

A Typology of Number Systems

Greville G. Corbett
LIS, University of Surrey, Guildford, Surrey, GU2 5XH

One of the categories most frequently associated with the noun phrase is number, a category which turns out to be much more difficult and interesting than it at first appears.¹ It is of particular importance at present for work in natural language processing; a recurring difficulty in parsing is that of local ambiguity, and it is clear that number (in particular, agreement in number) has an important role in resolving these ambiguities. I will propose a general typology of number systems, and suggest how the languages of Europe in particular fit into this typology.²

I will first outline the distinction between nominal and verbal number (section 1), then move on to the semantics of number (section 2), and finally discuss the ways in which number is expressed (section 3).

1. Nominal and verbal number

For speakers of Indo-European languages, it seems natural to consider number to be a nominal category, distinguishing *book* from *books*, *cat* from *cats* and so on. When we talk of verbal number (if at all), then often it is simply nominal number expressed (by agreement) on the verb:

- | | |
|-------------------------|--------------------------|
| (1a) the goat jumps ... | (1b) the sheep jumps ... |
| (2a) the goats jump ... | (2b) the sheep jump ... |

The essential difference between (1) and (2) is, of course, the number of goats or sheep involved. But as we shall see, verbal number encodes a rather different opposition.

1.1. Nominal and verbal number: meaning

Now consider this sentence:

¹ Thanks to Alan Cruse, David Gil, Dick Hayward, Sasha Kibrik, Nina Koval', Velma Leeding, Nick Reid, Bruce Rigsby and John Roberts for useful information or helpful discussion of particular languages and comments at various times, and to Jim Hurford, Edith Moravcsik and Neil Smith for discussing markedness. Usual disclaimers. Earlier versions of the paper were read at La Trobe University (April 1991), the Australian National University (May 1991) and the University of Manchester (November 1991).

² I am in the early stages of investigating number, so I shall be particularly grateful for comments and suggestions.

- (3) the girl ran round the track

This could mean that the girl ran round the track once, or several times. We could express the latter meaning with *round and round*. This possible difference is retained if the noun phrase is plural:

- (4) the girls ran round the track

Again, one circuit of the track may be intended, or several. In other words, the running might be thought of as singular or plural, and there are languages which mark this distinction - this is what is meant by verbal number. It might also be said that there is a difference between a lone runner going round the track (once or several times) and a several runners - the latter might form a group or they might be racing. Differences resulting from the number of participants in an action may also be encoded in a language as a slightly different type of verbal number.

The following example is from Rapanui (the language of Easter Island: data from Veronica Du Feu):

- (5) ruku rukuruku
 'dive' 'go diving'

Here reduplication is used to indicate verbal plurality since the event is in a sense plural.³

As noted above, verbal number may also mark the number of participants (Durie 1986: 363).⁴ This can be illustrated from Georgian (Aronson 1982: 243, 406-7, quoted in Durie 1986):

- (6) ivane šemovi-d-a da daʃ-d-a
 John enter-PAST-3.SG and sit.SG-PAST-3.SG
 'John entered and sat down'

- (7) čemi mšobl-eb-i šemovi-d-nen da dasx-d-nen
 my parent-PL-NOM enter-PAST-3.PL and sit.PL-PAST-3.PL
 'My parents entered and sat down'

The interesting point is the verb 'sit', which has different forms according to whether one person sits (*daʃ-*), or more than one

³ Durie (1986: 363) refers to the 'number of event'; Arnold Zwicky (p. c.) calls these 'internally complex events'

⁴ Zwicky calls this 'multiple action'.

(*dasx-*) - unlike the verb 'enter'. But it also agrees in number with the subject (as does 'enter'). So are there any grounds for claiming that we have a case of verbal number? Could we not simply say that this is a complex type of agreement? We can isolate the difference if we examine an example where there is a numeral phrase. Numeral phrases in Georgian control singular agreement:

- (8) čemi sami megobar-i šemovi-d-a da
 my three friend.SG-NOM enter-PAST-3.SG and
 dasx-d-a
 sit.PL-PAST-3.SG
 'My three friends entered and sat down'

Numerals require a singular noun (*megobari*) and singular agreement. This singular agreement is found on both verbs. Yet the second, which has two forms according to the number of participants, shows the plural verbal form *dasx-*, since there is more than one participant in the action. This is a special form in that a plural verb occurs with singular agreement. For more examples of verbal number see Dressler (1968), Durie (1986), Mithun (1988) and Newman (1990:53-87).⁵

1.2. Nominal and verbal number: form

We should now ask where nominal and verbal number are expressed. The expected answer would be that nominal number should be expressed on an element of the noun phrase, typically the noun, while verbal number should be expressed on the verb. While both can be expressed in this way, this is not the only possibility. As we noted in looking at examples (1) and (2), nominal number may be expressed on the verb (through agreement). Let us consider all the possibilities for expressing number (Figure 1):

	nominal number		verbal number	
(a) on noun phrase		Lezgian		never (?)
(b) on both		Russian		never (?)
(c) on verb		Amele		always (?)

Figure 1: Place where number is primarily expressed

⁵ See also Cook (1974), Jensen (1952:17-20), and Foley (1986:128-31).

Let us consider the cases in turn. Not surprisingly, we find languages in which nominal number is expressed primarily on the noun phrase. In Lezgian (a North-Central Caucasian or Nakh language), nouns mark number by suffixation, but finite verbs do not (Martin Haspelmath, p. c.). And in French, as a result of attrition, number marking on verbs and on nouns is largely lost in the spoken language; the clearest marker of nominal number is the article (*le* and *la* singular, versus *les* plural). These languages are examples of type (a). In Russian, which is typical of many Indo-European languages in this respect, nominal number is clearly signalled on the noun. It is also marked elsewhere in the noun phrase by agreement. In addition, it is clearly marked on the verb by agreement with the subject noun phrase (type b)

If the noun phrase and the verb both mark nominal number, then the differentiations made are normally the same: thus Russian distinguishes singular and plural on both. But this is not always the case. Modern Hebrew distinguishes three numbers in the nouns phrase (three controller numbers), namely singular, dual and plural, but only two in the verb phrase (David Gil p.c.):⁶

- (9) hayom šavar maher
 day.DEF pass.PAST.3.SG.MASC quickly
 'the day passed quickly'
- (10) hayomayim šavru maher
 day.DEF.DUAL pass.PAST.3.PL quickly
 'the two days passed quickly'
- (11) hayamim šavru maher
 day.DEF.PL pass.PAST.3.PL quickly
 'the days passed quickly'

The same verbal form (called 'plural' by convention) is used with both dual and plural subjects.⁷

⁶ The dual in Modern Hebrew is facultative - it need not be used for every instance in which there are just two referents, see section 2.3; for details of its use see Tobin (1988; 1990: 100-24).

⁷ The Hebrew data suggest that the three types of language separated for the expression of nominal number form part of a more gradual cline. In Lezgian, nominal number is expressed only in the noun phrase, while in French, primarily so. In Modern Hebrew it is expressed both in the noun phrase and in the verb, but the noun phrase has the more important role because it distinguishes three numbers, while the verb distinguishes only two. Russian gives greater weight to the verb since it and the noun phrase both distinguish both numbers found in Russian. In Amele, discussed next, the primary locus of nominal number is the verb, since number must be

The most interesting type (c) is exemplified by the Papuan language Amele (Roberts 1987: 162, 201, 203 and personal communication):

(12) Dana (uqa) ho-i-a
man 3.SG come-3.SG-TOD.P
'The man came.'

(13) Dana (ale) ho-si-a
man 3.DU come-3.DU-TOD.P
'The two men came.'

(14) Dana (age) ho-ig-a
man 3.PL come-3.PL-TOD.P
'The men came.'

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

When we turn to verbal number, we note that in all the examples we have considered, verbal number is expressed on the verb, and I have been unable to find examples of verbal number number being expressed on the noun phrase (a claim made independently by Gil 1991: 8). An example would be a language in which one could say something like: *goats jump* meaning 'one goat jumps several times'.⁸

expressed there, but not necessarily in the noun phrase. Furthermore, the verb distinguishes three numbers, as do the pronouns; however, ordinary nouns may show plurality by reduplication but this is not a normal way in which the dual can be expressed (John Roberts, p.c.). Therefore duality is more readily expressed in the verb than in the noun phrase. It remains to be seen whether there is a language which expresses nominal number not just primarily in the verb but exclusively in the verb.

⁸ D. A. Cruse (p.c.) suggested Turkish, where one can say:

Why should there be such a disparity? I suggest that number marking is found originally 'where it belongs'; that is, nominal number is marked on the noun phrase and verbal number on the verb. As is well known, a common route for the development of verb agreement is from pronouns, which, being nominal, typically mark nominal number. When these become attached to the verb as agreement markers, we find nominal number marked on the verb. Thus the progression for nominal number is from type (a) in Figure 1, through type (b) to type (c). But there is no corresponding route for verbal number to become marked on the noun phrase. It starts as type (c) and has no means to move to any other type.⁹

While verbal number is found in Europe, nominal number is much more widespread. And since the noun phrase is the concern of the ESF group we will now concentrate on nominal number, and look first at the meaning distinctions involved and then at the ways in which they can be expressed.

2. The meaning of number

In this section we concentrate on the meaning of number, looking primarily at the possible meanings of the numbers rather than at the means used to express them. Too often the situation in languages like English is taken as normal, whereas it represents only one of the possible systems. We will first consider whether number needs to be expressed; we shall see that for some languages the expression of number is in a sense optional, while in others it is a central category, which speakers cannot avoid. To investigate these systems we shall first consider the notion of 'general' meaning and then base a typology upon it (section 2.1). We then narrow our attention to the cases where number is

(i) kapi vuruldu
door is.knocked
'there is a knocking at the door'

(ii) kapi-lar vuruldu
door-PL is.knocked
'there is a tremendous knocking at the door'

(The verb is singular in both; since the subject is inanimate it does not take the plural.) The interesting point is the plural noun in (ii): it suggests multiple knocking, not multiple doors. It may be that this example works because real world knowledge suggests there will be one door. How productive this construction is remains unclear.

⁹ Frans Plank's analysis of Chamorro (1989) should be considered here in that a nominal number (dual) is expressed by the combination of nominal and verbal number marking (provided Durie's account of the Chamorro data is accepted).

expressed, and establish the main types of distinction within the category (section 2.2). Thus section 2.1 is devoted to the place of number as opposed to 'non-number', while section 2.2 examines the possibilities within the number domain. We shall look especially at the way in which whole systems operate rather than at the peculiarities of small groups of nouns or other elements involved in the number system.

2.1 General number

In English we are generally forced to choose between singular and plural whenever we use a noun. However, there are languages for which number appears less dominant, languages in which the meaning of the noun can be expressed without reference to number. We shall call such uses 'general', by which we mean that they are outside the number system.¹⁰ Various other terms have been used: Hayward (1979) calls them 'unit reference', the German tradition is to label them 'transnumeral', as in Biermann (1982); we follow Andrzejewski (1960) in using the term 'general'. Most importantly, we are labelling possible meanings and not morphological forms at this stage. We shall first consider number systems as a whole (section 2.1.1), and then look at more restricted examples of general number (section 2.1.2).

2.1.1 The place of general meaning

Given our definition of general meaning, let us analyze its place in the number systems of various languages. In the Cushitic language Bayso, nouns have a form which represents the general meaning, that is, it is non-committal as to number. *Lúban* 'lion' refers to a particular type of animal, but the use of this form does not commit the speaker to a number of lions - there could be one or more than that. Other forms are available for indicating reference specifically to one or to more than one lion, when required. This situation can be diagrammed in the following way:

¹⁰ General meaning forms may be thought of as the controller equivalent of neutral agreement forms (see Corbett 1980). We might have assumed that they would match up, in the sense that general forms would take neutral agreement. But this is not so in Fula. A few nouns in Fula have a suffixless form which shows general meaning (section 1.2 below). This form is not usually used in contexts where agreement is found, but when this does occur it is not the neutral agreement form (the ^{ADUM} gender form) which is used Koval' (1979:12-13).

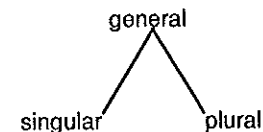


Figure 2:

The point is that the meaning of the noun may be expressed independently of number, as occurs with the general meaning, or it may be expressed within the number system, which at its simplest means there will be a choice of singular or plural. In Bayso these meanings all have independent forms: as we have already noted, *lúban* 'lion(s)' is the general form. For reference to one lion, especially for reference to a specific lion, the singulative *lúbantiti* 'a/the particular lion' is used. For reference to a small number of lions, two to about six, the paucal *lúbanjaa* 'a few lions' is available, while for larger numbers the plural *lúbanjool* 'lions' can be used. The system of Bayso may be diagrammed as follows:

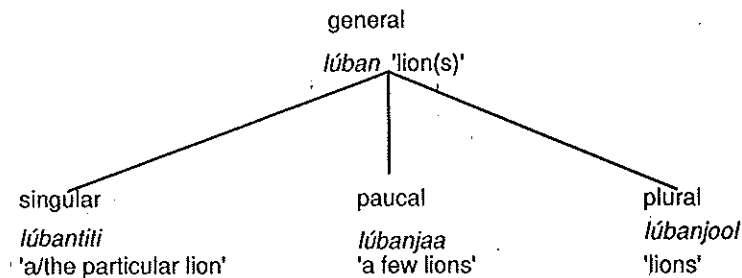


Figure 3: Bayso (Cushitic)

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

- (15) I saw dog (one or more)
- (16) I saw dog-a (singulative: exactly one)
- (17) I saw dog-i (plural: more than one)

The primary opposition is non-number (general) versus number, and within the number category singular is opposed to plural.

I believe that this system is rare in Europe, but according to Bechert (1990: 119-20) it is found in Basque. We find *gizon* 'man or men (absolute case)', as opposed to *gizon-a* 'the man (definite and singular)', *gizon-a-k* 'the men (definite and plural)', *gizon bat* 'a man (literally 'man one': indefinite singular).

The type of language in which the general meaning is regularly expressible by a form outside the number system, as in Bayso, is rare. However, there are several languages in which the general meaning is widely expressed, but by means of a form used also for one of the more restricted number meanings. If we limit ourselves to languages which distinguish singular and plural, this offers two possibilities. The first is that the general is combined with the singular, giving a general/singular versus plural system:

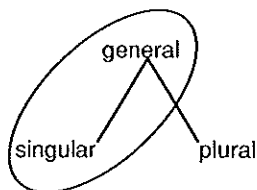


Figure 4

Let us consider how such a system would work. In artificial English the possibilities are these:

- (18) I saw dog (one or more, or exactly one)
- (19) I saw dog-i (plural: more than one)

In this system, examples like (18) are ambiguous between the meanings of examples (15) and (16) above; (18) could be used for seeing one dog or more than one (and so this system differs from that of English); however, (19) unambiguously indicates that more than one dog is intended. An example of this type is provided by Tagalog (data from David Gil). In Tagalog, a form like *aso* can mean 'dog' or 'dogs'. Number can be expressed in Tagalog, primarily by means of the interesting element *mga* [maga], which is probably best analysed as a clitic. It may occur before virtually any constituent, as in:

- (20) *mga aso*
PLURAL dog
'dogs'

Thus the presence of *mga* indicates plurality, but its absence leaves the possibility of singular or general meaning.¹¹

The converse of this system would be one in which the general and plural meanings shared a form:

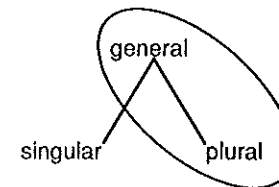


Figure 5

This system allows the following contrasts in artificial English:

- (21) I saw dog-a (singulative: exactly one)
- (22) I saw dog (one or more, more than one)

In this system, examples like (22) are ambiguous between the meanings of examples (15) and (17) above; (22) could be used either for non-committal reference to a dog or dogs, or to indicate more than one.

We must now consider the distribution of these two systems. While we have seen an example of the first (Tagalog, represented in Figure 4), as far as I am aware the second does not exist in the pure form; that is, no language employs the second system as the normal case, forming the basis of an entire number system. This gap may be explained by appealing to markedness considerations. It is generally accepted that the singular is the unmarked number as compared to the plural. When one member of an opposition is neutral with regard to the opposition, then this should be the unmarked form. Thus if one out of singular and plural can also be used as a number-neutral form (that is, having general meaning), then this should be the singular.

While the system of singular versus general/plural (Figure 5) has not been found in pure form, it is nevertheless well established for sections of the noun inventories of particular languages

¹¹ According to David Gil (personal communication), Vietnamese seems to have the same pattern.

(together with the system of general/singular versus plural). According to Dick Hayward (personal communication) Cushitic languages normally have a form which is number-irrelevant. However, we frequently find that this general form is the same as the singular for some nouns, and the same as the plural for others. This situation can be illustrated from the Cushitic language Arbore (Hayward 1984:159-83). Many Arbore nouns have a general form (unit reference in Hayward's terms) which is 'semantically non-specific as to the "singular: plural" distinction' (1984:161). We find pairs like the following:

<i>general</i>	<i>plural</i>
kér 'dog(s)'	kér-o 'dogs'
garlá 'needle(s)'	garlá-n 'needles'

But we also find pairs in which the general form contrasts with a singulative:

<i>singulative</i>	<i>general</i>
tiis-in 'a maize cob'	tiise 'maize cob(s)'
lassa-n 'loaf'	lássa 'bread'
nebel-in 'a cock ostrich'	nebel 'ostrich(es)'

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

<i>singulative</i>	<i>plural</i>
zann-en 'piece of charcoal'	zann-é 'charcoal'
farr-it 'a finger/toe'	farr-ó 'fingers/ toes'

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

The situation found in the Borana dialect of Galla (another Cushitic language) is comparable and equally interesting. When a noun has two forms these are normally singulative and general, or general and plural (singulative plus plural is very rare). Andrzejewski (1960:68) reports that: 'the vast majority of Nouns occur normally only in their *General Forms*. The *Plural* and *Singulative Forms* are seldom used and in fact it is possible to listen to conversations among the Borana for a whole day or even longer without coming across one *Plural* or *Singulative Form*. Nevertheless, there are Nouns whose *Plural* or *Singulative Forms* I

have found in common use.' [Italic in original - GGC.] This quotation makes it quite clear that the difference in forms is not simply a morphological one; in this Galla dialect, nouns have a general form, the one normally used, which gives no information as to number. A singulative or plural form may well be available when specificity as to number is required, but this need occurs infrequently.

Andrzejewski (1960:64n) gives an idea of the relative importance of singulatives as opposed to plurals in Galla: singulative forms are very rare and, with one exception (the word for 'young bull') are found only with nouns denoting persons. Thus the the general form appears to coincide with the singular more often than with the plural. That is to say, nouns following the model of Figure 4 outnumber those following Figure 5. This is what we would expect, if the singular is indeed unmarked with respect to the plural.

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

This leads us to the last possibility in our typology, and English is a good example. This is a system in which there is number must normally be expressed - there is no way of expressing general meaning (except by circumlocution), no forms outside the number system. We have the following picture:

(general)

Figure 6: singular — plural

Let us consider our examples again, this time in normal English:

- (23) I saw a dog (exactly one)
 (24) I saw some dogs (more than one)

There is no form which is ambiguous between the readings of (23) and (24). Indeed, this situation is taken by many people, including large numbers of linguists, to be completely normal and yet, as we have seen, there are many languages which employ rather different systems.

2.1.2 Restricted instances of general number

As mentioned earlier, we have been primarily concerned with the number systems which languages adopt. Larger or smaller groups of nouns may fall outside the dominant system and we now consider how the notion of general meaning may be of help here. Here we need to distinguish two cases. In the first, some nouns fall outside the number system of a particular language. In the second, a group of nouns may express general meaning but may also express specific number meaning.

As an example of nouns falling outside the number system, recall that we said that general meaning cannot be expressed in English, except by circumlocution. Normally this is true. Consider, however, the following example:

- (25) There's ice on the path.

This would be an appropriate utterance if there were one patch of ice or several - it would be normal in both cases. We appear to have an example of general meaning (expressed by the singular) in English. It is normal for languages to have a 'number-differentiability' threshold, below which nouns fall outside the number opposition. In English, 'mass' nouns like *ice*, *snow*, *water* and so on are below the threshold. Different languages put the threshold at very different points. When investigating the dominant number system of a given language we look at those cases where number is differentiated - the 'count' nouns of the particular language. However, it is normal, as in English, for some nouns to fall outside the system, in that they do not express the contrasts available, and so in this sense express general number.

The other possible restricted use of general meaning is found in languages where the majority of nouns are within the dominant number system, but where some nouns are also able to express general meaning. This situation is found in the Futa-Jallon dialect of Fula, which has about a million speakers in Guinea (Koval'

1979); it is not found in other dialects of Fula. Fula has an exceptionally large gender (noun class) system reflected in various agreeing elements; nouns also typically carry a marker which indicates gender and number: *can-ngol/can-di* 'river/s' (-*ngol* shows singular number and membership of a particular gender; -*di* is one of the plural markers - the relations between the singular and plural markers are complex, and need not detain us here). Further examples include *cikka-woll/cikkaa-ji* 'shawl/s', *gabb-ii/gabb-i* 'hippopotamus/es' (Koval' 1979:39). In most instances, when a noun is used the speaker is required to indicate singular or plural number, since one or other marker is required. But just a few nouns may be used without such a marker (Koval' 1979:11; and personal communication):

<i>general</i>	<i>singular</i>	<i>plural</i>
dawa 'dog(s)'	rawaa-ndu 'dog'	dawaa-di 'dogs' ¹²
toti 'toad(s)'	totii-ru 'toad'	totii-ji 'toads'
nyaari 'cat(s)'	nyarii-ru 'cat'	nyarii-ji 'cats'
gerto 'hen(s)'	gerto-gal 'hen'	gerto-de 'hens'
bofo 'egg(s)'	wofoo-nde 'egg'	bofoo-de 'eggs'
biini 'bottle(s)'	biinii-ri 'bottle'	biinii-ji 'bottles'

The forms which have no suffix express general meaning, that is they are used when number is irrelevant, as in the following saying (Koval' 1979:11):¹³

- (26) ko biini tun waawi marde beere
 PARTICLE bottle only can-PERF preserve beer
 'only a bottle/bottles can preserve beer'

This situation is rather different from that of English. In both languages nouns usually indicate singular or plural number. But a group of Fula nouns has the additional option available of showing general meaning: that is, they can show number, but can also in appropriate circumstances appear in a number-indifferent form. The English nouns considered were those outside the number system in the sense that they cannot show two opposing numbers.

¹² Some nouns preserve a system of initial consonant alternation (Koval' 1979:8): this noun shows the *d/r* alternation, and *bofo* 'egg' below illustrates *b/v*. The symbol *d̥* indicates a preglottalized *d*.

¹³ PERF is for perfective; note that in this dialect the verb gives no information on number.

2.2. Possible number systems

Here we examine the semantic distinctions which can be drawn within the number system:

SINGULAR - PLURAL

Numerous examples in Europe, such as English and French.

SINGULAR - DUAL - PLURAL

Several examples in Europe, like Upper Sorbian.

SINGULAR - PAUCAL - PLURAL

Bayso (plus general): no examples in Europe (?)

SINGULAR - DUAL - PAUCAL - PLURAL

Ungarinjin (Rumsey 1982)

Yimas (Foley 1986): no examples in Europe (?)

SINGULAR - DUAL - TRIAL - PLURAL

Kiwai (Ray 1933): no examples in Europe (?)

SINGULAR - DUAL - TRIAL - QUADRAL - PLURAL

Sursurunga (Hutchisson 1986), an Austronesian language with 2700 speakers in southern New Ireland (PNG):

- (24) gim-at k'aw'an
 I.EXCL-QUADRAL maternal.uncle:nephew/niece
 'we four who are in an uncle-nephew/niece
 relationship'

Note, however, that plural pronouns are never used with terms for relationships and so the quadral is then used instead for a minimum of four, and not just for exactly four (Hutchisson 1986:10).

SINGULAR - COLLECTIVE PLURAL - DISTRIBUTIVE PLURAL

As Lyons (1977:178) explains: 'Sometimes we refer to a class of individuals distributively* in order to ascribe a certain property to each of its members; on other occasions we do so collectively* in order to ascribe a property to, or assert something of, the class as a whole ...'. This distinction may serve as the basic of the number system. Indeed it is central to the number system of Papago, an Uto-Aztecan language of Arizona (Mathiot 1962; 1967; Cruse forthcoming). Nouns have up to three different forms, a

singular *daikuḍ* 'one chair', a collective plural¹⁴ *dadaikuḍ* 'several chairs (from a single household)' and a distributive plural *daddaikuḍ* 'several chairs (from several households - one or several chairs per household)'. Thus the collective and the distributive are both formed by partial reduplication. Mathiot (1977:209) suggests that there are two related oppositions here: there is the familiar opposition of single entity versus several entities, and the less familiar opposition of single locus versus several loci. (Note that the culturally determined locus is important: if several chairs were in the same room and belonged there, then the collective would be used; but if they were from different households, and so really belonged in different places, then the distributive would be used.)

Given an opposition between single entity and several entities, and a second opposition between single locus and several loci, we might expect to find four possibilities, as shown in Figure 7.





	single entity	several entities
single locus		
a	b	
several loci		
	c	d

Figure 7: Interaction of entities and loci

One entity in a single locus (cell a) represents the singular, and several entities in a single locus represents the collective (cell b). Cell c is a problem: having a single entity in several loci implies in

¹⁴ Mathiot calls them 'singular', 'plural' and 'distributive'.

fact that there are several entities (at least in the case of normal concrete objects). There is no noun in the language which distinguishes the situations represented in cells c and d - if different loci are involved, then the distributive is the required form. We noted earlier that the distributive plural *daddaikud* could indicate 'one chair from each of several households' (cell c) or 'several chairs from different households' (cell d).

It is interesting to note that some nouns do not have the full three-way distinction. Some distinguish only singular (cell a) from non-singular, for instance *ban* 'one coyote (singular)' versus *baaban* 'several coyotes'. Others distinguish single locus from several loci, in other words singular and collective (cells a and b) from distributive: *hivañ* 'one cow, or several cows from a single herd' versus *hahaivañ* 'several cows from several herds'. The full, three-way distinction was illustrated earlier with *daikud* 'chair'.

The distinction between collective and distributive is marked sharply in the American Indian language Sahaptin (which has under 500 speakers in Oregon). The data are from Bruce Rigsby (forthcoming; and personal communication). For nouns which denote humans, there is a three-way distinction: singular, dual and plural (though verb agreement is identical for dual and plural, thus it opposes singular and non-singular). An example of the oppositions marked on a (human) noun would be: *tlaaki* 'woman', *tlaaki-in* 'two women', *tlaaki-ma* 'women'. For other nouns (not denoting humans), number is restricted to certain cases (ergative, accusative and objective). In those cases where number is marked there are just two forms: one is identical to the inflectional root, the second may be marked by suffixation or by full or partial reduplication (*pšwápšwá* 'stones, rocks' from *pšwá* 'stone, rock'). Forms like *pšwápšwá* are distributive plurals, appropriate for distribution in place or time. In the case of rocks, the obvious use would be in a situation in which rocks are scattered. If several rocks are involved but they are not distributed (say they are in a single heap), then the form *pšwá* is used. Similarly *k'úsi* 'horse' could be used of a single horse or a group of horses (collective), but horses not in a group would be *k'úsi-ma* 'horses (distributive plural). Thus Sahaptin distinguishes distributive plural from other (for nouns denoting non-humans); the other form (the 'singular') is for multiple entities viewed as a collective as well as for single entities. (For nouns denoting humans, the plural is used for more than two irrespective of the collective versus distributive distinction -

alternatively one may say that humans are always treated as being distributed.)

2.3. Other distinctions

Number systems may differ not just in terms of the numbers of items denoted but along rather different dimensions:

1. the status of different numbers

Numbers may not all be of the same status. The use of a particular number may be facultative, just as number as a whole may be facultative (as in Bayso). Thus in Ngan'gityemerri (a Daly River language with two dialects, Ngan'gikurunggurr and Ngan'giwumirri, and with 100 speakers, 300 miles SW of Darwin) there is singular, dual, trial and plural. The dual must be used for two, the plural must be used for four and more. For three the trial is used when the fact of there being three is salient (e.g. introduction into discourse) but otherwise the plural is used for three. Note that the trial is strictly for three, and is not a paucal (Reid 1990:118-19 and personal communication). The trial is similarly optional in Anindilyakwa (Velma Leeding, personal communication).¹⁵

It follows, of course, that 'plural' means different things in different systems.

2. Determinate versus indeterminate

I suggest these terms to differentiate situations where, given the knowledge of real world which the speaker has, we can determine that only one form is appropriate (determinate number) from those where we cannot (indeterminate). Thus in a language with an obligatory dual, this would be an instance of determinate number (given two instances of the referent only the dual is appropriate). Indeterminate number covers:

- facultative number (e.g. the dual/plural choice in languages like Modern Hebrew with an optional dual)
- paucal (paucal versus plural choice)
- collective versus distributive choice

A similar opposition is found in quantifiers: numerals like *five* are determinate while quantifiers like *several* are not.

3. Different morphological patterns

¹⁵ It is similar in that the plural can be used in its place. It differs in that it used to be used for three or for four; younger speakers now tend to use it for three only.

These are instances where the morphology of number is organized in a less familiar way but one which does not lead to additional semantic distinctions. Two of these have been identified:

- augmented number, as in Rembarnga and Ndjébbana (McKay 1978; 1979)
- inverse number, as in Kiowa (Watkins 1984).

2.4 The meaning of number: summary

There are two main points arising from this section

1. the importance of general meaning for a full typology of number systems
2. the view (generally accepted) that the singular is the unmarked number (and as such is likely to have the general meaning function where there is no special form)

3. The expression of number

We now move on to the expression of number. The possibilities are considerable:

1. separate lexical item (number words), as in, for instance, the Austronesian language Yapese (see Dryer 1989).
2. syntactic means:
 - a) free grammatical items (clitics)
 - b) agreement (as mentioned above)
3. morphological means - the main method,
4. alternative lexical realizations (suppletion)

We will concentrate on morphology, the major means for expressing number. Number morphology varies from relative simplicity in some languages to great complexity in others. Even in languages where the main patterns are straightforward, we often find isolated examples of more interesting types. In this section we shall first set up a general model of the morphological means by which number can be signalled.¹⁶ This will then allow us to characterize the range of variation found and to uncover the regularities within the diversity.

3.1. The inflectional root

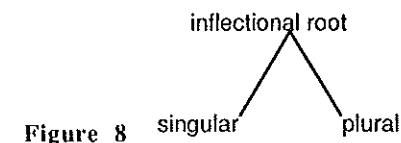
We start from the notion of 'inflectional root' (Matthews 1974:73-5). The inflectional root of a lexical item is the basic form, the one which cannot be further reduced as far as inflectional categories are concerned. It is indifferent to inflectional category, in the

¹⁶ Our typology draws insights from two rather different sources, namely Kibrik's work on the morphology of the noun in Dagestani languages, (Kibrik 1991) and Gazdar's work on DATR (Evans and Gazdar 1990, see especially the work by Gazdar and Cahill on German).

sense that all the inflectional forms normally available to lexical items of the particular word-class can be formed from it. The inflectional root *play* is the base for the forms *play-s* and *play-ed*. The inflectional root may, however, not be basic derivationally; *re-play* consists of prefix + root in terms of derivational morphology, but for inflectional purposes the root is *replay*.

3.2. The relation of number to the inflectional root

Consider a language which has at least two numbers, singular and plural. What are the possible relations between the number forms and the (inflectional) root for a given lexical item (or group of lexical items)? Let us start from a maximally general model:



To keep the initial examples simple, we shall consider a series of imaginary languages, which fortunately share two nouns with the following inflectional roots: *dog* 'dog' and *cat* 'cat'. Let us ask in what ways the singular and plural forms can differ from the inflectional root. First they may differ in inflection. In our first language there is a singular ending *s* and a plural ending *p*, hence we find these forms:

singular	plural
dog-a-s	dog-a-p
kata-s	kata-p

Here we have a familiar form of inflection, namely endings; inflection can, of course, appear in various other guises. The other way in which the number forms can vary from the basic stem is through stem formation. In our second language, number is marked by partial reduplication of the stem, initial for the singular and final for the plural,¹⁷ which gives the following forms:

singular	plural
----------	--------

¹⁷ Reduplication is in fact common for plural stem formation but not for singular stem formation.

do-doga	doga-ga
ka-kata	kata-ta

Here it is quite clear that we have a change of stem: it is the form of the inflectional root which determines the number forms, and not an independent inflection which is added. As we shall see later, there are cases which are not as clear-cut.

These two devices, inflection and stem formation may occur together. Our third language marks number by stem formation and inflection, as in the following forms:

singular	plural
do-doga-s	doga-ga-p
ka-kata-s	kata-ta-p

Here either the stem or the inflection carries sufficient information to indicate number.

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.¹⁸

Having established that inflection and stem formation may distinguish the singular and plural forms from the inflectional root, we see that our original model was at a high level of abstraction. To include the different means of differentiating number we can elaborate it as follows:

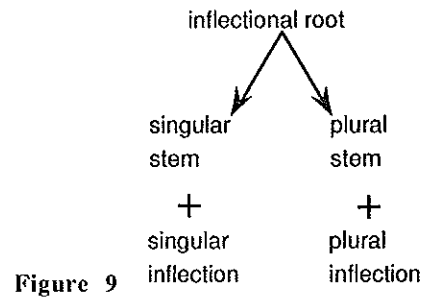


Figure 9

¹⁸ At first sight it appears that suppletion belongs here. However, as we shall see below, suppletion is generally found where two stems are not regularly related to each other.

We will now look at examples of how number is marked, considering in particular whether all the elements identified in the diagram are distinct or not in particular examples. It is important to note 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 norm for that language.

If we start with the relations between the (inflectional) root and the stems, the first logical possibility is that all are distinct. This possibility can be illustrated by the irregular Russian noun, *xozjain* 'landlord'. The inflectional root is *xozja(j)*¹⁹, the singular stem is *xozja-in-* and the plural stem is *xozja-ev-*. Both stems allow the normal addition of endings.²⁰ 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; their relation is purely semantic. An example is Russian *čelovek* 'person', plural *ljud-i* 'people'.²¹ Note that we are indeed dealing with stems here: *čelovek* 'person', takes normal singular inflections, and *ljud-i* 'people', takes plural inflections. A way of thinking of such cases is to say that the lexeme involved has alternative inflectional roots (Matthews 1974:74).

It is unusual for the root and the singular and plural stems all to be distinct, in Russian and more generally. Often we find that the root and the singular stem are identical, as in this diagram:

¹⁹ Since Russian orthography is largely morphophonemic, we use a transliterated form for simplicity here. The details of the use of *j* need not detain us.

²⁰ It is possible for there to be stem alternations which do not correlate with number; for example, in Sanskrit the noun for 'king' has two stems, *rājan-* and *rājā-*, both of which appear, for different cases, in the singular, dual and plural (Carstairs 1987:211).

²¹ The Russian suppletion is matched by the natural English translations *person* and *people*, though of course *persons* does exist in English.

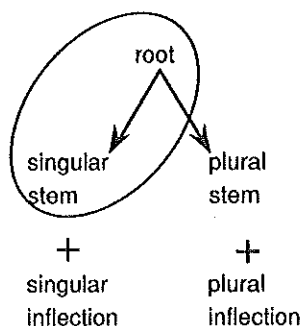


Figure 10

This pattern too can be illustrated from Russian. The noun *krylo* 'wing' has the inflectional root *kryl-*, to which the singular endings are added directly (*kryl-o*, *kryl-a*, *kryl-u* and so on). The plural stem is *kryl'j-* (the ' marks palatalization of the preceding consonant), as in the nominative plural *kryl'ja*. 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 *a* is a regular nominative plural ending, found on hundreds of nouns which do not have a separate plural stem. The plural endings for the remaining five cases of Russian are also found on other nouns; we would be missing an obvious generalization if we claimed there were special endings right through the plural paradigm while in fact 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 inflectional root:

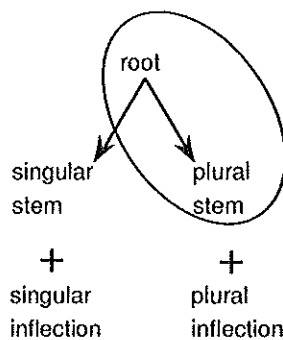


Figure 11

Again the pattern is found in Russian. The noun *bolgarin* 'a Bulgarian' has the inflectional root *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.

An interesting regularity which should be noted here is that if a language expresses general number (as discussed under meanings above) then it will do so (either solely, or as the normal case) through the maximally unmarked form, that is, the inflectional root. This is true whether general number is the sole number expressed by the inflectional root (as in Bayso), or whether the inflectional root also commonly functions as the singular stem (as in Galla).

Our next type of identity of form would be that illustrated in the next diagram:

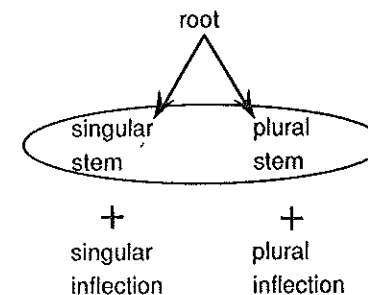


Figure 12

Given just the elements in the diagram, it is not at first obvious what the situation depicted would be like. If the singular and plural stems are identical, how do they differ from the root? To make sense of this situation, we need to consider possible additional elements. For example, if there were at least one additional number (say dual or paucal), then two or more number stems could be identical, while differing from a third.

This situation is found in the personal pronouns of Kiwai, a Papuan language (Kiwaiian family) spoken near the mouth of the Fly River (data from Foley 1986:72, following Ray 1933). The forms are these:

	singular	dual	trial	plural
1st person	mo	nimo-to	nimo-ibi	nimo
2nd person	ro	nigo-to	nigo-ibi	nigo
3rd person	nou	nei-to	nei-ibi	nei

Here we see that the singular and plural forms differ, though they are clearly related. The relation of the dual and trial forms to the plural is transparent: the plural stem is the base on which the other two numbers are constructed. This situation is analogous to that in the diagram, except that the identity is of the dual, trial and plural stems as opposed to the singular.

The last logically possible relation of root to stems is that all are identical, diagrammed as follows:

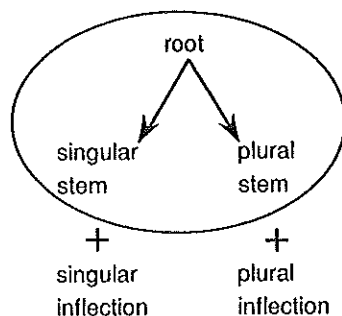


Figure 13

This situation is extremely common. Again in Russian we find many nouns like that for 'newspaper', which has the basic stem *gazet-*. The (nominative) singular is *gazet-a* and the (nominative) plural is *gazet-y*. Here stem formation has no role, and the entire burden of signalling a difference in number is carried by the inflections (endings in this instance).

We move on to look for identities elsewhere in the model. It may be that stem plus ending for a particular number forms is identical to the stem. Thus if we look at a Russian noun from a different paradigm form that to which *gazet-a* 'newspaper' belongs, we may find the following pattern: with the noun *student* 'student': the inflectional root and the singular and plural stems are all *student-*; the nominative singular is *student* as opposed to the plural *studenty*. The nominative singular for these nouns happens to have no phonetic realization (unlike nouns of other paradigms), but the absence of an ending contrasts with the remaining forms in the paradigm and signal number and case in this way. We are dealing with a null morph, which we indicate \emptyset ; the forms are *student \emptyset* (singular) and *studenty*. (plural).

There is a further, initially rather surprising type of identity, shown in the next diagram:

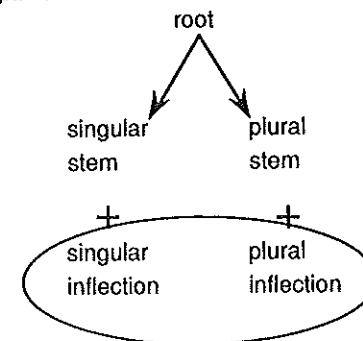


Figure 14

This pattern suggests that the inflections used for singular and plural could be identical. This situation regularly occurs in Dagestani languages for the majority of the large numbers of cases they distinguish (often just the nominative is an exception). The Akhvakh noun *nido* 'forehead' shows a clear case of identical endings: we take nominative (what some would call 'absolutive') and ergative endings to illustrate the point (Kibrik 1991):

	singular	plural
nominative	<i>nido-\emptyset</i>	<i>nido-di-\emptyset</i>
ergative	<i>nido-la-de</i>	<i>nido-di-le-de</i>

In this example the inflectional root is *nido*, and the singular stem is identical to it. The plural stem is *nido-di-*. The nominative case, in singular and plural, has the zero ending (\emptyset). 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 nominative, 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 ending have no role in the number system - their function is to make clear the case of the noun (the case system is extensive). This identity of form of endings in the singular and plural is quite general in Dagestani 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.

There is one final pattern of identity we should consider, that in which both stems are identical to the inflectional root, and where the stems are identical to the forms with endings (that is, there are no endings). This means that the noun is indeclinable - 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 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, *taksi* 'taxi' may denote one or more taxis (the ambiguity will often be removed by elements showing agreement in number).

3.3. Typological insights and the role of markedness

We saw in the last section how number may be marked by stem formation and by inflection. However, the full range of possible differentiation is not normally found. There are instances where, for example, one of the stems is identical to the inflectional root (equivalently, it is the inflectional root plus zero); we also saw instances where number inflections were not maximally different (and again there were examples of zero inflection). These possibilities are not randomly distributed, but conform to certain patterns.

Greenberg's (1963) paper includes several universals relating to number. Some of these connect to the notion of markedness, as worked out in this context particularly by Greenberg (1966:25-55), following ideas of Jakobson.²² Universal 35 (Greenberg 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 be expressed by zero (and the plural has additional material to form a plural stem or a plural inflection, or both), the plural will not be expressed regularly through a given language as zero. (Recall the mention earlier that certain types of number marking might be available for particular nouns

but that did not indicate that the pattern formed the basis for number-marking in the language as a whole. More generally, it is common for zero to be the regular marker of the singular, or at least a common marker for the singular; it is less usual for it to mark the plural, even for a smaller group of lexical items. We noted the example of Russian *bolgarin* 'a Bulgarian' and similar nouns; here the inflectional root 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 the null form for these nouns (*bolgar-Ø-y*). Conversely, in the nominative singular, the inflection is zero (*bolgar-in-Ø*). However, there are many nouns in Russian which mark the nominative singular by zero, while there are relatively few which have the *-in-Ø* singular/plural stem alternation shown by *bolgarin*. The English situation is odd in the extreme; while nouns behave normally, the verb morphology is most unusual in that the singular, third singular (present) has a marker, while the plural does not: compare *she writes* with *they write-Ø*. But though English is unusual in this respect it does not form a counter-example to Greenberg's universal since the plural has nonzero allomorphs (on nouns: *writer-Ø* but *writer-s*).

The use of zero for marking number on the verb was investigated by Bybee (1985:52-4), using a sample of 50 languages based on Perkins' sample. Of the 27 languages in the sample in which the verb marks number, 78% use zero to mark the singular; zero occurs as one of the means of marking the plural in only two languages (and never, in her sample, as the only means). There were no examples of the dual or trial being marked by zero.

This suggests that the singular is some way more basic than the plural, and indeed there are good reasons for believing this to be the case; and this is why it is considered to be the unmarked member of the singular-plural opposition. Not only is it frequently signalled by a zero form, it is used more frequently in texts, as we shall see. The relatively markedness of the different numbers can be inferred from Greenberg's universal 34, which concerns the presence or absence of the numbers in different languages:

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

From this universal about the existence of the different numbers we may infer their relative markedness:

singular < plural < dual < trial

²² See Zwicky (1978) for a helpful discussion of markedness.

The singular is the least marked, the others are all more marked. This claim receives support from the relative frequency of occurrence in texts. Greenberg (1966:31-2) gives the data based on counts of running text in four languages. In each case, only nouns were considered²³ (how many is indicated by 'size of sample') and the relative frequencies the use of singular and plural were worked out. Of the languages for which data were available, only Sanskrit has the dual number. Fortunately, these data can be supplemented by information on modern Slovene (literary prose Neweklowsky 1988), since Slovene has a dual:

language	size of sample	singular	plural	dual
Sanskrit	93,277	70.3%	25.1%	4.6%
Latin	8,342	85.2%	14.8%	
Russian	8,194	77.7%	22.3%	
French	1,000	74.3%	25.7%	
Slovene	2,182	75.3%	24.2%	0.5%

These data clearly support the scale of relative markedness proposed.

This scale of relative markedness also relates to Universal 35, quoted earlier; it allows us to make claims about the distribution of zero markers between other numbers in addition to the singular and plural. We would expect the dual and trial to be formed using additional material from one of the less marked numbers, rather than vice versa. And indeed, we saw in Kiwai how the plural stem was the base on which the dual and trial are constructed: it has a zero marker by comparison to the dual and trial.²⁴

It is important to bear in mind the distinction between the dominant system of a particular language and the behaviour of individual lexical items or groups of lexical items. Thus while the singular is the least marked number, Tiersma (1982) points to what he calls 'local markedness'. The idea is that there are groups of nouns for which the plural is the unmarked number. The first

²³ Data on number in verb forms (Greenberg 1966:37) also support the claim that the singular is the unmarked number and that the plural is less marked than the dual. This is further supported by the data on Slovene verb forms (Neweklowsky 1988:343).

²⁴ For further evidence for this scale of markedness see Foley (1986:132-3); he shows how in various Papuan languages the most marked numbers, paucal, trial and dual are marked by suffixes which signal the particular number as their only function, while the less marked numbers are indicated by portmanteau morphs which signal person as well as number.

type of evidence which Tiersma cites concerns dialects of Frisian spoken in the Netherlands. Some nouns show an alternation of diphthong between the singular and plural stems, for example:

alternation	singular	plural	gloss
/iə ~ jɪ/	stien	stjinnen	stone
/iə ~ jɛ/	beam	bjemmen	tree
/uə ~ wo/	stoel	stvollen	chair
/oə ~ wa/	doar	dwarren	door

Note that the plurals have the plural inflection *-en* in addition to the modified stem. The stem alternations are being lost. Since the singular is the unmarked number, we would expect the singular stem to be preserved, and for its use to be generalized to include the plural too. This is in indeed what happens in most instances of change. Consider the following alternatives, the old form maintained by conservative speakers, and the newer form as used by innovative speakers:

CONSERVATIVE		INNOVATIVE		gloss
singular	plural	singular	plural	
koal	kwallen	koal	koalen	coal
miel	mjillen	miel	mielen	meal
poel	pwollen	poel	poelen	pool

For the innovative speakers the singular stem has become the stem for singular and plural, and plurality is marked for these nouns simply by the productive plural inflection *-en*. There are a few items, however, for which the change has operated in the opposite direction (for some innovative speakers in certain dialect areas):

CONSERVATIVE		INNOVATIVE		gloss
singular	plural	singular	plural	
earm	jermen	jerm	jermen	arm
goes	gwozen	gwos	gwozen	goose
hoarn	hwarren	hwarne	hwarren	(animal) horn
hoas	vjazzen	vjazze	vjazzen	stocking
kies	kjizzen	kjizze	kjizzen	tooth
spoen	spwonnen	spwonne	spwonnen	shaving, splinter
toarn	twarnen	twarne	twarnen	thorn
trien	trjinnen	trjin	trjinnen	tear

The question is why precisely these nouns should be involved. Tiersma (1982:835) proposes the following principle:

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.

If indeed the noun meaning 'arm' is locally unmarked in the plural, then it is understandable that its plural stem resists levelling, and the singular stem is lost in favour of the plural. To support the claim about this and the other nouns which have generalized the plural stem, Tiersma gives statistical data, from a variety of sources, on the relative frequency of singular and plural with these nouns, as follows (instances of the corresponding noun being used in the singular are given before the colon, and the plural after the colon).

	Dutch	English	German	Spanish
'arm'	43 : 50	94 : 121	-	44 : 71
'goose'	1 : 1	4 : 3	20 : 18	12 : 15
'horn'	-	-	58 : 74	2 : 9
'stocking'	5 : 11	1 : 5	10 : 52	-
'tooth'	6 : 19	20 : 103	67 : 274	7 : 13
'shaving'	-	-	10 : 21	-
'thorn'	1 : 1	-	36 : 69	6 : 0
'tear'	1 : 27	11 : 34	36 : 142	4 : 23

The evidence is indirect, since it relates to languages other than Frisian. It is, nevertheless, convincing. The data given earlier from Greenberg show that in a small sample of Indo-European languages the singular occurs considerably more frequently than the plural, roughly in three instances out of four. Yet for the nouns in question here, where data on comparable lexical items are available, the plural is generally found more commonly than the singular. This strongly suggests that for these nouns the plural is the unmarked form.

Tiersma quotes further examples of levelling in favour of the plural stem.²⁵ But he also produces evidence of other types to confirm that the plural can be locally unmarked for nouns like those discussed. One type of evidence is closely related to the Frisian evidence, and concerns child language acquisition. If the

²⁵ Following Greenberg (1969:192-3) he points out that of a group of irregular neuter nouns, the only one to retain an extended stem in Russian is *koleso* 'wheel', which is much more frequently used in the plural than in the singular.

child hears the plural form much more frequently than the singular for particular nouns, then the plural form may be taken as indicating the inflectional root and then the singular may be reformed from that (Tiersma gives an example from Modern Hebrew). This acquisition strategy is, of course, what leads to the types of levelling we have already considered.

Then there are examples of double plurals, that is, instances of a plural marker being added to a form which is already plural. For example, the Middle English plural form *childre* gained a second marker of plurality, resulting in the modern form *children*. Tiersma gives several examples from English and Dutch, and again these are nouns for which the plural is locally unmarked.

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 give 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'.

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 considered the 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, feet, oxen, geese, lice* and *teeth*.²⁶

As mentioned earlier, if one member of an opposition takes a null marker, we expect this to be the unmarked one; hence, for English regular nouns, the singular is marked by zero. There are some English plurals which also have a null marker, like *sheep* and *fish*; these too are examples of locally unmarked plurals.

²⁶ 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 & Hayward (1987) involved several of the same items: the nouns for 'feet', 'oxen', 'teeth' and so on.

An extension to the notion of local markedness which Tiersma did not consider is that when larger numbers of oppositions are involved it may be a relative notion. We saw above that the plural is normally more marked than the singular and the dual is more marked than the plural. For certain nouns, those which typically occur in pairs, the dual may be locally unmarked as compared to the plural. Most of the Slavonic languages have lost the dual number, but they retain traces of the old dual in irregular plurals. For example, Russian *uxo* 'ear' has the plural *ušī* in place of the regular **uxa*, and *plečo* 'shoulder' has *pleči* instead of **pleča*. The irregular plurals go back to dual forms: it is more common and natural to talk of two ears or shoulders rather than, say, seven or twenty-three. The dual was locally unmarked as compared to the plural for these nouns and it is the dual forms which survived as irregularities. Their continuing survival can also be explained by appealing to local unmarkedness; with the loss of the dual, the plural *ušī* can denote two or more ears (usually just two), which is more common and natural than the use of the singular, and so the irregular plural survives.

We have seen that the relevant universals in Greenberg (1963) do indeed constrain the patterns of identities in the morphology of number, and that the notion of markedness, including local markedness, if of help. There is one further universal (number 39) which is relevant to the material examined so far:

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.

We met a confirming example in Akhvakh. Recall the form noun *nido-di-le-de* 'forehead' (ergative plural). The inflectional root is *nido*, the formative *-di-* forms the plural stem, *-le-* marks plurality again, together with obliqueness, and *-de-* marks the ergative case; thus the elements come in the order:

root - plural marker - plural + oblique marker - case marker

Such forms, of which many examples can be given from Dagestanian languages, support Greenberg's claim particularly well, since the marker which indicates number and case (oblique rather than direct) stands between the straightforward number marker and the straightforward case marker.

We have seen that there is considerable variety in the morphological expression of number, involving stems and their relation to the inflectional root, and inflections (including zero).

3.4 Expression of number: conclusion

1. It appears that when a language expresses general number meaning, its means of expression is likely to be equivalent to the inflectional root.
2. The expression of the singular (the unmarked number) will be simpler than that of the other numbers (by the use of \emptyset stem formants and/or \emptyset inflections). This claim holds true when the majority of nouns are considered, but it is quite possible for smaller groups of nouns to behave differently.

4. The EUROTYP perspective

Nominal and verbal number (section 1) are both present in the EUROTYP area; nominal number is widespread, and verbal number is found, in the Caucasus at least.

In section 2 we noted the wide range of possible semantic systems of number. Of these, a relative small subset is found in Europe. Many Indo-European languages have an opposition between singular and plural, which is an obligatory choice (that is, there is no general number). For this reason, this particular system has often been taken as 'normal' in some sense, whereas it is simply one possibility out of several. There are also in Europe examples of the dual, including a facultative dual, as in Slovene. The existence of general number is reported for Basque.

As far as the expression of number is concerned (section 3), Europe gives examples of almost all the possible systems. The probably exception is the use of number words, which appear not to occur in Europe.

References

- Andrzejewski, B. W. 1960. The categories of number in noun forms in the Borana dialect of Galla. *Africa* 30.62-75.
- Aronson, Howard I. 1982. *Georgian: a Reading Grammar*. Columbus, Ohio: Slavica.
- Bechert, Johannes 1990. The structure of the noun in European languages. In Johannes Bechert, Giuliano Bernini and Claude Buridant (eds) *Toward a Typology of European Languages*, 115-40. Berlin: Mouton de Gruyter.
- Biermann, Anna. 1982. Die grammatische Kategorie Numerus. In: Hansjakob Seiler & Christian Lehmann (eds) *Apprehension: Das*

- sprachliche Erfassen von Gegenständen: I: Bereich und Ordnung der Phänomene, 229-43. Tübingen: Narr.
- Bybee, Joan. 1985. *Morphology: a Study of the Relation between Meaning and Form* (Typological Studies in Language 9). Amsterdam: Benjamins.
- Carstairs, Andrew. 1987. *Allomorphy in Inflection*. London: Croom Helm.
- Cook, Eung-Do 1974. Internal evidence for the evolution of number categories in Sarcee. In John M. Anderson & Charles Jones (eds) *Historical Linguistics I: Syntax, morphology, internal and comparative reconstruction. Proceedings of the First International Conference on Historical Linguistics, Edinburgh 2nd - 7th September 1973*, 101-15. Amsterdam: North-Holland.
- Corbett, G. G. 1980. Neutral agreement. *Quinquereme - New Studies in Modern Languages* 3.164-70.
- Corbett, G. G. and Hayward, R. J. 1987. Gender and number in Bayso. *Lingua* 73.1-28.
- Cruse, D. Alan forthcoming. Number and number systems. To appear in *Encyclopedia of Language and Linguistics* edited by R. E. Asher, Pergamon and Aberdeen University Press.
- Dressler, Wolfgang V. 1968. *Studien sur verbalen Pluralität: Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen*. (Österreichische Akademie der Wissenschaften: Philosophisch-Historische Klasse: Sitzungsberichte, 259 Band, 1. Abhandlung). Vienna: Hermann Böhlau Nachf.
- Dryer, Matthew 1989. *Linguistics* 27.865-95.
- Durie, Mark 1986. The grammaticization of number as a verbal category. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 12.355-70.
- Evans, R. & Gazdar, G. eds. 1990. *The DATR Papers*. University of Sussex Cognitive Science Research Paper CSRP 139.
- Foley, William A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- Gil, David. 1991. *Nouns, verbs and quantification* (EUROTYP Working Papers, Series X, No. 1).
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed) *Universals of Language*, 73-113. Cambridge, Massachusetts: MIT Press.
- Greenberg, Joseph H. 1966. *Language Universals: with special reference to feature hierarchies*. The Hague: Mouton.
- Hayward, R. J. 1979. Bayso revisited: some preliminary linguistic observations - II. *Bulletin of the School of Oriental and African Studies, University of London* 42.101-32.

- Hayward, Dick [= R. J. Hayward] 1984. *The Arbore Language: A First Investigation: Including a Vocabulary*. (Cushitic Language Studies 2). Hamburg: Buske.
- Hutchisson, Don. 1986. Sursurunga pronouns and the special uses of quadral number. In Ursula Wiesemann (ed.) *Pronominal Systems* (Continuum 5), 217-55. Tübingen: Narr.
- Jensen, H. 1952. Die sprachliche Kategorie des Numerus. *Wissenschaftliche Zeitschrift der Universität Rostock, Reihe Gesellschafts- und Sprachwissenschaften* 1, 2.1-21.
- Kibrik, Aleksandr E. 1991. Organizing principles for nominal paradigms in Daghestanian languages: comparative and typological observations. In: Frans Plank (ed.) *Paradigms: The Economy of Inflection*, 255-XX. Berlin: Mouton de Gruyter.
- Koval', A. I. 1979. O značenii morfoložičeskogo pokazatelja klasa v fula. In N. V. Oxotina (ed.) *Morfonologija i morfoložija klassov slov v jazykax Afriki*, 5-100. Moscow: Nauka.
- Lyons, J. 1977. *Semantics: I*. Cambridge: Cambridge University Press.
- McKay, Graham R. 1978. Pronominal person and number categories in Rembarrnga and Djebbbana. *Oceanic Linguistics* 17.27-37.
- McKay, Graham R. 1979. Gender and the category *unit augmented*. *Oceanic Linguistics* 18.203-10.
- Matthews, P. H. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Mithun, Marianne 1988. Lexical categories and the evolution of number marking. In Michael Hammond & Michael Noonan (eds) *Theoretical Morphology: Approaches in Modern Linguistics*, pp. 211-34. San Diego: Academic Press.
- Moravcsik, Edith A. 1978. Agreement. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds) *Universals of Human Language: IV: Syntax*, 331-74. Stanford: Stanford University Press.
- Neweklowsky, Gerhard 1988. Zur Häufigkeit morphologischer Kategorien in slowenischen Prosatexten. In Boris Paternu & Franc Jakopin (eds) *Sodobni slovenski jezik, književnost in kultura: mednarodni simpozij v Ljubljani od 1. do 3. julija 1986* (=Obdobja 8), pp. 337-49. Ljubljana: Univerza Edvarda Kardelja v Ljubljani, Filozofska fakulteta.
- Newman, Paul. 1990. *Nominal and Verbal Plurality in Chadic* (Publications in Agrican Languages and Linguistics 12). Dordrecht: Foris
- Plank, Frans. ms 1989. Dual in Chamorro, mostly indirect. Unpublished paper

- Ray, S. 1933. *A Grammar of the Kiwai Language, Fly Delta, Papua. Port Moresby: Government Printer.* [Quoted from Foley 1986.]
- Reid, Nicholas J. 1990. Ngan'gityemerri: A Language of the Daly River region, Northern Territory of Australia. Unpublished PhD thesis, Australian National University, Canberra.
- Rigsby, Bruce forthcoming. Sahaptin Grammar. To appear in *Handbook of North American Indians: Volume 17: Languages.*
- Roberts, John R. 1987. *Amele.* London: Croom Helm.
- Rumsey, A. 1982. *An Intra-sentence Grammar of Ungarinjin North-Western Australia* (Pacific Linguistics, Series B, No. 86). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
- Tiersma, Peter M. 1982. Local and general markedness. *Language* 58.832-49.
- Tobin, Yishai. 1988. Sign: context: text - theoretical and methodological implications for translation: the dual number in Modern Hebrew: a case in point. In Reiner Arntz (ed.) *Textlinguistik und Fachsprache: Akten des Internationalen Übersetzungswissenschaftlichen AILA-Symposiums: Hildesheim, 13.-16. April 1987* (Studien zu Sprache und Technik 1), 449-68. Hildesheim: Georg Olms.
- Tobin, Yishai. 1990. *Semiotics and Linguistics.* London: Longman.
- Watkins, Laurel J. 1984. *A Grammar of Kiowa.* Lincoln and London: University of Nebraska Press.
- Zwicky, Arnold M. 1978. On markedness in morphology. *Die Sprache* 24.129-43.

EUROTYP
THEME 7: NOUN PHRASE STRUCTURE

WORKING PAPERS

1. Frans Plank, Suffix Copying as a Mirror-Image Phenomenon.
(February 1990)
Published in: *Linguistics* 28 (1990) 1039-1045.
2. Frans Plank, On the Selective Elaboration of Nominal or Pronominal Inflection.
(May 1990)
3. Greville G. Corbett, Gender and Gender Systems.
(June 1990)
4. Edith A. Moravcsik, Descriptors of Noun-Phrase-Internal Structure.
(August 1990)
5. Greville G. Corbett, Agreement: An Overview.
(September 1990)
6. Frans Plank, *Review of Agreement in Natural Language*, ed. by Michael Barlow & Charles A. Ferguson, Stanford: CSLI, 1988.
(October 1990)
7. Maria Koptjevskaja-Tamm, Action Nominal Constructions in the European Languages.
(November 1990)
8. Jim Hurford, An Inventory of Noun Phrase Universals: Format and Specimen List.
(November 1990)
9. Greville G. Corbett, The Head of the Noun Phrase: Evidence from Russian Numeral Expressions.
(January 1991)
10. Frans Plank, Inflection and Derivation.
(March 1991)
11. Frans Plank, On Determiners. 1. Ellipsis and Inflection; 2. Co-occurrence of Possessives.
(April 1991)
12. David Gil, Universal Quantifiers: A Typological Study.
(April 1991)
13. Frans Plank, From Cases to Adpositions.
(July 1991)
14. Kristina Jokinen, On the two Genitives in Finnish.
(July 1991)
15. Greville G. Corbett, A Typology of Number Systems.
(March 1992)