Greville G. Corbett

Gender in Russian: An Account of Gender Specification and Its Relationship to Declension

Introduction

A native speaker of Russian 'knows' the gender of many thousands of nouns. While the literature on gender in Russian is considerable — as our bibliography shows — the basic question of how it is that Russians use gender correctly has been largely ignored. It is this question we shall try to answer. One hypothesis would be that a native speaker remembers the gender of each item individually; that is to say, each noun is specified for gender in his internalized lexicon. The opposite hypothesis would be that gender is never remembered — it can always be derived from other information about a noun (such as its meaning or phonological form); in other words, no noun is specified for gender in the lexicon. Between these extremes there are various hypotheses according to which the gender of many nouns would be derived by some sort of procedure, while that of various exceptions would be specified in the lexicon. We shall see that gender and morphological class are strongly interrelated, but we shall claim that the few investigators who have considered this relationship in any detail have postulated a type of dependency which creates more problems than it solves.

1. Scope of the Study

Our aim is to discover how it is that Russian native speakers correctly assign nouns to gender classes. Therefore we shall focus on the regular nouns, which form the vast majority, though of course exceptions may be invaluable in indicating the preferable analysis when alternatives are available.

When referring to gender, we mean the three genders — masculine, feminine and neuter — traditionally recognised in Russian grammar. Various writers have tried to give a formal definition for gender. Trager's attempt (1940) was rather disappointing. More recent accounts include those of Marcus (1963), Zaliznjak (1964), Karpinskaja (1964, 1966); these are considered in Revzin (1967, 158–162; 1977, 145–51). Further contributions have been made by Gladkij (1969, 1973a, b). According to some definitions, animacy must be taken into account, giving six rather than three genders.
Animacy will not be considered in the present study because, as we have shown elsewhere (Corbett, 1980b), the animacy of a noun is predictable, given its meaning and its gender. The significance of the few exceptions and the way animacy can be handled in a grammar have also been discussed (Corbett, 1981b).

Returning to the three genders in question, we find that definitions depend on agreement phenomena. Thus a masculine noun is one which occurs in a particular set of syntactic frames. There are four such frames in Russian:

first, as the head noun of a simple noun phrase in the nominative case which also contains an attributive adjective ending in -j/y:

(1) interesnyj junoša
   (an) interesting youth

second, as the subject of a past tense verb ending in -o:

(2) junoša stojal
   (the) youth stood

third, as the antecedent of a relative pronoun ending in -j/y:

(3) junoša, kotoryj stojal...
   (the) youth, who stood...

fourth, as the antecedent of the personal pronoun on:

(4) junoša dolgo stojal, potom on...
   (the) youth a long time stood, then he...

In most instances, these four tests will produce the same result. There are two main classes of exceptions. The first involves nouns like vrač ‘doctor’. When used of a man there is no problem – all agreements are masculine. When such nouns refer to a woman, however, both masculine and feminine agreeing forms are found: attributive modifiers are usually masculine (though the feminine is possible), and both forms are found in the predicate, while the relative and personal pronouns are usually feminine (Corbett, 1979a, 14–16; 1979b). The second type of exception consists of nouns of common gender like plaksa ‘crybaby’. When they refer to a male, masculine and feminine forms occur: the feminine occurs most often in attributive position but it can also be found in the predicate, as examples given by Kopeliovič (1977) prove. This is an important article for the new data it provides. Another useful analysis is that of Iomdin (1980). The data available to date, in both types of exception, are consistent with the Agreement Hierarchy (Corbett, 1979b).

Both of these exceptional types will be excluded from our account. Common nouns of the plaksa type require further empirical study, and both types, it can be argued, are a problem of syntax rather than of gender assignment. Furthermore, the real issues of gender assignment depend on
the analysis of nouns like *djadjja 'uncle', which 'look like' feminines but are masculine, and of nouns like *nedelja 'week' and *metel' 'snowstorm'; both of these are feminine, and both have a stem ending in a soft consonant, but they decline differently. The facts here are undisputed and can be found in any reference work. But they have not been incorporated into an adequate theory. We therefore restrict our study to the cases which are uncontroversial.

Given, then, that in all instances under discussion the four tests described above yield the same result, three of them are redundant. Let us adopt the first test for convenience. Any noun which, when in the nominative case, can form a grammatical phrase when modified by a (semantically compatible) regular adjective ending in -yj is masculine; if the adjective must end in -aja the noun is feminine; neuter nouns are those for which the adjectives must end in -oe:

(5) interesnyj junoša (*interesnaja/*interesnoe)
    (an) interesting youth

(6) interesnaja kniga (*interesnyj/*interesnoe)
    (an) interesting book

(7) interesnoe leto (*interesnyj/*interesnaja)
    (an) interesting summer

From (5) and (6) it is clear that the ending of the noun itself is insufficient to determine the gender – an agreeing form is required. We might suppose that having established this test we have already completed our task. However, a moment's thought will establish that this is not so. To produce syntactically correct phrases such as (5)–(7) the native speaker must know the gender of the noun he is using. If he has nothing apart from a syntactic model, then this is equivalent to our null hypothesis – he has to remember the gender of each individual noun. Thus the test we have described will allow us to establish with certainty the gender of a given noun. The test depends on access to the intuition of native speakers or simply on observation of their linguistic behaviour, in speech or writing. Our object is to determine the mechanism by which they produce only the 'correct' variants in (5)–(7), in other words, how they know the gender of nouns.

2. THE LEXICON

According to the 'null-hypothesis' discussed above, a native speaker of Russian must remember the gender of every individual word. That is equivalent to saying that gender is recorded in the lexicon for every word. The lexicon is essentially a list of lexical items: the problem is to determine
exactly what information it includes. According to Chomsky (1965, 87): “In general, all properties of a formative that are essentially idiosyncratic will be specified in the lexicon”.

Let us begin with a clear case. No one would suggest that a Russian stores twelve forms for each regular noun; grammarians have always recognised the existence of regular classes and have assumed that it is sufficient to know to which class a noun belongs. We can be confident that the entry for the word razgovorčivost’ ‘talkativeness’ does not include the specific information that the accusative singular is razgovorčivost’, the genitive singular is razgovorčivosti, and so on. At most it would indicate that this noun belongs to declension III, or declines like, say, kost’ ‘bone’. However, even this information is redundant in that the declensional class is completely predictable from the suffix -ost’. As the noun is regular, no information on declension is required. In the case of a noun like doč’ ‘daughter’ it is necessary to include the information that in all forms except the nominative and accusative singular the stem is dočer’- and that it has an irregular instrumental plural dočer’mi; but there is no need to specify, say, the dative singular or nominative plural.

2.1. Minimal Specification in the Lexicon

The very least that the lexicon must specify for a lexical term is its phonological form and its meaning. With a few exceptions the combination is arbitrary; that is to say, given the form there is no way to determine the meaning, nor is the meaning sufficient to predict the form of the word. Exceptions are of two sorts. First there are onomatopoeic words. Second, words which are morphologically complex may allow their meaning to be determined as the sum of the parts. Even in this case, the meanings of the parts are still arbitrary. Furthermore, if the derivation is completely regular and predictable, it is at least arguable that such an item should not be listed in the lexicon.

For items that are listed in the lexicon, what form does their entry take? For the meaning, let us assume that this is recorded in terms of a set of binary semantic features (± concrete, ± human, ± male etc.). While this view is not uncontroversial, it is fairly widespread and is compatible with the account of gender we develop below. However, the formulation of the semantic entries in the lexicon is of great importance for our analysis.

On the other hand, the way in which the phonological form is entered in the lexicon is more critical. While the theoretical view of phonology adopted is not vital, the form to be entered in the lexicon is of significance. It is generally assumed that the stem is stored (e.g. Chomsky and Halle,
1968, 373); thus for the word komnata ‘room’ the lexicon will contain komnat-. Komnata is the nominative singular form, that is to say, one part of the paradigm.

The other decision we must take concerning underlying phonological forms is whether to consider softening part of the stem. In other words, does djadja ‘uncle’ consist of stem djad’- (soft) plus ending -a or of djad- plus ending -a. We shall adopt the former view, recognising underlying softening. It reduces the number of paradigms required and, for this reason, is the generally held view (there are dissenters, such as Lunt, 1975). The alternative we have rejected would complicate our account but would not require fundamental changes.

Given that both form and meaning must be specified in the lexicon, it is natural to ask whether gender could be derived from either of these. We might expect the meaning of a word to be sufficient. Indeed, mužčina ‘man’ is masculine while žensčina ‘woman’ is feminine. However, meaning does not help us when faced with the following data on the names of trees (Vinogradov, 1972, 58): vjaz ‘elm’, klen ‘maple’, jasen ‘ash’ and dub ‘oak’ are masculine, while lipa ‘lime’, osina ‘aspen’, bereza ‘birch’, sosna ‘pine’, iva ‘willow’, vetla ‘white willow’ and čeremuxa ‘bird-cherry’ are feminine. Yet the general word derevo ‘tree’ is neuter. (Van Schooneveld, 1977, attempts to analyse gender in semantic terms, but he gives few examples and is unable to account for data such as those given by Vinogradov.) The names of trees might lead us to believe that, if we stored the nominative singular form, any noun ending in a consonant would be masculine, any noun ending in -a, feminine, and in -o, neuter. But of course both mužčina ‘man’ and žensčina ‘woman’ end in -a in the nominative singular. There is, then, a ranking of factors involved and, as we shall see, there are further complications. It is therefore worth asking what other elements must or may be specified in the lexicon, as these will give us further potential sources for deriving gender.

2.2. Specification of Russian Nouns

We have seen that the lexical entry for a Russian noun must include its meaning and its phonological form. For syntactic, morphological and phonological rules to operate correctly it must be possible to obtain the following information about a given noun:

(a) case  (e) valency
(b) number  (f) declensional class
(c) gender  (g) position of stress
(d) animacy  (h) stress pattern.
We must consider which of these need to be included in the lexicon. Clearly case does not belong in the lexicon, for it depends on the slot the noun occupies in the syntactic construction of a particular sentence. Number is normally determined by semantic considerations, and will be part of a lexical entry only for those nouns which lack the singular or plural. Gender is the feature in question. As pointed out above, the animacy of a noun can be determined given the meaning, number and gender of a noun; there are a very few exceptions for which animacy must be specified in the lexicon (Corbett, 1980b, 1981b). Similarly, there are a few nouns whose valency is not predictable; for example, ljubov' ‘love’ requires the preposition k ‘to’ (ljubov' k otcu ‘love of the father’), while the corresponding verb ljubit’ ‘to love’ governs the accusative case (Comrie 1974, 40–41).

The main complication arises with declensional class. Sometimes this is derivable from a word’s stem structure. Given a noun’s gender, it is possible in many instances to predict its declensional class. But equally, given the declensional class, it is often possible to predict the gender. The relationship between gender and declensional class is the crux of the problem of gender in Russian, and so we shall devote the next section to the problem of declensional classes or paradigms. Before doing so, let us consider briefly the question of stress. While the place of the stress on a Russian noun can be predicted statistically to some extent (Papp, 1972), it must be specified in the lexicon in many cases. For most nouns this is all the information on stress which is required; there are, however, several patterns of shifting stress and the nouns affected must be so marked in the lexicon. As we shall see (§6.2.), this is of some relevance to the questions of gender and paradigm. First, however, we must consider the more basic problem of the nominal paradigms of Russian.

3. NOMINAL PARADIGMS

The reader not familiar with the literature will quite reasonably expect a straightforward account of the paradigms of Russian. However, there is no agreement as to the number of nominal paradigms in Russian. Tradition answers three, some writers claim four, and more recently it has been suggested that only two paradigms are required. Even the number of cases in Russian is in dispute; we shall retain the traditional six cases (for justification see Revzin, 1962, 108–110). There is no set procedure for determining the number of paradigms required given a set of data; there are certain vague intuitions and tradition plays a large part. We will first review the data, then consider the traditional account and go on to alternative suggestions.
3.1. Data

Any theory must account for forms like those given in the table below.

### TABLE I
Nominal forms (standard orthography)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>A</th>
<th>G</th>
<th>D</th>
<th>I</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>zakon</td>
<td>vino</td>
<td>škola</td>
<td>kost'</td>
<td>put'</td>
<td>vremja</td>
</tr>
<tr>
<td></td>
<td>zakon</td>
<td>vino</td>
<td>školu</td>
<td>kost'</td>
<td>put'</td>
<td>vremja</td>
</tr>
<tr>
<td></td>
<td>zakona</td>
<td>vina</td>
<td>školy</td>
<td>kosti</td>
<td>puti</td>
<td>vremeni</td>
</tr>
<tr>
<td></td>
<td>zakonu</td>
<td>vinu</td>
<td>škole</td>
<td>kosti</td>
<td>puti</td>
<td>vremeni</td>
</tr>
<tr>
<td></td>
<td>zakonom</td>
<td>vinom</td>
<td>školej</td>
<td>kost'ju</td>
<td>putem</td>
<td>vremenem</td>
</tr>
<tr>
<td></td>
<td>zakone</td>
<td>vine</td>
<td>škole</td>
<td>kosti</td>
<td>puti</td>
<td>vremeni</td>
</tr>
<tr>
<td></td>
<td>zakony</td>
<td>vina</td>
<td>školy</td>
<td>kosti</td>
<td>puti</td>
<td>vremena</td>
</tr>
<tr>
<td></td>
<td>zakony</td>
<td>vina</td>
<td>školy</td>
<td>kosti</td>
<td>puti</td>
<td>vremena</td>
</tr>
<tr>
<td></td>
<td>zkonov</td>
<td>vin</td>
<td>škoł</td>
<td>kostej</td>
<td>putej</td>
<td>vremen</td>
</tr>
<tr>
<td></td>
<td>zkonam</td>
<td>vinam</td>
<td>školami</td>
<td>kostjam</td>
<td>putjam</td>
<td>vremenam</td>
</tr>
<tr>
<td></td>
<td>zkonami</td>
<td>vinami</td>
<td>školami</td>
<td>kostjami</td>
<td>putjami</td>
<td>vremenami</td>
</tr>
<tr>
<td></td>
<td>zkonax</td>
<td>vinax</td>
<td>školax</td>
<td>kostjax</td>
<td>putjax</td>
<td>vremenax</td>
</tr>
<tr>
<td></td>
<td>'law'</td>
<td>'wine'</td>
<td>'school'</td>
<td>'bone'</td>
<td>'way'</td>
<td>'time'</td>
</tr>
</tbody>
</table>

**Abbreviations**: Sg = singular; Pl = plural; N = nominative; A = accusative; G = genitive; D = dative; I = instrumental; L = locative.

The standard orthography is largely morphophonemic (Stankiewicz, 1968a, 28–29). Note that palatalization is indicated by both ' and j.

In the following table, the same endings are given in morphophonemic transcription, together with variant forms shown by different nouns. The variants are explained in the notes.

### TABLE II
Nominal endings (morphophonemic transcription)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>A</th>
<th>G</th>
<th>D</th>
<th>I</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>φ</td>
<td>o</td>
<td>a</td>
<td>φ</td>
<td>φ</td>
<td>φ</td>
</tr>
<tr>
<td></td>
<td>N/G</td>
<td>o</td>
<td>u</td>
<td>φ</td>
<td>φ</td>
<td>φ</td>
</tr>
<tr>
<td></td>
<td>a [u]</td>
<td>a</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>u</td>
<td>e</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>om</td>
<td>om</td>
<td>oj: oju</td>
<td>ju</td>
<td>om</td>
<td>om</td>
</tr>
<tr>
<td></td>
<td>e [u]</td>
<td>e</td>
<td>e</td>
<td>i [i]</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>i (å)</td>
<td>a (i)</td>
<td>i</td>
<td>i</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>N/G</td>
<td>N/G</td>
<td>N/G</td>
<td>N/G</td>
<td>i</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>ov (ej, φ)</td>
<td>0 (ov, ej)</td>
<td>ej</td>
<td>ej</td>
<td>φ</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>am</td>
<td>am</td>
<td>am</td>
<td>am</td>
<td>am</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>am'i</td>
<td>am'i</td>
<td>am'i (m'j)</td>
<td>am'i</td>
<td>am'i</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>ax</td>
<td>ax</td>
<td>ax</td>
<td>ax</td>
<td>ax</td>
</tr>
</tbody>
</table>

**Abbreviations**: as in the previous table.

**Notes:**
(1) φ represents a null ending. Not all would accept the interpretation of the nominative of vremja as a null ending; Lightner (1967, 1187) considers this an example of the nasal ~ a alternations found elsewhere in Russian; in this case, the final a is clearly part of an alternating stem form. (This is the solution adopted by Coats, 1976, 55.)
(2) N/G indicates that the form of the accusative case will match that of the nominative for inanimate nouns, and that of the genitive for animates (all the examples spelled out in the previous table are inanimate). It has been shown elsewhere that we are dealing with an instance of syncretism which goes beyond the paradigm, and a formal solution has been proposed (Corbett 1980c, 1981b).

(3) [ ] indicates a variant available for certain nouns under certain syntactic conditions; e.g. the noun saxaru 'sugar' has the form saxaru in some quantified expressions. This form is in addition to the ordinary genitive form saxara. (Detailed information on the nouns involved in this instance, and in those which follow, is available in various sources, including Stankiewicz, 1968a; Švedova et al., 1980. The morphology of any particular noun can be established readily using Zaliznjak, 1977.)

(4) oj: oju, on the other hand, indicates a variant available to all nouns which show the endings of this column. Oj is the normal form, oju is now a literary form, found most commonly in poetry (see Ermlenko, 1976; Katlinskaja, 1977 and references there, for statistics on the usage of the two forms).

(5) ' indicates stress. It is marked on those forms which must carry the stress, independently of the stress pattern of the particular noun. Other endings may or may not be stressed.

(6) ( ) indicates a form which is used in place of the main listed form for certain nouns. For some nouns both forms may be used.

It can be seen that the oblique cases of the plural show little differentiation between the different types, and that elsewhere there is considerable overlap of forms.

3.2. The Three-Paradigm Approach

The traditional approach recognises three declensions. The title of first declension used to belong to the škola type, while zakon and vino made up the second declension. This scheme of numbering is preserved in Unbegaun (1957, 37–71), but most recent accounts reverse the numbering. Normally kost is assigned to the third declension. Thus we have:

decension I: zakon, vino

decension II: škola

decension III: kost

Unbegaun also assigns put and vremja to declension III. The Academy Grammar does not admit these two to the main declensions (but labels them raznosklojajme). Đurović (1964, 67–142) does likewise. In addition, the Academy Grammar recognizes three declensions in the singular, and only one declension in the plural (Vinogradov et al., 1952, 133). Isačenko (1962, 86–129) follows the Academy Grammar in excluding vremja from the main types. He combines it with various marginal types which he labels 'Heteroklitika'; together these form the fourth declension. In essence his solution is like that of the Academy Grammar though he includes put in declension III. Stankiewicz (1968a, 24–57) distinguishes three declensions in the singular only (one in the plural) and, like Unbegaun, admits put and vremja into declension III. In this respect he follows Trager (1940, 301–303), and Durnovo (1922, 250–255). Durnovo also treats singular and plural
forms separately, but has seven subclasses in the plural. The different positions described are summed up in Table III.

### Table III

<table>
<thead>
<tr>
<th>Author</th>
<th>Number of declensions</th>
<th>Declension to which each noun is assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>zakon</td>
</tr>
<tr>
<td>Durnovo (1922)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(sing., 7 in pl.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trager (1940)</td>
<td>3</td>
<td>IA</td>
</tr>
<tr>
<td>Academy Grammar</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(1952)</td>
<td>(sing. only)</td>
<td></td>
</tr>
<tr>
<td>Unbegaun (1957)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Isačenko (1962)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Đurović (1964)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Stankiewicz (1968a)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(sing. only)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is almost no discussion of why the different types should be assigned to paradigms in this particular way. Unbegaun writes (1957, 37): “Nouns may be divided into three declensions. This classification is based upon the various systems of case-endings”. This provides no basis to favour one solution rather than another. Trager makes the strong claim that his declensional classes are “determined entirely by formal criteria” (1940, 303). However, the technique he uses, ‘seriation’, hardly justifies his optimism. He defines it as follows:

Seriation establishes categories by the examination of the sets of related forms created by the modification of a base form by some element which changes it. The element of change may be something added to the base (before, after, within the base), or may be recognizable only by the fact that one of the component parts of the base is changed. The sets of forms or SERIES may be called paradigms. (1940, 301).

There is nothing in this definition to show why his criteria lead him to establish the traditional three paradigms, which he retains in a later paper (Trager, 1953).

Isačenko’s contribution is more helpful; he claims that the main criterion should be productivity:


Certainly, the zakon, vino, škola and kost’ types are all productive, though the vino type depends on the soft variant (in particular the verbal nouns in
-nie), and the kost' type almost exclusively on the suffix -ost'. It will be recalled that Isačenko's solution is zakon and vino – I; škola – II, kost' and put' – III, vremja – IV. The productivity criterion would justify the exclusion of vremja from the main paradigms, but some other criterion must be invoked to justify the inclusion of put' in declension III. The exclusion of put' and vremja (as in Vinogradov et al., 1952) can be justified, as they are not productive; the actual number of nouns involved is small, so that there would be no objection to labelling them all as exceptions. However, productivity provides no reason for combining the zakon and vino types into one paradigm. The reason seems to be purely historical – both go back to original o-stems. The tradition has been maintained for a considerable time with no argued justification.

If we compare the two declensional types zakon and vino, we find different forms in the nominative singular – a most significant distinction. Then there are alternative forms in the genitive and locative singular of some nouns of the zakon type, which never occur in any vino type noun. The apparent similarity of the nominative plurals is more apparent than real, for while many nouns of the zakon type take the -a form, this must always be stressed, which distinguishes it from the ending of the vino type. The three genitive plural endings are shared but, for the vast majority, -ov is the ending for the zakon type and -φ for the vino type. Thus, while there are definite similarities between the two types, if they are to be combined there must be a good reason for so doing.

If they are in some sense 'sufficiently similar' to be treated as a singular paradigm, then clearly so are kost', put' and vremja (as in Unbegaun's account). Different solutions depend on giving greater or less weight to productivity or to similarity, but this is not made explicit. Most accounts make a reference to "the close connexion which exists between declension and gender" (Unbegaun, 1957, 37), but this problem – our main concern – is not treated explicitly.

At this point it is worth recording the data relevant to the factors under discussion: productivity, the actual number of nouns (not a criterion invoked by those who take the traditional stance), and gender.

<table>
<thead>
<tr>
<th>Table IV</th>
<th>Data on the different types of noun</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>zakon</td>
</tr>
<tr>
<td>Productive</td>
<td>yes</td>
</tr>
<tr>
<td>Number of nouns</td>
<td>20,830</td>
</tr>
<tr>
<td>Gender</td>
<td>masc.</td>
</tr>
</tbody>
</table>
The bracketed forms show the less productive types. The numbers of nouns are derived from the Obratnyj slovar’ russkogo jazyka (Lazova, 1974, especially 942–943) and, for the first four types, are given to the nearest fifty. The gender correspondences have not been discussed so far. For the škola type, masculines are a minority (about 26) examples, but including some very common nouns).

The effect of combining the zakon type and the vino type nouns into one paradigm is to mix two genders. Such a move surely requires clear argumentation. We conclude that, while the traditional three-paradigm solution must be taken seriously, as it is widely accepted and has some highly distinguished scholars among its advocates, its justification has not been established.

3.3. The Two-Paradigm Approach

A more recent suggestion is outlined briefly by Stankiewicz:

The existence of two singular declensions... can be postulated for all inflected contemporary Slavic languages despite the differences in the degree of distinction between the two declensions and their gender affiliation. Thus we may subsume the so-called III declension of Russian substantives (with the exception of the single noun put') under the I declension, inasmuch as the distinctiveative, locative and instrumental singular endings of the feminine stems (i.e. the original -i stems) are predictable from their gender and those of the neuter stems from the morphological makeup of their stems... (1978, 666–667).

Though the criteria for setting up paradigms are not made explicit here, certain assumptions can be identified. The main one is that variation within a paradigm is acceptable if that variation is predictable. The second is that the gender of a noun can be taken as given. If the gender of a noun can be referred to in order to distinguish the forms like zakon from those like vino in declension I, then Stankiewicz is taking the principle to its logical conclusion in including nouns like kost’ (and vremja) in the same paradigm. Intuitively, however, his solution seems far-fetched. First because, though zakon and vino share several forms, they bear little resemblance to the forms of kost’. Furthermore, all nouns like kost’ have a stem ending in a soft consonant, while zakon and vino have both hard and soft variants (as indeed does škola).

A more attractive variant of the two paradigm solution is put forward by Zaliznjak (1967, 205–207), and this approach is adopted in the later Academy Grammar (Svedova, 1970, 367–377). Zaliznjak (1967, 14) specifically excludes justifying his paradigms from the aims of his book. (While the 1970 Academy Grammar includes some discussion of the nature of
paradigms—the two noun paradigms are seen as making up one full paradigm (Švedova, 1970, 367)—there is nothing to explain why the three paradigms proposed in the previous Academy Grammar have been replaced by two. The latest Academy Grammar goes back to three paradigms, as in 1952 except that put’ and vremja are now included in declension III (Švedova et al., 1980, 483). Again the change is not explained.) In Zaliznjak’s version of the two-paradigm solution the nouns are distributed as shown in Table V.

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Neuter</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declension I</td>
<td>zakon</td>
<td>vino</td>
</tr>
<tr>
<td>Declension II</td>
<td>put’</td>
<td>vremja</td>
</tr>
</tbody>
</table>

The advantages over Stankiewicz’s solution are as follows:

(a) škola at least shares the locative singular ending -e with both zakon and vino;

(b) put’ need not be excluded – though this is a minor point;

(c) the three types in declension I all have hard and soft stem variants, while there are no such variants in declension II types;

(d) all the major classes, in terms of the number of nouns, are in declension I;

(e) as a consequence of (d), one can imagine a solution involving markedness, whereby all unmarked nouns belong to declension I, all marked nouns to declension II;

(f) there is a pleasing balance in the system, each gender being represented in each declension.

This version of the two-paradigm solution is clearly preferable to that of Stankiewicz. However, its attractions are partly superficial. In particular, the balance depends on the existence of the single noun put’. Furthermore, to establish how any particular noun declines, it is necessary for both the declension type and the gender to be specified. Indeed it could be argued that in fact we have a six-paradigm solution rather than the claimed two-paradigm solution. Be that as it may, the major objections to Stankiewicz’s solution hold against this solution too. Completely dissimilar endings are artificially combined into a single paradigm (in this instance those of the škola type together with those of zakon and vino) and they are predictable not in phonological terms but only on the basis of gender.

While both solutions work, there is the feeling that by allowing such solutions we empty the term ‘paradigm’ of any content. A similar situation
arose over the definition of the phoneme. It was claimed that phonemes were made up of groups of allophones which were in complementary distribution. However, given that criterion as the only one, English [h] and [ŋ] could be treated as allophones of a single phoneme, as they are in complementary distribution. This solution works, but it was felt that it was an undesirable solution. A second criterion for phonemes was introduced – the sounds treated as allophones of a single phoneme must be phonetically similar (see Twaddell, 1935, for discussion). This criterion rules out the counter-intuitive suggestion for an English phoneme with completely dissimilar allophones. The two-paradigm analyses described above should be rejected on similar grounds. In those analyses the distribution of the different endings within a single paradigm can normally be predicted, providing a noun’s gender is given. But this solution too appears counterintuitive, because the different endings are not similar. We should therefore make similarity of forms a criterion for paradigms. This will rule out unnatural paradigms like those described above.

We have seen that there are objections to both versions of the two-paradigm solution. Nevertheless, if unlike endings are allowed in a single paradigm and are resolved on the basis of gender as in the three-paradigm solution, then it is logical to use this technique consistently and reduce the three paradigms to two. Thus the three-paradigm approach can be challenged on two counts: on the one hand it combines dissimilar forms, but on the other it does not do so consistently (as the two-paradigm approach does).

3.4. The Four-Paradigm Approach

This approach is suggested in Karcevskij (1932, 65–66). The argumentation is very brief, but nevertheless more clear than in many of the accounts discussed above. Starting from the nominative singular, Karcevskij divides nouns into two types: those with an ending and those with zero ending. Nouns with an ending are divided into those ending in -a (which he points out may be masculine or feminine) and those ending in -o (neuter). Those with zero ending are also divided into two groups: the first may end in a hard or soft consonant and are masculine, the second can only end in a soft consonant and are feminine. (He does not include the vremja type, nor put’.)

There are weaknesses in this account as it stands. First the distinction between -o and -a in the nominative is merely orthographical for many nouns. The effect of akan’e is to make the pronunciation of the final vowel of škola ‘school’ and delo ‘affair’ identical; only under stress is there a
difference. The two do indeed decline differently but it would be necessary to refer to another case form to show this. Similarly the reference to a type which may end in a hard or soft consonant is unsatisfactory. It is true that the zakon type has a soft variant (e.g. portfel’ 'briefcase') whereas the kost’ type can only be soft. However, given a noun ending in a soft consonant, it is again necessary to look at another case form to establish the paradigm for certain.

If reference to another form in the paradigm were included, then Karcevskij’s analysis would be well founded on the notion of similarity of forms. In his account, gender is not required as a criterion – the references to gender are simply attributes of the different paradigms, not defining characteristics. It appears to be a very reasonable position. Surprisingly, perhaps, Karcevskij later presented a traditional three-paradigm account (1948, 4–10; 1964, 168–173). When we survey the literature, we encounter the four-paradigm approach infrequently. Vaillant, considering Slavonic in general, talks of four types of nominal inflection, which match Karcevskij’s four paradigms (1958, 13). Mareš too, in his study of noun declensions in Slavonic, accepts the same four declensions for the East Slavonic languages (1967, 503).

In Jakobson’s account of Russian declension he refers to a masculine, feminine and neuter type of declension (1958, 162) and also to a secondary feminine type (nouns like kost’ and also to a secondary feminine type (nouns like kost’) and a secondary non-feminine type (put’ and vremja etc.) (1958, 166). We could claim that this is a five-declension system; it is, however, very similar to Karcevskij’s system, as Karcevskij did not include put’ and vremja (the nouns which make up Jakobson’s fifth category). Jakobson’s main concern is the extent and types of syncretism: he does not make an issue of the number of paradigms postulated. Nevertheless it is most interesting to note that his original approach to Russian declension leads away from the traditional three paradigms. To sum up: the four-paradigm analysis respects the condition of similarity of forms. It has relatively few champions, but these include some notable scholars.

3.5. The Number of Paradigms

We have seen that the number of noun paradigms to be postulated in a grammar of Russian is not at all straightforward. The answer will depend on which other factors are taken as given, in particular, whether the gender of a noun is assumed. That is to say, the question cannot be answered independently, but only as part of a more complete account of Russian grammar.

One of the prime criteria for judging an analysis is economy. If an
analysis could be found which eliminated a complete category (paradigm or gender) from the lexical specification of nouns, the economy achieved would make the number of paradigms recognised a matter of lesser importance. If it can be done, we would be strongly inclined to accept the number of paradigms required by such an account. We should consider, therefore, whether it is possible to produce an analysis in which the paradigm of a noun would be predictable, given the gender, or one in which the gender would be predictable, given the paradigm. The two possibilities are illustrated in Table VI.

<table>
<thead>
<tr>
<th>TABLE VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative accounts of lexical specification</td>
</tr>
</tbody>
</table>

- **Features specified in lexicon:**
  - Semantic features
  - Gender (m, f or n)
  - Phonological features

- **Rules to derive missing feature:**
  - Rules to derive paradigm

**Alternative A:**
- gender-to-declension

**Alternative B:**
- declension-to-gender

(In addition, some nouns have irregular features to indicate moveable stress, unpredictable valency, etc. These are required in both alternatives.)

In the following section we shall investigate Alternative A. It turns out that it cannot work, no matter which of the approaches described above is adopted – that is to say, no matter how many paradigms are assumed. In Section 5 we shall therefore consider Alternative B.

4. ATTEMPTS TO SPECIFY DECLENSIONAL CLASS IN TERMS OF GENDER

As mentioned earlier, it is often assumed that the gender of nouns should be specified, and the declensional class should be derived from the gender specification. Normally the position is not justified; however, two investigators, Thelin (1975) and Crockett (1976), have attempted an explicit account. We shall therefore consider their analyses in turn. Both assume the traditional three nominal paradigms but, as we shall see, the objections to their approach would apply equally strongly if they were to postulate two or four paradigms.
4.1. Thelin's Analysis (1975)

Thelin's account forms part of a more general work on morphology; he discusses the problem of inflectional class to illustrate the usefulness of markedness theory (1975, 30–33). His view of the relationship between the lexicon and the morphological component (1975, 21–25) is somewhat different to what we have assumed above. Thelin's lexicon is more abstract than ours, including only phonological representation, rule features (for derivational behaviour and information on accessional properties of the root), and semantic representation (1975, 21). Thus, given the root *bereg-* (the phonological representation need not concern us), the noun *bereg-* 'bank, shore, beach' is reached by a stem formation rule. The latter also introduces the feature [+ masculine].

In our account, following recent trends in theory, we have assumed a less abstract lexicon and a correspondingly diminished role for the morphological component; in this framework *bereg-* is listed in the lexicon as a noun. For the main point of debate this makes no difference: though our lexicon contains elements which in Thelin's account are derived by rule, once the items have the feature sets considered necessary in each account, the feature sets are fully comparable. Thelin argues that:

the inflectional class membership of Russian nouns can to a great extent be predicted from information on gender and/or the nature of the final consonant of the stem. (1975, 30).

The fact that the information on gender was inserted by a stem formation rule rather than being specified in the lexicon in his account is of secondary importance; for us the main point is that Thelin clearly believes gender should be specified, and declensional class derived from it.

Thelin outlines his solution as follows:

(1) *masculine stems*: generally assigned to the first declension (masculine paradigm); but certain words must be specially marked:

(a) words like *muččina* 'man', *sud'ja* 'judge', *djadja* 'uncle', *brodjaga* 'tramp', which are inflected according to the second declension;

(b) hypocoristics such as *Volodja*, *Sasha*, which are also inflected according to the second declension;

(c) diminutives and augmentatives like *klebuško* 'bread' (diminutive) and *domišče* 'house' (augmentative), which follow the neuter paradigm of the first declension;

(d) *put* 'way', which must be marked as following the third declension and as having the singular instrumental form according to the first declension.

(2) *Neuter stems*: assigned to the first declension (neuter); but nouns
like *vremja* ‘time’, *imja* ‘name’ belong to the third declension (neuter), which can “probably” be predicted from the combined information on gender and the phonotactic nature of the root.

(3) **Feminine stems.** This is the most interesting group for Thelin, who is concerned to show the usefulness of markedness. The feminine stems are distributed between two declensions, the second and the third. In some instances, the paradigm can be predicted from the stem, depending on whether the final consonant is hard or soft. Most consonants can be hard or soft, but *c, š, ž* are only hard, while *j, ě, šě (= šː*) are only soft. When the final stem consonant is the hard member of a hard-soft pair, then the noun follows the 2nd declension (e.g. *stran-a* ‘country’). This is also the case if the noun ends in *c* (ptic-a ‘bird’) or in -*j* (allej-a ‘avenue’). However, if the final stem consonant is the soft member of a pair (*t’, d’, n’, s’, z’*, etc.) or /š, ž, ě, šː/:, then the paradigm cannot be predicted. Thelin gives neat contrasting examples:

<table>
<thead>
<tr>
<th>2nd declension</th>
<th>3rd declension</th>
</tr>
</thead>
<tbody>
<tr>
<td>tělja (t’aː)</td>
<td>aunt</td>
</tr>
<tr>
<td>pesnja (p’n’aː)</td>
<td>‘song’</td>
</tr>
<tr>
<td>pulja (p’l’aː)</td>
<td>‘bullet’</td>
</tr>
<tr>
<td>gruśa</td>
<td>‘pear’</td>
</tr>
<tr>
<td>daća</td>
<td>‘dacha’</td>
</tr>
<tr>
<td>pišeća</td>
<td>‘food’</td>
</tr>
</tbody>
</table>

(In the last four the ′ is purely orthographical.) He demonstrates that it is more economical to mark those which belong to the third declension. Then all remaining feminine nouns will be treated as second declension.

Thelin does prove his main point, namely, the value of markedness theory. In his framework, markedness dictates that the third declension feminines be specified as such. However, in the more general context of gender and inflection specification there are two main objections to his account:

(1) his inflectional rules need access both to specification for gender and to specification for inflectional class. In view of his concern to eliminate redundancy, it is surprising that he does not consider whether one of these could be eliminated;

(2) he makes no mention of indeclinable nouns or of abbreviations. These are discussed by Crockett and so we shall turn to her analysis.
4.2. *Crockett’s Account* (1976)

Crockett is concerned with gender as an agreement category (1976, 1–137). She does not attempt to formalise the gender-declension relation in the way that Thelin does, but the main strength of her account is the accumulation of data from various sources. The extra data, not discussed in Thelin, raise serious problems for deriving declensional class specification from gender. Nevertheless, Crockett explicitly maintains the position of deriving declensional specification from gender:

...though the inherent gender of inanimate nouns may occasionally be predicted from the structure of their stems, it is essentially an unpredictable property and entries for inanimate nouns in the lexicon of Russian (in a model of linguistic competence) must include gender specifications.

Declensional paradigms, on the other hand, need not be specified for most inanimate nouns in the lexicon, for they are predictable from the gender specifications: masculine nouns follow the masculine variant of the First Declension, neuter nouns follow the neuter variant of the First Declension, and feminine nouns follow the Second Declension. (1976, 12).

As Crockett herself points out, nouns like škola and kost’ are both feminine but decline according to different paradigms. As in Thelin’s account, all the nouns of the kost’ type must be marked as such in the lexicon. The *Obratniy Slovar’* (Lazova, 1974, 942) indicates that there are over 5,100 such nouns. However, about 4,300 of these are formed with the productive suffix -ost’ ‘-ness’, which always declines according to the third declension. If we assume that these are not listed in the lexicon, then only about 800 words would bear the exceptional marker. It may be possible to reduce this figure slightly by reference to stem types. (Note that analyses like that of Ermolenko (1969) do not help us here. Ermolenko gives an interesting set of rules to determine the gender of nouns spelt with a soft sign. He therefore starts from the nominative singular and so does not have the problem of distinguishing third declension from second declension nouns.) Even if it proves possible to reduce the number of exceptional markers required, we are still dealing not with a few exceptions, but with a sizeable group. Crockett therefore requires both gender specification and specification for inflectional class in the lexicon.

As regards animate nouns, specifically those which are sex-differentiating, Crockett proposes that their gender be determined from the semantic feature of sex, which must in any case be specified in the lexicon. Nouns denoting males will be masculine, and so follow the first declension; those denoting females will be feminine and so follow the second declension (Crockett, 1976, 49–50). The main type of exception to this formulation involves nouns like djadja ‘uncle’ and mużčina ‘man’. Such nouns denote males and require masculine agreements in all syntactic positions. Never-
theless they follow the second declension, and this information must be specified in the lexicon (Crockett, 1976, 51). There is a large number of such nouns (over 250; Lazova, 1974, 942), and the problem is much worse than dictionaries reveal, for there are many hypocoristics such as Saša and Šura which decline in the same way. There are two other exceptional types: nouns of common gender such as plaksa 'crybaby' and nouns like vrač 'doctor' referring to a woman. In both cases Crockett proposes that their declensional type should be specified in the lexicon. Even given this exceptional feature she requires ad hoc switching of features to reach the required forms (Crockett, 1976, 69–74, 93–104). Both these types were excluded from our study; they are mentioned again here because if they had been included in another analysis, without the aid of exceptional features and feature-switching, then it would be unfair to make comparisons without reference to them. But as they are problematic, we will continue restricting ourselves to the clear evidence.

4.3. Reasons Why Declensional Class Cannot Be Derived from Gender

We conclude that the analysis of Thelin and Crockett, in which gender is specified in the lexicon, cannot avoid the need to specify declensional class as well. This is required for many hundred nouns. While these must be seen in the context of the total number of nouns – the Obratnyj Slovar', for example, lists over 56,000 (Lazova, 1974, 943) – the nouns in question are not isolated exceptions. Worse still, the groups described are not fossilized sets, but are productive. Both Thelin and Crockett assumed three declensional classes. However, the problems outlined above cannot be resolved by adopting a different number of paradigms. First škola and kost', both feminine, are always assigned to different paradigms in all accounts; therefore all the kost' type nouns must be labelled in the lexicon to prevent them being assigned the same endings as škola. Similarly, whether one adopts two, three or four paradigms, given a rule stating that masculine nouns are declined like zakon (whatever paradigm title this is given) some special label will be required to prevent roots like djad' being assigned the same endings as zakon. The problems involve the undisputed paradigm assignments; we conclude that it is impossible to derive declensional class from gender assignment without resorting to the strategy of having a large number of nouns labelled as irregular in the lexicon.

5. GENDER SPECIFIED IN TERMS OF DECLENSIONAL CLASS

Given that by specifying gender in the lexicon we cannot eliminate declensional class, we must now investigate whether by specifying declensional
class, we can eliminate gender. This approach was favoured by Scholz (1965, 289–290); he states that the gender of inanimates depends on declensional type, but he does not specify the rules required. This is the problem we tackle below.

5.1. The Gender of Non-Derived Nouns

We assume for the moment that each noun is marked for declensional class (as we shall see below there are instances where even the declensional class need not be specified in the lexicon). We recognise four regular declensional markers, as follows:

I. — nouns like zakon;
   II. — nouns like škola;
   III. — nouns like kost';
   IV. — nouns like vino.

Put' and the small group of nouns like vemfa will be marked as irregular: some sort of irregular marking is necessary in any account to handle the irregular forms and stress patterns; once the irregularities are dealt with, the remaining forms can be borrowed from the third declensional class.

Given these declensional markers, the gender of a noun can be established by a simple procedure. The rules are presented in the form of a flow chart; each will then be discussed in turn.

The flow chart is entered at E and the boxes represent decisions, with + for yes and — for no. (Traditionally diamonds are used for decision boxes; oblongs are substituted to make the diagram fit on a single page.) The oval represents a procedure box, which we shall meet in due course.
The first decision required, on entering the box, is whether the noun is sexed or not. Given the nature of human preoccupations, the majority of sexed nouns refer to humans. However, certain animals are sexed—basically those where the sex matters to humans, that is to say domesticated animals. For all sexed nouns, human and animal, the sex determines the gender, female nouns being feminine and male nouns masculine. Thus, *mat* ‘mother’, *devuška* ‘girl’ and *korova* ‘cow’ are feminine, while *otec* ‘father’, *mal’čik* ‘boy’, *djadja* ‘uncle’ and *Saša* ‘Sasha’ are all masculine (even though the last two are declined according to the second declension).

For the remaining nouns, we must first separate those which decline from those which do not (a question to which we return in more detail below). Of those which decline, we first assign *put’* to the masculines. This can be done either as in the diagram by reference to the item itself, or by reference to its exceptional feature. Then nouns marked as belonging to the first declension, like *zakon*, are assigned to the masculine gender. Nouns belonging to declensional types II and III are feminine. Remaining declinable nouns—type IV nouns like *okno*, and the irregular nouns like *vremja*—are neuter; there is no need to include them separately in the diagram as they form the ‘left over’ category which is assigned to the neuter gender.

We must retrace our steps to consider nouns which are neither sexed nor declinable. Of these, we isolate the acronyms. Given an indeclinable acronym such as *MGU* we substitute for it the head noun of the phrase from which it is abbreviated and assign gender on that basis. (This is the effect of the ‘consider head’ box in the diagram.) *MGU* represents *Moskovskij gosudarstvennyj universitet* ‘Moscow State University’. The head noun is *universitet* which follows declension I. The diagram shows that this noun will be assigned to the masculine gender and so we can assign the acronym to the same gender.

For non-sexed, indeclinable nouns which are not acronyms, the deciding criterion is animacy. Nouns in this category which are animate are masculine, while inanimates are neuter. This can be illustrated by the following contrast: *boa* ‘a boa’ can be neuter, and means something to put round one’s neck; it can also be masculine, in which case it is animate and definitely not to put around one’s neck. Less critical examples are the following: *taksi* ‘taxi’, *kino* ‘cinema’ and *pal’to* ‘coat’ are all neuter, while *gnu* ‘gnu’, *kenguru* ‘kangaroo’ and *marabou* ‘marabou’ are masculine.

The simple algorithm presented above correctly assigns Russian nouns to their gender class. It is worth pausing to see how this occurs in the less obvious cases. Let us begin with a consideration of acronyms. The diagram makes specific reference to acronyms only if they are indeclinable. This is because acronyms which are declined are assigned to a gender class in
exactly the same way as ordinary declined nouns. Take for example the
acronym *vuz* (vyššee učebnoe zavedenie 'higher educational institution').
*Vuz* declines like *zakon* and by this criterion it is assigned to the masculine
gender. Note that the head noun is neuter; this fact demonstrates the
need to distinguish between declinable and indeclinable acronyms. Some
acronyms belong in both categories. Comrie and Stone (1978, 79) quote
the interesting case of *ŽÈK* (žilišno-èksesual'nostnaja kontora 'housing-
exploitation office'). This acronym can be declined, in which case it is
masculine: or it can remain indeclinable, in which case it is feminine (the
head noun *kontora* is feminine). Given the two possibilities, our algorithm
predicts the correct gender in both instances. For further information on
acronyms see Šanskaja (1964), Crockett (1976, 45–59), Comrie and Stone
(1978, 78–82) and Graudina et al. (1976, 83–90).

In the case of *ŽÈK* it is necessary to specify in the lexicon that it may
or may not be declined. This is an idiosyncratic fact about *ŽÈK* and some
other acronyms, and so correctly belongs in the lexicon. But this does not
mean that all indeclinable acronyms, nor indeed all indeclinable nouns,
need to be specifically marked. There is a simple generalization which covers
all instances of nominal declension in Russian:

*Any stem ending in a vowel cannot take declensional endings.*

Thus in our acronym example above, *MGU* need not be labelled as
indeclinable; the fact that it ends in a vowel will guarantee this. The
same holds for ordinary nouns which are indeclinable. The noun *pal'to* is
entered in the lexicon in this form; that is to say, the *o* is part of the stem.
It cannot then acquire inflectional endings. This is the view taken by Worth
(1966); he points out that the idea goes back at least as far as Grot. Įsačenko
(1969) prefers to postulate a ‘non-implemented inflectional morpheme’, as
this allows him to use a single stem for derivational and inflectional
morphology. If this suggestion is followed, then the presence of the special
morpheme at the end of the stem is sufficient to signal that a noun is
indeclinable. The main point is that *okno* is entered in the lexicon as the
stem *okn*- and the -*o* is an ending supplied by the inflectional rules. In the
case of *pal'to*, however, the *o* is part of the stem.

We can therefore identify the indeclinable nouns: some are labelled as
such in the lexicon, but in the majority of instances their indeclinability is
a consequence of their stem structure. Let us consider the different ways in
which they can be assigned to gender classes by our rules. It is possible
for an indeclinable noun to be sexed, thus *attuše* ‘attaché’ is masculine, while
*ledi* ‘lady’ is feminine. Indeclinables referring to animals may also be treated
as sexed, though this is rare. In the oft-quoted example:
The straightforward assignment rules for indeclinables are a relatively recent development; for details of the earlier position see Unbegaun (1947). In Modern Russian, exceptions to the rules given are extremely rare, and the nouns involved must be marked as exceptional in the lexicon (for examples see Šanskaja, 1965). The most commonly occurring exception is kafe ‘coffee’ which is indeclinable and inanimate and so, by our rules, should be neuter. In the standard language it is masculine; the reason is that earlier the forms kafej and kofij existed – both declined regularly as first declension nouns and so were masculine. In Modern Russian only the indeclinable form survives. It is significant, however, that in spoken Russian kafe is neuter as our rules would predict (for examples, see Graudina et al., 1976, 79). There is another type of exception, which is more apparent than real. Certain forms, such as rono (rajonnyj otdel (masc.) narodnogo obrazo-
vaniya, literally ‘district department of people’s education’), may be treated as neuter, even though the head word is not neuter. According to Šanskaja (1964, 68), this happens mainly in the spoken language; Graudina et al. (1976, 88) say that it arises with frequently occurring forms. It is particularly likely with forms ending in -o or -ê; for these there is the model of indeclinable nouns like pal’to ‘coat’. What is happening in the case of forms like rono is simply that speakers are treating them as nouns in their own right; it is natural that forms similar in structure to existing nouns should lead the way. Once they are perceived as ordinary nouns, then the fact that they are acronyms in origin becomes irrelevant. Rono has a stem ending in a vowel, and so it will be indeclinable; for speakers who do not perceive it as an acronym it will therefore be assigned to the neuter gender.

Declinable neuter nouns (those belonging to declension IV, and the vremja type) are not mentioned in our diagram. The reason is simply that nouns that are not assigned to the other two genders will automatically fall into the neuter category. This approach is justified for two reasons: first, the rules are simpler if written in this form; second, items other than nouns and pronouns, when pressed into service as nouns, take neuter agreements (except when special neutral forms are available; see Corbett, 1979a, 16–31, 1980a). Thus the neuter serves as the gender for items which do not fit into the other two genders.

We have claimed that nouns can be assigned to a gender class, providing
their entries in the lexicon include a marker for declensional class; the other information required – the meaning of the item and the form of the stem – is not in dispute as every theory requires that these be specified. We must now consider how the gender of nouns is to be determined if the noun is not listed in the lexicon, but is derived by a morphological rule.

5.2. The Gender of Derived Nouns

There are different views as to the relationship between the lexicon and the morphological rules; some would list more words in the lexicon, thereby reducing the role of the morphological rules, while others attempt to restrict the size of the lexicon by the greater use of morphological rules. As long as we are prepared to accept that some words are morphologically derived synchronically, and not just diachronically, then we need to demonstrate how the analysis we have proposed will assign gender to these nouns. We will consider separately the use of noun-forming suffixes (§5.2.1) and of affective suffixes (§5.2.2).

5.2.1. Noun-forming suffixes. Noun-forming suffixes are those used to form nouns from other parts of speech. For example, the suffix -ost' is used to form abstract nouns from adjectives; given the adjective razgovorčiv-ýj ‘talkative’ we can form the abstract noun razgovorčivosť ‘talkativeness’. We shall consider this suffix because it is the clearest case in Russian of a noun-forming suffix; if one does not accept it as a noun-forming suffix, then there are no such suffixes in Russian and so no potential problem for our analysis to handle. In modern Russian, -out' is both productive and common – the Obratnyj slovar' lists about 4,300 examples.

The suffix itself must be labelled as following the third declension. The nouns formed from it will then be assigned to the feminine gender by the rules given above. There is no possibility of any clash, as adjectives are not marked for declension according to a noun paradigm. There is, therefore, no other source for gender assignment, apart from the declensional class of the suffix. It is clear that noun-forming suffixes pose no problems for our analysis. We can leave the question of which suffixes productively form nouns, and which do not, for our rules will operate correctly in either case. Thus even if all words like razgovorčivosť were to be listed in the lexicon, our rules would still operate correctly: either each word of this type would be labelled as belonging to the third declension, or a redundancy rule would specify that all nouns containing the suffix -ost' belong to third declension. Then the rules above assign the correct gender.
5.2.2. Affective suffixes. Nouns formed with noun-forming suffixes like
-os't fit easily into our account, because there is only one possible source
for gender assignment. Affective suffixes are more complex as they form
nouns from nouns. If a noun belonging to one declensional class gains a
suffix belonging to another declensional class, then there are two potential
sources for gender assignment. The two sources are both used, which leads
to a confusing situation. This has been further complicated by the fact that
many of the accounts dealing with affective suffixes are normative in
nature; data are hard to find as the relevant examples occur rarely, and
what data are available suggest that the normative accounts are misleading.
For example, Gorbačevič (1978a, 147) demonstrates that statements in
dictionaries as to the gender of nouns formed with the suffix -in-a are belied
by the examples which occur both in speech and writing. We shall not
attempt a detailed description of the gender usage with each of the affective
suffixes in Russian; that must wait until many more factual studies have
been made (and these are hampered by the infrequency of the examples).
Our main purpose is to demonstrate the mechanism by which gender is
assigned: we shall therefore restrict ourselves to the principles involved. We
shall see that the model we have described can handle the data available.
The most common affective suffix is -išk-, which is a diminutive suffix
and belongs either to declension II or to declension IV. Consider the
following example:

(9) ...stroju odin nebol'šoj zavodiško
    (I am building one small factory) (diminutive)
    (Vojnovič, Putem vzainnoj perepiski)

Zavodiško is masculine, as the agreeing modifiers odin and nebol'šoj demonstrat. In form, however, it has the appearance of a declension IV noun,
which one would expect to belong to the neuter gender. As the ending is
never stressed in such formations, the nominative singular is indistinguish-
able phonetically from the -a ending of declension II and indeed, forms
from both declensions occur. This is seen in the oblique cases, as in the
following examples (taken from a useful set given in Gorbačevič, 1978a, 154):

(10) S etim gorodiškom (Astařev, Kraža)
    With this town (diminutive)
(11) Štim gorodiškoj (Leonov, Belaja nóž')
    This town (diminutive)

The instrumental ending in (11) is unambiguously that of the second declen-
sion. That in (10) could theoretically be a declension I or a declension IV
ending: as the null ending is impossible in the nominative, the ending in
question must signal the fourth declension. We therefore have a suffix which
can take endings of either the second or the fourth declension. What is
important for our purposes is that neither of these declensional types is associated with the masculine gender: while the agreements in (10) and (11) could theoretically be masculine or neuter, the form in (9) indicates that such forms are masculine. The masculine gender is evidently derived from the source word gorod, which belongs to the first declension and is therefore masculine. How can these facts be accommodated in the model we have proposed? We claim that forms such as gorodiško in (10) and (11) above are not listed in the lexicon. Instead they are morphologically derived. This is very reasonable, for -išk- is a productive suffix. It does not alter the meaning of the noun to which it is added but rather signals the speaker's attitude. We maintain, therefore, that at the point at which the rules for determining gender operate, a form like that to be realised in (10) above has the following structure:

(12) zavod-išk-

I IV

As the rules stand, such a form could be assigned to either the masculine or the neuter gender. We must make the stipulation therefore that in such instances the gender is to be determined by reference to the head noun: in the case above, the fact that zavod belongs to the first declension will ensure that the derived form is assigned to the masculine gender.

Why then should there ever be any variation in usage with such suffixes? The reason is that morphologically derived words become lexicalised, that is to say they are stored in the lexicon as fixed lexical items. Once this occurs, the noun will be treated like any other ordinary noun and its gender will be determined in the normal way. This can be seen particularly clearly in the form toporišće (quoted by Šanskaja, 1961). Topor 'axe' belongs to the first declension and so is masculine. -išć- is an affective suffix; it is declined according to the fourth declension but this does not affect the gender of the derived form. However, toporišće can also mean 'axe helve'; this is a different meaning, not an expression of the speaker's attitude.

Naturally the form and the meaning must be stored in the lexicon (just as with the large number of other nouns which are morphologically derived in diachronic terms only). This form is treated as an ordinary noun and, being declined according to the fourth declension, it is neuter. This is an extreme example; of course there are many transitional cases. Some forms are morphologically derived for some speakers, but already lexicalised for others. The forms which will be lexicalised first are those where there is any change in meaning and those formed with the less productive suffixes. The important thing for our purposes is that we have demonstrated that such forms can be accounted for in a natural way in the model we have proposed. We merely have to specify that, in the case of synchronically derived forms,
gender is determined by reference to the source noun. When the gender of an item is not that which would be assigned on the basis of the suffix, this means that we are dealing with a morphologically derived form; if, however, the gender is that of the suffix, then the form has already been lexicalised. For more examples of the use of affective suffixes see Šanskaja (1961), Zaliznjak (1967, 148–149), Kačevskaja (1969), Maksimov (1969, 1971), Crockett (1976, 9–12), Graudina et al. (1976, 74–75) and Gorbačevič (1978a, 147, 153–156).

5.3. Gender from Declensional Class

We have shown in this section that it is possible to derive the gender of Russian nouns, providing their declensional class is specified in the lexicon. (In some cases even the declensional class can be derived from elsewhere.) The algorithm required is a simple one. In order to set up such an algorithm it is necessary to recognise four, rather than the traditional three, declensional classes. Given that we can thereby eliminate gender from the lexicon, whereas in the alternative approach it was impossible to eliminate declensional class, this demonstrates that our postulated declensional classes are justified. Indeed, it was observed earlier that, using morphological criteria only, we could not decide on the number of declensions required; different suggestions have been made in the past and there was no clear-cut argument to decide the issue. The fact that the lexicon can be simplified if four declensions are postulated, provides such an argument. In this section we have concentrated on showing that gender can be derived in the way we claim; in the next section we present arguments to support the stronger claim that gender not only can be derived in this way but that it must be derived in this way.

6. ARGUMENTS FOR DERIVING GENDER FROM DECLENSIONAL CLASS

We have seen that gender can be derived from declensional class. We now consider arguments showing that this is not only possible but necessary. The arguments are of three types: those based on simplicity, those which indicate that the declensional class must be specified in the lexicon in any case, and those relating to the different predictions made by the two models.

6.1. Simplicity

In the discussion of indeclinables it was pointed out that the indeclinability of pal'to 'coat' and gnu 'gnu' can be derived from their stem structure;
given the fact that they are indeclinable, and given their meaning, their
gender can also be derived. If, however, their gender is specified in the
lexicon, this is completely redundant – nothing can be derived from it which
is not available from elsewhere. Similarly, nouns like djadja ‘uncle’ and
Saša ‘Sasha’, which follow the second declension but are masculine, would
all have to be labelled for gender and for declensional class, in the model
which specifies declension from gender. It is worth reiterating that nouns
of this type are numerous (this is partly concealed by the fact that names
like Saša do not appear in standard dictionaries and so are omitted from
the counts made of the different lexical classes). And it is their frequency
which removes the main attraction of the model deriving declensional class
from paradigm – namely the desire to predict declensional class from sex:
as malčik ‘boy’ refers to a male, it is masculine, and therefore first declen-
sion. The problem with this approach is that there are just too many
exceptions like djadja and Saša. Finally, the gender-to-declension approach
must label all nouns of the kost‘ type (otherwise a stem ending in a soft
consonant and labelled feminine could be assigned either to the second or to
the third declension). Again a large number of nouns is involved. On
grounds of simplicity we must favour the declension-to-gender approach.

6.2. The Necessity for Marking Declensional Class

There are reasons for believing that a noun’s declensional class should be
specified in the lexicon. The most basic reason is as follows: the declensional
class of a noun is required every time it is used, but the gender is not. Thus,
whatever the function of a noun in a sentence, it must be in a case form,
and so its declensional class is required. On the other hand, there are many
instances where a grammatical sentence can be constructed without refer-
cence to the gender of the nouns included in it. Furthermore, there are
numerals such as pjet ‘five’, which follow the third declension but have no
gender. In any analysis these must therefore be specified for declensional
class.

The declensional class of nouns is also required for the formation of
possessive adjectives. From first declension nouns, possessive adjectives are
formed with -ov, and from second declension nouns with -in, irrespective of
gender. Thus from ded ‘grandfather’ (first declension, masculine) the pos-
sessive adjective is dedov ‘grandfather’s’. From mama ‘mummy’ (second
decension, feminine) the adjective is namin ‘mummy’s). The important
case is papa ‘daddy’ (second declension, masculine) from which the pos-
sessive adjective papin ‘daddy’s’ is formed. This last example demonstrates
that the suffix used to form the possessive adjective depends on the declen-
sional class of the noun and not on its gender. It is reasonable to suppose that the morphological rules operate on the features available in the lexicon (in this case the declensional class): in the alternative analysis, the morphological rules would not use the feature available (gender) but would first derive the declensional class and then operate on that. The latter analysis seems implausible.

The last argument supporting the view that the declensional class of nouns must in any case be specified in the lexicon comes from stress. As was mentioned earlier, the stress pattern which a noun follows must also be listed in the lexicon – at least for all moveable types. This is not the place for an account of the complexities of Russian stress, for which the reader is referred to Zaliznjak (1967), Halle (1973, 1975) and Coats (1976), and references there. However, a study of any adequate taxonomy of Russian nominal stress will soon convince the analyst that the generalizations possible refer to declensional class and not to gender. For example, the nouns like *kost′* are divided between two stress patterns only. The class which shows the greatest number of possibilities is the second declension (and the nouns of masculine gender in this class do not behave differently from the feminines). There are no generalizations referring to feminine nouns as a class. It is significant too that masculine nouns declined like *zakon* and neuter nouns declined like *okno* are distributed differently among the types of stress pattern (a significant proportion of the latter show retraction of stress in the plural, which is rare in the first declension). This is a further argument in favour of recognising four declensional paradigms.

6.3. Predictions Made by the Different Models

It is generally accepted that the lexicon contains the arbitrary facts relating to individual lexical items. In claiming that gender is derived from declension, we are claiming that the former, rather than the latter, is an arbitrary feature of Russian nouns. For example, *Saša* happens to follow the second declension. However, the stem *Saš-* could equally well be assigned to the first declension. Those who claim that gender is specified in the lexicon should be able to point to examples of arbitrary gender assignment – for example, nouns referring to males but arbitrarily marked as feminine (the inanimates favour neither side on this point, as their gender and paradigm can both be viewed as arbitrary). But such exceptions do not occur. The argument is, therefore, that the exceptions are of the type predicted by the paradigm-to-gender approach and not of the type predicted by the gender-to-paradigm analysis.

Let us now reconsider acronym in terms of what the different accounts
would predict about their behaviour. Earlier we considered ŽÈK, formed from žilišno-eksplotaciomajka kontora 'housing-exploitation office'. The head noun of the phrase is the obvious source for the gender of such an acronym. If we were to take the view that gender is specified in the lexicon, then ŽÈK would be marked as feminine. We would further predict that, if it were to be declined, it would decline according to declension II (unless irregularly marked for declension III), i.e. *ŽÈKu, *ŽÈKu, *ŽÈKi etc. But this does not happen. If the form is declined, then it is first assigned to a declensional class, in this instance the first, and from that the masculine gender is assigned. Again the gender-to-paradigm account makes the wrong prediction.

There is a rather unusual final piece of evidence to support the view that the gender-to-declension model makes the wrong predictions. In a short story by L. Lenč, called Tanja i Tanja (in the collection Dušespasatel'naja beseda, Moscow, 1977), a little girl called Tanya makes friends with a young bear. The girl's aunt calls the bear Tanya too. However, the bear is a male. It is referred to with the pronoun on 'he' and all agreements are masculine, e.g.:

(13) ...ja nikak ne mogla zabyt' svoego (masc.) Tanyu

I could not forget my Tanya

Tanya is normally, of course, a girl's name, and it follows the second declension. Such forms constitute an area in which the gender-to-declension hypothesis appears to work well: nouns of female gender will be feminine and therefore, normally, second declension. In the declension-to-gender hypothesis they are marked as second declension, even though their gender is derived from their sex. Now consider what predictions the two models make about such a noun when its gender is changed. The gender-to-declension model predicts that if such a noun refers to a male, then it will be masculine and will decline according to the first declension; in other words it will decline: nominative *Tan', accusative *Tanja, genitive *Tanja, dative *Tanju... . The model which specifies declensional class predicts that changing the sex of such a noun will involve a change in its syntactic gender, but that its declension will not be affected, as that is specified in the lexicon. The change will, in fact, make the word analogous to a word like djadja 'uncle'. This is the correct prediction.

To conclude this section we may say that there is evidence to support the view that we not only can derive gender from morphological gender in Russian, but also that the data force us to adopt this position.

7. CONCLUSION

We have argued that gender need not be part of the lexical entries of Russian nouns, but that it can be derived, normally from the declensional
class of the noun, by a simple algorithm. This algorithm requires that we recognise four declensional types in Russian. As morphological data provide no clear grounds for recognising any particular number of declensions (tradition seems to have been the main motivating factor in most accounts) then the fact that with four declensional classes we can eliminate gender from the lexicon is itself sufficient reason for postulating four classes. We have seen that there are arguments to support this move. We have reached an analysis in which the majority of nouns have their gender specified from their declensional class (sex takes precedence in the case of sexed nouns). This makes Russian interesting typologically. Worsley (1953–54) gives examples of gender systems based on semantic classification, while Melčuk (1974) and, independently, Tucker et al. (1977) demonstrate that French gender is phonologically based (sexed nouns are excepted as in Russian). Such languages may not be so radically opposed as it appears at first, because Jones (1978) claims to have demonstrated a correlation between phonological features and semantic categories. Russian, we claim, is an example of a language where gender is morphologically based. It is a reasonable working hypothesis that gender is never a regular lexicon entry but that it is derived from an item's semantic, morphological or phonological features, or from a combination of these.

ACKNOWLEDGEMENTS

This paper was completed during the tenure of a Research Fellowship, for which I wish to thank the Council of the University of Melbourne. I am also grateful to Roland Sussex for his helpful comments.

University of Surrey

BIBLIOGRAPHY

This bibliography includes both items which bear directly on the problem of gender in Russian and other items referred to in the paper. The former are indicated by an asterisk and, when they are not devoted entirely to gender, the relevant pages are indicated in brackets. The standard grammars are included only if they are referred to in the text.

*Anon.: 1969, 'Rod sušćestvitil'nxy s osnovoj na mjagkij seglaśnyj i na šipjaččij', Russian Language Journal 85, 49–52.


REFERENCES


Janko-Trinickaja, N. A.: 1966, 'Naimenovanie lekenskogopolasuschestvitelnimennenskb}-


Kitaigorodskaja, M. V.: 1976, 'Variativnost' v vyraženii roda suschestvitelnogopri obozna-

Kopelovitch, A. B.: 1970a, 'K voprosu ob opredelenii kategorii obšego roda v sovremenno-

Kopelovitch, A. B.: 1970b, 'K voprosu o kategorii obšego roda v sovremenno-


Krylov, N. A.: 1962, 'Bessuflknye suschestvitelnyny obšego roda, soostnositel'nye s gla-


Vinogradov, V. V.: 1972, Russkij jazyk (grammatičeskoj učenije o slove), izdanie vtoroe, Vyššaja škola, Moskva (pp. 55–124).

Vinogradov, V. V., Istrina, E. S., and Barxudarov, S. G. (eds.): 1952, Grammatika russkogo jazyka I: Fonetika i morfoloģija, AN SSSR, Moskva.


