

# THE SLAVONIC AND EAST EUROPEAN REVIEW

Volume 56, Number 1 — January 1978

---

## Problems in the Syntax of Slavonic Numerals

G. G. CORBETT

IN this paper a structure for phrases involving simple cardinal numerals<sup>1</sup> is proposed along with a set of rules required for deriving the basic constructions. This allows a fruitful approach to the complexities of numeral phrases in different Slavonic languages and highlights some of the remaining problems. These theoretical devices will be introduced with reference to Russian (I) and their effectiveness tested on Serbo-Croat (II) and Polish (III).

### I

The basic syntactic facts about Russian numerals may be summarized as follows. *Odin* agrees with its noun in gender and case, and also in number (*odni sani*). *Dva* agrees in gender in the nominative (and accusative = nominative) and agrees in case in the oblique cases. *Odin*, *dva*, *tri*, *četyre* all agree in animacy with the following noun. *Dva*, *tri*, *četyre*, when in the nominative case are followed by a dual survival (almost always equivalent to the genitive singular) while *pjat'*, *šest'* etc. are followed by a genitive plural when in the nominative and agree in case in the oblique cases. Turning to the higher numerals, while *tysjača* and *million* may be used in all cases in the plural, *sto* is restricted to the oblique cases and to fixed expressions (e.g. *neskol'ko sot*). While *tysjača* and *million* take a determiner agreeing

G. G. Corbett is a Lecturer in Russian at the University of Surrey. [This paper is to be presented to the VIIIth International Congress of Slavists, Zagreb, 1978.]

<sup>1</sup> J. R. Hurford, *The Linguistic Theory of Numerals*, Cambridge, 1975, is a recent major contribution to the study of numeral systems; the syntax of numerals in Slavonic languages has still received insufficient attention, though there is valuable material in J. Šerech [G. Shevelov], *Probleme der Bildung des Zahlwortes als Redeteil in den slavischen Sprachen*, Lund, 1952, and in A. E. Suprun, *Slavjanskije čislitel'nyje*, Minsk, 1969 (and references there).

with them in gender and number, all others require a plural determiner. *Million* always takes a following noun in the genitive plural while *tysjača* may behave this way or behave like *þjat'* and *sto*. The features discussed so far are summed up in Table 1.

TABLE 1

	<i>odin</i>	<i>dva</i>	<i>tri</i>	<i>þjat'</i>	<i>sto</i>	<i>tysjača</i>	<i>million</i>
1. agrees with N in syntactic number	+	-	-	-	-	-	-
2. agrees in case throughout	+	-	-	-	-	-	-
3. agrees in gender	+	(+)	-	-	-	-	-
4. marks animacy	+	+	+	-	-	-	-
5. has own plural	-	-	-	-	(+)	+	+
6. takes agreeing determiner	-	-	-	-	-	+	+
7. takes N in genitive plural throughout	-	-	-	-	-	±	+

*Četyre* behaves like *tri*; *šest'* etc. like *þjat'*; *milliard* etc. like *million*. Compound numerals are omitted. Brackets indicate restriction in case, ± indicates an alternative.

The table is more than a convenient summary of the data. Features 1-4 may be considered 'adjective tests' and 5-7 'noun tests'. The first four tests, rather than splitting the numerals into two clear groups (adjectives and non-adjectives), show that the numerals are more or less adjectival — the more adjectival standing to the left. The higher tests are 'choosier' than the lower. The resultant matrix has a clear area of positive results (bottom left), of clear negatives (top right) and a small transitional area in the middle. A matrix of this type shows that instead of a discrete inventory of syntactic categories (adjective and noun in this case) we have a quasi-continuum or, to use the term introduced by Ross,<sup>2</sup> a 'squish'.

If we now turn to tests 5-7 we have the choosier tests at the bottom and clear results bottom left and top right with a less clear central area. We may combine the two parts of the table by using the single standard of 'adjectiveness'; this is done by adding 'it is not the case that' to tests 5-7. It then becomes clear that we are dealing with a single squish as Table 2 shows.

<sup>2</sup> J. R. Ross, 'The Category Squish: Endstation Hauptwort' (*Papers from the Eighth Regional Meeting, Chicago Linguistic Society*, Chicago, 1972, pp. 316-28); id., 'Nouniness', in O. Fujimura, ed., *Three Dimensions of Linguistic Theory*, Tokyo, 1973; id., 'A Fake NP Squish', in C.-J. Bailey and R. Shuy, eds, *New Ways of Analyzing Variation in English*, Washington D.C., 1973; id., 'Clausematiness' in E. L. Keenan, ed., *Formal Semantics of Natural Language*, Cambridge, 1975.

TABLE 2

	<i>odin</i>	<i>dva</i>	<i>tri</i>	<i>pjat'</i>	<i>sto</i>	<i>tysjača</i>	<i>million</i>	
1. agrees with N in syntactic number	+	-	-	-	-	-	-	
2. agrees in case throughout	+	-	-	-	-	-	-	
3. agrees in gender	+	(+)	-	-	-	-	-	
4. marks animacy	+	+	+	-	-	-	-	
not the case that	{	5. has own plural	+	+	+	+	(-)	-
		6. takes agreeing determiner	+	+	+	+	+	-
		7. takes N in genitive plural throughout	+	+	+	+	+	±

The squish is remarkably 'well-behaved'. Note too that 'nouniness' increases with numerical value. This is obvious and probably surprises no one, yet logically we could just as well have found 1, 3, 5 and 1000 showing noun-like qualities and 2, 100, 1 million behaving more like adjectives.

It should be pointed out that the squish obtained for modern Russian is no fluke. At earlier stages of the language's development a similar squish obtained, though the behaviour of certain elements was different. For example, in the seventeenth century in certain documents<sup>3</sup> we find *pjat'* etc. marking animacy. This feature soon disappeared. The point of interest is that this does not involve 'stepping out of line' in the squish.

In earlier times, up to the fourteenth century,<sup>4</sup> all numerals from *pjat'* upwards were like nouns — they took agreement in gender and number both of the determiner and the predicate, and were followed by a noun in the genitive plural. (*Million* was not borrowed until 1705.<sup>5</sup>) The historical development shows the loss of nouniness of the higher numerals while a squish has been maintained at each stage. At present *sto* is losing the last vestiges of 'nouniness' and *tysjača* is under pressure.<sup>6</sup>

It has been shown elsewhere<sup>7</sup> that the Russian picture is a particularly clear case of a more general, perhaps universal situation.

<sup>3</sup> L. N. Drovnikova, 'Konstrukcii tipa "vstretit' pjati čelovek" v xvii veke (k istorii sklonenija čislitel'nych)' (*Filologičeskije nauki*, Moscow, 1962, pp. 206-9).

<sup>4</sup> T. P. Lomtev, *Očerki po istoričeskomu sintaksisu russkogo jazyka*, Moscow, 1956, pp. 442-48; J. Šerech, op. cit., p. 146.

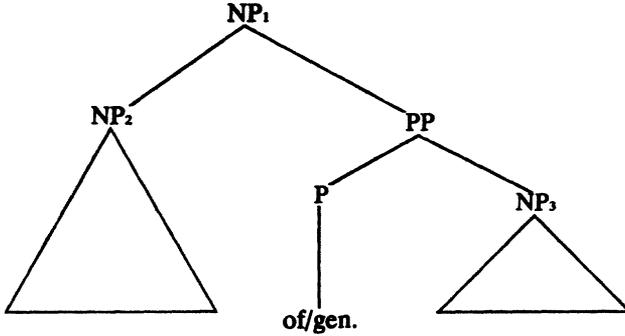
<sup>5</sup> V. Kiparsky, *Russische historische Grammatik*, vol. 3: *Entwicklung des Wortschatzes*, Heidelberg, 1975, p. 134.

<sup>6</sup> V. F. Ivanova, 'Čelovek s tysjač'ju lic ili čelovek s tysjač'ju licami?' (*Russkij jazyk v škole*, 2, Moscow, 1969, pp. 66-73).

<sup>7</sup> G. G. Corbett, 'The Numeral Squish' (paper read at the Spring Meeting of the Linguistics Association of Great Britain, Walsall, 1977).

The syntactic behaviour of simple cardinal numerals falls in the range between adjective and noun, and when there is a variation between numerals, then higher numerals are nounier than lower ones.

Clearly it is desirable to establish a single underlying structure for all the constructions with cardinal numerals. Perlmutter and Orešnik, working with data from Slovene, English and French,<sup>8</sup> suggested the following structure:



A typical derivation would reduce 'three ones' (NP<sub>2</sub>) 'of', 'boats' (NP<sub>3</sub>) to 'three boats'. This structure has a great deal to commend it. Firstly it can handle phrases which exhibit two cases at once, e.g. *pjat' bol'sich stolov*; *pjat'* (NP<sub>2</sub>) stands in the nominative, *bol'sich stolov* form NP<sub>3</sub>. Similarly, appositional phrases, normally genitive: *use sem' pyl'nych i polnych komnat, vyrastivšykh molodykh Turbinych*<sup>9</sup> can be analysed as modifying NP<sub>3</sub>. However, they may also be nominative: *150 rabot Sabatury, otnjatye na obyske*<sup>10</sup> in which case they are modifiers to NP<sub>1</sub>. Secondly, the same structure can be used for phrases involving numeral classifiers<sup>11</sup> like *čelovek, štuka, mesto* (in *mesto багаža*).

However, the preposition node is hard to justify (it would be appropriate for expressions with *iz*). It would cause great problems when trying to generate oblique case forms like *pjati stolam*. Equally, the adjective form in *dve krasivye rozy* would require an *ad hoc* genitive→nominative rule.<sup>12</sup> More generally, the presence of the

<sup>8</sup> D. M. Perlmutter and J. Orešnik, 'Language-Particular Rules and Explanation in Syntax' in S. R. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*, New York, 1973, pp. 44<sup>8</sup>-57.

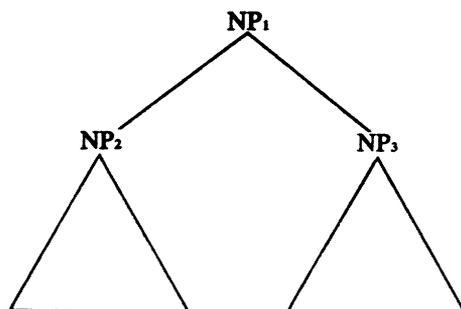
<sup>9</sup> Bulgakov, *Belaja Gardija*, Chapter 1.

<sup>10</sup> *Chronika tekušich sobytij*, vypusk 41.

<sup>11</sup> Cf. R. Sussex, 'The Numeral Classifiers of Russian' (*Russian Linguistics*, 3, Dordrecht, 1976, pp. 145-55).

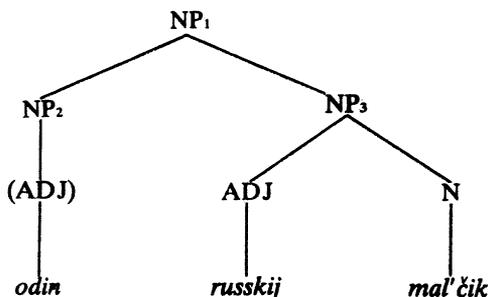
<sup>12</sup> As in M. Hums, 'Die Verwendung abstrakter Elemente bei der Strukturbeschreibung der Quantitätskonstruktionen' (*Zeitschrift für Slawistik*, 17, Berlin, 1972, p. 430).

genitive case is quite arbitrary as becomes apparent when one considers the situation in other languages; in numeral expressions some include it in surface structure while others, like English, do not; in other quantified expressions some include it (English: 'a glass of water') while others do not (German: 'ein Glas Wasser'). It seems illogical to include it in underlying structure; it should rather be inserted when required. This permits us to work with a simpler underlying structure:



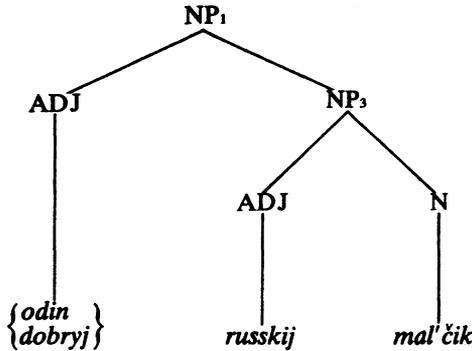
This structure copes easily with oblique cases; the case is marked on  $NP_1$  and copied onto the dominated nodes. The agreements with *dva*, *tri*, *četyre* can also be handled.<sup>13</sup> In addition a genitive insertion rule is required, and this must be constrained in its operation. The main factor determining it is the nouniness of the node under  $NP_2$ . The items labelled as nouny in the squish (rightmost in our Tables) will require genitive insertion throughout; the less nouny (like *pjat'*) require it only to avoid an unwelcome combination of two nominatives. The least nouny will not allow genitive insertion at all.

On the other hand they will require a rule to reduce the structure above to reach an appropriate surface structure. Fortunately this can be achieved using devices already required in the grammar. Consider a phrase involving *odin*:



<sup>13</sup> See G. G. Corbett, 'Problems of Adjectival Agreement in Russian' (Ph.D. Thesis, Birmingham, 1976, pp. 148-51, 154-57).

(the ADJ node under NP<sub>2</sub> is bracketed as *odin* is not a true adjective). NP<sub>2</sub> does not branch and so can be pruned.<sup>14</sup> This leaves a structure analogous to that which occurs with stacked prenominal adjectives:



The rule required to give a single surface structure NP *dobryj russkij mal'čik* will also reduce our numeral phrase to a single NP (*odin russkij mal'čik*).

There is still the problem of forms with *dva*, *tri*, *četyre*. The use of the genitive singular is an unmotivated archaism — while we may advance reasons why it has survived there is no justification for it in the modern language. The numerals involved must be labelled as exceptional, say [+dual].<sup>15</sup> In addition a dual-copying rule is required:

*Copy [+dual] onto the N in any sister NP*

(the sister NP to NP<sub>2</sub> is that marked NP<sub>3</sub> in the diagrams above). This rule operates after case assignment; in the oblique cases the [+dual] feature on NP<sub>2</sub> is reassessed as [+plural] and so copying cannot occur. After dual-copying a morphological prediction rule will give the form of the noun:

[+N, +dual] → [+N, +sing., +gen.]

Worlds like *čas*, *šag*, *rjad* which have special stress in the 'dual' form (*dva časá v. okolo časá*) will be marked as exceptional to this rule.

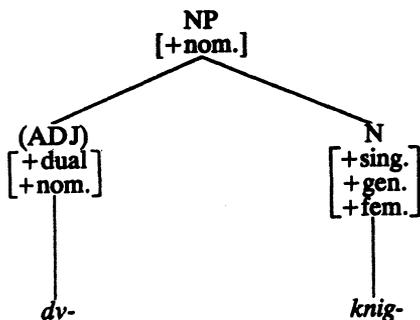
The problems of animacy with *dva*, *tri*, *četyre* like those of adjectival agreement can be handled within this framework. For the sake of brevity they will not be reviewed here.<sup>16</sup> However, the agreement of *dva/dve* is important for comparative purposes and so cannot be

<sup>14</sup> This is an extension of the convention described in J. R. Ross, 'A Proposed Rule of Tree-pruning' in D. A. Reibel and S. A. Schane, eds, *Modern Studies in English: Readings in Transformational Grammar*, Englewood Cliffs, N.J., 1969.

<sup>15</sup> As in W. O. Dingwall, 'Government, Concord and Feature-Change Rules' (*Glossa*, 3, Burnaby, B.C., 1969, p. 227); cf. K. E. Naylor, 'A Note on the "Quantification Form" in Russian' (*Folia Slavica*, 1, Columbus, Ohio, 1977, pp. 89-95).

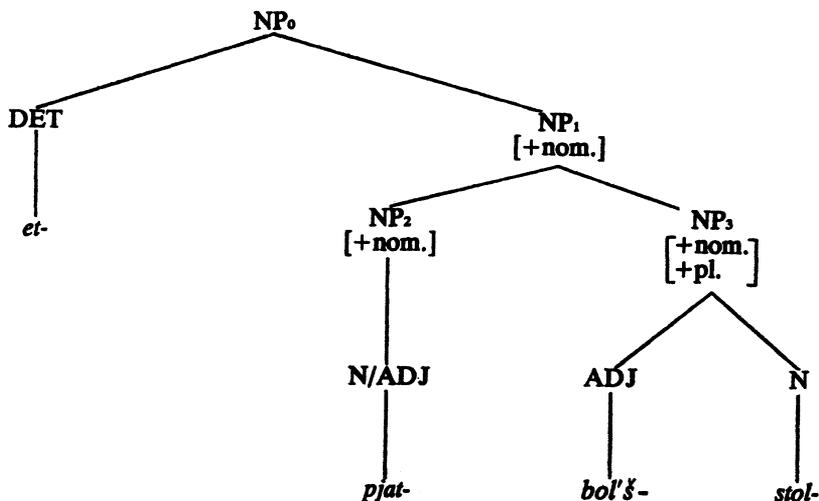
<sup>16</sup> G. G. Corbett, op. cit., pp. 151-54.

excluded. The devices described above enable us to reach the stage in the derivation given below:



The pruning of the two NPs which failed to branch has brought *dva* into a configuration similar to that of an ordinary adjective. However, the only feature 'vacant' is that of gender, which is copied at this point and the phrase *dve knigi* results. Note that agreement here occurs with a singular noun; Russian has no syntactic gender in the plural.<sup>17</sup>

The last problem we shall tackle in Russian is that of pre-numeral modifiers. These fit well into our framework, and provide further evidence in favour of genitive insertion. These modifiers, usually determiners, modify the whole group:



<sup>17</sup> The instances quoted by A. A. Zaliznjak, 'K voprosu o grammatičeskich kategorijach roda i oduševlennosti v sovremennom russkom jazyke' (*Voprosy jazykoznanija*, 13, 1, Moscow, 1964, pp. 25-40) as constructions involving gender agreement with plural nouns all involve deleted singular nouns. The other apparent counter-example is the oblique case forms of *oba*; however, this in fact supports our contention, as the feminine forms are being lost in spoken Russian.

(the node above *pjat-* indicates its intermediate position, made clear in Tables 1 and 2). The determiner *et-* should agree with the head of the complex NP<sub>1</sub>, i.e. the N in NP<sub>2</sub>; however, numerals like *pjat'* are no longer specified for gender and number in modern Russian. As NP<sub>2</sub> cannot provide the necessary features, NP<sub>3</sub> is scanned and the feature [+pl.] is copied from there. This operation occurs before genitive insertion. The operation of the latter rule gives the final phrase *eti pjat' bol'sich stolov*. The agreement of determiners also occurs before dual copying; at that stage, the singular number has not been imposed onto the N in NP<sub>3</sub> when the numeral is *dva*, *tri* or *četyre* and again a nominative plural determiner results.

This concludes our discussion of Russian data. We will now consider how relevant the main devices proposed — the two NP structure with a 'squishy' node under NP<sub>2</sub>, genitive insertion and dual copying — are in analysing two other Slavonic languages.

## II

In Serbo-Croat we can see clearly the same processes at work. Numerals form a squish, though this is not as extensive as in Russian, partly because Serbo-Croat numerals are losing their declension. The two NP structure and rule of genitive insertion will still apply. However, two areas require careful consideration: the range of the dual and the case of pre-numeral modifiers.

While dual agreement has considerable scope in Serbo-Croat, it should be stressed that the dual survives solely as an agreement feature, i.e. as a syntactic category. As a semantic category it is long since quite dead, as Belić has shown.<sup>18</sup> In Serbo-Croat the dual copying rule will read:

*Copy [+dual] onto the N in any sister NP providing it is marked  
[+masculine] or [+neuter].*

Feminines are not affected; with these the numeral behaves like an adjective and the noun and modifiers stand in the nominative plural. Adjectives in NPs qualifying a noun marked as dual will take this feature by attributive agreement. Items carrying a dual marker will, as in Russian, require a morphological predication rule. This could refer to the genitive, as in Russian. However, there is the further complication that it would be necessary to specify the indefinite form for adjectives. Even this would fail to accommodate determiners like *ovaj* which has no indefinite form yet takes the form *ova* with these numerals. This leads us to prefer a simpler prediction rule:

[+dual] → [−a]

<sup>18</sup> A. Belić, *O dvojini u slovenskim jezicima*, Belgrade, 1932.

Such a rule has the disadvantage of disregarding the fact that *-a* is the genitive singular ending; on the other hand it avoids reference to the indefinite adjectival form which many speakers do not use in oblique cases, though they do of course use the *-a* form with numerals. With pre-numeral modifiers, the node NP<sub>2</sub> affords no specification and so agreement is with NP<sub>3</sub>. This gives phrases like: *ova dva dobra čoveka*.

The difference, compared with Russian, is seen more strikingly when we turn to the other syntactic positions affected by the dual, for in Serbo-Croat dual agreement extends well beyond the numeral phrase. Adjectives in apposition to numeral phrases with *dva*, *tri*, *četiri* may show dual or plural agreement: *dva nestalna oka*, *plava i bistra*, *kao oči u male dece*,<sup>19</sup> cf. *dva visoka i crna čoveka*, *slični kao braća*.<sup>20</sup> Similarly predicative adjectives and participles (including past active participles forming the past tense) may occur in either form: *gde su se*, *pored vode*, *opraštala dva konsula*,<sup>21</sup> cf. *tu su se dva čoveka rastali i oprostili*.<sup>22</sup> The relative pronoun usually stands in the plural: *dva čoveka*, *koji bi mogli da jedan drugog razumeju*.<sup>23</sup> Informants state that this is the only possible form, but examples with dual agreement do occur, if rarely: *dva eskadrone koja vodi sa sobom*.<sup>24</sup> The personal pronoun must be plural. The occurrence of dual agreement is summarized in Table 3.

TABLE 3

	attributive modifier	appositive adjective	agreeing predicate	relative pronoun	personal pronoun
dual agreement	+	+	+	(+)	-
plural agreement	-	+	+	+	+

This pattern of agreement is an example of a much more general phenomenon. When two forms of agreement may occur with a given item, 'strict' or 'syntactic' agreement (in this instance dual agreement) and 'loose' or 'semantic' agreement (here plural), then the nearer the agreeing item is to the controller in terms of syntactic distance (the further left on the Table) the more likely is strict agreement, and the further away (further right on the Table) the more likely is loose agreement. This 'agreement hierarchy' accounts

<sup>19</sup> I. Andrić, *Travnička hronika*, Chapter 9.

<sup>20</sup> *Ibid.*, Chapter 16.

<sup>21</sup> *Ibid.*, Chapter 19.

<sup>22</sup> *Loc. cit.*

<sup>23</sup> *Ibid.*, Chapter 5.

<sup>24</sup> *Id.*, *Mara Milosnica II*.

for a wide range of data both within and beyond the Slavonic languages; a fuller account is given elsewhere.<sup>25</sup>

We must now turn to the second outstanding problem in Serbo-Croat, that of pre-numeral modifiers. Russian forms like *eti pjat' čelovek* have the Serbo-Croat equivalent *ovih pet čovekâ*. Here again, the modifier, finding no specification on NP<sub>2</sub>, agrees with NP<sub>3</sub>. However, it agrees fully, in case as well as number. This can be dealt with easily in the framework described. While in Russian this agreement precedes genitive insertion, in Serbo-Croat it follows it. The same rules operate in both languages — it is merely the order which is different.

### III

This is an appropriate point to move to a consideration of Polish, which allows pre-numeral modifiers in the nominative or genitive — i.e. it allows constructions as in Russian and as in Serbo-Croat, e.g. *wszystkie pięć pociągów*, *wszystkich pięć pociągów*.<sup>26</sup> Polish therefore exhibits the third logical possibility — both rule orders are acceptable: agreement and then genitive insertion (giving our first example) or genitive insertion and then agreement (second example). This freedom is typical: Polish numerals show a surprising variety of forms and constructions. The two other constructions in which Polish is most original, and therefore most likely to cause difficulties in the scheme outlined above, are the use of genitive subjects and the *dwaj* series of numerals.

The problem of genitive subjects concerns expressions of the type: *przyszło pięciu panów*.<sup>27</sup> If *pięciu panów* is the subject, as the semantics suggest, we must explain the form of the verb. This is third singular, neuter, by default — the typical result when a subject cannot be agreed with (i.e. when it does not stand in the nominative). In this respect the sentence is analogous to Russian expressions of the type *Ivana tam ne bylo*. However, the case of *pięciu* is a more serious problem. The account so far suggests an underlying structure with two NPs, both labelled [+nom.] as a result of copying from the dominating NP. We would expect genitive insertion to apply, giving the form \**pięć panów*. We must now postulate an agreement rule, whereby *pięć* agrees with NP<sub>3</sub>, taking the genitive case.

<sup>25</sup> G. G. Corbett, 'The Agreement Hierarchy', unpublished article, 1976.

<sup>26</sup> See Z. Klemensiewicz, 'Liczebnik główny w polszczyźnie literackiej' (*Prace filologiczne*, 15, 1, Warsaw, 1930, p. 122); H. Grappin, *Les Noms de nombre en Polonais*, Cracow, 1950, pp. 83-84; E. Decaux, 'L'Expression de la détermination au pluriel numérique en Polonais' (*Revue des études slaves*, 40, Paris, 1964, pp. 70-71); A. Bogusławski, 'Nazwy pospolite przedmiotów konkretnych i niektóre właściwości ich form liczbowych i połączeń z liczebnikami w języku polskim', in Z. Topolińska and M. Grochowski, eds, *Liczba, ilość, miara*, Wrocław, 1973, pp. 30-31.

<sup>27</sup> For the history of this construction see Grappin, op. cit., pp. 102-16.

This derivation is plausible, as the intermediate stage (numeral in the nominative, noun in the genitive) is the correct surface structure for non-personal nouns. We must ask, however, why personal nouns should behave differently. This is not difficult to answer — personal nouns have formed a clear agreement class which regularly requires special forms. It is therefore a question of the type of agreement involved when one item appears to govern another (though genitive insertion is not government in the strict sense) and then agree with it. This is indeed agreement of an unusual sort, yet it is analogous to the situation with *dwa/dve* in Russian: the numeral imposes the dual/genitive singular and then agrees with the singular noun in gender.

The personal category is also what determines the use of the *dwaj* series of numerals. The numerals 2-4 have become more adjective-like in that they have a special form for personal nouns. This agreement can be handled by the agreement rules required for ordinary adjectives and no special rule of dual-copying is required in Polish. Given that the numeral agrees completely with the noun in the nominative case, verb agreement can operate and we reach expressions like: *dwaj panowie przyszli*.

A more difficult problem remains; alongside the example just quoted, an alternative form exists: *przyszło dwóch panów*. According to Decaux<sup>28</sup> there is a semantic distinction between the two expressions, the first being definite and second indefinite. Our problem is how to derive the syntactic form. The obvious way would be to allow genitive insertion with 2-4 in the case of personal nouns only. This would be followed, as with *pieć* and upwards by agreement of the numeral.

Thus from an underlying phrase consisting of a numeral (2-4) and a personal noun we have two possible derivations: the first (definite) involves immediate agreement of the numeral; the second (indefinite) requires genitive insertion followed by numeral agreement; this latter leaves no nominative case, therefore verb agreement is blocked and the third singular neuter results. The effect of personal nouns on the numerals 2-4 is thus to make them more like their neighbours, giving a 'squishier' picture; when definite, the numeral is more adjectival, when indefinite, genitive insertion operates making the numeral more nouny, more like *pieć* etc.

It has been shown that the constructs proposed to describe Russian numerals are applicable to languages of the other branches of the Slavonic group. Numerals form a squish, going from those like adjectives to those like nouns and this variation can be dealt with

<sup>28</sup> Op. cit., pp. 66-70.

using an underlying structure with two NPs. A rule of genitive insertion is required and, for Russian and Serbo-Croat, an exceptional marker to cope with the remnants of the dual system. We have seen how the same basic rules, given different scope or ordering, produce very different surface forms.

While certain underlying regularities have been discovered many problems remain. The complex area of compound numerals has not been considered, nor has the status of collectives. The general question of why larger numerals are nounier requires an explanation. More particularly it is puzzling that remnants of the dual have survived so long in Russian and Serbo-Croat; that Serbo-Croat, which forces declension on the most unlikely loan-words does not decline the majority of its numerals; that Polish numerals are apparently so chaotic. It is hoped that the questions raised just as much as the questions answered show that the syntax of numerals is a promising area for further comparative work in the Slavonic languages.