989), and on their word-order properties (Dryer 1992: 104–5). In Dogon (a Voltaic language, within Niger-Kordofanian, spoken in Mali), the number word is a clitic found at the end of the phrase (Plungian 1995: 9–10):

(11) \( \epsilon n \epsilon mbe \)
    goat PL
    ‘goats’

(12) \( \epsilon n \epsilon g \epsilon mbe \)
    goat DEF PL
    ‘the goats’

(13) \( \epsilon n \epsilon wo mi \eta \ ob-i-Ø \ g \epsilon mbe \)
    goat 3 1 OBJ give-AOR-3.SG DEF PL
    ‘the goats which he gave me’

Number marking with *mbe* is not obligatory.\(^2\) In Dogon the personal pronouns mark number (by suppletion) and a small number of nouns denoting humans have a plural marker (and this fits the pattern according to which different means of marking can follow the Animacy Hierarchy, see §3.5). With such nouns, the plural clitic is often used as well:

(14) \( nde-m \ mbe \)
    person-PL PL
    ‘people’

Further examples of plural words can be found in Dryer (1989).\(^3\) Diachronically, number words are a potential source of number morphology (§9.1.1). Of course, languages typically have quantifiers, with meanings such as ‘many’ or ‘few’, and various numerals (§6.7.1), but our interest here is in the purer number words, equivalent to plural.\(^4\)

New Guinea (Bunn 1974: 55). For the complex interaction of number words with pronouns in Vietnamese, see Luong (1987).

\(^{2}\) Our focus here is on the means of marking. As discussed elsewhere, these different means may have different uses. Thus the Dogon *mbe* can have a sort reading (§3.7.2), can be used with associative meaning (§4.3) and as an approximative plural (§7.3.3); here these uses are of secondary interest while we concentrate on means.


\(^{4}\) The remarkable phenomenon of ‘promiscuous’ number marking is found in languages of the northern north-west coast of North America, namely Haida, Eyak, Aleut and to a
The expression of number

5.2 Syntax

The way in which number is marked syntactically is through agreement. Often agreement is in addition to other morphological means of marking. For instance:

(15) Those goats are eating the washing.

Here we find plurality marked on the head noun *goats*, the controller of the agreement, and by agreement on two agreement targets: the demonstrative *those* within the noun phrase, and the verb *are* outside the noun phrase. However, we may also find a comparable situation when the controller itself does not mark number:

(16) Those sheep are doing nothing about it.

Here *sheep* is plural, though this is not indicated on the noun itself. We noted in §3.3 that where the morphological marking and control of agreement do not coincide, it is agreement which will be more regular in terms of the Animacy Hierarchy. It is important to note that the number marked on the verb is nominal number, indicating the number of sheep, rather than the number of eating events. We return to this distinction in chapter 8. Thus in a sense we have nominal number marked ‘in the wrong place’; it is this displaced information which is the essential ingredient of agreement. Cross-linguistically, demonstratives and verbs are relatively frequent agreement targets, showing agreement in number either uniquely or combined with other categories, notably gender. Other targets can include articles, adjectives, pronouns, nouns (especially in possessive constructions), numerals, adverbs, adpositions and complementizers (see Corbett 1991: 106–15 for examples).

While agreement is often an additional means of marking number, as in (15) above, there are languages where the situation in (16) is prevalent. This particularly interesting situation is found in the Papuan language Amele (J. Roberts 1987: 162, 201, 203 and personal communication). Amele belongs to the Madang subgroup within the Trans New Guinea phylum. Here are some relevant examples:

Footnote 4 (cont.)

lesser extent Tlingit (Leer 1991). Here a clitic which marks number may be associated semantically with different pronominal forms.

In Tlingit, for example, the number marker *has*# pluralizes animate third-person pronouns. Morphosyntactically, it is a proclitic; semantically, it associates with an animate third person in the word to which it is proclitic. If this word is a transitive verb, either subject or object (or both) may be animate third-person pronouns. In this case, the number marker is free to pluralize such a pronoun whether it is subject or object; and if both subject and object are animate third-person pronouns, the marker may pluralize either or both. (Leer 1991: 160).
5.2 Syntax

(17) Dana (uqa) ho-i-a
    man 3.SG come-3.SG-TODAY'S.PAST
    ‘The man came’

(18) Dana (ale) ho-si-a
    man 3.DU come-3.DU-TODAY'S.PAST
    ‘The two men came’

(19) Dana (age) ho-ig-a
    man 3.PL come-3.PL-TODAY'S.PAST
    ‘The men came’

In Amele the verb must agree in number with the subject, as shown in our examples by the formatives -i-, -si- and -ig-. (The formative -a indicates today’s past tense; and the unmarked NP has a definite referent (J. Roberts 1987: 203).) The noun may show plural number by reduplication (dana-dana ‘men’), but this is optional.\(^5\)

Number may also be indicated by pronominal copy, but this too is optional, as indicated by parentheses in our examples. Thus we have nominal number which must be indicated on the verb and which may optionally be indicated on the noun.

We have dealt with the cases of matching between the features of controller and target. There are rather different situations where mismatches between nominal and verb create a new number value, and these are discussed in §5.7. Problem cases of agreement which do not produce new number values are considered in chapter 6, where agreement is the central concern. The interplay of the syntactic and morphological expression of number is a recurring theme, and so we now turn to morphological expression.\(^6\)

\(^5\) William Foley (personal communication) suggests Arafundi as an even clearer example. Arafundi is spoken in Papua New Guinea and is an isolate or may be distantly related to the Piawi family. In six weeks of fieldwork Foley found no number marking of any kind on nouns, thus even more than in Amele the verb is the main locus of number marking. The pronoun does, however, mark number. Nichols (1992: 148–149) shows that in general number is ‘highly prone to be drawn off the noun and marked by agreement, primarily on the verb’.

\(^6\) A truly remarkable situation is found in Korean. According to Seok Choong Song (1975), plurality may or may not be marked on the noun by the marker -tul. The choice depends in part on definiteness (definite noun phrases are likely to have the head noun marked for number) and on the Animacy Hierarchy (§3.2): nouns denoting humans and other animals are much more likely to be marked for number than are inanimates. The plural marker may also occur in the predicate (Song calls it the ‘ubiquitous plural marker’), and it is here that the special interest of Korean is to be found. Consider the following example (Lee 1991: 81):

(i) ai-tul-i Tom-eyke-[ ] ppang-ul-[ ] manhi-[ ] cwoesseyo-[ ]
    child-PL-NOM Tom-to bread-ACC a.lot gave
    ‘the children gave Tom a lot of bread’
5.3 Morphology

We have seen many examples of the morphological expression of number in earlier chapters. It is widespread in the world’s languages and very varied. Indeed one could describe a substantial part of the morphological resources of natural language while taking the examples exclusively from number. We shall consider the range of possibilities briefly. Given the necessary brevity an important distinction to bear in mind is that between the system of number marking in a particular language and that found with individual lexical items. In the book so far we have tried where possible to concentrate on whole systems: for instance, we have said that nouns in English distinguish singular and plural (and then discussed special cases like sheep and friendliness). In this section we shall instead concentrate on variety, even if small numbers of items are involved. We do this for two reasons. First, relatively little has been done on the distribution of morphological subsystems within languages. There are few languages for which we can state the proportion of nominals which mark number according to different methods. However, we did note the generalizations based on the Animacy Hierarchy in §3.5. The other reason for this approach is that the coexistence of different types of morphology is so common that it seemed more appropriate to emphasize this coexistence. Thus in §5.3.1 we are able to take almost all the different examples from a single language. The variety can become so extensive that in some languages number becomes almost a lexical matter (a possibility we return to in §5.4). We first consider the relation between stems and inflections in an abstract way (§5.3.1) and then look in turn at the range of variety within inflections (§5.3.2) and within stems (§5.3.3) and at zero

Footnote 6 (cont.)

The subject ai-tul-i is marked as plural. The plural marker -tul can also appear at any of the points indicated [ ]. The -tul used in the predicate is not redundant. It indicates that the constituent to which it is attached represents new information (Lee 1991: 83). Furthermore it guarantees a distributive reading, as can be shown by the next two examples:

(ii) haksayng-tul-i phwungsen hana-lul sasseyo
student-PL-NOM balloon one-ACC bought
‘the students bought a balloon’

(iii) haksayng-tul-i phwungsen hana-lul-tul sasseyo
student-PL-NOM balloon one-ACC-PL bought
‘the students bought a balloon each’

In sentence (ii) the students may have bought just one balloon between them, or one each. Sentence (iii), with -tul marking the object constituent has only the second, distributive reading (see §2.5 and §4.4.1). The distributive role of -tul is stressed by Jae Jung Song (1997). For other research on this interesting construction and on number in Korean more generally see Kuh (1987), Unterbeck (1993), Kang (1994), Kim (1994), Prost (1992) and Kwak (1996).
expression (§5.3.4). Next we return to the issue of phrasal affixes as a type of clitic (§5.3.5) and finally look at multiple marking (§5.3.6).

5.3.1 Relations between stems and inflections

We start from the notion of ‘base’ (or ‘basic inflectional stem’). The base of a lexical item is the form which cannot be further reduced as far as inflectional categories are concerned. The Russian noun komnata ‘room’ (nominative singular) has the plural komnaty. The base is komnat- and the inflections (endings in this instance) are -a (nominative singular) and -y (nominative plural). Why do we make the division there, rather than claiming, for instance, that the base is komn- and the inflections -ata and -aty? There are thousands of other nouns which can be analysed as taking the inflections -a and -y (such as golov-a ‘head’, plural golov-y, or sten-a ‘wall’ plural sten-y) but relatively few for which the other segmentation makes any sense. Let us consider a language which has at least two numbers, singular and plural. What are the possible relations between the number forms and the base for a given lexical item (or group of lexical items)? We first take a maximally general model, as in figure 5.1. How can the singular and plural forms differ from the base? First they may differ in inflection. Or they can vary from the base through stem formation. These two devices, inflection and stem formation may occur separately or together. That gives three possibilities: difference just in terms of inflection, just in terms of stem, or in both. The fourth logical possibility is that neither inflection nor stem formation is employed. If this means that the singular form, plural form and basic stem are all identical, then clearly number is simply not marked morphologically for the items in question (as in the case of English sheep or Russian kenguru ‘kangaroo(s)’). We therefore elaborate the model as in figure 5.2, allowing for different stems. Let us take examples of number marking, and consider in particular whether or not all the elements identified in the diagram are distinct in particular examples. Remember that different patterns often coexist within a single language; if an example is given from a particular language this does not mean that the pattern is the major one for that language.

If we start with the relations between the base and the stems, the first logical possibility is that all are distinct. This possibility can be illustrated by the irregular

![Figure 5.1 Potential distinctions for number marking](image-url)
Russian noun, \textit{xozjain} `landlord’. The base is \textit{xozja}(j)-,\(^7\) the singular stem is \textit{xozja-in-} and the plural stem is \textit{xozja-ev-}. Both stems allow the addition of endings. Thus singular and plural are each indicated both by the stem and by the ending.

The extreme type of difference is found in cases of suppletion, where there are different stems which are not related by any regular or irregular type of stem formation (though they show a regular grammatical opposition). An example is Russian \textit{čelovek} `person’, plural \textit{ljud-i} `people’.\(^8\) Note that we are indeed dealing with stems here: \textit{čelovek-} `person’ takes normal singular inflections, and \textit{ljud-} `people’ takes plural inflections. We return to suppletion in §5.4 and §9.2.2.

It is unusual for the base, the singular stem and the plural stem all to be distinct, in Russian and more generally. Often we find that the base and the singular stem are identical, as in figure 5.3 (there are interesting parallels with the systems of values in §2.1). A Russian illustration of this pattern is the noun \textit{krylo} `wing’. This has the base \textit{kryl-}, to which the singular endings are added directly (\textit{kryl-o, kryl-a, kryl-u} and so on). The plural stem is \textit{kryl’j-} (the ’ marks palatalization of the preceding consonant), as in the nominative plural \textit{kryl’j-a}. Why should we say that there is a distinct plural stem here, rather than that the nominative plural ending is palatalization plus -ja? The point is that -\textit{a} is a regular nominative plural ending, found on thousands of nouns which do not have a separate plural stem. (We return in §5.3.2 to the point about -\textit{a} being a singular ending in \textit{komnat-a} ‘room’ but a plural ending here.) The plural endings for the remaining cases of Russian are also found on other nouns (see table 5.2 below); we would be missing an obvious generalization if we claimed there were special endings right through the plural paradigm while in fact

\(^7\) Since Russian orthography is largely morphophonemic, we use a transliterated form for simplicity here. The details of the use of \textit{j}, being present in the root but not spelled in the stems, need not detain us.

\(^8\) The Russian suppletion is matched by the natural English translations \textit{person} and \textit{people}, though of course \textit{persons} does exist in English.
nouns like *krylo* ‘wing’ differ from other nouns only in having a different stem for the plural.

The next possibility is that the plural stem should be the same as the base as in figure 5.4. Again the pattern is found in Russian. The noun *bolgarin* ‘a Bulgarian’ has the base *bolgar-* and the plural stem is identical, as in forms like the nominative plural *bolgar-y*. The singular stem differs, and is *bolgarin-*. Several nouns denoting nationalities and other social groupings behave in this way.

A final relation of base to stems is that all are identical, diagrammed as in figure 5.5 (p. 142). This situation is extremely common. Again in Russian we find many nouns like that for ‘room’, which has the basic stem *konmat-* . The (nominative) singular is *konmat-a* and the (nominative) plural is *konmat-y*. Here stem formation has no role, and the entire burden of signalling a difference in number is carried by the inflections.

We move on to look for identities elsewhere in the model. There is a further, initially rather surprising type of identity, shown in figure 5.6. This pattern suggests
that the inflections used for singular and plural could be identical. This situation regularly occurs in Daghestanian languages for the majority of the large numbers of cases they distinguish (often just the absolutive is an exception). The Akhvakh noun *nido* ‘forehead’ shows a clear case of identical endings: we take absolutive and ergative cases to illustrate the point in table 5.1. In this example the base is *nido*, and the singular stem is identical to it. The plural stem is *nido-di*. The absolutive case, in singular and plural, has no ending. In both numbers there is an oblique stem, distinct from the basic singular or plural stem; in the singular it is formed with *-la*, and in the plural with *-le*. The various oblique case endings are added to this stem; in our example the ergative case is given, and the appropriate ending is *-de*. As with the absolutive, the ending is the same for singular and plural. The point is that information about number is signalled by the differences in the stems: *-di* indicates plurality for this noun, *-la* shows singular oblique, and *-le* plural oblique. Thus a form like *nido-di-le-de* indicates plurality twice. The endings have no role in the number system, their function is to mark the case of the noun. This identity of form of endings in the singular and plural is quite general in

Figure 5.5 *Both stems match the base*

![Diagram](image)

Figure 5.6 *Inflections not sensitive to number*

![Diagram](image)
Daghestanian languages. It is to be distinguished from occasional syncretisms of form involving small numbers of nouns in languages where the coincidence of form is not systematic (we see an example from another language family in table 5.4 below).

There is a final pattern of identity we should mention again, that in which both stems are identical to the base, and where the stems are identical to the forms with endings (that is, there are no endings). Under these conditions the noun is indeclinable, hence number is not marked morphologically. There are numerous examples of this situation, both of languages where number is not marked morphologically on particular word classes (English adjectives, for example) or not marked morphologically at all. But more interestingly it may be found for a subset of a word class within a system where number is usually marked morphologically. Thus in Russian, the majority of nouns distinguish two numbers but some, especially foreign borrowings, do not. For example, *kenguru* may denote one or more kangaroos, and *taksi* may denote one or more taxis (the ambiguity will often be removed by elements showing agreement in number, see §5.2).

### 5.3.2 Inflections

In some languages a particular ending signals immediately what the number is. Thus in Central Alaskan Yup’ik, as we saw in §4.3.3, the ending -t indicates a noun in the plural. Compare this with the situation we met in Russian (§5.3.1), where -a signals the singular in *komnat-a* ‘room’, but plural in *kryl’j-a*. ‘wings’. This is the first important typological distinction, between languages which have different inflectional classes (of nouns in this instance), like Russian, and those which do not, like Central Alaskan Yup’ik. In a language with distinct inflectional classes, we may know the significance of a particular ending only in respect of the particular stem it is attached to. Typically we find a restricted number of main inflectional classes, each with a sizeable number of nouns. Russian has arguably four inflectional classes for nouns (Corbett 1982; Corbett and Fraser 1993: 114–16), shown in table 5.2.

### Table 5.1 Number marking in Akhvakh

*(Kibrik 1991: 260)*

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolutive</td>
<td>nido</td>
<td>nido-di</td>
</tr>
<tr>
<td>ergative</td>
<td>nido-la-de</td>
<td>nido-di-le-de</td>
</tr>
</tbody>
</table>

9 Establishing paradigms is not always straightforward; here we accept the traditional view for Russian of six cases and two numbers, hence twelve cells in all (see Comrie 1986, 1991 for discussion of the issues).
This table shows clearly that inflections have a different significance for different nouns. There are several instances of the use of the bare stem, as in zakon ‘law (nominative singular)’ and komnata ‘room (genitive plural)’. We shall return to the different inflectional cases shortly.

Most nouns, as in the examples in table 5.2, have a single stem, but some have stem modifications, as discussed in §5.3.1. These four inflectional classes cover the vast majority of Russian nouns, but there are some further variants and exceptions, for which see Timberlake (1993: 837–41), and several hundred indeclinable nouns;10 the relation of exceptional number marking to frequency will be discussed in §9.3.3. We can get a picture of the distribution of the nouns over these four main classes by calculating the membership of around 44,000 nouns. Table 5.3 gives the number in each class, to the nearest fifty. Each class has a substantial number of nouns, but the dominance of class I is increasing over time.

We should return to the question of case, since this brings us to a second major type of difference. If we take a form like zakon-a ‘law (genitive singular)’ the information that it is singular is encoded in the inflection -a, but so is the fact that it is genitive. There is no way to pull apart the information about the two categories:

10 Particular inflections can present great complexity of choice; a good instance is the Polish nominative plural for nouns denoting males, as shown by Dunaj (1993), also discussed by Wierzbicka (1988: 455–9).
this phenomenon is known as ‘cumulation’. Number is often marked morphologically in this way, cumulated with case, or gender or person, depending on the word-class (see §9.2). On the other hand number can be expressed morphologically without cumulation, and in two different ways. In English, nouns such as law ~ laws mark number only, and so there is no cumulation. Alternatively, in languages like Uzbek (Mugdan 1994: 2550), number and case have separate markers, as in table 5.4.

In such languages, unlike in Russian, it is possible to identify the suffix -lar as the marker of plural number. The Uzbek type is said to have agglutinative morphology, and the Russian type fusional morphology. But agglutinative and fusional are poles of a continuum, with the great majority of systems lying somewhere between (Comrie 1989: 42–52). In this example, the case marker is outside the number marker, which is in accord with Greenberg’s universal number 39 (1963):

Table 5.3  Membership of the inflectional classes of Russian

<table>
<thead>
<tr>
<th>Class</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>zakon</td>
<td>komnata</td>
<td>kost’</td>
<td>boloto</td>
</tr>
<tr>
<td></td>
<td>‘law’</td>
<td>‘room’</td>
<td>‘bone’</td>
<td>‘marsh’</td>
</tr>
<tr>
<td>No. of nouns</td>
<td>20,700</td>
<td>13,600</td>
<td>3,950</td>
<td>5,750</td>
</tr>
</tbody>
</table>

Table 5.4  Number and case in Uzbek

<table>
<thead>
<tr>
<th>(olma ‘apple’)</th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>olma</td>
<td>olma-lar</td>
</tr>
<tr>
<td>ACC</td>
<td>olma-ni</td>
<td>olma-lar-ni</td>
</tr>
<tr>
<td>GEN</td>
<td>olma-niŋ</td>
<td>olma-lar-niŋ</td>
</tr>
<tr>
<td>DAT</td>
<td>olma-ga</td>
<td>olma-lar-ga</td>
</tr>
<tr>
<td>ABL</td>
<td>olma-dan</td>
<td>olma-lar-dan</td>
</tr>
<tr>
<td>LOC</td>
<td>olma-da</td>
<td>olma-lar-da</td>
</tr>
</tbody>
</table>

In such languages, unlike in Russian, it is possible to identify the suffix -lar as the marker of plural number. The Uzbek type is said to have agglutinative morphology, and the Russian type fusional morphology. But agglutinative and fusional are poles of a continuum, with the great majority of systems lying somewhere between (Comrie 1989: 42–52). In this example, the case marker is outside the number marker, which is in accord with Greenberg’s universal number 39 (1963):

Brown et al. (1996: 57), calculated from Zaliznjak (1977). These statistics give a good first impression, but they do not take account of the effect of derivational morphology. Thus the majority of the nouns in class III have the suffix -ost’ which forms abstract nouns from adjectives (star-yy ‘old’, star-ost’ ‘old age’). If the suffix is labelled as belonging to class III, then the number of distinct members of the class is substantially reduced. Similarly nominalizations in -aniel-enie (like razrušenie ‘destruction’, derived from razrušit’ ‘destroy’) inflate the figure for IV (see Schupbach 1984 for discussion). Apart from via these suffixes, classes III and IV are barely productive.

11 Brown et al. (1996: 57), calculated from Zaliznjak (1977). These statistics give a good first impression, but they do not take account of the effect of derivational morphology. Thus the majority of the nouns in class III have the suffix -ost’ which forms abstract nouns from adjectives (star-yy ‘old’, star-ost’ ‘old age’). If the suffix is labelled as belonging to class III, then the number of distinct members of the class is substantially reduced. Similarly nominalizations in -aniel-enie (like razrušenie ‘destruction’, derived from razrušit’ ‘destroy’) inflate the figure for IV (see Schupbach 1984 for discussion). Apart from via these suffixes, classes III and IV are barely productive.
The expression of number

Where morphemes of both number and case are present and both follow or precede the noun base, the expression of number almost always comes between the noun base and the expression of case.

The evidence we saw in Akhvakh (table 5.1) fitted even better. Recall the noun form nido-di-le-de ‘forehead’ (ergative plural), where the elements come in the order:

base plural marker [plural + oblique marker] case marker

The marker which indicates number and case (oblique rather than direct) stands between the straightforward number marker and the straightforward case marker.

5.3.3 Stems

Stems can show changes from minor stress alternations to major restructuring. Cross-cutting this variation is the distinction between stem distinctions which alone signal number and those which occur in conjunction with inflections which distinguish number. We shall start from the smaller modification and progress to the larger; not surprisingly, the larger the modification the more likely it is to be the sole marker of number.

We start with prosodic differences in stems. Even these can be the only marker of number. Hence in Shilluk (discussed in more detail in §5.4) we find examples such as: kiływ (with low-high tone) contrasted with kiıy (high-low) ‘plant(s) with edible roots’. The opposition of stress, on the stem as opposed to not on it (hence on the inflection) is found frequently. We can illustrate this from Russian, in order to fit with the inflectional information given earlier. Basically the stress can be on the stem or on the ending, and it can be the same in the singular and plural or different. This gives four main classes, as shown in table 5.5. Clearly when singular and plural stems differ, this is only a secondary marker, since the inflection is normally a clear marker. Interestingly these are abstract patterns, which are available for different inflectional classes. Not all possibilities are found, but the majority are (given the four types of nominal stress and the four inflectional classes of table 5.2, we might expect sixteen combinations and in fact fifteen are found). The distribution of these stress patterns is also of interest: a count of approximately 44,000 nouns shows that the vast majority (43,000) have fixed stress (40,300 in type A and 2,700 in type B); of the remainder, type C has under 450 nouns, type D under 350, and the subtypes together have around 250).

12 For details on the interesting but relatively small subtypes see Brown et al. (1996).

13 The missing one is inflectional class III with stress pattern D. Note that the figure of fifteen is correct provided the small stress subtypes are counted with the main types. If only the main types are counted, then twelve combinations are found, out of the theoretically possible sixteen.

14 Figures again from Brown et al. (1996: 57).
A somewhat more substantial difference between stems is represented by alterations. For instance in the South Slavonic language Macedonian, there is a set of nouns where we normally find the bare stem in the singular and the stem plus -i in the plural, e.g. koren ‘root’, plural koreni. However, in nouns ending in -k, -g or -x we find an alternation to c (an unvoiced dental fricative), z or s respectively (Friedman 1993: 258), as table 5.6 illustrates. The point is that this alternation occurs in specific morphological environments. It is possible to find ki, gi and xi elsewhere. What was once a regular phonological alternation (the Slavonic second palatalization of velars) is retained as a morphological one.

Table 5.6  Examples of stem alternations in Macedonian

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>učenik</td>
<td>učenici</td>
<td>pupil (school)</td>
</tr>
<tr>
<td>parking</td>
<td>parkinzi</td>
<td>parking space</td>
</tr>
<tr>
<td>uspeh</td>
<td>uspesi</td>
<td>success</td>
</tr>
</tbody>
</table>

A more substantial modification is the addition of augments in the singular, or plural, or both. We saw several examples in §5.3.1, for instance Russian bolgar-y ‘Bulgarians’, singular bolgar-in, where the singular has the augment -in- before the regular endings. We also saw the case of Akhvakh. Nakh-Daghestanian languages are complex in this respect; there is a good deal of information available on the systems found, with data on how prevalent particular patterns are in particular languages: for an overview, see Kibrik (1991, forthcoming) and for more detail see

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>stem</td>
<td>stem</td>
</tr>
<tr>
<td>kómnata</td>
<td>kómnaty</td>
</tr>
<tr>
<td>‘room’</td>
<td>‘rooms’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>stem (initial)</td>
<td>ending</td>
</tr>
<tr>
<td>věčer</td>
<td>večerá</td>
</tr>
<tr>
<td>‘evening’</td>
<td>‘evenings’</td>
</tr>
</tbody>
</table>

| singular               | plural                 |
| stem (predesinential)  | ending                 |
|                        | vinó                   |
| ‘wine’                 | ‘wines’                |
Kibrik and Kodzasov (1990: 251–310). We shall look briefly at another of them, namely Lak, whose nouns, like those of Akhvakh (table 5.1), may have up to four distinct stems as shown in table 5.7.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolutive</td>
<td>muIh</td>
<td>muIh-ru</td>
</tr>
<tr>
<td>genitive</td>
<td>muIh-li-l</td>
<td>muIh-a-l</td>
</tr>
</tbody>
</table>

In terms of our analyses earlier, the absolutive singular is the same as the base. The absolutive plural is formed from the base with the formant -ru, and the absolutive has no ending, singular or plural. There is an oblique singular stem, formed from the base with -li-, and there is a plural oblique stem formed with -a-; then different case inflections are added, in this instance the genitive ending -l. This is not the only pattern in Lak. The important point here is to show the wealth of means available for stem formation. To form the oblique singular stem there are over thirty formatives available, and, though there are some recurring patterns, it has not proved possible to predict which one will be found with a particular noun. For forming the absolutive plural there are again more than thirty possibilities. Finally, to form the oblique plural there are over a dozen possibilities.

A rather different type of stem modification is known as reduplication. Instead of adding new material (as with augments) or modifying particular segments, reduplication primarily involves use of material from the specific lexical item. This may involve the whole as in Dyirbal (Dixon 1972: 242–3). Thus we find \( \text{nal}g\text{a-} \text{nal}g\text{a-} \text{nal}g\text{a-gu} \) ‘girl-ERG’, with the plural \( \text{nal}g\text{a-nal}g\text{a-gu} \) ‘girls-ERG’; note that the ergative inflection varies according to the number of syllables of the stem. Alternatively it may involve only a part (partial reduplication) as in these examples from Ilocano (table 5.8), an Austronesian language of the Philippines (Hayes and Abad 1989: 357–1; McCarthy and Prince 1995: 319). These and similar data have been extensively discussed, particularly within Prosodic Morphology. The central point is that reduplication does not specify a segment to be copied; rather it sets a template, to ‘aim’ at as it were: in the case of Ilocano this is a heavy syllable. The content of the reduplication is then dependent on the base. The last example \( \text{tra}:-\text{trak} \) suggests that copying the entire stem is excluded in Ilocano and so the vowel is lengthened in order to fit the heavy syllable tem-

--

15 Kibrik and Kodzasov use ‘nominative’ rather than ‘absolutive’.
plate. We have already seen several examples of reduplication. This morphological resource is used for a variety of purposes, of which signalling plurality is only one; it is particularly frequent in distributive use (§4.4.1) so cases must be analysed carefully.16

We now come to a dramatic type of difference, internal modifications to the stem. These are common in some of the Germanic languages (see table 5.9), where particularly for German the plural form is referred to as an ‘umlaut’ form.

Table 5.9  Stem internal modification in Germanic languages

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>goose</td>
<td>geese</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse</td>
<td>mice</td>
<td></td>
</tr>
<tr>
<td>German</td>
<td>Apfel</td>
<td>Äpfel</td>
<td>apple</td>
</tr>
<tr>
<td></td>
<td>Sohn</td>
<td>Söhne</td>
<td>son</td>
</tr>
<tr>
<td>Frisian</td>
<td>beam</td>
<td>bjemen</td>
<td>tree</td>
</tr>
<tr>
<td></td>
<td>stoei</td>
<td>stollen</td>
<td>chair</td>
</tr>
</tbody>
</table>

16 The literature on reduplication as a morphological and phonological problem is extensive: see McCarthy and Prince (1995: 332–3) and Spencer (1991: 150–6, 169–71) for references; for a typology see Moravcsik (1978). Discussions of reduplication which include reference to its use for marking plurality include Wilkins (1984) and Henderson (1990) on Arrernte, Kiyomi (1995) on Malayo-Polynesian languages, Gonda (1942) on the languages of Indonesia, El-Solami-Mewis (1988) who gives statistics on the use of reduplication as opposed to other means of plural marking in Somali, Haeberlin (1918) on Salish languages. In a survey of reduplication in the languages of Australia, Fabricius (1998: 14, 59) points out that it is used there for various sorts of plurality but not for reference to two. Recall that distributives typically are inappropriate for reference to two (§4.4.1), which serves as a reminder that we should be wary of suggested reduplicated plurals in case they are in fact distributives.
Note that sometimes the alternation is the only marker of number (as in the English examples and in German *Apfel ~ Äpfel*) but frequently there is a regular inflection too. (For the complexity of German see Köpcke 1988, 1993, 1994, 1998, and for Old High German see Salmons 1994.)

One of the systems that looks most complex at first sight is that found in various Semitic languages, where we find ‘broken’ plurals (which have stem-internal modification) as opposed to the ‘sound’ plurals which take a suffix. The different patterns of broken plurals are certainly complicated, as the Arabic examples in table 5.10 show (Ratcliffe 1998: 23). There are many subpatterns, and yet there is a good degree of predictability. This is also a case where there is statistical data on the distribution of the different types of plural marking (see Ratcliffe 1998).\textsuperscript{17} We return to Arabic plurals in §6.6.

It is worth asking whether number could ever be marked by subtractive morphology, in other words whether number could be indicated by removing material. The examples we saw in §5.3.1, for instance Russian *bolgar-in* ‘Bulgarian’ with plural *bolgar-y* do not fit the bill. Here the singular is formed from the plural by the addition of the suffix *-in*, which can be attached to various plural stems in the same way. To find an example of number marking by subtractive morphology, one would need different phonological material to be removed, say the last segment, so that in imaginary English, the plural of *cat* would be *ca* and the plural of *dog* would be *do*. There is one instance which at first sight comes close to this, namely the Hessian dialect of German (see Golston and Wiese 1996 for sources). In Hessian German we do indeed find examples like the following: *hond* ‘dog’ plural *hon*, *viend* ‘wind’ plural *viem*, *berk* ‘mountain’ plural *ber*, and *vek* ‘way’

\textsuperscript{17} Broken plurals have been important in the development of phonological theory; see for instance: McCarthy (1982, 1983), Hammond (1988), McCarthy and Prince (1990), McCarthy and Prince (1995: 345–6), and see Idrissi (1997) and Ratcliffe (1998) for discussion; for the development of the phenomenon, see Murtonen (1964) and Ratcliffe (1998); see also Wallace (1988) on Biblical Hebrew, Palmer (1955) on Tigrinya and Palmer (1962) on Tigre. For a connectionist model of the Arabic plurals see Plunkett and Nakisa (1997).
plural *ve*. This seems to be what we were looking for. However, it is always the same two consonants involved (underlying *d* or *g*), which suggests that we are dealing with phonological conditioning. Golston and Wiese (1996) argue that there is a constraint that Hessian plurals must end in a sonorant; this constraint is satisfied in various ways, including in some instances the loss of the final obstruent. Thus for some nouns (and which ones is phonologically determined) the plural is indeed a reduced form of the singular, but there is not a general rule ‘subtracting’ a segment to form the plural. Holsinger and Houseman (1999) offer an alternative analysis, again without the need for subtractive morphology. We conclude that while Hessian was the best candidate, it appears that number is not marked by means of subtractive morphology.

### 5.3.4 Zero expression

We saw in the last three sections how number may be marked by inflection or by the stem. However, the full range of possible differentiation is not normally found. We saw instances where one of the stems was identical to the base (in other terms it was the base plus zero); we also saw instances where number inflections were not maximally different (and again there were examples of stems with no inflection). These possibilities are not randomly distributed, but conform to certain patterns. Greenberg’s universal number 35 (1963) states that: ‘There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.’ This universal conflates stems and inflections; it claims that whereas the singular may have zero expression (and the plural has additional material to form a plural stem, or a plural inflection, or both), the plural in a given language will not regularly have zero expression, though this may be the means for some nouns (see also §2.1). It is common for the singular to have zero expression; it is less usual for the plural to be marked in this way, even for a smaller group of lexical items.\(^{18}\)

Finally we should note how odd English is in this respect. In the verb, the third singular (present) has a marker, while the plural does not: compare *she writes* with *they write*. But though unusual, English is not a counter-example to Greenberg’s universal since elsewhere in the system the plural has non-zero allomorphs (on nouns: *writer* but *writer-s*).

---

\(^{18}\) It is interesting to separate out stem and inflection in this regard. We noted the example of Russian *bolgarin* ‘a Bulgarian’ and similar nouns; here the base is *bolgar-* and the plural stem is identical, as in forms like the nominative plural *bolgar-y*. The singular stem is *bolgarin-*. In terms of stems the plural has zero expression for these nouns; there are relatively few which behave in this way. In terms of inflection, in the nominative singular there is zero expression (bolgar-*in* with no inflection); there are many nouns in Russian which mark the nominative singular by zero (see table 5.3).
5.3.5 Clitics revisited

Clitics should be mentioned again here, since the term covers a range of items which fall between being independent words and being inflections. Clitics are like independent words in that they are not bound to particular stems, but they are phonologically dependent on a host (for instance, they do not bear their own stress). One type of clitic attaches to a particular type of phrase, hence the more specific term for these, namely ‘phrasal affixes’. We saw an example from Dogon in §5.1 above:

(20) **ene ge mbe**

goat DEF PL

‘the goats’

Here mbe is part of the phrase, but it is not required to attach to, say, a noun (as a Russian inflection might be), rather it simply appears last in the noun phrase and attaches to the preceding element. The diachronic development from independent plural word, to clitic, and finally to bound inflection may go through various intermediate stages, so that in a particular language the analysis can be difficult. Thus there are languages like Basque where various markers (including number) attach to the last element of the noun phrase, and there are several different items which that can be. But in a language where the noun phrase is noun-final, deciding whether a following item is a phrasal affix or an inflection on the noun is trickier, and in some instances probably undecidable.\(^\text{19}\)

5.3.6 Multiple marking

Number is often marked in more than one way. It may be by two different means: by morphological means and by agreement, by morphological means and by a number word as in (14), or very commonly by two or more morphological means. We have seen numerous instances of stem changes together with inflection. Some languages show a profusion of means of marking number even on a single item.\(^\text{20}\)

Within morphological marking there are also more systematic instances, where having two markers is the norm. One such is Breton, where the formation of diminutive plurals involves two plural markers, giving the structure: stem–PL–DIMINUTIVE–PL, as in the examples in table 5.11 (see Stump 1990: 104–5, 1991:

---

\(^\text{19}\) A somewhat different type of complexity in the marking of number within noun phrases with different elements present is found in Dagaare (a Gur language spoken in parts of Ghana and Burkina Faso), for which see Bodomo (1997: 47–50).

\(^\text{20}\) Turner (1976) shows how Seri and Chontal have between them eleven means of pluralization and some nouns show four of them at once (compare also Moser and Moser 1976 and Marlett 1990); for Ket see G. D. S. Anderson (1996) and for Breton see Trépos (1957); there are relevant points in Plank (1985: 61–5, 67–8, 77–80).
The diminutive plurals have two plural markers, whether the same one repeated as in *bagoùigoù ‘small boats’* or two different ones, as in *labousedigoù ‘small birds’* (-ed is the default for animates and -où for others; the remaining two nouns given have irregular plurals). It should be stressed that none of the instances in this section affect number values. The Russian *kryl’ja ‘wings’*, which has an augment, a plural inflection and a stress shift as compared to the singular is not ‘more plural’ than *komnaty ‘rooms’*, which has only a plural inflection. This multiple marking is therefore to be distinguished from the rare composed numbers (as in Breton, §2.2.7) where the additional markers have an additional semantic effect.

An interesting facet of multiple marking is usually known as ‘double plurals’, which are instances where an old plural gains a new plural marker ‘on top of’ the old one. For instance, the Middle English plural *childre*, gained a second marker of plurality, resulting in the modern form *children*. There are many examples of double plurals in a variety of languages. Tiersma (1982) showed the particular interest of one type of double plural, for which we need the fruitful but not unproblematic notion of markedness. Originally developed in phonology, it was later applied to grammatical categories (see in particular Jakobson 1932, Greenberg 1966, Zwicky 1978, Tomić 1989, Battistella 1990, Croft 1990, Fradkin 2000).

These are to be distinguished from the instances where particular nouns have two distinct plural forms, typically with an accompanying difference in meaning. For example, in Italian there is *l’ossa ‘the bone’* (masculine singular). There are two plurals: *le ossa ‘the bones’* (feminine), sometimes called the ‘collective plural’, which would be used of bones which belong together, a particular person’s bones, a skeleton, and a second plural *gli ossi ‘the bones’* (masculine), sometimes called the ‘distributive plural’, for bones which do not belong together: it would be used, for instance, when buying bones for a dog. Few nouns are of this type in Italian; for examples and discussion see Rocchetti (1968), Brunet (1978: 30–76), Santangelo (1981), Vincent (1988: 289) and Ojeda (1995). Note that the difference in meaning between the plurals is somewhat separate from number, in contrast to the greater numbers discussed in §2.2.6.

---

### Table 5.11 Formation of diminutive plurals in Breton

<table>
<thead>
<tr>
<th>singular</th>
<th>diminutive</th>
<th>plural</th>
<th>diminutive</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>bag</td>
<td>bagig</td>
<td>bagoù</td>
<td>bagoùigoù</td>
<td>boat</td>
</tr>
<tr>
<td>labous</td>
<td>labousig</td>
<td>laboused</td>
<td>labousedigoù</td>
<td>bird</td>
</tr>
<tr>
<td>bugel</td>
<td>bugelig</td>
<td>bugale</td>
<td>bugaligoù</td>
<td>child</td>
</tr>
<tr>
<td>kazh</td>
<td>kazhig</td>
<td>kizhier</td>
<td>kizhierigoù</td>
<td>cat</td>
</tr>
</tbody>
</table>

---

696–7, 1996 and references there; for comparable data on Yiddish diminutives see Perlmutter 1988).
The expression of number

1991 on Arabic and Janda 1991 on German). The idea is that many oppositions are not balanced: one member of a grammatical category is the normal or unmarked member, the other is marked. We see this in terms of semantics, form and frequency. If we take number, it is claimed that the singular is unmarked and the plural marked. In semantic terms, where number is irrelevant or unimportant it is the singular which is used (this is not fully clear, see McCawley 1968 and Ojeda 1992c for dissenting voices). In terms of form, we certainly see many instances where the singular has no marker but the plural has one, as discussed in §5.3.4. And in terms of frequency, the singular occurs considerably more frequently than the plural (§9.3.2).22

The notion of markedness is typically applied to categories as a whole, but it was known that there were interesting effects with particular items. Tiersma researched some of them and made the claim that: ‘When the referent of a noun naturally occurs in pairs or groups, and/or when it is generally referred to collectively, such a noun is locally unmarked in the plural’ (Tiersma 1982: 835). Several things follow, one being that nouns with double plurals are often those where the plural is the locally unmarked form. Table 5.12 provides examples from Dutch. Here we see clearly how the modern plural is formed with a suffix on the basis of what was formerly a plural in its own right.23 (For further examples, from Lezgian, see §5.8.3.)

<table>
<thead>
<tr>
<th>Table 5.12 Double plurals in Dutch (Tiersma 1982: 438)</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
</tr>
<tr>
<td>blad</td>
</tr>
<tr>
<td>ei</td>
</tr>
<tr>
<td>kind</td>
</tr>
<tr>
<td>lamm</td>
</tr>
</tbody>
</table>

Borrowings also illustrate the local unmarkedness of the plural for certain types of noun. Thus the Russian for ‘rail’, borrowed in the context of railways, is rel’s,
clearly borrowed from the plural form: rails come in twos after all. In Russian rel’s is singular and has its own regular plural rel’s-y. Tiersma gives examples from Karok, Cahuilla, Tetelcingo Aztec, Acoma, Yokuts, Chamorro and Czech. A particularly interesting example he gives concerns the Latin folium ‘leaf’. Since leaves generally come in large numbers, the plural form folia was the unmarked one, and this developed into a new singular (with plural foliae). In Spanish, the descendant form was hoja (plural hojas). This was borrowed into Chamorro but the plural form was taken, and so ohas (singular) means ‘leaf’.24

5.4 Lexical means
We now consider whether number can be a purely lexical matter, that is, whether there are languages where number forms would be specified for each nominal separately. In §5.3.1 we noted individual examples of this such as Russian čelovek ‘person’ ~ ljudi ‘people’, which is an example of suppletion. Suppletion is the relation between two stems when a regular grammatical opposition is expressed with maximum irregularity (see Mel’čuk 1994 and further discussion in §8.6). Within an inflectional system there may be larger or smaller numbers of items which are suppletive in respect of number, and these are instances of number being lexically marked. Suppletion is common with the pronominal system (as in I ~ we), but as we noted in §3.5 there are also languages in which the relations between singular and plural forms of personal pronouns are inflectionally regular.

It is interesting to ask how much of the nominal inventory can be specified lexically in a given language. In Obolo (a Niger-Congo language, Lower Cross sub-branch, spoken on islands in the Niger Delta; data from Faracías 1984: 10–11, 34) most nouns do not change for number, thus úwù ‘house(s)’. Only a few nouns, all denoting persons, mark number, such as ògwú ‘person’, plural èbí ‘people’, and gwún ‘child’, bón ‘children’. It appears, then, that Obolo nouns mark number only lexically.

In many instances, however, although there is not full suppletion, different number forms are paired in ways which are not regular (this is sometimes called partial suppletion). An example is English tooth ~ teeth, where the forms are

24 Unmarked forms are known to be more tolerant of irregularity than marked forms (Tiersma 1982: 841, following Greenberg 1966: 29). Locally unmarked noun plurals can therefore maintain irregularities which are regularized with marked plurals. When we consider the native irregular plurals of English, it is striking that they involve nouns for which a case can be made for local unmarkedness of the plural: men, women, children, oxen, geese, mice, lice, feet and teeth. It is interesting to note that in an analysis of the Cushitic language Bayso, which has a complex number system, the irregularities investigated by Corbett and Hayward (1987) involved several of the same items: the nouns for ‘feet’, ‘oxen’, ‘teeth’ and others.
clearly similar but are not related by any synchronically active rule (they show stem-internal modification §5.3.3). These must be lexically marked. They have a different status from ordinary singular–plural pairs, in that the plural is available for compounds, as in teeth marks whereas for ordinary plurals it is not *claws marks (Kiparsky 1982: 137). 25 Once we consider forms which are phonologically similar but not related regularly, then we find languages where a substantial portion of the noun lexicon marks plurality in this way. They are unlike the English type of pattern, which has one dominant type of plural formation, and several irregular types with relatively few members. In some languages it is far from obvious whether there is one dominant type of plural.

Perhaps the extreme point in this regard is the particularly interesting patterns of number marking found in Nilotic languages. 26 We shall consider Shilluk, within the Luo group, as described by Gilley (1992). Shilluk is spoken by about a million people in Southern Sudan, along the Nile River in the vicinity of Malakal. Gilley quotes Kohnen, who in his Shilluk grammar published after thirty years of study states that ‘A general rule for the formation of plurals in Shilluk cannot be given’ (1933: 19). After detailed analysis Gilley confirms this and goes on to claim that each noun has two independent representations: singular and plural. These two forms are not totally dissimilar in most cases. For many nouns one form is the base to which a suffix is added (whether to form the singular or plural), while for some nouns both forms have a suffix. 27 Three patterns can be distinguished:

- type A: base versus plural
- type B: singulative versus base
- type C: singulative versus plural

‘Singulative’ here indicates a singular which has a morphological marker, rather than being a base (where singular is indicated by the absence of a marker). The three patterns may be illustrated as follows:

<table>
<thead>
<tr>
<th>Type A</th>
<th>Type B</th>
<th>Type C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ñâŋ ‘crocodile’</td>
<td>ñâŋːi ‘crocodiles’</td>
<td>ñâŋ ‘termite (type of)’</td>
</tr>
<tr>
<td>wâr ‘shoe’</td>
<td>wâr ‘shoes’</td>
<td>ñ’i ‘termites’</td>
</tr>
</tbody>
</table>

25 Booij (1996: 6–7) argues however that in Dutch and some other languages regular plural forms can feed compounding, as in Dutch /student-en/ team ‘students’ team’.


27 Note however that this morphologically basic form seems to have singular or plural meaning, depending on the form it is opposed to, rather than having ‘general’ meaning, from the information given by Gilley (1992: 63–5).
Since these patterns can be established, why should it be claimed that each noun has two representations? The problem is that the relations between the two forms are so varied and complex that, it is claimed, no plausible set of rules can be given.

In terms of affixation, the situation is not too difficult: some nouns add a plural suffix (type A), some a singulative suffix (type B), some have both (type C), while some nouns have neither. But the difficulties arise within the stem: typically this is of the shape consonant–vowel–consonant. The initial segment may be consonant plus glide or just consonant. While the initial segment is usually the same in singular and plural, some nouns have consonant plus glide in the singular opposed to a single consonant in the plural, for example pyēn ~ pē:nī ‘sleeping skin(s)’ (Gilley 1992:84). The final consonant is generally the same; however, it may alternate: singular [l] or [r] may give [t], for example pāl ~ pā:t ‘spoon(s)’. This does not hold for all nouns, as shown by acwi:l ~ acwi:l ‘brown cow(s)’. There are also other erratic alternations.

Most problems arise with the vowel in the consonant–vowel–consonant pattern of the stem. Here there are four sources of variation:

1 **vowel height**: the most common alternations in height involve two front vowels, as in rè:jɔ ¯ ~ ric ‘fish(es)’, two back vowels, as in cùŋ ~ cɔŋ ‘knee(s)’, or two unrounded vowels, as in kwéy ~ kwà:y ‘grandfather(s)’.

2 **vowel quality**: Shilluk has a set of ‘breathy vowels’, which may be described as bearing the feature [+expanded larynx]: these are: i, e, a, ɔ, u. Their [−expanded larynx] counterparts are i, e, a, ɔ, o. The vowel may or may not carry the feature [+expanded larynx] in both singular and plural, thus giving four possibilities:
   a. both [−expanded larynx]: kél ~ kè:l ‘cheetah(s)’
   b. singular [−expanded larynx], plural [+expanded larynx]: ƞàŋ ~ ƞàŋ:i ‘crocodile(s)’
   c. singular [+expanded larynx], plural [−expanded larynx]: pyēn ~ pē:nī ‘sleeping skin(s)’
   d. both [+expanded larynx]: yēp ~ yē:p ‘tail(s)’

The types which have the same value of the feature in singular and plural are the more common; however, the quality of the vowel and the possibility of change in the plural are not predictable.

3 **tone**: there are three tones: high (‘), mid (’) and low (‘); in addition sequences of tones can be found on a single vowel, such as high-low (‘). While some tone patterns are restricted to the singular, and some to the plural, the majority of patterns are found on both. Some of the
many possibilities are illustrated on the examples already given. Others include kùl (low tone) ~ kûl (low-high) ‘pig(s)’, kî (low-high) ~ kî (high-low) ‘plant(s) with edible roots’. According to Gilley (1992:90) ‘It is a rare experience to find a singular/plural pair which has the same tone.’ She claims that the tone cannot be predicted and so must be listed in the lexicon for both singular and plural forms.

4 length: Shilluk has short vowels, and long vowels (indicated by :). The data are complex, since the whole question of Shilluk syllable structure is involved. Suffice it to say that vowel length may vary between singular and plural: the vowel may be long in the singular and short in the plural: bü:r ~ bûr ‘grave(s)’, or the converse may be found: acwi:l ~ acwi:l ‘brown cow(s)’. (There may also be no variation between singular and plural, and this is commonly the case.)

Given the degree of variation according to these four factors, which cannot be predicted, Gilley concludes that singulars and plurals in Shilluk cannot be derived from a single underlying representation but that two forms must be stored for each noun. The claim is not that the two forms are totally unrelated: they are not fully suppletive (though Shilluk does have some fully suppletive singular–plural pairs, such as gin ~ jäm:j ‘thing(s)’). Normally the initial and final consonants are shared. It is the vowel, and its attendant features, which accounts for the differences. It is as though the English plurals of the foot ~ feet and mouse ~ mice type made up a significant proportion of the lexicon (and with more possible types of variation). Shilluk therefore represents a language which marks number by means of a high degree of non-predictable stem variation. Of course, it is difficult to prove a negative, to prove that no set of rules can cover the data. But Gilley makes a plausible case that two forms of nouns must be remembered by the speaker rather than one being derivable from the other.

There are interesting implications for child language acquisition. Gilley (1992:190–1) suggests that if Shilluk children must store two forms for nouns then, if children learn different languages at the same rate, this implies that ‘children learning other languages may be memorizing more than had been thought’. However, though Shilluk is indeed near the extreme in its use of stem formation, it does not necessarily present a qualitatively different task for the language learner. Note first that Gilley reports a suggestion that Shilluk children generalize the use of the -i suffix for plural and then later learn the more complex forms, thus learning the most regular forms, overgeneralizing the rule, and later learning lesser regularities and exceptions. But there is a second important point about Shilluk: like other Nilotic languages, it appears to be moving towards becoming monosyllabic; thus the amount of phonological information which the child has to acquire for a single
lexical entry – even if this involves two related forms for a noun rather than just one – may turn out to be no more than for many other languages which have considerable numbers of polysyllabic words. Of course, we do not know the relative load of storing information such as tone as compared to segmental phonemes, but it is likely that there is a trade-off of length as opposed to complexity; thus the Shilluk child, having to learn two complex one-syllable forms for many nouns may not face a harder task than the child learning a language in which plural formation is regular but in which many nouns are polysyllabic. Shilluk then represents almost the ultimate point in the lexical marking of number.

Outside Nilotic, there are other languages where there is a wide variety of plural forms. In a discussion of Hausa, Haspelmath (1989) argues that there are several competing patterns and none of them is the dominant one. He describes them using ‘schemas’, following work by Köpcke (1988) on plural formation in German (see these two references for other work on schemas).

5.5 Inverse number (and polarity)

In Shilluk we saw instances where the marking of singular number in one noun could be by the same formal means as are used for marking plural in another. We now look at languages where that is regularly the case. An unusual and very interesting system is found in Kiowa (a language of the Kiowa-Tanoan family, which in 1984 had 400 speakers in south-western Oklahoma). The main source is Watkins (1984: 78–100), following earlier work by Wonderly, Gibson and Kirk (1954), who introduced the term ‘inverse’, and Merrifield (1959). There is a marker -gɔ` (with various variants including -dɔ`) , which may be simply attached to nouns or may replace another suffix. It is termed an ‘inverse’ suffix, because it appears to change the basic number meaning of the stem to which it is attached. This switch can go in either direction, as shown in table 5.13.

Table 5.13 Noun number markers in Kiowa

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>dual</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>cê˛·</td>
<td>cê·</td>
<td>cê·gɔ`</td>
<td>horse</td>
<td></td>
</tr>
<tr>
<td>á·dɔ`</td>
<td>á·</td>
<td>á·</td>
<td>pole</td>
<td></td>
</tr>
</tbody>
</table>

The suffix -gɔ` ~ -dɔ` appears to have opposing effects, according to the type of noun.28 We should first ask, however, why it is appropriate to label the columns

28 The related language Jemez has a comparable but interestingly different system of number marking on nouns (Kenneth Hale 1956–57 and personal communication), as the table shows. In Jemez the inverse marker marks the dual, together with either plural or singular
The expression of number

‘singular’, ‘dual’ and ‘plural’. There are two reasons: first, this matches the meanings of the forms, and second, the marking on the verb indicates number. For nouns like cêː ‘horse’, the verb shows a singular-dual-plural system, for example (Watkins 1984: 84):

(21) cêː gyà-tʰɔn
    horse 1.SG.AGT/SG.OBJ-find.PERFV
    ‘I found a horse’

(22) cêː nèn-tʰɔn
    horse 1.SG.AGT/DUAL.OBJ-find.PERFV
    ‘I found two horses’

(23) cêː-gɔ́ dé-tʰɔn
    horse.INV 1.SG.AGT/INV.OBJ-find.PERFV
    ‘I found some horses’

Note that the verbal prefixes mark subject and object, and it is the distinctions for objects that are of interest here, since they separate singular, dual and inverse. There are some complications, involving different inverse or plural markers, but typically the marking on the verb distinguishes three numbers. The two nouns in table 5.13 represent the main classes of Kiowa nouns. All animates behave broadly like cêː ‘horse’, in that the number marker, whether -gɔ́ or some alternative, is found only in the plural. Other examples include tɔ́l ‘father’. (Interestingly, just one noun tʼáp ‘deer’ does not decline, but it takes the same agreements as cêː ‘horse’: it is like English sheep.) There are also some inanimates which behave like cêː ‘horse’, such as dèn ‘tongue’ and pʼɔ́ ‘river, stream’. However, most nouns denoting inanimates behave like áː ‘pole, stick’; it is not clear whether it can be predicted which inanimates behave like the animates.

There are smaller groups of inanimates with different behaviour, which are probably best viewed as irregulars. There are just four nouns (‘orange’, ‘tomato’, ‘plum, apple’, ‘hair (on head)’) which take the inverse marker -gɔ́ for singular and

Footnote 28 (cont.)

(depending on the noun, in the main animates follow the first pattern and inanimates the second). For Taos see Trager (1961), for Tewa see Speirs (1972), and for comparisons across the family see Watkins (1995).

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>dual</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>veːla</td>
<td>veːlæš</td>
<td>veːlæš</td>
<td>man</td>
<td></td>
</tr>
<tr>
<td>tʼetibæš</td>
<td>tʼetibæš</td>
<td>tʼetiba</td>
<td>box</td>
<td></td>
</tr>
</tbody>
</table>
plural but not for the dual. And then there are various other inanimates which do not mark number on the noun at all: some like c’ó· ‘rock’ are simple indeclinables for which the verb marking establishes number in a straightforward way. For a second group, the verb distinguishes only the dual. And the third group consists of those for which number is not distinguished; these include kʰɔ·dé ‘trousers’ and tó· ‘tepee’. These always take plural markers on the verb (they are pluralia tantum, as in §5.8.2 below).

If we concentrate on the main classes of nouns we see that each has a basic form, without -gɔ and a ‘less expected’ form with it. Those denoting animates are singular/dual in their basic form, with -gɔ signalling a shift from that number, while nouns in the main class for inanimates are treated as basically dual/plural, with -gɔ and variants signalling a shift to singular. It looks as though we have a new number value, singular/dual (for nouns denoting one or two) in the case of those denoting animates. This would be a new number system, a possibility not allowed for in our typology of number values (§2.3.2). However, the verbal system resolves this apparent new value into the normal values singular and plural. As we saw in §4.5.1, such ‘conflated numbers’ can exist, but they occur lower on the Animacy Hierarchy with a regular system higher on the hierarchy. There could not be a language like Kiowa but in which the singular/dual conflation were not resolved by the verbal system and still occurred higher on the hierarchy than the dual/plural conflation (which is of course a normal plural). Conflated forms which give an irregular number value must occur lower on the hierarchy than a regular system of number values. Thus we are dealing with unusual noun morphology here: Kiowa does not give us new number values unknown elsewhere. The overall system is singular–dual–plural. Thus inverse systems are a matter of morphological arrangement and not of semantic values.

There are two interesting points of connection with the Animacy Hierarchy. First, the two main classes of noun in Kiowa, one with the inverse marker for plural and the other with inverse marking for singular, conform broadly with the Animacy Hierarchy, since the first contains all the animates. And second, we noted in §3.3 that there can be a mismatch between the evidence of marking on the noun and the evidence of nominal number marked on the verb. The latter is the more regular. We have a similar situation here, in that the marking of number on the animates gives an unexpected plural versus singular/dual system; however, the verb is quite regular in showing singular–dual–plural. The opposite cannot occur; we do not find the converse of Kiowa, with the nouns showing straightforward number values and the verb having an ‘illegal’ system (we are concerned with systems here and not with individual irregular items within a regular system). The marking of one number for some nouns and another for others also recalls the discussion of the marking of plural versus general and singular versus general in §2.1. What is
remarkable about Kiowa is that the same marker is used for apparently opposing ends.

The notion of an inverse marker which indicates the less expected number gains some support from a quite different part of the world, from the Nilotic language Maa (Bernd Heine, personal communications). Various suffixes indicate singular or plural, depending on the noun; among them, the suffix -i may be thought of as signalling the unexpected number. In the Maasai dialect, for instance, we have:

\[
\begin{align*}
(24) & \quad \text{o-sínkirr-î} & \quad \text{i-sínkir} \\
& \quad \text{GN-fish-SG} & \quad \text{GN-fish} \\
& \quad \text{‘a fish’} & \quad \text{‘fish’ (plural)}
\end{align*}
\]

There is a prefixed gender/number marker, and a number suffix. The Maasai have a taboo against eating fish. Thus their singling out one fish would be less expected, and the singular is marked with the suffix -î. However, the speakers of the Camus dialect, who live at Lake Baringo in Kenya, do eat fish. And for them, the forms are:

\[
\begin{align*}
(25) & \quad \text{sinkir} & \quad \text{sinkir-î} \\
& \quad \text{fish} & \quad \text{fish-PL} \\
& \quad \text{‘a fish’} & \quad \text{‘fish’ (plural)}
\end{align*}
\]

The gender–number prefix vowel is lost in the Camus dialect. For these speakers the noun behaves as anticipated; the plural is the less expected number value, and has the special marker. This is a singular–plural system, not one with general number.

If we consider other means of marking number from the same point of view, then we find, elsewhere in the world, systems which are similarly surprising. Ross (1988: 293–305)\(^{29}\) discusses the systems found in two groups of Oceanic languages. Nine are spoken in southern New Ireland: Lihir, Lamasong, Madak, Tangga, Bilur, Kandas, Ramoaaina (previously called Duke of York), Siar and Tomoip, and a further nine in northern Bougainville: Nehan, Solos, Petats, Halia, Taiof, Hahon, Tinputz, Teop and Papapana. These languages divide their nouns into two classes. The first class is readily identifiable as the class of nouns which head count noun phrases, as in this Ramoaaina example:

\[
\begin{align*}
(26) & \quad \text{a pap a kum pap} \\
& \quad \text{ART dog ART PL dog} \\
& \quad \text{‘a/the dog’ ‘some/the dogs’}
\end{align*}
\]

\(^{29}\) I am grateful to Malcolm Ross for bringing this phenomenon to my attention and for generously supplying fieldnotes and references.
The second class, which we would expect to be those which head mass noun phrases, includes nouns denoting fish, fruit, birds and trees; we will here call them simply 'non-count'. The noun pika, which might be glossed 'bird' is a non-count noun (Ross glosses it 'poultry'). Together with the article, it denotes a portion or usual unit (§3.7.2):

(27) a pika
    ART poultry
    ‘a plate of poultry’

For reference to an individual bird, what Ross terms a ‘quantity marker’ (QM) is required, here functioning as a singulative:

(28) a ina pika
    ART QM poultry
    ‘a bird’

In this language, plurality is signalled as for the count nouns:

(29) a kum pika
    ART PL poultry
    ‘some/the birds’

The eighteen languages in question make a similar distinction of nouns into the two classes. However, the formal means used to make the distinction vary considerably, and there is some syntactic variation too. For our purposes, the languages of particular interest are those like Tangga, which shows the following pattern (Ross 1988: 295, 298):

(30) fel am-fel taña fel
    house QM-house some house
    ‘the house’ ‘the houses’ ‘some houses’

(31) man an-man taña man
    bird QM-bird some bird
    ‘the bird’ ‘some birds’

Here we see that the quantity marker anlamlag, which indicates plural for count nouns, is also the singulative marker for non-count nouns. Hence we have a situation analogous to the Kiowa inverse system (though here only singular and plural values are involved).

Further examples are found in north Bougainville languages. In Nehan we find the following pattern (Ross 1988: 299, 301; this supersedes the account in Todd 1978):

5.5 Inverse number (and polarity)
The expression of number

(32) a um[a] o um[a]  
ART house ART house  
’a/the house’ ‘some/the houses’

(33) o dok[i] a dok[i]  
ART tree ART tree  
’a tree, a stick’ ‘a collection of trees’

The bracketed final vowels are not realized in citation forms, nor phrase-finally. Some nouns can function as count and non-count, with related meanings, as is the case for pos[o]:

(34) a pos[o] o pos[o]  
ART banana ART banana  
’a/the banana’ ‘some/the bananas’

(35) o pos[o] a pos[o]  
ART banana ART banana  
’a banana tree’ ‘a collection of banana trees’

Thus both a and o signal singularity for some nouns and plurality for others. The other north Bougainville languages, apart from Hahon, share elements of inverse number systems, though with additional complications (Ross 1988: 299–301); for Halia see Allen (1987). An example of the type of complication which may arise is provided by Teop (Mosel and Spriggs 1993: 43–51). Here there are three classes of noun, which we may call the e-class, a-class and o-class, according to the article taken. The types of noun which typically belong in each class are summarized in table 5.14. Now consider the singular and plural forms of the articles taken by the nouns of these different types of noun in Teop (see table 5.15). Here two classes (the a-class and the o-class) exhibit inverse number, while the e-class does not. This means that we now have nouns which take different agreements and which can reasonably be said to belong to different genders (see Corbett 1991: 145–50 for defini-

| Classes of noun and their semantics in Teop (Mosel and Spriggs 1993) |
|-----------------|---------------|---------------|---------------|
|                 | e-class (+)   | a-class (+)   | o-class (+)   |
| humans          | +             | +             | –             |
| animals         | (+)           | +             | –             |
| trees           | –             | –             | +             |
| fruit           | –             | +             | –             |
| manufactured items | –             | +             | +             |
A gender system has arisen in an unusual way and in an unexpected part of the world, as Mosel and Spriggs (2000) point out.

We might have expected that inverse systems would arise as mere accidents of phonological change. However Ross (1988: 301–5) gives some insight into the origins of these inverse systems which indicates that this may not be the case. He suggests that the o article which occurs with count nouns and that which occurs with non-count nouns have different origins; from comparative evidence he reconstructs an u form of the article, used to mark plurality with count nouns. However, there is no phonological reason for the falling together of the original u and o articles. Thus the limited evidence from Oceanic suggests that inverse systems are not merely the result of chance.

Once we are dealing with agreement (of the article and other targets) and with gender as well as number, then we come to the interesting phenomenon of polarity: the situation in which two markers are exponents of two features (gender and number) and when the value of one feature is changed the marker changes, but if both values are changed the form stays the same. The polar opposites are identical, hence the term ‘polarity’. This phenomenon can be found in the Cushitic language Somali (data from Serzisko 1982: 184–6; see also Bell 1953: 12–13 and Saeed 1987: 114–16):

<table>
<thead>
<tr>
<th></th>
<th>e-class</th>
<th>a-class</th>
<th>o-class</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>e</td>
<td>a</td>
<td>o</td>
</tr>
<tr>
<td>plural</td>
<td>o</td>
<td>o</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 5.15 Articles taken by different noun types in Teop

(36) inan-kii baa y-imid
    'the boy (!) came’
    (the (!) indicates that boy is focused)

(37) inán-tii baa t-imid
    'the girl (!) came’
    SG.FEM-came

(38) inammá-dii baa y-imid
    'the boys (!) came’
    PL-came

(39) ináma-hii baa y-imid
    'the girls (!) came’
    PL-came
The postposed definite article has various morphophonologically determined variants: after any vowel except \(i\), \(kii\) becomes \(hii\), and after any vowel \(tii\) becomes \(dii\). Given this, in the examples above the article used for the masculine plural might be considered the same as that for the feminine singular, while that for the feminine plural is the same as that for the masculine singular. The basic forms are as in table 5.16. Here masculine singular and feminine plural are the same (as are feminine singular and masculine plural); of course, the case would be more convincing if these basic forms were not subject to variation. While table 5.16 suggests a clear and surprising picture, things are actually more complicated. First, Somali has polarity only in noun-phrase-internal agreement. Examples (36)–(39) show that the verbal agreement forms are different: there the plural for both genders is the same as the masculine singular. And second, not all nouns fall into the pattern shown in (36)–(39). Some masculine nouns form their plural by partial reduplication and take the same article in the singular and the plural, for example \(nin-kii\) ‘the man’, \(niman-kii\) ‘the men’. Thus not all targets show polarity, nor are all nouns included in the polarity system. (Conversely, a small number of nouns is exceptional in taking polarity type agreements for predicate agreement too: see Hetzron 1972, Zwicky and Pullum 1983.) The importance of polarity should not be overrated, as Speiser warned (1938).

Inverse number and the special case of polarity are surprising and interesting phenomena. However, they are interesting only as means of expression. They do not add to the semantic possibilities of number systems.

### 5.6 Minimal-augmented systems

In descriptions of certain languages of the Philippines, starting from Thomas (1955) on Ilocano, there was a perception that the data did not fit well into conventional accounts of person and number. Conklin (1962: 134–6) proposed a different analysis for Hanunoo, which also proved applicable to various languages of Arnhem Land in Australia. We shall follow the account of one of the latter, namely Rembarrnga, as described in McKay (1978, 1979). First consider the Rembarrnga forms given in table 5.17 (McKay 1978: 28). This analysis captures the facts, but it seems unsatisfactory: first the paradigm looks disjointed, and
second there is a marker -bbarrah, marked in bold in table 5.17, which appears in an odd selection of cells. If we try to characterize -bbarrah in absolute terms, we find no solution: in one instance it is used of three individuals, in other instances it is for two. But if we treat it (and the entire system) in relative terms (Evans §7.1.1 in forthcoming) then a more elegant picture emerges. The form -bbarrah is used when there is one entity more than the logical minimum. For most cells that view makes no difference. However, for the first person inclusive, the logical minimum is two (otherwise it would not be inclusive). Thus yu kku is a minimal form (we can label it 1/2 to suggest it represents another person value) and ngakorrbbarrah is used where there is one more than that minimum, that is, three. We redraw the paradigm from this relative perspective (McKay 1978: 28) in table 5.18. This is a more satisfy-

Table 5.17 Rembarrnga dative pronoun forms: traditional categories

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>dual</th>
<th>trial</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 inclusive</td>
<td>—</td>
<td>yu kku</td>
<td>ngakorrbbarrah</td>
<td>ngakorru</td>
</tr>
<tr>
<td>1 exclusive</td>
<td>ngunu</td>
<td>yarbbbarrah</td>
<td>yarru</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ku</td>
<td>nakorrbbarrah</td>
<td>nakorru</td>
<td></td>
</tr>
<tr>
<td>3 masculine</td>
<td>nau</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 feminine</td>
<td>ngadu</td>
<td>barrbbarrah</td>
<td>barru</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.18 Rembarrnga dative pronoun forms: minimal-augmented analysis

<table>
<thead>
<tr>
<th></th>
<th>minimal</th>
<th>unit augmented</th>
<th>augmented</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ngunu</td>
<td>yarbbbarrah</td>
<td>yarru</td>
</tr>
<tr>
<td>1/2</td>
<td>yu kku</td>
<td>ngakorrbbarrah</td>
<td>ngakorru</td>
</tr>
<tr>
<td>2</td>
<td>ku</td>
<td>nakorrbbarrah</td>
<td>nakorru</td>
</tr>
<tr>
<td>3 masc</td>
<td>nau</td>
<td>barrbbarrah</td>
<td>barru</td>
</tr>
<tr>
<td>3 fem</td>
<td>ngadu</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We have forms for the minimal number of the pronoun, for one more than that (the unit augmented form) and and for more than that again (augmented). The simplest system of this type would have just minimal and augmented. In such systems, with only two number values, the difference from conventional systems is in one form only. We see this clearly by comparing Ilocano
analysed in the two different ways. In the traditional analysis (table 5.19) there is just one dual form; from our discussion in §4.1, if there is an additional number value in one place only, the first person is exactly the place in which we would expect to find it.

<table>
<thead>
<tr>
<th>singular</th>
<th>dual</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 exclusive</td>
<td>-ko</td>
<td>-mi</td>
</tr>
<tr>
<td>1 inclusive</td>
<td>-ta</td>
<td>-tayo</td>
</tr>
<tr>
<td>2</td>
<td>-mo</td>
<td>-yo</td>
</tr>
<tr>
<td>3</td>
<td>-na</td>
<td>-da</td>
</tr>
</tbody>
</table>

On the other hand, though the evidence is not as convincing as in Rembarrnga, there are grounds for favouring the minimal-augmented account for Ilocano given in table 5.20. The main difference is in the status accorded to -ta. The possible ambiguity of the first person inclusive dual or 1/2 minimal form in such cases is discussed in an interesting exchange (Greenberg 1988, McGregor 1989, Greenberg 1989, McKay 1990). Besides being found in Hanunóo as noted earlier, and other languages of the Philippines, minimal-augmented systems have been identified in various languages of Arnhem Land, starting with Burrara (Burera) (Glasgow 1964);30 others are listed

Table 5.20 *Minimal-augmented analysis of Ilocano pronominal forms*

<table>
<thead>
<tr>
<th>minimal</th>
<th>augmented</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 -ko</td>
<td>-mi</td>
</tr>
<tr>
<td>1/2 -ta</td>
<td>-tayo</td>
</tr>
<tr>
<td>2 -mo</td>
<td>-yo</td>
</tr>
<tr>
<td>3 -na</td>
<td>-da</td>
</tr>
</tbody>
</table>

30 It is important to note how similarly these number systems behave to those we have seen before. Thus Gurr-goni is one of the four Manigrida group languages (the others are Burrara, as just mentioned, Nakkara and Ndjebbana), all of which have systems similar to that of Rembarrnga. In Gurr-goni the number distinctions are available only for reference to humans and higher animates (as in the systems discussed in chapter 2). Furthermore, in the third person the minimal form is unmarked even for reference to humans. However, if an augmented form is used, it must be the appropriate one. In other words, the minimal form acts rather like a general number form (partially similar to Kaytete, §2.3.3); the choice is minimal/general or number-specific: if the latter is chosen it must be the appropriate form, that is, unit augmented versus augmented is not a facultative choice in Gurr-goni. The information on Gurr-goni is from Rebecca Green (personal communication).
in McKay (1978:29). Following from these alternative analyses, the essential point is that minimal-augmented systems represent an alternative way of organizing the morphology of person and number. They do not give additional semantic distinctions in number. This is why we consider them here as an alternative means of expression, and not as a set of additional number values.  

5.7 ‘Constructed’ numbers

Constructed numbers appear where there is a mismatch between number marking of different elements which produces additional number values. Consider the following data from the Uto-Aztecan language Hopi (Hale 1997: 74). The pronominal and verbal forms each make a two-way distinction:

(40) Pam wari
    that.SG run.PERFV.SG
    ‘He/she ran’

(41) Puma yùtu
    that.PL run.PERFV.PL
    ‘They (plural) ran’

However there is a third possibility:

(42) Puma wari
    that.PL run.PERFV.SG
    ‘They (two) ran’

The combination of plural pronoun and singular verb gives a dual (‘they two ran’). This dual is ‘constructed’ from the number on the pronoun and that on the verb; we have a singular–dual–plural system, ‘constructed’ from the two parts. If we retain notional labels, then we could say that the pronoun distinguishes singular from dual/plural while the verb distinguishes singular/dual from plural. It must be stressed, however, that this is only a part of the system: animate nouns in Hopi have a straightforward singular–dual–plural system, indicated by three distinct markers.

Now consider Zuni, a language isolate with some 8,000 speakers in north-west New Mexico (data from Lynn Nichols, personal communications). No pronoun is found in the third person, and there is a dual marker available. The verb has a marker for the plural (as in (47) below), otherwise it takes no number marker.

31 Similarly, the ‘quasi-duals’ of Bantu represent interesting combinations of person, but do not extend the semantic possibilities of number; for the complex system of Ngembaon-Bamileke, a Grassfield Bantu language of Cameroon, see S. C. Anderson (1985); for consideration of other complex systems using the ‘augmented’ notion see Noyer (1997: 148–54).
(43) ʔa:či ʔa:-kya  
DUAL go-past  
‘they (two) went’

A pronoun is needed for first or second person:

(44) hon ʔa:či ʔa:-kya  
1.PL.NOM DUAL go-past  
‘we (two) went’

Here we have the dual marker, but the pronoun is plural, and the verb has no marker (hence is singular). Moreover, the dual marker may be dropped:

(45) hon ʔa:-kya  
1.PL.NOM go-past  
‘we (two) went’

This is an alternative to (44), and is similar to the Hopi equivalent form. It is not possible to reverse the markers:

(46) *hoʔʔ ʔa:w-a:-kya  
1.SG.NOM PL-go-past  
‘we (two) went’

This is ungrammatical in any interpretation. If both are plural, then not surprisingly the interpretation is plural:

(47) hon ʔa:w-a:-kya  
1.PL.NOM PL-go-past  
‘we went’

Where Zuni goes beyond Hopi is in using this system with nouns too. The plural prefix here is ʔa:- before consonant-initial stems, ʔa:w- before vowel-initial stems; the final ʔi of ʔa:w-akcek(ʔi) drops before ʔa:

(48) ʔa:w-akcek(ʔi) ʔa:-kya  
PL-boy go-past  
‘two boys went’

(49) ʔa:w-akcek(ʔi) ʔa:či ʔa:-kya  
PL-boy DUAL go-past  
‘two boys went’

Here again the dual ʔa:či can be included as in (49) or omitted as in (48). When it is omitted, we have a constructed dual, this time with nouns as well as with pronouns. Hopi and Zuni are both found in North America, as is Kawaiisu, a Uto-Aztecan
language of the Numic branch, which also shows the phenomenon (see Zigmond, Booth and Munro 1991: 76). Elsewhere there is the complex case of Mele-Fila, as we saw in §2.2.6.

Before leaving constructed numbers there are several general points to make. The first is that while the phenomenon is surprising and interesting, it tends to be limited in terms of its extent within a number system. For instance in Hopi only the pronouns were affected: nouns had full marking. Second, and related to this: in all of the languages pronouns are involved; constructed number appears to affect the top of the Animacy Hierarchy, but we have rather few languages and so should be cautious about this claim. There are interesting differences between these systems and conflated numbers (§4.5.1) in that constructed numbers involve a top segment of the hierarchy, and the conflation (which is not in the nominal system) is resolved, in a way which gives an additional number value. And finally, the nature of the verbal forms requires careful scrutiny: at least in some cases we are dealing with verbal number, which we shall consider in chapter 8; see also the discussion of Chamorro in Durie (1986: 364–5). In such cases the number value is constructed from the value of nominal number of the pronoun or noun together with that which can be inferred from the participant number value of the verb.

Given what we have seen in Hopi, we might consider whether to treat British English as having a constructed number, shown in examples like:

\[(50) \text{This committee have decided . . .} \]

Since the agreements are singular and plural, does this constitute a constructed number? There are differences which should make us suspicious: first the number value which arose in the constructed systems (typically the dual) was one found in regular systems. The English construction does not give rise to a normal number (from the inventory in §2.3.2). Nor does the construction affect a top segment of the Animacy Hierarchy. What we have rather is a set of nouns which vary in the number agreement they take (from target to target, from speaker to speaker, and most importantly from noun to noun). They are ‘hybrids’ in this respect, and will be discussed more fully in §6.2. Finally, note that these, like constructed number cases, are instances of a wider problem of mismatches in number: Kiowa (§5.5) showed a mismatch in the system found in the noun phrase and on the verb, there is also the Hebrew type (where the values on the verb are a subset of those on certain nouns, §4.2.1), and that type in turn can be seen in the context of differential marking on noun phrase and verb discussed in §6.1.1.

5.8 Reduced expression of number

We now turn to those nouns which for various reasons do not have the full range of number possibilities. Let us start with a basic singular–plural system as in English
and consider the possible combinations for distinguishing number morphologically, syntactically and semantically. Note that nouns can move from type to type, through intermediate stages. Let us begin at the right of table 5.21 with the non-number-differentiable noun *friendliness*. This noun cannot straightforwardly distinguish number in semantic terms. We can talk of *instances of friendliness*, *kinds of friendliness*, but nouns like *instances* and *kinds* indicate the semantic shifts or recategorizations (§3.7.2) involved. It shows no syntactic or morphological signs of number-differentiability (that is, no agreement or marking opposition). *Friendliness* then is simply off the scale of number-differentiability for English, being low on the Animacy Hierarchy. We might conclude that number is therefore irrelevant and no more need be said. But there is still the question of which number it is: its form and especially the agreements it takes show that it is singular, so it is called a *singulare tantum* (singular only) noun. Nouns which are non-number-differentiable are not always singular: even in English we have the contrast between *wheat* and *oats*. We will therefore ask how languages deal with nouns which are off the scale of number-differentiability (§5.8.1).

The noun *scissors* is rather different; it is exceptional in that it denotes a countable entity (there is no problem about counting pairs of scissors) but expressing the opposition in number typically requires a classifier, namely *pair*. The countability of *scissors* is not reflected in its syntax and morphology, since it has only plural forms (it is a *plurale tantum* noun). *Sheep* is number-differentiable as its syntax (agreement) shows, but lacks the appropriate morphology (a mismatch discussed in §3.3). It is defective just in terms of its morphology. We will consider various types of defectives in §5.8.2.

There is a second, cross-cutting distinction here. There are several nouns like *scissors*, which denote objects made of two similar parts (*trousers, goggles* and so on) and their defectiveness is motivated. *Sheep* on the other hand is an example of an individual defective noun, which does not represent a class of similar nouns (hence we have *goats, pigs* and *cows*). Surprisingly, *aircraft* behaves like *sheep*;

---

**Table 5.21 Different indicators of number-differentiability**

<table>
<thead>
<tr>
<th></th>
<th>dog</th>
<th>sheep</th>
<th>scissors</th>
<th>friendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantics</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>syntax</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>morphology</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

§32 Where by distinguishing number semantically we mean that we distinguish individuals and collections of individuals (as shown by the ability to enumerate them; see the discussion of Allan’s work in §3.6 for a more delicate analysis).
these appear to be unmotivated exceptions. We discuss motivation further in §5.8.3.

Finally, *dog* is a typical noun of English; it may refer to an individual or to a set of individuals, and in the latter case the agreements it takes and the morphology are both plural (*the dogs are barking*). Thus *dog* is a fully number-differentiable noun, with all the possibilities (just two in English). Such nouns have been a main focus of attention, particularly in chapters 2 and 3. This section concerns those nouns not so favoured in terms of number.

### 5.8.1 Nouns which are not number-differentiable

Typically there is a set of nouns, very large in some languages, very small or non-existent at the other extreme (§3.7.2), which are off the end of the scale for number-differentiability, nouns like *friendliness* in English. As we noted in §3.6, the point at which the ‘split’ comes can vary slightly and in interesting ways between relatively closely related languages. *Friendliness* is a *singulare tantum* noun, but it is not defective. For a typical noun of its type in English, given where it stands on the Animacy Hierarchy, it has all the forms expected, namely one. Typically in English such nouns are singular. Their form usually suggests this and agreements show it clearly (*his friendliness was particularly appreciated that day*). Not all behave this way in English, thus we have *oats* in contrast to *wheat* (for discussion see Wierzbicka 1988: 459–60, 1991a, 1991b; Palmer 1990; Moravcsik 1991). There are many languages like English in this respect. Some with larger systems still work in the same way: in Mansi, which has a singular–dual–plural system, nouns which are not number-differentiable have just the singular (§4.1). However, this is not the only solution. In Manam (Lichtenberk 1983: 269) all mass nouns are plural, (as marking on the verb shows). In Turkana, some are singular and some are plural (Dimmendaal 1983: 224). In Bantu languages too, mass nouns are frequently split between singular and plural (Guthrie 1948: 851).

There may be other items which do not fully qualify as nouns: nominalizations of verbs in some languages come into this category. Such items usually remain outside the number system, and take default agreement (§6.1.2), usually singular, which makes them similar to *singulare tantum* nouns. Differences between such items and genuine *singulare tantum* nouns may be evident in conjoined structures; there is variation from language to language, but plural agreement is more likely, I suggest, with conjoined *singulare tantum* nouns than with conjoined nominalizations.33

---

33 The reader may wish to investigate judgements on sentences such as:

```
wheat and barley make/makes a good combination
persistence and intelligence make/makes a good combination
running and cycling make/makes a good combination
```
5.8.2 Defectives

Nouns like English *scissors* are defective, in that they have only one number form, and the other is missing, even though in semantic terms there would be no difficulty (for extensive examples see Wickens 1992). Since these instances involve twosomes, an obvious question is what would happen in a language with a dual.

We turn to Central Alaskan Yup’ik (data here are from Marianne Mithun, personal communication, elicited from Elizabeth Ali). We saw in §3.7.2 that there are few mass nouns (which have just singular number); there are also relatively few nouns which are defective in terms of number in other ways. There are some examples, however: *uskurak* ‘dog harness’ is dual in form, appears with dual pronominal suffixes on the verb, and can be used for one dog harness or for two. Three or more dog harnesses would be *uskurat*. There is no singular form *uskuraq. Atasuak* ‘summer trousers’ functions similarly. (For comparative data on the dual in Eskimo languages see Hammerich 1959.) As so often, larger systems are valuable indicators. *Uskurak* is defective, in that it lacks a singular, but it is not *duale tantum*, it has two forms out of three. Consider this example, with *niicugnissuutet* ‘radio’:

(51) niicugni-ssuut-et nipe-s-ki
    listen-INST.NOMINALIZER-PL go.out-CAUS-OPTATIVE.2.SG/3.PL

‘Turn off the radio!’

This is the only number form available, so this noun is a *plurale tantum* (missing two forms out of three). Recalling the evidence from §3.7.2, we may summarize the noun types of Yup’ik as in table 5.22. The great majority of nouns are normal count nouns; of the mass nouns, most allow recategorization (§3.7.2); just a few like *meq* ‘water’ are not number-differentiable, and have just the singular. There are two sorts of defective nouns (within the *scissors*-type), those which have dual and plural, and those which have just the plural.

<table>
<thead>
<tr>
<th>Table 5.22 Main noun types in Central Alaskan Yup’ik</th>
</tr>
</thead>
<tbody>
<tr>
<td>description</td>
</tr>
<tr>
<td>normal count noun</td>
</tr>
<tr>
<td>mass noun (with recategorization)</td>
</tr>
<tr>
<td>mass noun</td>
</tr>
<tr>
<td>missing singular</td>
</tr>
<tr>
<td>pluralia tantum</td>
</tr>
</tbody>
</table>
Continuing with the different possible types of defective nouns, let us move on to the ‘sheep-type’, those lacking morphological forms but being regular in terms of agreement. An interesting example is from the Nakh-Daghestanian language Tsez (Bernard Comrie, personal communication). It is xex-bi ‘child(ren)’; this noun is plural in form (the -bi is a regular plural marker) and it has a full plural paradigm of case forms (thus the genitive is xex-za-s). It may denote one or more children, and takes the appropriate agreements, singular for one and plural for more than one. The noun y'ana-bi ‘woman/women’ behaves similarly.

There are still further possible types of defectiveness to be explored. Thus in Arbore, which has singular and plural (usually with one form also available for use as a general form; see §2.1) there are nouns with just one form, which depending on the noun may match the typical form of general number, singular or plural. And for the surprising defectives of Bayso, the adventurous reader should tackle Corbett and Hayward (1987), for which §6.1.1 is gentle preparation.

If we look at nouns which do not have the full range of values, whether because they are non-number-differentiable or because they are defective, then the data we have discussed suggest that the values line up broadly with the scheme of §2.3.2. Generally singularia tantum are the most common; we find instances with just the plural or with dual and plural but lacking the singular; dualia tantum are quite rare.

5.8.3 Motivation for defectives
The noun scissors is not a haphazard exception. It is an example of what Quirk et al. call ‘summation plurals’, which ‘denote tools, instruments, and articles of dress consisting of two equal parts which are joined together’ (1985: 300; see Wierzbicka 1988: 514–16 for discussion); Quirk et al. point out that the plural -s is often not found when such nouns are used attributively as in a trouser leg. Similarly the motivation for Yup‘ik atasuak ‘summer trousers’ having dual and plural seems obvious, since we are used to trousers being defective. And this can be a persistent defective-ness: Russian has borrowed the word jeans, and created džins-y; the English -s is made part of the stem, and the Russian plural inflections are added, making the word plurale tantum. The duality of uskurak ‘dog harness’ may not be so immediately obvious to us, but then most of us do not have regular experience of harnessing huskies. Sometimes the problem is the outsider’s lack of knowledge; but for speakers too, of course, there may be no retrievable motivation. We noted niicugni-suutet ‘radio’, which is plurale tantum and a similar example is alliritet ‘one-piece trousers with attached fur socks’, also found only in the plural. Such nouns typically denote objects consisting of multiple parts. It does not follow that every noun

34 This motivation carries over into galife ‘riding breeches’ which is indeclinable but also a plurale tantum noun in that it takes only plural agreements.
denoting objects of this type is *plurale tantum* in Yup’ik, any more than paired objects are of that type in English (for example, *bicycle* is quite regular). There are fascinating groups and oddities to be investigated. For instance, names of festivals and celebrations are often plural, as in Estonian, Finnish, Latvian and Lithuanian; examples include Estonian *kihlad* ‘engagement’, *varrud* ‘christening party’ (Tauli 1973: 79). Another common grouping is names of diseases, originally perhaps from the multiple signs on the body (for example, Estonian *sarlad* ‘scarlet fever’). The nouns in such groupings can motivate each other, even when the original motivation is lost.

There are other interesting cases, from small groups down to individual lexical items. In the Nakh-Daghestanian language Archi, *buwa* ‘father’ and *dija* ‘mother’ are both *singularia tantum* (Aleksandr Kibrik, personal communication). Even in English there are some surprising hidden cases. Quirk et al. (1985: 300–1) point out that there seem to be very few defectives in English, because generally the other form exists, though with slightly different meaning. For instance we have both *fund* and *funds*, but these do not match up completely as the singular and plural of a single lexical item (see earlier discussion of similar cases in §3.6 note 24). Conversely, suppletion (§5.4) may be viewed as two defectives functioning as a single lexical item.

Since defective nouns may be synchronically unmotivated, they can be re-integrated into the full number system. *Pluralia tantum* nouns are therefore a potential source for ‘double plurals’ (§5.3.6); if they are potentially countable, their plural form may be used as a singular, and a new plural suffix added; this appears to have happened for instance in the Nakh-Daghestanian language Lezgian (Haspelmath 1993: 81–2), where there are nouns like *gurar* ‘stairs’ and *purar* ‘saddle’, with the plural suffix -ar, but which now behave like normal count nouns and have the plurals *gurar-ar* ‘staircases’ and *purar-ar* ‘saddles’. On the way to this point, nouns may take mixed agreements; the noun *sur* ‘grave’ gives the plural *surar*, which means ‘cemetery’, but this is itself countable. *Surar* ‘cemetery’ does not take a (double) plural suffix, but has alternative agreement possibilities:

(52) Či šeher.đi-n surar jeke-bur / jeke-di 1.PL.GEN town-GEN cemetery big-SBST.PL / big-SBST.SG ja.

COP

“Our town’s cemetery is large.”

In terms of the types discussed earlier, *surar* ‘cemetery’ has moved beyond the *scissors* type in that it allows singular agreement, but does not yet belong to the *sheep* type in that it does not yet always take the number agreement matching the meaning.
5.9 Conclusion

We have examined the ways in which number is expressed, and found that they are many and varied, with morphological means providing the greatest variety. We also looked in detail at systems which are unusual in terms of the marking of number, yet which do not give new semantic distinctions. These were a useful reminder of the need to be clear whether claimed generalizations relate to number values or to the means of their expression, and more generally to be careful about comparing like with like. Having looked at the various means of expression, with morphology taking centre stage, in the next chapter we concentrate on syntactic questions.