

AN OFFPRINT FROM

---

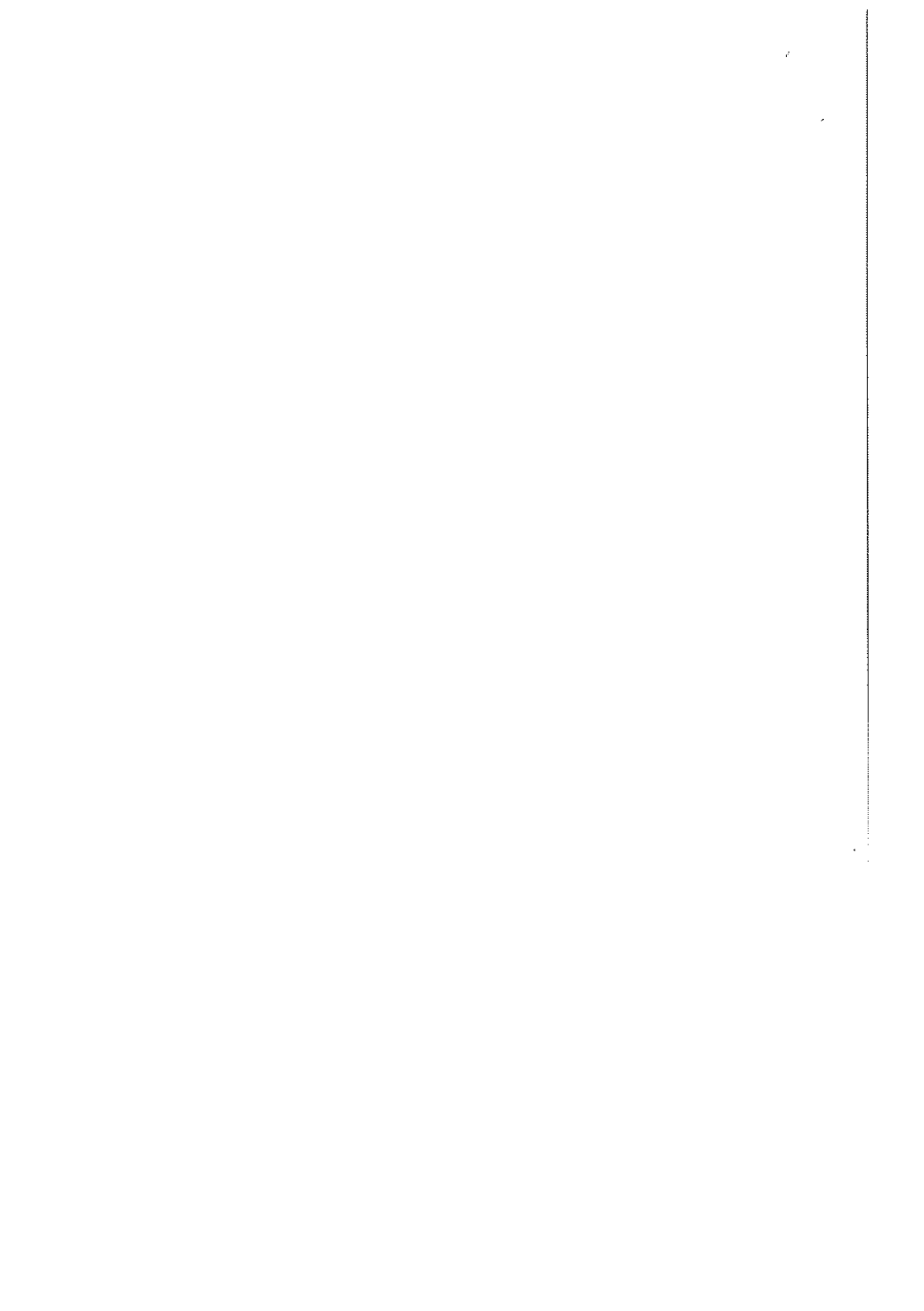
Essays on  
Grammatical Theory and  
Universal Grammar

EDITED BY

Doug Arnold  
Martin Atkinson  
Jacques Durand  
Claire Grover  
Louisa Sadler

CLARENDON PRESS · OXFORD

1989



## An Approach to the Description of Gender Systems\*

*Greville G. Corbett*

### 1. Introduction

In several of the more familiar languages, the gender pattern is straightforward and the way in which the system is analysed is taken as self-evident. In other languages, linguists may present the pattern as though it were equally uncontroversial, but we find that similar situations are described differently by those working on different language families. This causes problems, particularly for those involved in typology. In contrast there are a few languages in which the number of genders has been the subject of interminable dispute.

The purpose of this paper is to develop a consistent approach to gender. After a brief discussion of terminology (section 2), we move on to the central notion of 'agreement class' (section 3). Then, in section 4, we investigate how the nouns in an agreement class may make up a 'controller gender'. Controller genders must be distinguished from 'target genders', and this distinction allows us to solve one of the perennial problems in the study of gender, namely the

© Greville G. Corbett 1989

\* This article is based on research funded by the Economic and Social Research Council (ESRC), reference number C00232218. This support is gratefully acknowledged. Versions of the paper were read at the Spring Meeting of the Linguistics Association of Great Britain, University of East Anglia, Apr. 1986; at the Taylor Institution, Oxford, May 1986; at the Graduate Faculty Seminar in Linguistics, University of Sussex, May 1986; and at the Typology Seminar, School of Oriental and African Studies, University of London, May 1986. I would like to thank those present for helpful discussion. I am also very grateful to Professor Bernd Heine, Professor Dee Ann Holisky, Dr Ewan Klein, Dr Graham Mallinson, Professor F. R. Palmer, Professor Rebecca Posner, Professor Malathi Rao, Professor Roland Sussex and two anonymous referees for their comments, and to Doug Arnold and Dr Louisa Sadler for editorial suggestions. This does not mean, of course, that they agree with the views expressed here.

number of genders in Romanian, and to describe other complex systems in a consistent way. For all its advantages, the agreement-class approach runs into what may be termed the 'maximalist problem', where more agreement classes are identified than the intuitively satisfying number of genders in a given language. Principled reasons are proposed for determining which types of agreement class should be recognized as genders (section 5). Zaliznjak's original insight is considerably extended as a result of analysing a wide range of languages. We shall see that, in many cases, deciding the number of genders in a language is a genuinely difficult problem. In some instances we reach solutions similar to traditional ones but we are able to put these traditional solutions on a sounder theoretical base.

## 2. Terms

The use of terms in this area has become confused (see the discussion at the end of *Classification nominale* 1967: 391–7). There seems little point in trying to maintain a strict distinction between 'gender' and 'noun class' since comparable systems are described as genders in one language family and as noun classes in another. For example, there are several points of similarity between Tamil and Karata; Tamil is said to have three genders, since it is a Dravidian language, while Karata is described as having three noun classes, because it is a North-east Caucasian language (Andi subgroup).<sup>1</sup> We shall follow the widely accepted view that the existence of gender can be demonstrated only by agreement evidence. (The point is that, for instance, the presence of class markers on nouns, as prefixes or suffixes, does not of itself indicate that a language has genders/noun classes; if we accepted that it had, then we could equally claim that

<sup>1</sup> Certain attempts to distinguish gender and noun class depend largely on a contrast between Indo-European and Bantu languages (part of the Bantoid subgroup of Benue-Congo, within Niger-Congo, which is the major part of Niger-Kordofanian). These attempts prove less convincing when Dravidian languages and languages of the Caucasus are taken into account. An alternative approach, however, is that adopted in the UNITYP project, in which gender and noun class are seen as 'focal instances' on a 'scale of classificatory techniques' (as in Serzisko 1982a, for example). A discussion of different approaches can be found in De Wolf (1971: 26–46). He uses gender as a higher level term than noun class (thus a singular and plural noun class may form a gender) but he does not consider agreement to be a prerequisite.

English had a gender comprising all nouns ending in *-tion*.) The view we are following is implicit in most work on gender, and it is stated explicitly in, for example, Fodor (1959: 2), Greenberg (1978b: 50), and Heine (1982: 190). The fact that agreement is the defining factor reinforces the point that gender and noun classes should not be distinguished, but at the same time allows us, like Dixon (1982: 213–18), to separate them from nominal classifiers (in languages like Thai), which we shall not consider here.

### 3. Agreement classes

The approach based on agreement classes is usually associated with the name of Zaliznjak (1964); the approach is partly foreshadowed by Schenker (1955). Gladkij (1969, 1973a, 1973b) gives a more formal, mathematical treatment and reaches results similar to those of Zaliznjak; Kibrik (1972) uses the method in an interesting account of Archi. In a series of papers and in a section of his book, Marcus (1962, 1963, 1967: 115–55, 1970), following ideas of Revzin, also proposed a formal definition of gender, but this works less well than that of Gladkij. His notion of 'class of distribution' proves too wide, and leads to the postulation of large numbers of genders—eight or possibly more for Romanian (1967: 151). By treating, for example, the difference between proper and common nouns as a question of gender, he moves away from the central, well-established problems of gender. Most seriously, it is not clear that there is necessarily any upper limit to the number of genders to be found in a given language. Nevertheless, his contribution to the study of gender has been considerable; several scholars working on Indo-European languages have adopted his approach, probably the most recent being Cosmas (1981, 1982).

Let us now return to Zaliznjak and the agreement-class approach. An agreement class may be defined as follows (this definition is a reworking of the idea found in Zaliznjak (1964: 30):

An agreement class is a set of nouns such that any two members of that set have the property that whenever

- (i) they stand in the same morphosyntactic form, and
- (ii) they occur in the same agreement domain, and
- (iii) they have the same lexical item as agreement target, then their targets have the same morphological realization.

The intuitive content of the definition is that two nouns are in the same agreement class provided that given the same conditions they will take the same agreement form. The three numbered clauses of the definition spell out what is involved in 'the same conditions'. Being in 'the same morphosyntactic form' (clause (i)), or the same 'grammatical' form, means that the nouns have the same specifications for all relevant syntactic features. The features most commonly involved are number and case. We rely on the notions of number and case being given (Gładkij, on the other hand, defines case); this is reasonable since they are simpler notions, which can often be justified simply on morphological evidence, without reference to agreement. It must be stressed that identity of morphosyntactic form does not imply morphological identity. Two nouns may both be specified as, say, accusative singular, and yet differ morphologically; for example, Russian *mat'- $\phi$*  'mother' and *sestr-u* 'sister'. Such nouns have different morphological features in their lexical entries, indicating that they belong to different declensional paradigms. Moreover, two morphosyntactic forms may have a single morphological realization; for example, Russian *okno* 'window' may be the nominative singular or the accusative singular.

Clause (ii) requires that the nouns occur in the same agreement domain. This means that the configuration in which agreement applies must be identical in each case: it might be the agreement of modifiers with the head of a noun phrase, subject-verb agreement, and so on.

Clause (iii) requires that the lexical item which stands as the agreeing element or target must be the same. The point is that not all lexical items have the same agreement possibilities. It would not do to use in one instance an adjective which distinguished gender and in the other an adjective which did not. Of course, we are interested in items which have the largest number of forms, and by specifying that identity must be found 'whenever' these conditions are met, we include the case of the most differentiated target.

The definition depends on the notion of agreement, which we understand to be a matching of features between controller and target. (The reference to matching excludes government where, for example, a verb may require a particular case on a noun phrase but is not itself in that case.) As is generally, but not universally, accepted, the agreement relation covers anaphoric pronouns, whose agreement features depend on features of the controller. Our

definition of agreement class requires two additions. For technical reasons we should add a clause that the nouns in question must share at least one grammatical form. This is to cover the limiting case in which two defective nouns could never stand in the same grammatical form and so the conditions would be filled vacuously. The second addition, which is found in Zaliznjak's definition, is that the target may take not only the same morphological realization but also the same set of stylistically variant forms. This is to cover the situation in which a given combination of syntactic features may have alternative realizations (e.g. Russian adjectives mark the feature set instrumental singular feminine with *-oj* or, archaically and poetically, with *-aju*).

Let us first consider French, as a straightforward illustration of the basic notion of agreement classes.

- |     |                  |                        |
|-----|------------------|------------------------|
| (1) | un grand garçon  | cf. *une grande garçon |
|     | a big boy        |                        |
| (2) | un grand jardin  | cf. *une grande jardin |
|     | a big garden     |                        |
| (3) | une grande femme | cf. *un grand femme    |
|     | a big woman      |                        |
| (4) | une grande fleur | cf. *un grand fleur    |
|     | a big flower     |                        |

In these examples we have ensured that the nouns being tested occur in identical conditions: they stand in the same morphosyntactic form (the relevant feature specification is singular), in the same agreement domain (agreement of modifiers within the noun phrase), and the lexical items involved as agreement targets are identical. The nouns *garçon* and *jardin* require the article and the attributive adjective to stand in the same form ((1) and (2)). If we consider other possible agreement targets, or if we change to the plural, we still find that the agreements required by *garçon* and *jardin* are identical. They therefore belong to the same agreement class. Similarly, the nouns *femme* and *fleur*, while differing from *garçon* and *jardin*, require the same agreements as each other ((3) and (4)). They belong to the second agreement class. There are many thousands of nouns in each of these two classes, which are the traditional masculine and feminine genders.

Zaliznjak was concerned with a more difficult case—Russian.

Before looking at Russian, let us see what the notion of agreement class can contribute to one of the established conundrums of gender, the Romanian system.

#### 4. Controller genders and target genders

As mentioned above, the literature on gender in Romanian is extensive; it includes contributions by Jakobson (1959), major Romanian linguists such as Graur (1937), Rosetti (1965: 83–92, 1983: 382–404), and notably Marcus (1967: 115–55), and many more; for further references see R. A. Hall (1965: 421–2), Marcus (1967: 153–5), Wienold (1967: 75, 170), Luxt (1970: 88–9), Windisch (1973), Priestly (1983), and Mallinson (1984: 450–1).

Consider the following data (from Mallinson 1984: 441) concerning the nouns *bărbat* 'man', *fată* 'girl', and *scaun* 'chair'.

- (5) *bărbatul e bun*  
man-the is good
- (6) *scaunul e bun*  
chair-the is good
- (7) *fata e bună*  
girl-the is good

Note that the definite article is postposed; in nouns like *fată* its effect is to change the quality of the final vowel (mainly by lowering) to *fata*. The evidence so far demonstrates the existence of two agreement classes, one including nouns like *bărbat* and *scaun* and the other comprising nouns like *fată*. There is a second case (genitive-dative), but in the singular *bărbat* and *scaun* again take identical agreements while *fată* differs. But the situation is more complex, as emerges when we consider the same phrases in the plural:

- (8) *bărbatii sînt buni*  
men-the are good
- (9) *scaunele sînt bune*  
chairs-the are good
- (10) *fetele sînt bune*  
girls-the are good

If we had only the data of (8)–(10), then we would postulate two agreement classes—one for nouns like *bărbat* and one for nouns like

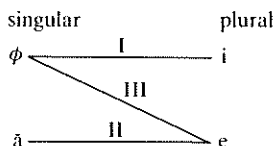
*scaun* and *fată* (the oblique case in the plural shows the same pattern, separating *bărbat* from the other two). The argument, which has gone on for decades, is whether we have two genders or three. The problem is that nouns like *scaun* have no agreement forms which are used uniquely for them. In terms of agreement classes, however, the situation is clear: we must set up three classes as follows:

- I. nouns taking  $\phi$  in singular and  $i$  in plural (*bărbat*)
- II. nouns taking  $\phi$  in singular and  $e$  in plural (*scaun*)
- III. nouns taking  $\bar{a}$  in a singular and  $e$  in plural (*fată*)

Thus we have an unambiguous answer: there are three agreement classes, and there is no reason not to recognize each as a gender. However, this analysis is exactly what we would obtain for Latin, Tamil, or German, even though in each of these languages, intuitively, the situation is rather different. All of them have some agreement forms which are unique to each gender. The point is that the agreement-class approach leads us to the number of sets into which nouns are to be divided or, in a feature-based approach, to the number of different feature specifications which are required on nouns to enable gender agreement to operate correctly. It is certainly the case that *bărbat*, *scaun*, and *fată* (and the hundreds of other nouns in each class) must be labelled differently. Nevertheless the morphology of agreeing forms (targets) is simpler than is implied by the statement that Romanian has three genders. We should therefore differentiate *controller genders*, (the genders into which nouns are divided) from *target genders*, the genders which are marked on adjectives, verbs, and so on (depending on the language).

The distinction is illustrated in Table 3.1. It can be seen that Romanian has two target genders in both singular and plural, but three controller genders (indicated by the lines). I is usually called masculine, II is the feminine, and III is the disputed class sometimes called 'neuter', sometimes 'ambigeneric'; the latter is perhaps the better term, provided it is used not to imply that there is no distinct

TABLE 3.1. *The gender system of Romanian*



gender but rather that the situation is different from the more common Indo-European three-gender system.

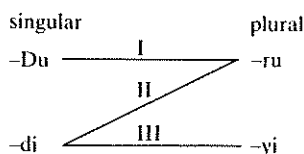
Diagrams like that above can be labelled in various ways; the alternatives deserve brief consideration. The first controller gender is designated  $\phi$  on the basis of adjectives like *bun* 'good'. However, not all adjectives take this form: *aspru* 'rough' has *-u*, as shown by comparison with *aspră* corresponding to *bună*. We have chosen to give a typical allomorph for each target gender. This method avoids the danger of premature naming of genders; on the other hand, problems can arise when the typical forms chosen suggests similarities which are not general through the system. (Taking Latin adjectives, we might suppose that the feminine singular *-a* is equivalent to the neuter plural *-a*; however, not all adjectives have identical forms in these two positions.) A way of avoiding the latter problem is to list all the allomorphs, but this can become unwieldy. Since it is hardly practical to keep referring to strings of allomorphs, names tend to be preferred. Let us consider French again in this form (Table 3.2).

TABLE 3.2. *The gender system of French*

singular		plural
	masculine	
masculine	—————	masculine
	feminine	
feminine	—————	feminine

In a language of this type, it is natural to use the same labels for the classes into which nouns are divided (controller genders) and the sets of agreeing forms (target genders). It is when this usage is carried over into more complex systems that difficulties arise. Indeed, although the distinction between controller and target genders may seem an obvious one, there are numerous examples in the literature of the number of genders being given for a particular language, in cases where the situation is complex, without any indication as to what is meant. It is true that there are many languages where the number of controller and target genders are the same, but mismatches of the type we have seen in Romanian are not uncommon. Another example is found in Telugu, a Dravidian language (South

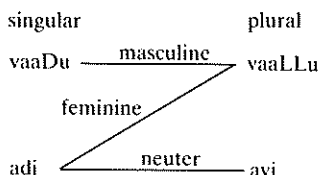
TABLE 3.3. *Verbal agreement forms in Telugu*



Central Group). Table 3.3 gives the verb-agreement forms (Malathi Rao, personal communication).<sup>2</sup>

Again there are three controller genders and two target genders (in both singular and plural). However, I is masculine, II feminine, and III irrational or neuter. The comparison between Romanian and Telugu illustrates the point made in the introduction that similar situations are described differently in different geographical traditions; while there has been interminable debate about Romanian, there has, to my knowledge, been no such debate about the status of the third (controller) gender of Telugu. However, unlike Romanian where the personal pronouns follow the same pattern as the other agreement targets, in Telugu the pronouns are more complex. There is a set which corresponds closely to the verb agreements (see Table 3.4). These are not the only forms. Besides

TABLE 3.4. *Telugu personal pronouns*



*vaaDu*, which is an informal form for referring to a man or a boy, there is also *atanu* (partly formal) and *aayana* (formal) for referring to an adult. Telugu uses the plural to show respect (when referring to someone who merits respect, not just in direct address); *atanu* can take plural agreement and *aayana* must do so in Koosta Telugu (in some other dialects *aayana* can take a singular). The use of the plural

<sup>2</sup> I am most grateful to Malathi Rao for the data on Telugu; her Koosta dialect varies slightly from that described in Krishnamurti and Sarma (1968), where the transliteration scheme is also somewhat different.

to show respect requires a complication of the grammar, but it does not affect our account of the gender system, nor does the existence of additional masculine pronouns. When we turn to the pronoun *adi* the picture changes; *adi* is used for reference to neuters (animals and things) and sometimes for reference to girls. For reference to a woman, there is a different form *aaviDa* (partly formal or formal) which can take singular or plural agreement. It is not used for inanimates. The existence of this pronoun means that pronoun targets distinguish feminines from neuters. This confirms the analysis which claims three controller genders (cf. section 5.2); it also breaks the parallelism with Romanian.

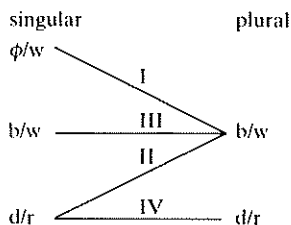
A more complex system is found in Lak, a North-east Caucasian language (Lak-Dargva group). Agreement is marked on some attributives, verbs, adverbials, and pronouns, as illustrated in the next example:

- (11) *k'i-j-a ars ša-w-a φ-ušar* (Xajdakov 1980: 206)  
 two son at home be  
 'Two sons are at home.'

The italicized elements all mark agreement with the Class I noun *ars* 'son'.

The pattern of typical markers is given in Table 3.5 (from Kibrik 1979: 4; the forms like *-j-* found in numerals are not included). By convention, forms before the slash are prefixal, those after are infixal or suffixal. There are four controller genders in Lak. Then there are three target genders in the singular and two in the plural. (In view of the warning about representing target genders by typical markers, it should be said that the table gives typical verb agreement markers which undergo numerous alternations, but that these alternations do not provide grounds for subdividing the target genders as given.) It

TABLE 3.5. *The gender system of Lak*



is noteworthy that there are only three different sets of forms ( $\phi/w$ ,  $b/w$ , and  $d/r$ ) covering eight theoretical forms.

We have established the value of distinguishing the notions 'controller gender' and 'target gender', in our analysis of complex gender systems. The distinction leads to other fruitful lines of enquiry, as we shall see in the sections which follow.

#### 4.1. The relation of gender and number

As far as controller genders are concerned, gender and number are typically independent: a noun has a particular gender, irrespective of the number it stands in. Gender is inherent to the noun and is in a sense prior to the number in which it occurs in a given sentence. (There are occasional exceptions which must be labelled as such in the lexical entry of the nouns involved; see section 5.3 for some examples.) Target genders, on the other hand, may be found in the realizations of feature bundles and so may vary according to the other features involved, most notably number. And there are interesting areas to investigate in these interactions. In particular, there are the patterns of matching between singular and plural.

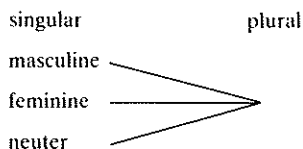
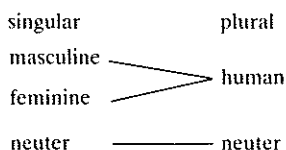
The most straightforward system is that in which each singular target gender matches one plural gender and vice versa. Gender in one number determines gender in the other and vice versa: there is a one-to-one mapping of the target genders in one number on to the target genders in the other. We call such systems 'parallel' systems. French can serve as an example of this type (see Table 3.6).

TABLE 3.6. *Target genders in French*

singular		plural
masculine	—————	masculine
feminine	—————	feminine

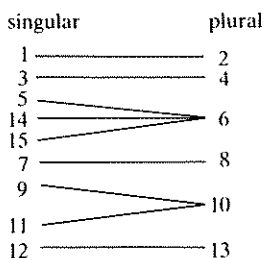
Then there are 'convergent' systems, in which gender in one number determines gender in the other, but not vice versa: there is a many-to-one mapping of target genders in one number on to target genders in the other. (Parallel and convergent systems taken together are what Heine terms 'paired systems', 1982: 196-7.)

A clear example of a convergent system is found in German (see Table 3.7). In Tamil, a Dravidian language, three singular target

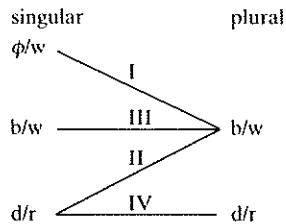
TABLE 3.7. *Target genders in German*TABLE 3.8. *Target genders in Tamil*

genders correspond to two plural genders (Table 3.8). An identical system is found, as mentioned earlier, in the North-east Caucasian language Karata (Magomedbekova 1967: 324–5, 327).

In ChiBemba, a Bantu language, (data from Givón 1972: 14–17) there are considerably more target genders. The numbers given in Table 3.9 are the traditional Bantu noun-class numbers. We omit here the locative genders, which are discussed in relation to another Bantu language in section 4.3.

TABLE 3.9. *Target genders in ChiBemba*

The third possibility is what Heine calls a 'crossed' system (1982: 197). In such a system, gender in neither number determines the gender in the other; there is what some would call a many-to-many mapping between the target genders in the different numbers. Examples are Romanian (Table 3.1), Telugu (Table 3.3), and Lak (given again for convenience as Table 3.10).

TABLE 3.10. *The gender system of Lak*

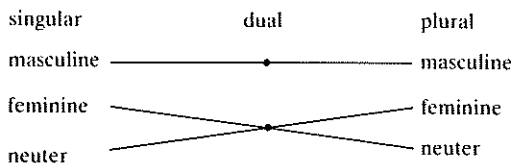
Here we see a system with four agreement classes, but with three target genders in the singular and two in the plural. The fact that we find such systems, as well as convergent systems like those above, with more singular target genders than plural, but not the converse (with more plural than singular genders), is, of course, a reflection of Greenberg's universal number 37: 'A language never has more gender categories in nonsingular numbers than in the singular' (1966: 112). If we adopted just the agreement-class approach, there would be no way in which to state Greenberg's universal, since differences caused by change in number lead to the setting up of additional agreement classes. In terms of agreement classes, Table 3.10 simply provides evidence for four classes. Clearly then Greenberg's universal must be stated in terms of target genders.<sup>3</sup> The Lak data also suggest how, once controller and target genders are distinguished, we can approach the question of syncretism of target genders; we also have the framework for an account of the associated question of polarity, notably in Cushitic languages (see Serzisko 1982b).

Before leaving the question of gender and number we consider languages with more than two numbers. For simplicity we have discussed languages with only singular and plural. When we turn to fuller number systems, we find that the types established, 'parallel',

<sup>3</sup> Though this is not required for Greenberg's universal, we should consider the justification for linking particular target genders in the singular with others in the plural. In languages like French, the point may seem too obvious for discussion; but there are instances where the analysis is not so clear-cut, and in any event, we must recognize that it is a matter of analysis, not a fact. To justify associating particular singular and plural target genders (and target genders in different agreement targets, such as attributive modifiers and predicate verbs), we need the notion of 'consistent agreement pattern'. This concept is discussed in section 5.5; it gives a principled basis for associating particular agreement forms.

'convergent', and 'crossed', hold between each pair of numbers, rather than of the system as a whole. This can be seen from the gender system of Slovenian, a South Slavonic language (we omit here two subgenders, which are just as in Slovenian's closest relative Serbo-Croat, described in section 5.1) (see Table 3.11). We can see that the feminine and neuter share forms in the dual only. Thus there is a relation of convergence between singular and dual and between plural and dual, but of parallelism between singular and plural. (See Corbett and Hayward 1987 for an analysis of a four-number system.)

TABLE 3.11. *The gender system of Slovenian*



#### 4.2. Relation to semantics

The semantic aspect of gender is most evident in controller genders; the way in which nouns are assigned to genders (assignment systems) is a well-established area of research with a correspondingly large literature. There is always a semantic core to gender systems, but the degree of semantic justification can vary from almost complete to much less than that. Thus in Dravidian languages such as Tamil and Telugu it is normally sufficient to know the meaning of a noun in order to assign it to a gender (see e.g. Asher 1985: 136–7). In many other instances, in Indo-European languages like French and German, in Bantu languages like Swahili and Chichewa, and in a large proportion of the languages of the Caucasus like Lak, many nouns can be assigned to a gender according to their meaning (typically those denoting humans) but many cannot.

It is much less frequently observed that the question of semantic justification can also be pursued with regard to target genders. If the controller genders of a language are based solely on semantics, then the appearance of a particular gender agreement form will provide unambiguous information. Thus a feminine agreement form in Tamil implies that a female human is referred to at some point, while

a feminine agreement form in French or German does not necessarily do so. When we turn to convergent and crossed systems, we find that the syncretic forms may have varying degrees of semantic justification. Compare Lak (Table 3.10) with Archi (Kibrik 1972) (Table 3.12). The plural target-gender forms of Lak have no direct semantic correspondence. In Archi, their semantics are clear, once the controller genders are given: Class I comprises male humans, God, and other spiritual beings considered male, II is similarly for females, III is for most animals and some inanimates, and IV includes the remainder (some animals and the remaining inanimates). Thus the target gender *b/ib* in the plural implies reference to humans, while  $\phi/ib$  is for non-humans. (For other examples see Xajdakov 1980.)

TABLE 3.12. *The gender system of Archi*

singular	plural
w/w	I
d/r	II
b/b	III
$\phi/t$ or $t'$	IV
	<i>b/ib</i>
	$\phi/ib$

### 4.3. The relation of controller genders to target genders

In many languages, as noted earlier, there is a straightforward one-to-one relation between the two types of gender. We have also seen several examples of languages in which the number of controller genders exceeds the number of target-gender forms distinguished in either number. For example, Romanian has three controller genders, but only two target-gender forms in both singular and plural.

It is also possible to find languages in which the number of target genders exceeds the number of controller genders, if we include within the target genders what we shall term 'non-lexical genders'. These are of two types; the first, the neutral gender, is a target-gender form which cannot have a noun phrase headed by a noun as its controller. A good example is found in the Surselvan dialect of Romansh (Haiman 1974: 130-4). The agreement-class approach leads us to postulate two genders; one takes the ending *-s* on

predicative adjectives and participles (masculine), sometimes with a change of root vowel, while the other takes *-a* (feminine). There is, however, a third agreement marker, namely  $\phi$ ; this occurs with controllers which are not specified for gender and number: sentential subjects, the demonstrative pronoun *quei* 'that', and the impersonal pronoun *igl*. For example:

(12) Sgarscheivla ei la paupradad  
Wretched (Fem) is the poverty

(13) In urezi ei sesalzaus  
A storm is arisen (Masc)

(14) Igl ei sesalzau in urezi  
There is arisen (Neutral) a storm

The form which is found in (14) cannot occur with a noun phrase headed by a noun as controller; we term this target gender the 'neutral' gender.

The second type of non-lexical gender (and so the second way in which target genders may exceed controller genders) involves 'minor target genders'. These differ from the neutral gender in that they can have a nominal controller. They differ from ordinary target-gender forms in that there are no nouns regularly assigned to a corresponding controller gender.<sup>4</sup> Examples would be the locative genders of Bantu; consider the following example from Chichewa, a Bantu language spoken in Malawi:

(15) munyumba mu-kutentha  
in house is hot  
'It is hot in the house.'

<sup>4</sup> Greenberg (1978b: 53) and Heine (1982: 198) both use the term 'minor gender' in this sense, and further examples are given in Greenberg (1978b: 79). The term 'minor gender' has also been used to signify simply an agreement class with few nouns in it (e.g. by De Wolf (1971: 42 and Watters (1980: 135). An alternative term for an agreement class with few members is 'minority gender' (Voorhoeve 1980: 68). We shall suggest in section 5.3 that the status of classes with few members depends on the agreements they take; the particular cases analysed by Watters and Voorhoeve are inqorate genders in our terms, while some classes with few members are recognized as genders (as we shall see in the case of Lelemi).

The potential for confusion over the different uses of the term 'minor gender' is not unduly serious since Greenberg's use can be seen as relating to target genders (minor genders are, in our approach, target-gender forms with no nouns regularly assigned to a corresponding controller gender) while De Wolf's use involves controller genders (minor genders are agreement classes with relatively few members). To be unambiguous we can use the terms 'minor target gender' and 'minor controller gender'.

*Nyumba* 'house' belong to gender 9/10 in the traditional Bantuist numbering and would normally take the prefixed agreement forms *i* (singular) and *zi* (plural). However, in (15) it has the prefix *mu*, meaning 'in', and the verb takes the same prefix *mu* as an agreement marker. No other gender has this marker so we could treat it as a separate target-gender form. But there are no nouns which must take *mu* as their agreement form; apart from semantic restrictions, any noun could take the prefix *mu* and become part of this locative gender. Thus the agreement form *mu* represents a minor target gender; such genders are clearly rather different from the more familiar genders.

Thus far the agreement-class approach appears to give useful results, provided it is recognized that it gives an analysis of one side of the problem, the controller genders, and that it must therefore be supplemented by the notion of target genders.

## 5. The maximalist problem

A problem which arises with the agreement-classes approach is that the number of classes may be considerably larger than the traditional (and often intuitively satisfying) number of genders generally accepted for a given language. Zaliznjak and Gladkij were both aware of the problem and addressed it directly. Some later investigators carefully establish agreement classes but take the analysis no further; in some cases the agreement classes are claimed to be genders. Dalewska-Greń and Feleszko (1984) find fifteen agreement classes in Serbo-Croat and thirteen in Polish. Quirk, Greenbaum, Leech, and Svartvik (1984: 314) propose nine gender classes for English. While the value of identifying all the classes of noun which differ in their agreement possibilities should not be underestimated, the raw analyses which result tend to be unsatisfactory for two reasons: first, they miss generalizations (some of which are captured in more traditional accounts); and second, they make similar systems appear more different than they really are.

We shall, therefore, investigate how the number of agreement classes may be reduced, in principled ways, to give a lower number of genders. Again we shall find the notion of controller and target genders helpful.

## 5.1. Subgenders

The notion of subgenders is useful and intuitively clear, though it proves difficult to formalize. Consider the data from Serbo-Croat given in Table 3.13. Serbo-Croat has seven cases, but the cases not given do not provide further distinctions. (More agreement classes can be set up on different evidence, as in Dalewska-Greń and Feleszko 1984, but these are more marginal; similar agreement classes will be considered later in the paper.) The accusative singular gives evidence for four agreement classes. We could simply say that there are four genders in Serbo-Croat. Given appropriate agreement rules this solution would work, but it appears unsatisfactory. The alternative is to treat the agreement class including nouns like *student* and that including nouns like *zakon* as subgenders of a single main gender. We would then recognize three genders (as is the tradition in Serbo-Croat grammars): masculine being divided into masculine animate (*student*) and masculine inanimate (*zakon*).<sup>5</sup>

TABLE 3.13. *Gender agreement in Serbo-Croat*

	ovaj student 'this student'	zakon 'law'	škola 'school'	vino 'wine'
<i>Singular</i>				
Nominative	ovaj student	ovaj zakon	ova škola	ovo vino
Accusative	ovog studenta	ovaj zakon	ovu školu	ovo vino
Genitive	ovog studenta	ovog zakona	ove škole	ovog vina
<i>Plural</i>				
Nominative	ovi studenti	ovi zakoni	ove škole	ova vina
Accusative	ove studente	ove zakone	ove škole	ova vina
Genitive	ovih studenāta	ovih zakonā	ovih školā	ovih vinā

The intuition we wish to capture is that the difference between the behaviour of words like *student* and *zakon* is less significant than that between both of them and those like *škola* or *vino*. Serbo-Croat provides a particularly clear case for a subgender; it proves difficult to give a definition which will cover more complex cases as well. The following is an attempt to specify when agreement classes can be analysed as subgenders rather than as full genders:

<sup>5</sup> There is a considerable literature on the place and development of animacy in Slavonic; see e.g. Hjelmslev (1956) and Huntley (1980).

Subgenders are agreement classes which control minimally different sets of agreements, that is, agreements differing for a small proportion of the morphosyntactic forms of the controller (typically a single one), not including the most basic form (usually the nominative singular).

This definition fits the Serbo-Croat data well. The agreement classes represented by *student* and *zakon* differ in the accusative singular only, that is, in one morphosyntactic form out of fourteen, if we consider only attributive agreement (though it should be said that syncretisms reduce the number of independent morphological forms to less than fourteen for any given declensional type). If we apply the same procedure to Romanian, we find that masculine and ambigeneric (and also feminine and ambigeneric) nouns take different agreements in two forms out of four; this is not 'minimally different'. It follows that subgenders will be found only in languages with relatively complex inflectional systems. Note that a minimal difference cannot include the nominative singular; the most basic nominal form is traditionally (and with good reason) considered to provide sufficient evidence to justify a separate gender. Returning to the Serbo-Croat data, we have already said that the use of subgenders allows us to reflect more accurately the relative similarity of *student* and *zakon* as opposed to the others. In particular, some rules can refer to the feature [+masculine] without any account of the feature [+/-animate]. There is a second advantage, which is that the assignment rules are rather different for genders and subgenders. For the main genders there is a semantic rule (basically males are masculine, females feminine); the remaining nouns (not referring to sex-differentiable beings) are assigned according to morphology (the declensional class). On the other hand, the distinction between animates and inanimates is, with a very few exceptions, purely semantic.

Table 3.13 reveals a further peculiarity of the subgenders in Serbo-Croat, namely that the agreement forms involved are not independent but syncretic.<sup>6</sup> The question arises as to how we should count the

<sup>6</sup> It should be said that a syntactic analysis, which would involve changing the case of the whole noun phrase, would be inadequate, as is shown by masculine nouns ending in *-a*, which themselves take the accusative in *-u*:

- (i) vidim ovog gazdu  
(I) see this boss

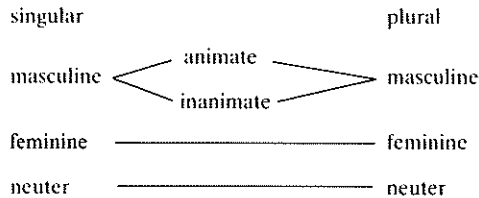
*Gazdu* is unambiguously accusative; the attributive modifier *ovog* is masculine and is accusative-genitive.

target genders of Serbo-Croat. For this we require the notion of 'dependent target gender'.

A dependent target gender is a target gender consisting of a set of morphological realizations which mark agreement with members of a given agreement class by an opposition involving only syncretism (and no independent form).

The sets of morphological realizations which mark accusative singular masculine in Serbo-Croat fit this definition. As was illustrated in Table 3.13, when, say, the determiner *ovaj* 'this' is in agreement with an accusative singular masculine inanimate noun like *zakon* 'law', its morphological realization is the same as for the nominative singular. This is a systematic syncretism; other agreeing elements take a different ending from *ovaj*, but always the same as for the nominative. No target has an independent form for this feature specification. Similarly, the corresponding forms which agree with animates are syncretic with the genitive. The opposition animate/inanimate is marked only by syncretism and so both sets of morphological realizations, those for animate and those for inanimate, constitute dependent target genders. We can then say that Serbo-Croat has three controller genders, one of which has two subgenders, three (independent) target genders (singular and plural mark the same distinctions), plus two dependent target genders. This situation is presented graphically in Table 3.14.

TABLE 3.14. *The gender system of Serbo-Croat*



Let us now examine Russian, the language which both Zaliznjak and Gladkij analysed. The singular yields the same distinctions as Serbo-Croat: masculine, feminine, and neuter, with two subgenders in the masculine—animate and inanimate. However, while Serbo-Croat differentiates masculine, feminine, and neuter in the plural, Russian has lost this distinction and replaced it by animacy, as shown in Table 3.15.

In some examples there are changes of stress, but this need not concern us here. We find three controller genders, each with two

TABLE 3.15. Gender agreement in Russian

	ètot student 'this student'	dub 'oak'	sestra 'sister'	sosna 'pine'	čudovišče 'monster'	okno 'window'
<i>Singular</i>						
Nominative	ètot student	ètot dub	èta sestra	èta sosna	èto čudovišče	èto okno
Accusative	ètogo studenta	ètot dub	ètu sestru	ètu sosnu	èto čudovišče	èto okno
Genitive	ètogo studenta	ètogo duba	ètój sestry	ètój sosny	ètogo čudovišča	ètogo okna
<i>Plural</i>						
Nominative	èti studenty	èti duby	èti sestry	èti sosny	èti čudovišča	èti okna
Accusative	ètix studentov	èti duby	ètix sester	ètix sosny	ètix čudovišč	ètix okna
Genitive	ètix studentov	ètix dubov	ètix sester	ètix sosen	ètix čudovišč	ètix okon

TABLE 3.16. *Genders and subgenders in Russian*

masculine	animate	<u>student</u>
	inanimate	<u>dub</u>
feminine	animate	<u>sestra</u>
	inanimate	<u>sosna</u>
neuter	animate	<u>čudovišće</u>
	inanimate	<u>okno</u>

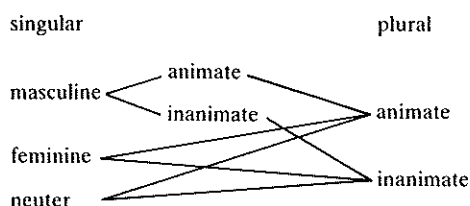
subgenders (Table 3.16). This is the solution favoured by both Zaliznjak and Gladkij (though terms vary and both also deal with problems of number, which we will address in section 5.4). Zaliznjak (1964: 28) considers that setting up two layers of features is an analysis which should and can be argued for, but that it does not reflect an objective division of nouns in the way that agreement classes do. Gladkij (1973b: 193–4) argues that: ‘if two case-number forms coincide for a given noun, and if they have the same segments in agreement with them . . . these forms should, in a definition of gender, “fuse” into a single form’. He then includes this idea in his formal definition. Unfortunately, the idea is flawed, since there are nouns which are clearly masculine animate but whose case-number forms do not coincide:

- (16) Nominative *moj djadja* ‘my uncle’  
 Accusative *mojego djadju*  
 Genitive *mojego djadi*

Our definition, which is fully in the spirit of Zaliznjak’s agreement-class approach, appears adequate here. For the feminine and neuter, the two agreement classes for each (animate and inanimate) differ only in one form out of twelve (excluding syncretism), namely the accusative plural. For the masculine, there is a difference in the accusative singular (as in Serbo-Croat) and the accusative plural. Since it is the same case in both instances and the syncretism is the same, it is reasonable to accept the two agreement classes as minimally different.<sup>7</sup>

<sup>7</sup> There is a little evidence showing that neuter animates are beginning to take animate agreements (that is, identical to the genitive) in the singular as an alternative to the inanimate forms described (Icković 1980: 88). They are thus becoming less like feminines and more like masculines. The change represents a further stage in the integration of animacy into the gender system.

Given that the type of syncretism found with the animacy features is always the same in Russian, and that it has a role elsewhere in the grammar (see Corbett 1981), the subgender analysis is clearly superior to one recognizing six genders. A six-gender scheme would allow agreements for animate masculines to be completely different from other animates, whereas in all examples it is syncretism of accusative and genitive agreeing forms which is involved. Russian also provides the justification for the last part of our definition. If the agreement form for the accusative singular is analysed as equal to nominative or genitive, depending on subgender, then the only remaining difference between masculine and neuter is the agreements with the nominative singular. According to our definition, this fact prevents their being reanalysed as subgenders of a non-feminine gender. We can thus retain the traditional three genders (Zaliznjak and Gladkij were also keen to retain the insights of the grammatical tradition as far as possible). The system is shown in Table 3.17. Russian then has three controller genders, each with two subgenders. It has three target genders, plus two dependent target genders.

TABLE 3.17. *The gender system of Russian*

## 5.2. Overdifferentiated targets

We continue with other methods which, like the use of subgenders, allow us to restrict the number of agreement classes which are recognized as genders. In 'maximalist' analyses, which produce large numbers of agreement classes, marginal targets may be included. Zaliznjak (1964: 30) discusses briefly Russian collective numerals such as *dvoe* 'two', which can be used with some nouns but certainly not all. He does not use them to set up further agreement classes, rightly I believe, since the use of *dvoe* rather than *dva* 'two' can be considered a matter of co-occurrence restrictions rather than a

question of agreement.<sup>8</sup> Several examples of such items being used greatly to complicate gender patterns can be found in the literature.

There are, however, more serious problems with apparent agreement targets which permit more distinctions than do typical targets in the language: we term such targets 'overdifferentiated' (using a term of Bloomfield 1933: 223-4). A dramatic example is found in a group of closely related Dravidian languages: Kolami (Emeneau 1955: 56), Ollari (Bhattacharya 1957: 19), Parji (Burrow and Bhattacharya 1953: 9-10), and Naiki (Emeneau 1955: 141). Each has two genders, basically male human and other. Yet some lower numerals (2, 3, and 4 in Kolami) have forms for female human additionally, as the following Kolami data show:

- (17) iddar ma·sur  
two men
- (18) i·ral pillakul  
two women
- (19) indiq siql  
two buffaloes

I would argue that we should maintain the traditional account that these languages have two genders (as demonstrated by verb agreement and the forms of personal pronouns) and that the three exceptional numerals (overdifferentiated targets) should be labelled as having special forms which co-occur with nouns denoting female humans.

It is worth considering the difference between subgenders and agreement classes induced by overdifferentiated targets. In the former case, agreement classes are separated by a distinction which is widely and consistently marked, but only for a small proportion of

<sup>8</sup> A related problem is Polish *dwa* 'two' (a normal numeral with no collective meaning), which distinguishes feminine gender from others in the plural; other attributive modifiers do not do so. Schenker (1955) bases a good deal of his analysis on *dwa*, while Zaliznjak, considering the similar case of *oba* 'both' in Russian, dismisses it in a footnote (1964: 31), saying that the feminine forms are being lost. The fact that such forms are being lost (speakers have some problems with Polish *dwa* as well) confirms their marginal status. When a small number of targets (particularly if they are in some other way unusual—*dwa* and *oba* both have syntactic quirks) serves to establish agreement classes in excess of those found on the basis of typical targets, these additional agreement classes will not be recognized as genders. Rather the overdifferentiated targets should be marked lexically.

the morphosyntactic forms of the controller. The distinction is therefore not so significant as that between the main genders. In the Serbo-Croat case almost all agreeing attributive modifiers (as well as the relative pronoun) make the animate-inanimate distinction,<sup>9</sup> but as we saw this distinction is found only in the accusative case. In Kolami, however, it is not a whole target type (like attributive modifiers or agreeing predicates) which is involved, nor is it a whole word class (such as all numerals). There are just three irregular numerals, and these should be lexically marked as such.<sup>10</sup>

We can now approach the gender system of English. It might be argued that the personal and reflexive pronouns of English are overdifferentiated and that therefore English has no gender system. But personal pronouns are central to the syntax in a way which the numerals 2, 3, and 4 are not. They are the basis of a separate target type. And given that pronominal systems which mark gender typically repeat gender systems found elsewhere, either in the same language (e.g. in the verb) or in gender systems in other languages, it seems more reasonable to accept gender systems based on the personal pronouns. In addition, those languages which have gender reflected both in pronouns and elsewhere (verbs or attributive modifiers) typically require similar machinery to account for the appropriate forms of each. It may be significant that pronouns may be used by native speakers to ask about gender, as in the following Russian example (the gender of *kofe* is changing):

- (20) *Kofe, èto on, da?*  
 coffee that (is) he yes  
 ‘‘Coffee’’ is masculine, isn’t it?’

More important is the fact that our definition of agreement class depends of course on a definition of agreement. This has caused considerable controversy, but most recent attempts to define

<sup>9</sup> An exception would be the numeral *dva* ‘two’, which distinguishes only feminine from non-feminine and so is underdifferentiated.

<sup>10</sup> We can therefore distinguish three cases: (i) agreement classes are induced for a sizable proportion of the morphosyntactic forms of the controller but only by a particular target type (e.g. attributive modifiers): separate genders; (ii) agreement classes are induced by a particular target type but only for a small proportion of the morphosyntactic forms of the controller: subgenders; (iii) agreement classes are induced by a small number of targets (not a complete target type or word class): overdifferentiated target.

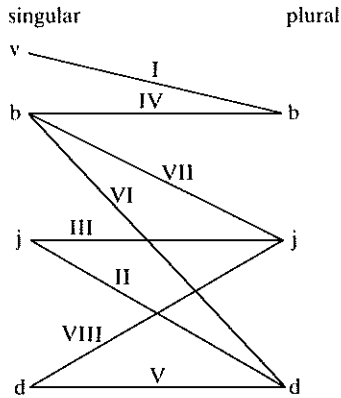
agreement include personal pronouns as a possible target. For this reason, we should recognize English (and similar languages such as Zande) as having a gender system; however, to make clear that it is not a system which penetrates the syntax to the extent of the gender system of, say, Romanian or Serbo-Croat, we will refer to systems like that of English as 'pronominal gender systems'.

### 5.3. Inquorate genders

Inquorate genders are the controller counterpart to overdifferentiated targets. While the latter might artificially raise the number of genders on the basis of a small number of targets, inquorate genders are those postulated on the basis of an insufficient number of nouns, which should be lexically marked as exceptions.

For example, it was stated above that Lak has four genders. Yet there is a single noun which does not fit into the four genders, namely *qāā* 'house'. Žirkov (1955: 21) immediately dismisses the idea of a fifth gender. There is in fact an easy solution: this noun can be labelled as class III in the singular and class IV in the plural. Similarly in Archi (Kibrik 1972: 126), there are two nouns which do not fit into the four main classes (there are also common-gender nouns of a type to be dealt with under section 5.5.1). These two nouns can be marked as class III (singular) and class I/II (plural). Even in French, which everyone 'knows' has two genders, there are three nouns (*amour* 'love', *délice* 'delight', and *orgue* 'organ') which are masculine in the singular and feminine in the plural. These too should be marked as lexical exceptions (in any case they are not straightforward singular-plural pairs). The situation is different from Romanian where there are thousands of nouns which take the same agreements.

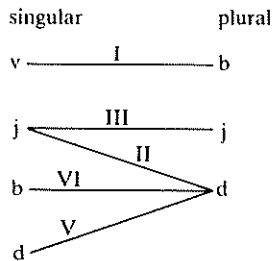
While these cases may appear uncontroversial, it is worth pointing out that if the first published analysis takes a different approach, the existence of a larger number of genders can be perpetuated through the literature. For example, it is regularly stated that Tsova-Tush, formerly called Bats, which is a North-central Caucasian or Nakh language, has eight genders. This view can be found, for example, in the grammar by Dešeriev (1953). Tsova-Tush has an apparently complex crossed system; the eight genders may be represented as in Table 3.18.

TABLE 3.18. *Gender in Tsova-Tush*

However, the membership of the classes is unequal.<sup>11</sup> Class VIII (*d-j*) contains only three nouns, 'lip', 'ear', and 'hand'; we should label them as [V singular/III plural]. Class IV is also small: it has only one native word, meaning 'a knit shoe', together with the word for 'boot', a borrowing from Russian. These should be labelled: [VI singular/I plural]. There are also some singularia tantum which Dešeriev treats as class IV, but these could equally well be treated as class VI. Class VI is the largest class: it appears to be absorbing the last remnants of class IV, which would leave the plural marker *b* as a marker exclusively for class I (male human).

The most interesting class for our purposes is class VII (*b-j*), which consists of the words for 'wing' (of a bird), 'mouth', 'lung', 'leg', 'arm', 'fist', 'finger', 'nail' (of finger), 'eye', and 'throat, gorge'. We find ten nouns in all (Dešeriev lists two others which now appear to have moved into class III). Like class VIII, class VII is strong on body parts, particularly paired body parts. There appear to be good grounds for declaring class VII 'inquate'; that is to say, we should mark the nouns listed as lexical exceptions and not recognize a class VII. The system would then be as in Table 3.19 (with the original class numbers retained, but with the layout

<sup>11</sup> Information on class membership is given in Dešeriev (1953: 138-45). This has been revised and updated in the light of data kindly provided by Dee Ann Holisky based on her fieldwork in Zemo-Alvani, on lecture notes from a course by Rusudan Gagaa, and on Kadagidze and Kadagidze (1984).

TABLE 3.19. *Gender in Tsova-Tush (excluding inquorate genders)*

rearranged). This leaves the interesting question as to when a gender becomes quorate. There can be no simple answer: we should look to evidence such as productivity, changes affecting the class of nouns as a whole (rather than individual members), and so on. But it is important that when the number of nouns in an agreement class is known to be small, this should be stated, particularly if it is claimed that the agreement class should be recognized as a gender.<sup>12</sup>

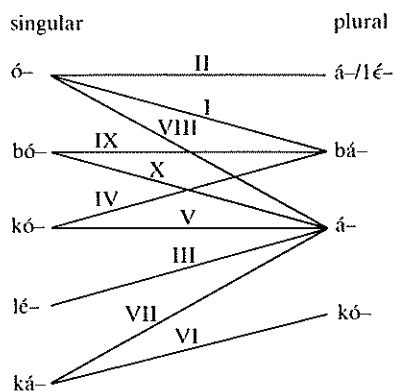
It is important to note, however, that in all the cases discussed the nouns could be given an exceptional marker (for an unusual pairing of singular and plural target-gender forms) which would allow the agreement rules to assign the required markers. It does not follow that any agreement class with a small number of members is necessarily inquorate, since it may not be possible to give all the nouns an irregular marker in this way.

Let us consider the very interesting case of Lelemi (a Togo Remnant language, in turn part of Niger-Kordofanian, spoken in the Volta region of Ghana by 14,900 people at the 1960 census). Heine (1982: 197–8) presents it as an example of a complicated crossed system. There is more information in Heine (1968: 114–15) and it is that source which we shall follow. There is evidence for ten agreement classes, which Heine labels I–X (see Table 3.20).

Agreement classes II and VIII normally both take the plural agreement marker *á-*, but class II has the optional alternative *lé-* for the demonstrative pronoun – *mɛ* ‘this’ only. We should mark this

<sup>12</sup> Voorhoeve (1968: 587) adopts the useful convention of indicating classes with few members by broken lines in representations like Table 3.20. Unfortunately the broken line, like the star, has a different meaning in synchronic and diachronic description; De Wolf (1971: 49) uses a broken line to indicate an innovative pairing while a solid line indicates one retained from the parent language.

TABLE 3.20. Agreement classes in Lelemi



pronoun as overdifferentiated and so not recognize VIII as a separate gender. When we look at the examples given both in the section on Lelemi and in the general word list (Heine 1968: 212–57), it is clear that the agreement classes shown in Table 3.21 account for most of the nouns.

TABLE 3.21. Main agreement classes in Lelemi

Agreement class	Agreements	
	Singular	Plural
I	ó-	bá-
II	ó-	á-
III	lé-	á-
V	kó-	á-
VI	ká-	kó-

Provided they are as uncommon as appears to be the case, the nouns in agreement classes IV (*kó-bá-*) and VII (*ká/á-*) could be lexically marked (as V singular and I plural and as VI singular and III plural respectively). What then of the remaining classes IX and X? Heine gives just one example in IX (agreements *bó-/bá-*), and two examples in X (agreements *bó-/á-*). Let us assume the worst case, that there are no other nouns in either agreement class. We would

like to treat them as inqorate. However, unlike the cases considered so far, there is no exceptional marker available, since no other gender has *bó-* for singular agreement. Furthermore, we cannot look to an irregularity in a particular target, since there are seven possible agreement targets in Lelemi and it appears that all seven take the marker *bó-*. Thus although we do not need to recognize Heine's IV, VII, and VII as genders, we must accept IX and X. These agreement classes are not inqorate, since the nouns cannot be lexically marked as exceptional, and they are not induced by overdifferentiated targets since the target gender form *bó-* is found throughout the agreement system. Lelemi has a genuinely complicated gender system, including genders which contain few nouns yet are not inqorate.

#### 5.4. Defective nouns

There are many instances of nouns which lack forms for certain morphosyntactic specifications; most commonly they have no singular or no plural. Intuitively, where in all the morphosyntactic forms available these nouns take the same agreements as normal nouns (in the forms they share), then they are of the same gender. Thus Russian *razgovorčivost'* 'talkativeness' is found only in the singular. Its agreements are the same as those of the feminine inanimate noun *sosna* 'pine', when singular. We can therefore label *razgovorčivost'* as feminine inanimate. There can similarly be nouns which have no singular. In languages like Russian where gender distinctions are not found in the plural, there is a slight complication (this is considered in detail by Zaliznjak (1964: 32–40), who introduces an additional gender class, which is something we shall avoid). The noun *nožnicy* 'scissors' occurs only in the plural; its agreements match those of *duby* 'oaks' (masculine inanimate), *sošny* 'pines' (feminine inanimate), and *okna* 'windows' (neuter inanimate) in Table 3.14. Since Russian distinguishes animacy but not the main genders in the plural, it is sufficient that *nožnicy* be marked as plural and inanimate. There are insufficient grounds for creating an extra gender. This much follows from the definition of agreement class given in section 3. Whenever nouns like *nožnicy* meet the conditions of the definition; they will be found to be in the same agreement class as inanimate plural nouns. The best solution is therefore to have them underspecified in this way (that is, as inanimate, but with no

specification as to main gender). If, however, in a particular theory it was necessary for every noun to have a gender feature, then *nožnicy* could be assigned arbitrarily to any gender, or preferably it could be assigned to the feminine since its morphology is like that of a feminine noun in the plural (other nouns would be masculine or neuter accordingly). Peculiarities of the syntactic behaviour of such nouns, notably with quantifiers, are a problem of number rather than of gender, and so need not concern us. Thus the fact that nouns lack forms for particular morphosyntactic specifications does not give rise to new gender categories.

### 5.5. Consistent agreement patterns

The question of differences in targets has already been mentioned. In some languages all targets mark the same distinctions so that we may take any one target type and use it to establish the agreement classes. In other languages different targets make a greater or lesser number of distinctions; we noted examples of animacy being reflected in attributive modifiers but not in the predicate. In such cases we include the target type which marks more distinctions. In both instances we must establish 'consistent agreement patterns'. This is usually done without comment, but it crucially determines the analysis of the gender systems.

Russian can again serve as an example. We established three main target-gender forms. For an attributive adjective we have the nominative singular endings *-yj*, *-aja*, and *-oe*. The past tense verb has three singular endings:  $\phi$ , *a*, *o*. The relative pronoun has the same endings as the attributive adjective, and the personal pronouns are *on*, *ona*, and *ono*. For the vast majority of nouns, the agreements are as in Table 3.22. (This is of course a simplified table; the full version would include the other cases, the plural number, and

TABLE 3.22. *Agreement patterns in Russian*

Attributive adjective	Predicate	Relative pronoun	Personal pronoun	Traditional gender
yj	$\phi$	yj	on	masculine
aja	a	aja	ona	feminine
oe	o	oe	ono	neuter

therefore the animate and inanimate subgenders.) The question is how this analysis is done, given that there are nouns which take other combinations of agreements. For example, *vrač* in the meaning 'female doctor' can occur in the following constructions:

(21) *novyj* (Masc) *vrač*

(22) *novaja* (Fem) *vrač*  
       new                  doctor

Examples like (21) are more common than (22). In the predicate, the feminine is somewhat more common:

(23) *vrač* *rabotal* (Masc)

(24) *vrač* *rabotala* (Fem)  
       doctor worked

The relative pronoun is usually feminine and the personal pronoun is almost always feminine.

How then is the table constructed? In this case there are phonological clues (*-a* in the feminines, for example) but this is not always a reliable indicator. There are two important factors. The first is that, as mentioned earlier, the vast majority of nouns which take *-yj* also take *-φ* and *on*; the second point is that these are nouns for which we can give absolute rules—they always take the same agreements. Each horizontal line of Table 3.22 represents a 'consistent agreement pattern', which we define as follows:

A consistent agreement pattern is a set of target-gender forms such that:

- (i) the agreement class it induces is as large as possible;
- (ii) agreement rules relating to this agreement class will be simple and exceptionless.

Bear in mind that the number of target genders in the singular may not be the same as that in the plural and that different targets may distinguish different numbers of target genders. However, when different targets make the same number of distinctions, then forms which occur within the same consistent agreement patterns are realizations of the same target gender. The notion of consistent agreement pattern gives us a principled way of capturing the intuition that, for example, a feminine marker on an attributive modifier is 'the same as' a feminine marker on a verb, even if they are phonologically different. A point often overlooked is that it is

insufficient to require that a particular type of agreement, say feminine, should occur with one and the same noun, since there are examples like Russian *vrač* 'doctor', which can take different agreements at the same time:

- (25) naš            vráč    prišla  
       our (Masc) doctor came (Fem)

It is the notion of consistent agreement pattern which allows us to differentiate the agreements in (25), and so to give an account of nouns like *vrač*.<sup>13</sup> Of course, the majority of nouns will belong to agreement classes which have a consistent agreement pattern and which are recognized as genders or subgenders. But we must now consider those which do not. We shall distinguish between nouns which take all the agreements of more than one consistent agreement pattern and those which have fewer possibilities than that. An example from the pronominal gender system of English will make the distinction clear. We find three consistent agreement patterns (see Table 3.23). The word *baby* can occur with all the forms of all

TABLE 3.23. *Gender in English*

who	he	(masculine)
who	she	(feminine)
which	it	(neuter)

three patterns. *Ship* on the other hand may take the neuter agreement pattern:

- (26) The Canberra, which has just docked, is a fine ship. It sails again on Friday.

<sup>13</sup> It is perhaps worth spelling out why the notion of consistent agreement pattern is necessary, since it might appear to duplicate apparatus already given. In particular, we must consider whether agreement patterns which are not consistent are simply those which induce agreement classes which are inqorate genders. Unfortunately things are not that simple. The notion of inqorate gender is necessary for cases like Tsova-Tush (section 5.3) where there would otherwise be no principled reason for choosing between possible singular-plural pairings. (We were able to exclude some since the classes involved were small and the nouns could be marked as lexical exceptions.) But nouns like *vrač* in Russian are too numerous to be simply labelled an inqorate gender. Given their ability to take two types of agreement, and especially the variability involved, we need the notion of consistent agreement pattern to ensure that such nouns do not form the basis of genders, rather than the more straightforward ones.

*Ship* (and other expressions denoting ships) can also take the personal pronoun *she*, but not the relative *who*:

(27) I sailed on the Canberra. She's a fine ship.

(28) \*The Canberra, who is a fine ship . . .

Thus *ship* takes agreements from two consistent agreement patterns but does not take all the forms of both.<sup>14</sup> We shall consider these two types of noun in turn.

### 5.5.1. Double and multiple gender nouns

These are nouns, like English *baby*, which can take all the agreements of more than one consistent pattern. Thus the Lak noun *ḥakin* 'doctor' can take the pattern associated with class I, class II, or class III, depending on whether the doctor is a man, an older woman, or a younger woman (Xajdakov 1963: 50). Such nouns are often called nouns of common gender. This term is useful, as long as it does not imply the existence of a new gender. (After all, when we find verbs which govern the accusative or the dative, we neither postulate a new archicase accusative-dative, nor do we set up a special class of verbs—those governing accusative or dative.) Clearly nouns like *ḥakin* and *baby* must be specially marked in some way. One possibility is to treat them as homonyms: *baby*<sub>1</sub> (feminine), *baby*<sub>2</sub> (masculine), *baby*<sub>3</sub> (neuter). This is the solution adopted by Zaliznjak (1964: 27). The alternative is to label them as having more than one gender, just as we label certain verbs as permitting more than one complement type.

The notion of double or multiple gender is wider than common gender. Common-gender nouns take different agreements for semantic reasons; other nouns may take alternative agreements without any semantic motivation and so still have double gender. For example, in Babanki (one of the Ring languages, which form part of the Western Grassfields group within Bantu, spoken in north-west Cameroon), there are several nouns which can take the agreements either of gender 3/13 or of gender 5/13. This forms part of a more general change in which the target-gender form labelled class 3 is supplanting class 5 (Hyman 1980: 231–3, 236). Other languages may show double gender with borrowings which do not fit into the established gender system, and so are not unambiguously

<sup>14</sup> For further discussion of 'boat nouns' see Malone (1985).

assigned to a gender by the prevailing assignment rules. Such genuine examples of double gender are to be distinguished from those where related nouns are morphologically distinct; the difference in morphology determines the different genders and two separate lexical items are involved.

### 5.5.2. Hybrid nouns

These are nouns which neither simply take the agreements of one consistent agreement pattern nor belong to two or more genders. A good example is the German word *Mädchen* 'girl'. It can take agreements exactly as a neuter noun. It may also take the feminine personal pronoun *sie*; but it cannot take the other agreements of the consistent agreement pattern associated with feminine nouns. Examples of this type typically arise when assignment rules are in conflict; in this case, nouns referring to females are usually feminine in German, but nouns formed with *-chen* are normally neuter. Such nouns must be lexically marked.

A more complex example is Russian *vrač* 'doctor', which was mentioned earlier. When *vrač* refers to a male doctor, then it takes the consistent agreement pattern associated with masculine animate nouns. When it refers to a female doctor it takes both masculine and feminine patterns as follows:

attributive modifiers	usually masculine, feminine possible
predicate	both possible
relative pronoun	normally feminine, masculine rare
personal pronoun	almost always feminine

The agreement required is variable for four different types of target. Thus *vrač* referring to a female is a 'hybrid' noun as it does not take consistently feminine agreements, nor consistently masculine agreements, nor both. Again this results from a conflict of the assignment rules, since it is female (and so should be feminine) yet its declensional class is such that it should be masculine. Taking *vrač* as a whole, it is a curious composite with one half being masculine, apparently half of a double gender noun while the other half is a hybrid noun.

There are several nouns like *vrač*. However, the frequency with which they take the different agreements varies. The Panov survey of nearly 4,000 speakers (see Corbett 1983: 32 for details and source) produced the data presented in Table 3.24.

TABLE 3.24. *Agreement with vrač and similar nouns in Russian*

	Attributive modifiers (% of informants favouring feminine agreement)	Predicate agreement (% of informants favouring feminine agreement)
<i>vrač</i> '(female) doctor'	16.9 (N = 3,835) <sup>a</sup>	51.7 (N = 3,806) <sup>a</sup>
<i>buxgalter</i> '(female) accountant'	25.5 (N = 3,835) <sup>a</sup>	— <sup>b</sup>
<i>upravdom</i> '(female) house manager'	— <sup>b</sup>	60.7 (N = 3,806) <sup>a</sup>

<sup>a</sup> N indicates the total number of informants.

<sup>b</sup> — indicates a combination not included in the survey.

It is evident that *vrač* and *buxgalter* do not take exactly the same agreements; no more, we must assume, do other nouns of the same type. This fact leads to an explosion in the number of agreement classes (Zaliznjak does not discuss nouns of this type). Naturally we do not wish to propose a separate gender for each noun; rather they must be lexically marked as masculine and feminine, with an indication as to the relative weight of the two features.<sup>15</sup> This solution puts the high degree of idiosyncratic variability into the lexicon, where it belongs. Nevertheless, there are generalizations to

<sup>15</sup> The position tentatively adopted here is that hybrids cannot form a full gender, but must be marked as lexical exceptions. An alternative would be to abstract away from the frequency of particular agreements and to allow a consistent agreement pattern to include alternatives for one or more targets. Since in the latter solution the nouns involved would still have to be lexically marked to allow for differences in frequency of the different forms of agreement, the first solution appears preferable.

There are more complex hybrids than Russian *vrač*. Serbo-Croat has a group of nouns like *gazda* 'master' which are consistently masculine in the singular, but can take masculine or feminine agreements in the plural. These nouns demonstrate that the notion 'consistent agreement pattern' cannot be restricted to establishing connections between different target positions but may also be required to establish connections within target positions. Thus masculine nouns in Serbo-Croat take one set of agreement forms and feminines another; hybrids like *gazda* take the same agreements as the masculines when singular, and the same as masculines or feminines when plural (provided the agreement target is an attributive modifier, a predicate, or a relative pronoun, but not a personal pronoun: data in Corbett 1983: 14–17; further examples in Corbett forthcoming).

be made concerning the behaviour of hybrid nouns: a set of such generalizations are grouped under the constraints of the Agreement Hierarchy (Corbett 1979, 1983: 8–41). It is also worth noting that hybrid nouns are one route for change in the gender system. (An alternative is for anomalous nouns to lead to a change in the semantics of the gender system, as has happened in Konkani, see Miranda 1975.)

## 6. Conclusion

We have seen that in establishing the number of genders in a particular language, Zaliznjak's approach, based on the notion of agreement class, is a useful starting-point. It is important to recognize, however, that this approach leads us towards controller genders; the other side of the coin is the system of different forms of the agreeing elements: we term these forms target genders.

In some languages there are numerous agreement classes: following Zaliznjak we do not recognize every agreement class as a gender. Some are subgenders—a notion we have made more precise. Other agreement classes are based on overdifferentiated targets and so are not accepted as genders; nor indeed are inquotate genders. Nouns belonging to the latter require a feature in their lexical entry, to mark their irregularity, as do defective nouns and nouns with double or multiple gender and hybrid nouns.

Thus, to find how many genders a language has, we begin with agreement classes but eliminate as many as possible. This minimalist position has been justified at each point. It is interesting to note that in many cases it leads us to traditional analyses; like Zaliznjak we have tried to improve the foundations of some shaky edifices.

