32. Systems of Gender Assignment

Greville G. Corbett

In chapters 30 and 31 we have seen how we as linguists analyze gender systems, and establish how many genders there are. The remaining question is how the *speaker* assigns nouns to those genders. In other words, if a speaker of Russian uses the word *kniga* 'book' or *djadja* 'uncle', how does he or she "know" the gender? Clearly speakers must know the gender in order to be able to make the appropriate agreements. A model of the way in which speakers allot nouns to genders is called a *gender assignment system.*

1. Defining the values

Assignment may depend on two sorts of information about the noun: its meaning and its form. We start with what we shall call **strict semantic systems.** In some languages the meaning of a noun is sufficient to determine its gender, for all or almost all nouns. This type is found in Dravidian languages like Kannada (Karnataka, southern India; Sridhar 1990: 198). In Kannada, nouns denoting male humans are masculine, and those denoting female humans are feminine. There are also deities, demons and heavenly bodies in these genders. All remaining nouns, including those denoting infants and animals, are neuter. Thus *appa* 'father', and *candra* 'moon' are masculine, *amma* 'mother' is feminine, and *na:yi* 'dog' is neuter.

Many languages have semantic assignment rules which do not cover the noun inventory as completely as do the rules of Kannada. We shall call these **predominantly semantic assignment systems**. An example is found in Bininj Gun-Wok, which was introduced at the beginning of chapter 31. The semantic categories found in each gender are given in Table 1.

Table 1. The semantics of gender in Bininj Gun-Wok

Masculine	**Vegetable**
• Male higher animates • Overall default for animates • Some lower animates • Rain • Compass points • Some items used in painting • Trade items, esp. Macassan and European • Some types of honey	• Plants and their products, including life-form terms • Sexual and excretory body parts • Song, ceremony, and custom • Fire (both bush and domestic) • Food, vegetable and otherwise • Some types of honey • Boats, planes and cars • [Drink, water, well] • [Camp nexus] • [Landscape features with water associations]
Feminine	**Neuter**
• Female higher animates • Some lower animates • Sun	• Most parts of animals and plants • Some parts of the landscape • Weather and sea • Time measures • Languages and speech • Country; place-based social categories

This table is from Evans et al. (2002), and full details of gender assignment in Bininj Gun-Wok can be found there; the items in square brackets represent categories which have moved into the vegetable gender from the neuter in the speech of younger speakers. The important points for our typology are that the semantic assignment rules are considerably more complex than those of Kannada, and yet the coverage is less good. For example, lower animates are split between the masculine and feminine genders, and it is hard to be more specific; nouns

denoting reptiles, birds, fish are found in both categories. There may well be principles of categorization here of which we are still unaware, but it seems likely that for at least some nouns there is no longer a principle for assignment which is still "live" for current speakers.

The genders of Bininj Gun-Wok have a semantic core, like those of Kannada. However, the rules of Bininj Gun-Wok are more complex, and still leave more nouns unaccounted for. This typological distinction applies equally well to languages where the dominant semantic principle involves animacy rather than sex. We noted in chapter 31 how Eastern Ojibwa (Algonquian; Ontario, Canada) assigns nouns to gender according to animacy, but how some nouns do not fall readily under the rule. Thus in languages with semantic assignment systems, the meaning of the noun determines gender. In the strict assignment systems, the rules are obvious and cover (virtually) the entire noun inventory. In the predominantly semantic systems, there is a minority of exceptions; these exceptions have been claimed to be largely only apparent in some languages, once the cultural setting of the language is taken into account. When we ask which are the semantic criteria on which semantic systems can be based, we see recurring patterns and occasional surprises (Corbett 1991: 30-32). For the present map we treat strict semantic and predominantly semantic systems together.

In many languages, however, assignment by semantic rules would leave many nouns without an assignment to a gender. In languages like Kannada, the nouns not assigned by the semantic rules (the "remainder" or "semantic residue") all belong to a single gender. In the languages we consider next, these residue nouns are distributed over more than one gender. Here we find additional rules for assigning nouns to genders according to their form. There is a significant asymmetry: languages may base their assignment system on semantic rules, or on semantic *and* formal rules, but not just on formal rules. Formal assignment rules may in turn access two types of

information: phonological and morphological. There may be combinations of such rules. We shall take a clear instance of each, considering languages from the sample.

A good example of assignment depending on phonological information is provided by Qafar (East Cushitic; north-eastern Ethiopia and Djibouti; Parker and Hayward 1985). In Qafar the semantic assignment rules are fairly standard, namely, for sex-differentiable nouns, those denoting males are masculine and those denoting females are feminine. It is the nouns which fall outside these semantic rules, the residue, which are of interest. For them there are the following phonological assignment rules: nouns whose citation form ends in an accented vowel are feminine (for example, *karmà* 'autumn'), while all others are masculine (for example, *gilàl* 'winter', which does not end in a vowel, and *tàmu* 'taste', which does end in a vowel, but not an accented one). These rules operate with few exceptions. Moreover, nouns denoting males and females typically accord with them too (for example, *bàqla* 'husband' and *barrà* 'woman, wife'). It might seem that we could dispense with semantic rules for Qafar. However, while the phonological rules give the right result in almost all cases, there are some nouns which show the role of the semantic rules. We find *abbà* 'father', which is masculine, even though it ends in an accented vowel. Conversely, *gabbixeèra* 'slender-waisted female' is feminine, though the accent is non-final. In such cases of conflict, the semantic rules take precedence (as is the normal situation in gender assignment systems). Qafar has remarkably simple phonological assignment rules, which assign semantically heterogeneous nouns to the appropriate gender by reference to their form.

The second type of formal assignment rule accesses morphological information. Here Russian is a good example. Once again for sex-differentiables, nouns denoting males are masculine and those denoting females are feminine. But unlike the situation in languages like Kannada, the residue is shared

between the three genders, with the neuter gender not even receiving the majority. We might think that further semantic rules would be sufficient, but this turns out to be at best highly unlikely; see the data in Table 2, where the nouns in each row are semantically similar yet belong to three different genders.

Table 2. Russian nouns belonging to the semantic residue

masculine	feminine	neuter
žurnal 'magazine'	*gazeta* 'newspaper'	*pis´mo* 'letter'
stul 'chair'	*taburetka* 'stool'	*kreslo* 'armchair'
dom 'house'	*izba* 'hut'	*zdanie* 'building'
čaj 'tea'	*voda* 'water'	*vino* 'wine'
ogon´ 'fire'	*peč´* 'stove'	*plamja* 'flame'
dub 'oak'	*bereza* 'birch'	*derevo* 'tree'
avtomobil´ 'car'	*mašina* 'car'	*taksi* 'taxi'
flag 'flag'	*èmblema* 'emblem'	*znamja* 'banner'
glaz 'eye'	*ščeka* 'cheek'	*uxo* 'ear'
lokot´ 'elbow'	*lodyška* 'ankle'	*koleno* 'knee'
nerv 'nerve'	*kost´* 'bone'	*serdce* 'heart'
večer 'evening'	*noč´* 'night'	*utro* 'morning'
čas 'hour'	*minuta* 'minute'	*vremja* 'time'

Thus the nouns of the semantic residue are scattered across the three genders in Russian. This situation is presented in table 3.

Table 3. Gender assignment in Russian

masculine	feminine	neuter
Sex differentiables denoting males PLUS part of semantic residue	Sex differentiables denoting females PLUS part of semantic residue	Part of semantic residue

In order to see how the remaining nouns are assigned, rather than looking at their meaning we should look instead at their morphology. There are four main inflectional classes in Russian,

each with several thousands of nouns (for justification of this view see Corbett 1982: 202–211). There are six cases and two numbers (though no paradigm has twelve distinct forms because of various syncretisms). We give just the singular forms in Table 4.

Table 4. Inflectional classes in Russian

	I	II	III	IV
Nominative	*žurnal*	*gazeta*	*kost´*	*pis´mo*
Accusative	*žurnal*	*gazetu*	*kost´*	*pis´mo*
Genitive	*žurnala*	*gazety*	*kosti*	*pis´ma*
Dative	*žurnalu*	*gazete*	*kosti*	*pis´mu*
Instrumental	*žurnalom*	*gazetoj*	*kost´ju*	*pis´mom*
Locative	*žurnale*	*gazete*	*kosti*	*pis´me*
gloss	'magazine'	'newspaper'	'bone'	'letter'

Given information about the inflectional class of nouns, the assignment rules are straightforward. Nouns in class I are masculine, those in classes II and III are feminine, and those in IV are neuter. (Further rules are required in Russian for indeclinable nouns, like *taksi* 'taxi', which is neuter; however, indeclinability is itself a morphological property.) In view of the coverage of these rules, we might be tempted to think that we could dispense with the semantic assignment, since *mal´čik* 'boy' is in class I, while *sestra* 'sister' is in class II, and *mat´* 'mother' is in class III. In other words, many of the sex-differentiable nouns would be assigned to the appropriate gender by the morphological assignment rules. But there are also instances where this is not so, for instance, *djadja* 'uncle', which denotes a male but is in class II, whose nouns are typically feminine. *Djadja* 'uncle' is masculine. Nouns like this confirm, once again, that we do not find languages where formal assignment rules are sufficient.

Of course, there are languages where the rules are more complex than these, but for languages where careful research

has been undertaken, gender is always predictable from a set of assignment rules, for at least 85% of the noun inventory and usually for a substantially larger proportion than that. For more details on these assignment systems see Corbett (1991: 7–69). For the purposes of the map we shall treat phonological and morphological assignment together. Thus the division will be between languages which have semantic assignment rules on the one hand, and those which have both semantic and formal assignment rules on the other. The values are as follows:

@	1.	No gender system		144
@	2.	Semantic assignment		53
@	3.	Semantic and formal assignment		59
			total	256

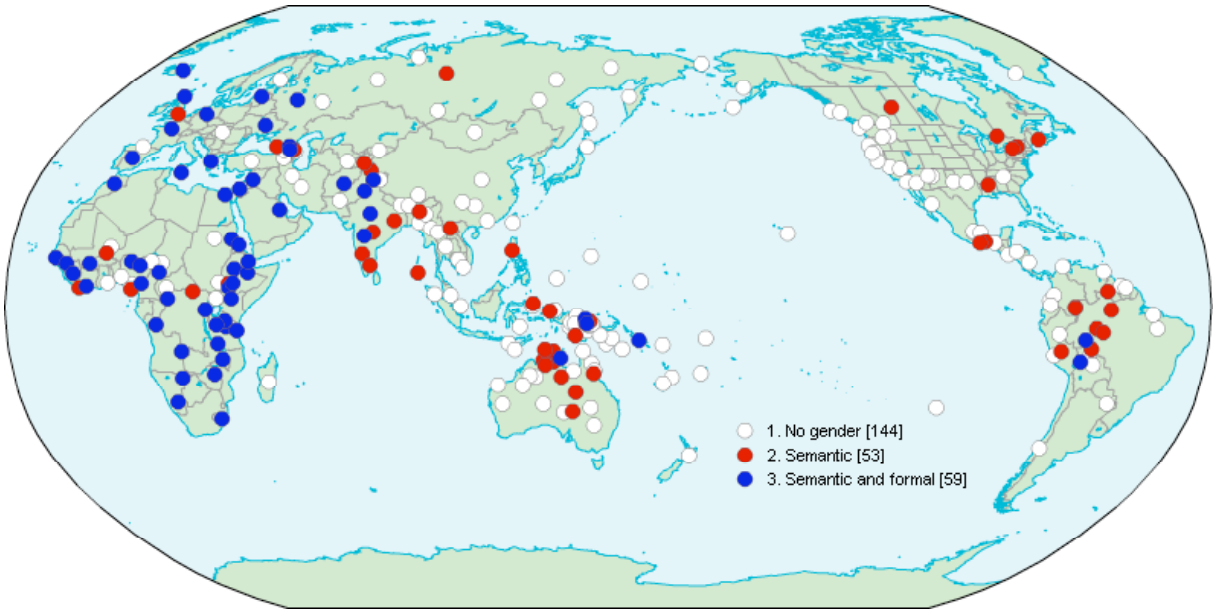
2. Geographical distribution

Of the languages in our sample with gender systems, there is a roughly even split between the two types of assignment system: 53 have semantic assignment while 59 have semantic and formal assignment. The distribution is interesting. Semantic and formal assignment is found mainly in Eurasia and Africa, in the Indo-European, Afroasiatic and Niger-Congo families. The convincing accounts of the rise of gender systems provide paths leading to systems with semantic assignment. And there are ways in which such systems may further develop into systems with semantic and formal assignment. Note that there is no necessity for this development to occur: the Dravidian systems have remained as semantic systems for a substantial period. However, one might reasonably expect that if a system is of the semantic and formal assignment type, this is likely to indicate an “old” gender system, since there must have been sufficient time for the system to develop from an earlier semantic assignment system. By contrast, if the system is of the predominantly semantic type we can make no prediction, as

such systems can arise at any time. It would follow that “old” gender systems are found in Eurasia and Africa.

3. Theoretical issues

There are several theoretical issues which arise. Given the typology of assignment systems, it is natural to ask what other features might be correlated with the assignment system. There have been several instances of modelling assignment systems, particularly those of the more difficult types, where it is helpful to be able to demonstrate that the proposed system does indeed account for the vast majority of the nouns in the lexicon (see Fraser and Corbett (1995) on Russian, Fraser and Corbett (1997) on Arapesh, both discussed in Corbett and Fraser (2000), and Evans et al. (2002) on Bininj Gun-Wok). Since it has been shown that gender is always largely predictable, this raises an interesting issue for lexicologists: what is the status of a lexical feature which is predictable? Psycholinguists too are beginning to tackle the issue of the place of gender in lexical entries. There has been some interesting work on how children acquire gender systems, for example, Mills (1986) and Müller (2000). Such studies may also help us to see how such systems change over time, as shown by the work of Polinsky and Jackson (1999), on Tsez; see also Comrie and Polinsky (1998); for development of the work on modelling change in assignment systems see Polinsky and van Everbroeck (2003). There are interestingly different gender systems to investigate, and so it is important that we are careful about definitions, in order to ensure that our comparisons are valid.


1. No gender [144]
2. Semantic [53]
3. Semantic and formal [59]