30. Number of Genders

Greville G. Corbett

In some languages gender is evident in almost every phrase, while in other languages it is absent. This contrast raises interesting questions as to its function. Equally, in some languages it is relatively easy to determine the type of gender system and to establish how many genders the language has, while in others it takes careful analysis.

1. Defining the values

The defining characteristic of gender is **agreement**: a language has a gender system only if we find different agreements ultimately dependent on nouns of different types. In other words, there must be evidence for gender outside the nouns themselves. These Russian sentences illustrate the point:

(1) Russian

a. *Žurnal ležal na stole.*
magazine lay.M on table
'The magazine lay on the table.'

b. *Kniga ležal-a na stole.*
book lay-F on table
'The book lay on the table.'

c. *Pis´mo ležal-o na stole.*
letter lay-N on table
'The letter lay on the table.'

We find *ležal* 'lay' in (1a) as opposed *to ležal-a* or *ležal-o* because of the gender of the noun *žurnal* 'magazine'. The difference in the form of the verb in our three examples results just from the type of noun; there are no other differences in number, case or syntactic structure. We are therefore dealing

with a gender system. If in (1a), instead of the noun *žurnal* ‘magazine’ we had the noun *mal´čik* ‘boy’, *djadja* ‘uncle’ or some other noun denoting a male, the same form of the verb would be used. Hence nouns like *žurnal* ‘magazine’, as indeed nouns like *mal´čik* ‘boy’ and *djadja* ‘uncle’, are said to belong to the masculine gender. Nouns like *sestra* ‘sister’ would take the agreement as in (1b), and so nouns like *kniga* ‘book’ and *sestra* ‘sister’ are said to belong to the feminine gender. That leaves nouns like *pis´mo* ‘letter’ as the members of the neuter gender.

To avoid confusion note that the mere existence of nouns like *djadja* ‘uncle’ and *sestra* ‘sister’, denoting males and females, is not enough to constitute a gender system. There must be syntactic evidence, in agreement. Kanuri (Nilo-Saharan; Nigeria) does not have a gender system, but does have lexical contrasts such as *tádà* ‘boy, son’ versus *férò* ‘girl, daughter’ (Hutchison 1981: 11, 38, 45). This is a matter of lexical semantics, and not a gender system. Such lexical oppositions may be instantiated through derivational morphology (English shows examples like *poet* and *poetess*). Again this of itself does not give a gender system. The reason is that there can be numerous similar oppositions, concrete versus abstract for example, none of which would be counted as grounds for postulating a grammatical category in the language in question. Similarly, inflectional markers on the nouns themselves are insufficient to ground a gender system: in our Russian examples, the inflection *-a* might seem to indicate feminine gender, but in fact *djadja* ‘uncle’ and *papa* ‘daddy’ are masculine, which we prove by the agreements they take, irrespective of their form. We shall limit ourselves to true gender systems which can be demonstrated on the basis of agreement evidence.

Our examples have involved agreement of the verb, but there are various other targets which may agree in gender, such as adjectives, determiners, numerals and even focus particles.

Most scholars working on agreement include the control of anaphoric pronouns by their antecedent (*the girl* ... *she*) as part of agreement. If this is accepted, as we do here, then languages in which free pronouns present the only evidence for gender will be counted as having a gender system. Of course, such languages with **pronominal gender systems** have a much less pervasive system than those like Russian. Including them, however, makes little difference to the overall picture, since they are rare (the best known example is English, which is typologically unusual in this respect; another is Defaka (Niger-Congo; Niger Delta, Nigeria; Jenewari 1983: 103-106)).

Since agreement is the defining characteristic, gender can be distinguished from other classification systems, such as **classifiers**. Classifiers are of various types, a common one being numeral classifiers analogous to *head* in the English expression *eighty head of cattle*. Distinguishing gender from classifiers is justified in Dixon (1982) and Corbett (1991: 136-137), and examples can be found in Aikhenvald (2000). It is possible, if unusual, for a language to have both gender and classifiers. One such language is Tariana (North Arawakan; Brazil; Aikhenvald 1994), which has a gender system and three sub-types of classifier. Coexistence of gender and classifiers is found in other languages of our sample: in Retuarã (Tucanoan; Columbia; Strom 1992: 10-11, 34-36, 45-47) and in Tidore (West Papuan; Halmahera, Indonesia; van Staden 2000: 77-81). Ngan'gityemerri (Daly; northern Australia) shows the development from generic classifiers into genders (Reid 1997). A further consequence of the definition is that differences in use of language which depend on the sex of the speaker (lexical choice, voice quality and so on) are not treated here; an example is the difference between men's and women's pronunciation in Chukchi (Dunn 2000).

We should note that often there is no substantive difference between what are called "genders" and what are called "noun classes"; the different terms may be merely the

products of different linguistic traditions. Thus we find systems with three genders, to which nouns are assigned by similar rules, in both Kannada (Dravidian; India) and Godoberi (Nakh-Daghestanian; eastern Caucasus). By tradition the first is said to have three genders, and the second three noun classes. We shall treat both as having gender.

We have established our criteria for deciding whether a language has a gender system. It is equally important to be clear on definitions when we ask how many genders particular languages have. Our approach starts from Zaliznjak (1964). Basically, two nouns are in the same gender provided that, however we change the environment (treating both the same), then both will take the same agreements. Again traditions vary. The earlier Bantuist tradition treated nouns as being in different noun classes when singular and plural; we consider the total behaviour of a noun, including both its singular and its plural, with the result that a typical Bantu language may have 7–10 genders rather than around 20 noun classes. More generally, while in many languages there is no dispute as to the number of genders, there are a few where the question is far from straightforward. The analytical problem of determining the number of genders and the tests for deciding the gender of a given noun depend on separating out the classes into which nouns are divided (the controller genders) from the number of different genders marked on agreement targets (the target genders). Frequently the two match up, but in several languages they do not. A full treatment of the subject with extensive references can be found in Corbett (1991). Based on the analysis there, the analytical decisions made for this chapter are that the number of genders given on the map is the core system, the number of controller genders. (Hence neutral genders, locative genders, subgenders, overdifferentiated targets, inquorate genders, hybrid nouns and those with double or multiple gender are all left out of account here; the interesting detail can be found in Corbett 1991: 145–188.)

The values for the number of genders are as follows:

@	1.	None	144
@	2.	Two	50
@	3.	Three	26
@	4.	Four	12
@	5.	Five or more	24
		total	256

As the figures show, the map is based on 256 languages, of which somewhat over half (144) have no gender system. A minimal gender system requires two genders, and such systems are common, with 50 examples in this sample. Three genders is around half as common (26 examples) and four genders, about half as common again (12). Larger systems, with five or more genders, are treated together, and represent a substantial minority (24 languages in the sample).

2. Geographical distribution

We look first at the distribution of gender languages versus non-gender languages. The two largest families differ in this regard. Niger-Congo has almost 1500 languages, a large proportion of which have gender systems (of interesting types, as we shall see in chapter 31). This family is found in western, central and southern Africa. The sample has 23 Niger-Congo languages, including Chichewa, Kisi and Swahili. Just three in our sample have lost gender, following radical simplification of the morphological system; these are Ewe, Igbo and Yoruba. The other very large family is Austronesian, with some 1250 languages spread across the Pacific, including Rapanui (the language of Easter Island), Indonesian and Maori. 21 Austronesian languages figure in the sample, none with a gender system, apart from the curious exception of Tagalog, which has partially borrowed the Spanish system (Schachter and

Otanes 1972: 197–198). Elsewhere in Austronesian, gender has arisen in Teop (North-West Solomonic, Oceanic; Papua New Guinea; not in our sample) by three spatial demonstratives forming a system of gender-distinguishing articles (Mosel and Spriggs 2000).

While Niger-Congo is the major source, other families of Africa are gender "hotbeds" in Nichols' terms (1992: 130–132). Khoisan languages in the south (represented by Ju|'hoan in our sample) have interesting gender systems (see Güldemann 2000), and Afroasiatic in the north contributes many gender languages (including Amharic, Modern Standard Arabic, Hausa, Hebrew and Qafar). Nilo-Saharan is mixed, but none of the five languages in the sample shows gender. To the north, Europe is a predominantly gender area, with Indo-European languages such as French, German and Russian. In the Caucasus, the Nakh-Daghestanian family is a stronghold of gender. Indo-European extends to the South Asian sub-continent, with gender languages such as Hindi and Marathi. In southern India, Dravidian languages typically show gender (examples in our sample are Kannada, Tamil and Kolami). Elsewhere gender is less strong. Austroasiatic presents a mixed picture, with gender in Khasi and Nicobarese, but not in Khmer or Vietnamese. In New Guinea, several families show gender, and of widely differing types. In Australia there are several gender languages, such as Maung and Bininj Gun-Wok, mainly clustered in the north. In the Americas, gender languages form a minority. In North America, there are Algonquian languages like Plains Cree and Eastern Ojibwa (discussed in chapter 31). There are a few gender languages in Central America, such as Chalcatongo Mixtec and Lealao Chinantec (both Oto-Manguean; Mexico). In South America the picture is mixed, with several of the families of Amazonia including gender languages. From the opposite perspective, the main areas *without* gender are the Pacific, most of Asia (notably the Sino-Tibetan family), including Siberia (notably the families grouped under Uralic and Altaic), together

with several families of North America, especially those to the west.

Our discussion has been in terms of languages; if we think of numbers of speakers, then the major gender family would be Indo-European, which includes English, French, German, Hindi, Portuguese, Russian, Spanish, each with many millions of speakers. On the other, non-gender side would be Sino-Tibetan, in which Mandarin and Cantonese alone account for a substantial proportion of the world's population. Again, then, there is a substantial split. Many speak gender languages and many do not.

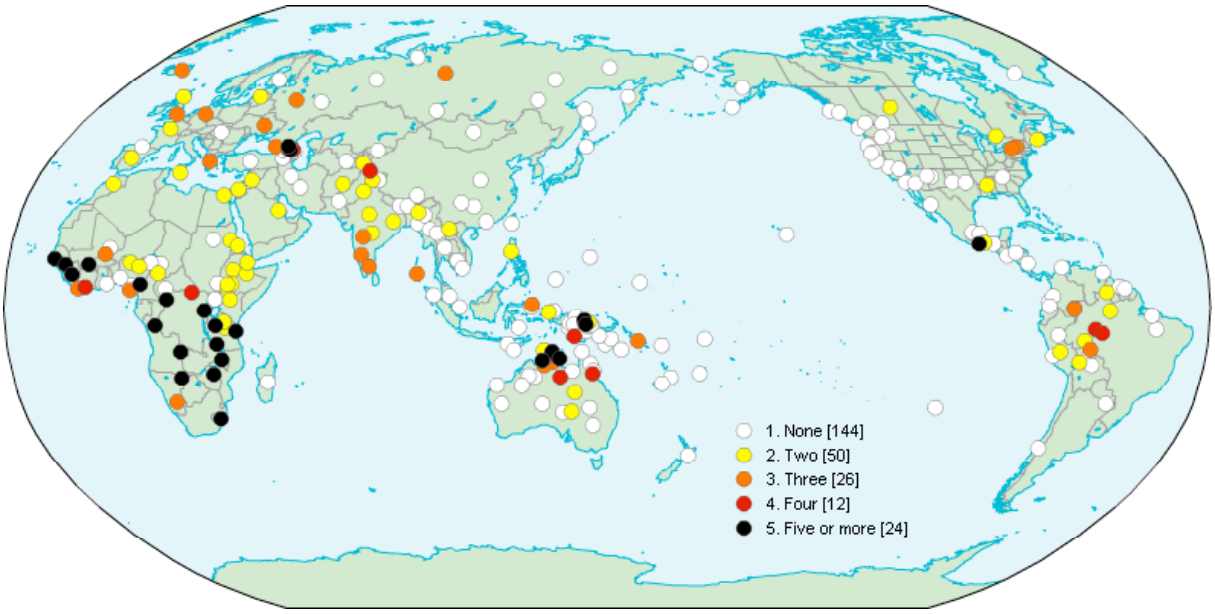
Given a gender system, the most common number of genders is two. Such languages are found in most gender areas. Afro-Asiatic systems regularly have two genders, though there are complex interactions with number (see Corbett and Hayward 1987). In Indo-European many languages retain three genders (like Icelandic and German), while many others have reduced to two (like French and Spanish); a minority have lost gender altogether (e.g. Eastern Armenian). Four-gender systems are particularly prevalent in Nakh-Daghestanian languages (our sample includes Archi, Lak and Tsez), though they occur elsewhere too, as in the isolate Burushaski. For larger systems the major source is Niger-Congo, where systems in excess of five genders are common. Nigerian Fula is exceptional, having around twenty genders, depending on the dialect (Arnott 1967; 1970: 67-75; Koval´ 1979; Breedveld 1995: 295-460). Other large systems can be found in Papua New Guinea, where Arapesh has 13 genders (Fortune 1942; Aronoff 1994: 89-114; Fraser and Corbett 1997). In northern Australia, too, Ngan'gityemerri arguably has 15 genders (Reid 1997).

3. Theoretical issues

One substantial issue, noted earlier, is the function of gender systems, given that they pervade some languages, are of some

importance in others, and are missing in yet others. This and related issues are discussed in Corbett (1991). Besides that source, the questions of the way in which gender systems develop, and the way in which they are represented, are also treated briefly in chapters 31 and 32.

Acknowledgement: This research was funded in part by the ESRC (UK) under grants R000271235 and RES051270122. This support is gratefully acknowledged.


1. None [144]
2. Two [50]
3. Three [26]
4. Four [12]
5. Five or more [24]