in press, JOURNAL OF LINGUISTICS

Countability, agreement and the loss of the dual in Russian[1]

ARTHUR STEPANOV & PENKA STATEVA

Center for Cognitive Science of Language,
University of Nova Gorica

(Received 9 March 2016; revised 4 December 2016)

In this work we investigate the internal syntax and semantics of quantifier phrases (QP) involving cardinal numerals. Concentrating on a set of previously documented puzzles concerning Case and number agreement within the numeral phrase in Russian, we argue that these agreement patterns follow naturally if one recognizes three structural layers in a numeral-based QP: the countability layer, the number layer and the quantificational layer. Our central theoretical claim is that the countability layer is implemented as a (pseudo-)classifier structure whose morphological manifestation obeys a principle of syntactic 'visibility'. Our specific claim for Russian is that, diachronically, this countability layer has emerged as a result of the loss of the dual number in the course of transition between Old and Modern Russian. We strengthen our conclusions with psycholinguistic evidence from a sentence completion study that tests Russian speakers' sensitivity to the countability layer.

agreement, atomicity, classifier, countability, dual number, numeral, Russian, Bulgarian

1. INTRODUCTION

The starting point of this study is the well-known agreement asymmetry in Russian quantifier phrases (QPs) involving numerals. With numerals greater than four, the noun appears in Genitive plural form, and if there is an accompanying adjective, the latter agrees with the noun in the relevant features. With numerals *dva* 'two', *tri* 'three' and

četyre ‘four’, however, the noun appears in Genitive singular, but the adjective nevertheless receives Genitive plural, resulting in a ‘broken’ noun–adjective agreement pattern.[2]

(1) pjatj star-**yx** gorod-**ov** / krovat-**ej** / jaic-∅
five.NOM/ACC old-GEN.PL town-M.GEN.PL bed-F.GEN.PL egg-N.GEN.PL
‘five old towns/beds/eggs’

(2) tr-i star-**yx** gorod-**a** / krovat-**i** / jajc-**a**
three-NOM/ACC old-GEN.PL town-M.GEN.SG bed-F.GEN.SG egg-N.GEN.PL[3]
‘three old towns/beds/eggs’

In oblique Cases, all elements of the NumP follow a ‘homogeneous’ Case pattern (regardless of the value of the numeral itself):

(3) vladet’/s pjatj-ju / tre-mja star-**ymi** gorod-**ami**
to.own/with five-INS three-INS old-INS.PL town-INS.PL
‘to own/with five/three old towns’

These data present non-trivial puzzles, set out in (4).

(4) *Puzzles for the theories of numerals and agreement*

Q1. How does the cardinality of the numeral affect the morphological ending on the noun, and what is the relevant difference between the higher and lower numerals?

Q2. What is the source of the 'broken' agreement pattern or number mismatch in noun–adjective combinations with lower numerals?

Q3. What is the source of the structural change from the 'heterogeneous' to 'homogeneous' agreement pattern in oblique Cases as in (3) above?

Below we review some of the major accounts of Q1–Q3 proposed in the generative transformational literature to this day and point out their strengths and weaknesses. We argue that although these accounts make a good progress in understanding the morpho-syntactic aspects of numerosity, they cannot be considered fully complete and satisfactory.

In this work, we advance the thesis that a proper understanding of the agreement pattern in Russian numeral phrases must take into consideration two important factors, or dimensions. One factor, synchronic, emphasizes the fundamental role of COUNTABILITY, whose syntactic realization may affect computation of syntactic agreement in numeral phrases. The second factor, diachronic, connects the above agreement puzzles with the loss of the dual number value in Old Church Slavonic. The ensuing grammatical change in this area led to a profound reorganization of the numeral phrase into the shape we see in present-day Russian. This reorganization was associated with a morpho-syntactic process of NUMERALIZATION, that is, changing the syntactic status of numerosity words from nominal or adjectival in Old Russian to proper numeral in Modern Russian (Section 5.2). This process resulted in splitting a (composite) number feature into two separate syntactic heads, with corresponding changes in their lexical entries. One of these newly emerged syntactic heads became responsible for atomization and countability. This change, accompanied by morpho-phonological

changes in the numerals themselves, led to the development of a (pseudo-)classifier system in Russian, akin to those featured in East Asian languages such as Japanese or Chinese.

Our goal in this work is thus threefold: (a) to place the Russian higher/lower numeral distinction in agreement and the 'broken' agreement phenomenon in a larger synchronic and diachronic context; (b) to suggest a likely syntactic mechanism behind the 'broken' agreement pattern with the 'lower' numerals, as well as a possible path of syntactic changes that led to it; and (c) to better understand the role and relevant syntactic and semantic aspects of countability in numeral-based QPs. To reinforce our claims, we also report results from a psycholinguistic sentence-completion study that tests the key role of countability and its syntactic realizations. The puzzles in Q1–Q3 will be central to our discussion below from the perspective of a particular syntactic structure that emerged as a result of a grammatical change at the level of featural makeup. Ultimately, we see our proposals as part of a larger theory in which the observed syntactic and semantic properties of numeral-based QPs follow from the way countability manifests itself in the synchronic and diachronic aspects at the syntax–semantics interface.

2. A SELECTIVE OVERVIEW OF THE PREVIOUS ACCOUNTS [4]

2.1 *The 'paucal' line of thought (Q1)*

Typologically, paucal is a member of the set of values for number, along with singular, plural and dual. It typically denotes small quantities in languages such as Bayso, e.g. *lubántiti* 'a/the particular lion', *lubanjaa* 'a few lions' ($2 \leq \text{Num} \leq 6$); *lubanjool* 'lions' (Corbett 2000: 11). Given that in Russian the switch to singular happens with 'lower' numerals from 2 to 4, it seems appealing to view the Genitive singular in Russian as a

phenomenon related in some sense to the paucal form. Many researchers acknowledge the superficial similarity and use the term 'paucal' in the descriptive sense only, to refer to the form assigned by the lower numerals. Some authors, e.g. Bailyn & Nevins (2008), Pereltsvaig (2010) and Despić (2013), go further to claim that what looks like Genitive singular with the lower numerals is actually a paucal number in the above or similar sense. In a similar vein, Rappaport (2002) argues that the lower numerals assign paucal Case. An argument often put forth in support of this view is based on the well-known sub-class of masculine, mostly monosyllabic, nouns in the modern language that, when used with the numeral *dva* 'two', *tri* 'three' and *četyre* 'four', appear Genitive singular, but, unexpectedly, change the stress pattern from the penultimate (as would be typical for Genitive singular) to the final syllable:

(5) (a) do poslednego čás-a / šág-a / rjád-a
til last hour-GEN.SG step-GEN.SG row-GEN.SG

(b) dva čas-á / šag-á / rjad-á
two hour-GEN.SG step-GEN.SG row-GEN.SG

The rest of the morphological exponence for this special 'paucal' number, presumably, coincides with Genitive singular.

While this argument goes some way towards explaining the special agreement form on the noun with the lower numerals, in its present form, it eventually leads to an ontological problem. The alleged paucal cannot be a number value, simply because it is not part of the number system of Russian or generally Slavic (see also Corbett 2000). For instance, *čas-á* does not mean 'a few hours' in a manner similar to (5) above, nor does it trigger clause-level (e.g. subject–verb) agreement. We are led, then, to conclude

that this particular number feature shows up only in the presence of a lower numeral, thus in an extremely local context, in a drastic contrast with the singular and plural markings. This looks like a case of missing a generalization, which raises doubts about treating the paucal ending as an instance of the category number.

2.2 *Halle 1990 (Q2)*

Halle (1990) proposes an explicit readjustment rule that converts the plural number on the noun in cases like (2) above to singular in the context of a lower numeral, and applies only in numeral phrases (see also Halle & Marantz 1993). The rule implies that the number value on the noun is underlyingly plural. Its structural description does not include the modifying adjective; the adjective thus remains in the plural form.

The descriptive and explanatory potential of this rule is limited, as Halle himself notes. The rule does not explain the 'homogeneous' agreement in cases like (3) above, where plural is retained. It also predicts non-existing forms, e.g. with collective numerals **dvoe mužčiny* 'two man.GEN.SG'; compare *dvoe mužčin* 'two men.GEN.PL'. In addition, it does not explain why the conversion is limited to the lower numerals, as well as why it happens in the first place. Nevertheless, it serves as a good starting point for investigating the morpho-syntax of these constructions.

2.3 *Babby 1987 (Q3)*

Babby (1987) offers an account of the switch from the heterogeneous to the homogeneous Case agreement pattern based on a particular timing of assigning certain kinds of Case over others: in particular, lexical/oblique Cases are assigned prior to configurational Cases. The latter include clause-level Cases such as nominative or accusative, as well as Genitive of quantification assigned to the head noun. There is also a sub-hierarchy of assignment among the configurational Cases. If the entire phrase is in

the syntactic position of Nominative (e.g. subject) or Accusative (e.g. object), the numeral receives that Case from 'outside' the NP but, before that, it assigns Genitive of quantification to its sister, a projection of the head noun. If the outside Case assigner is lexical (e.g. a preposition or a verb), then that Case is distributed down to all constituents of the NP, and Genitive of quantification is not assigned. This is because, for Babby, Genitive of quantification in NumPs is a configurational Case. In the oblique Case contexts, its application is suspended because the lexical/oblique Case assignment takes over.

Babby's implementation of this proposal requires some non-trivial complications. For instance, he assumes a version of the X-bar theory which implies that a maximal projection of the numerical quantifier assigns Case to an intermediate projection of the head noun. Another non-trivial assumption is assigning Case by percolation: for Babby, Case is assigned to the head noun's maximal projection and then percolated down to all the available lexical and phrasal categories in the phrase. In addition, Babby's postulated hierarchy in timing of assignment of lexical and configurational Cases is claimed to follow from the representational distinction between D-structure and S-structure, no longer maintained in the current versions of the theory. Nevertheless, Babby's account contains several important insights. One such insight is in the fact that the numeral does not head the phrase, despite the appearances. Another important idea is that the lexical Case causes a structural change in the numeral phrase.

2.4 *Rappaport 2002 (Q3)*

Rappaport (2002) is essentially a Minimalist version of Babby's (1987) account (see also Franks 1994). A strong point of Rappaport's account is that it derives the hierarchy of timing of application of different types of Case, largely stipulated in Babby's system,

from the Minimalist architectural considerations. In particular, the 'quantificational' Case is valued on the noun prior to assigning a clause-level Case to the NumP itself, simply because, on Minimalist assumptions, by the time 'quantificational' Case is assigned, the higher context does not exist yet (e.g. if the NumP is an object of the Accusative-assigning verb, that verb has not been Merged yet). In other words, the sub-hierarchy of Genitive of quantification and clause-level configurational Cases follows automatically from the cyclic nature of Merge operating bottom up (see this work for details). The priority of application of lexical/oblique Case over configurational Case is similarly encoded in the presence vs. absence of the valued Case feature in the way outlined above. At the same time, additional stipulations are still required in Rappaport's account in order to rule out unwanted combinations of numerals with valued (or unvalued) Case features and different structural Case positions. For instance, the mechanism of 'veering off' the configurational Case (nominative, accusative) in NumPs is left at the level of descriptive analogy with Genitive of Negation, without providing specific details that lead to testable predictions in this area. In addition, postulation of the separate 'quantificational' Case only seems to serve a specific purpose of the numeral not being assigned Genitive itself, along with its NP complement. In Section 5.7 below we discuss an alternative way of ordering the applications of different kinds of Case in numeral-based QPs.

2.5 *The 'hybrid' accounts (Q3)*

Some researchers attribute the switch from the heterogeneous to the homogeneous patterns to some hybrid, or dual categorical status to the numerals themselves. For instance, Franks (1995), argues that numerals are QPs (or Qs) in the Genitive-assigning environments and adjectivals in oblique Case environments. Similarly, Bailyn (2012)

assumes that the numeral may be generated either in Q or in Spec–Q. The idea is that in the homogeneous, but not in the heterogeneous pattern, the numeral generated in Q naturally 'absorbs' the Genitive-assigning property of Q. Among other things, this solution seems to predict that in oblique Cases the numeral cannot be phrasal, whereas in direct Cases (nominative, accusative) it can. This may potentially be problematic because of grammaticality of expressions like the following:

(6) vladetj dvumja s polovinoj gektarami (lesa)
own two-INS with half-INS hectar-INS forest-GEN
'to own two and a half hectars (of forest)'

2.6 *Pesetsky 2013 (Q2, Q3)*

The account in Pesetsky (2013) employs an intricate set of assumptions that can be summarized (with certain simplifications) as follows:

(i) Genitive marker (N_{GEN}) is the default part-of-speech marker on N, Nominative (D_{NOM}) on D, Accusative (V_{ACC}) on V, Oblique on P (P_{DAT}). These markers are pre-specified in the lexical entries.

(ii) 'Feature Assignment': a feature of the merging head is copied onto ALL constituents of its sister (this is abstractly similar to Babby's (1987) percolation mechanism; see above). This mechanism is activated if and only if the merging head qualifies as an appropriate 'feature assigner', and only after all selection requirements are met.

(iii) Merger of another Case-bearing head such as D copies that head's (e.g. Nominative) Case feature ON TOP of the default Genitive on the N, in a 'Case-stacking' manner.

(iv) The 'one-suffix' rule eliminates all stacked Cases except the outermost one.

(v) Nouns may combine with the number feature in syntax or in the lexicon. The number feature is realized as a non-projecting head Number (NBR) merged with N, below all adjectives, by a version of a structure building operation termed UNDERMERGE.

(vi) Higher numerals (as well as some quantifiers) are instances of the category QUANT merged above the N+NBR complex and all adjectives (in Spec–NP).

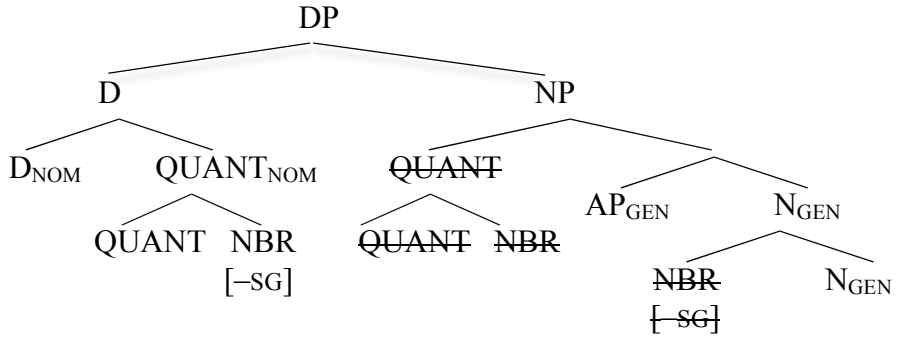
(vii) Nouns marked Genitive singular, used with the lower numerals (recall (2) above), are actually numberless, despite the appearance.

(viii) In NumPs with lower numerals, the number specification of the N derives from the numerals themselves, so that *dva* 'two', *tri* 'three' and *četyre* 'four' are number markers of dual, trial and quadral, respectively.

Dékány (2015) presents a concise summary of the major derivations utilizing these assumptions in Pesetsky's framework, which we largely follow here. With higher numerals (five and beyond), the noun bearing N_{GEN} first Merges with the non-projecting head NBR specified [–singular]. Then the adjective and QUANT are Merged, successively, each receiving Genitive by Feature Assignment (see (ii) above). NBR raises to QUANT (this ensures the correct word order numeral–adjective–noun). Next D is Merged probing for QUANT and triggering the movement of the QUANT+NBR complex, as a result of which D assigns Nominative to QUANT on top of Genitive. The

One-suffix rule (see (iv) above) deletes the Genitive and leaves Nominative. These steps are illustrated in (7).

(7))



The 'broken' pattern with the lower numerals (Q2 in (4)) is derived similarly with one notable difference. The lower numerals *dva*, *tri* and *četyre* are morphologically free exponents of NBR which by hypothesis are instances of [–singular] (see (vi) above), whereas the N is/remains numberless. The derivation proceeds as before, except that QUANT in this case is morphologically null.

The homogeneous Case pattern observed in oblique Cases (Q3) is explained as follows. Pesetsky assumes that oblique Case environments are associated with a P head, either explicit or implicit. When the DP merges with P, the oblique Case is uniformly assigned via Feature assignment to all of the DP's constituents. By the One-suffix rule, Oblique will override the Nominative assigned by D, and default Genitive of the N and A. The fact that there is no number mismatch even in the case of lower numerals, as (3) shows, is accounted for by postulating that P has an unvalued *u*NBR feature that gets its value [–singular] from the (moved copy of) NBR and later transmits it to all members of the DP via Feature assignment.

Our approach outlined below shares a number of ideas and concepts with Pesetsky (2013). We believe that separating number from the noun, on the one hand, and from the numeral itself, on the other, is a step in the right direction, as it highlights different structural layers of a numeral-based QP that can be shown to interact in non-trivial ways. We also agree that principles such as Feature assignment and 'one suffix rule' may be relevant to these QPs (see also below). However, in our view, Pesetsky's mechanism of agreement computation may be overly complicated. Pesetsky uses the Russian QPs as a case study for a novel theoretical approach dispensing with Case theory as a module of grammar altogether. In this approach, morphological case markers are primarily realizations of the parts of speech or syntactic categories, and the notion of Case (with a capital letter) as a syntactic licensing mechanism used in traditional syntactic theory can be seen as redundant. This approach also raises some non-trivial concerns regarding the nature of morpho-syntactic markers of numerosity in Russian. For instance, treating low numerals as number markers of dual, trial and quadral puts Russian alongside languages that are typologically very rare (Corbett 2000: 26) and suggests that Russian may have no lexical means to express the respective numbers, a seemingly strange state of affairs cross-linguistically. Another non-trivial issue lies in the conjecture that singular marking is some kind of default marking. There are reasons to believe that if there is a default number marking in Russian denoting 'one or more', it is actually plural (see e.g. Sauerland, Anderssen & Yatsushiro 2005, Spector 2007). For instance, when the context emphasizes the kind, not the token, of countable entities, the plural is chosen:

(8) U nas xorosho rastut ogurcy / *rastet ogurec
at us well grow-PL cucumber-PL grows-SG cucumber-SG
'Cucumbers grow well in our area.'

Mel'chuk (1985: 259)

Despite their intricate character, we believe the set of facts in (1)–(3) can be accounted for without the need to reevaluate the Case theory. In the present work we therefore adopt a more conservative stance, compared to the perspective in Pesetsky (2013).

3. THE 'DUAL' DIACHRONIC CONNECTION

Consider again the 'broken' agreement pattern observed with the lower numerals 2, 3, and 4, exemplified in (2) above, repeated here:

(2) tr-i star-**yx** gorod-**a** / krovat-**i** / jajc-**a**
three-NOM/ACC old-GEN.PL town-M.GEN.SG bed-F.GEN.SG egg- N.GEN.PL
'three old towns/beds/eggs'

To reiterate the solutions offered in the literature, the Genitive singular on the noun can be regarded as one of the following: (a) a misleading homophony for the special nominative 'paucal' ending reserved for nouns selected by lower numerals; (b) an output of a morphological rule that operates on the underlying Genitive plural (in that case the underlying morphology on the nouns occurring with lower and higher numerals is consistent); or (c) an indicator of numberlessness. Each solution is potentially problematic for various empirical and conceptual reasons. We would like here to

explore an alternative, diachronic, route that traces the Genitive singular ending to the dual markings in the older stages of Russian.

The view that Genitive singular on the nouns diachronically stems from the older dual morphology is adopted by the vast majority of researchers investigating the structure of Russian and Slavic numeral phrases, although some exceptions exist (see below). However, to the best of our knowledge, this diachronic connection has not so far been made formally explicit in terms of common mechanisms underlying the two endings, as well as their limited application with the lower numerals. We are thus interested in exploring the following questions related to the change from the dual to Genitive singular: (a) How did the change take place? (b) How did it spread to numerals for 3 and 4 (and not others)? (c) Why did the change not spread to adjectives?

3.1 *Dual in the numeral system of Old Russian*

The rise and fall of the dual number in Russian is documented in the historical Russian texts and discussed in the descriptive and theoretical literature (e.g. Shahmatov 1957, Ivanov 1983, Mel'chuk 1985, among others). Old Russian has inherited the dual from Old Church Slavonic, which itself preserved it from the early Indo-European stage. The recorded instances of dual can be traced back to the written texts dated as early as the 10th century. The loss of the dual number is a paramount phenomenon in the transfer from the Old Church Slavonic to the common Slavic varieties, on the way to the modern Slavic languages. Aside from Russian, the dual number is retained in the number systems of Slovenian and Sorbian, and lost in other Slavic languages, although some residues of the dual paradigm still remain throughout (see Corbett 2000 and Section 5.1). As seen in (9) below, the dual was part of the number system of Old

Russian, as it did not require an explicit numeral and triggered clause-level agreement with the verb (Ivanov 1983):

(9) *sę* *brat-a* *razloučist-a*
REFL brothers-DU separated-DU[5]
'the two brothers have separated'

(Tale of Igor's Campaign, 12th century)

From about the turn of the 13th century, the dual begins to decline and be gradually replaced by the plural, as evidenced by instances such as those in (10) and (11):

(10) na svoi rouky (compare *roucě* 'hands-DU')
on self's hands-PL
'On one's hands'

(Kliment's will, 13th century)

(11) pomozi rabomŭ svoimŭ Ivanou o Oleksiju
help servants-PL self's-PL Ivan and Oleksiy
'Help his servants Ivan and Oleksiy'

(St. Niphon, 13th century)

The gradual loss of the dual first targeted the forms in oblique Cases (Old Russian has inherited the seven-case system from the Old Church Slavonic, with subsequent loss of the vocative), leaving out the nominative. This was because the morphological case distinctions available in the plural were neutralized in the dual, possibly due to a markedness-triggered impoverishment process (Nevins 2011). According to Ivanov

(1983), some of the pressure for the loss of dual came from the speakers' motivation to distinguish different oblique Case forms, something that the plural system could furnish much more efficiently than the dual one. It is believed that the loss of the dual number has started around the 13th century with nouns with inherent duality in their lexical entries, such as 'hands' that could be used without the overt quantity word (cf. (10) and (11)) and consequently spread to other nouns, with parallel emergence of *dva* 'two' as a full blown numeral. The nominative dual ending was the last one to disappear, being retained at that stage in the numeral phrases using the overt numeral *dva* 'two'. After the disappearance of the dual from the number system of Russian, that ending 'naturally lost the dual meaning and was naturally merged with the morphologically identical Genitive singular forms' (Shahmatov 1957: 213 – our translation S&S). This merger has started from the masculine gender paradigm, extending to neuter, and finally, to feminine genders. Table 1, showing partial declination paradigms, illustrates the homophony in all three genders.

	Number		
Case	Singular	Dual	Plural
	Masculine (*rab* 'slave')		
Nominative	*rab-ŭ*	***rab-a***	*rab-i*
Genitive	***rab-a***	*rab-u*	*rab-ŭ*
	Neuter (*město* 'town')		
Nominative	*měst-o*	***měst-a***	*měst-a*
Genitive	***měst-a***	*měst-ou*	*měst-ŭ*
	Feminine (*glava* 'head')		
Nominative	*glav-a*	***glav-ě***	*glav-y*
Genitive	***glav-y***	*glav-u*	*glav-ŭ*

Table 1
Old Russian number inflection in Nominative and Genitive.

Some evidence for this transition comes from the change of the stress pattern in certain monosyllabic masculine nouns, as shown in (5), repeated here:

(5) (a) do poslednego čás-a / šág-a / rjád-a
til last hour-GEN.SG step-GEN.SG row-GEN.SG
(b) dva čas-á / šag-á / rjad-á
two hour-GEN.SG step-GEN.SG row-GEN.SG

A similar change of stress is also observed for feminine nouns, e.g. *réki* (rivers.PL) vs. *rekí* (river.DU). This change in the stress pattern is hard to explain if both instances involve the same grammatical form. But recognizing that a different, though closely related, form is used in the numeral phrase, resolves this tension. We consider the shifted stress form a 'frozen' dual ending that for some reason has resisted the overall transition to Genitive singular. Another piece of evidence comes from a series of regular plurals deriving from masculine nouns that are historically duals, more specifically, anatomical and natural duals, and thus have retained the original dual ending: *rogá* 'horns', *boká* 'body flanks', *glazá* 'eyes', *rukavá* 'sleeves', *beregá* 'shores' (the regular nominative plural ending for masculine nouns is *-i*/*-y*).

The final loss of the dual in the grammatical system of Russian is thought to have taken place in the 14th–15th century, when three separate East Slavic language varieties (Russian, Ukrainian and Byelo-Russian) were formed.

3.2 *Changes in the agreement patterns for lower and higher numerals*

An important fact about Old Russian numerals is their cross-categorical nature. There was no separate category for numerals in the old language. Words *dva* 'two', *tri* 'three'

and *četyre* ‘four’ (as well as *jedin-* ‘one’ which we put aside for the moment) are historically adjectives and as such, agree with the noun in Case, gender and number: thus, with *dva*, the noun shows up in dual, and with *tri* and *četyre* the noun is in plural, as Table 2 illustrates.

Gender	Num = 2		Num = 3	
	Old Russian	Modern Russian	Old Russian	Modern Russian
M	*dŭv-**a*** *stol-**a*** two-DU chair-DU	*dv-a* *stol-a* two table-C	*tr-ije* *stol-i* three-PL chair-PL	*tri* *stol-a* three tables-C
F	*dv-**ě*** *ryb-**ě*** two-DU fish-DU	*dv-e* *ryb-y* two fish-C	*tr-i* *ryb-y* three-PL fish-PL	*tri* *ryb-y* three fish-C
N	*dŭv-**ě*** *sel-**ě*** two-DU village-DU	*dv-a* *sel-a* two village-C	*tr-i* *sel-a* three-PL village-PL	*tri* *sel-a* three village-C

Table 2
Old and Modern Russian agreement patterns (masculine, feminine, neuter) with lower/adjectival numerals (DU=dual; PL=plural; C=count).

The numerals greater than four (e.g. *pjatj* ‘five’) are feminine singular nouns featuring the *-j/ĭ* stem (similar to e.g. *bolj* ‘pain’ or *branj* ‘fight’) which could inflect in Case and number, and agree with the determiner. These nouns always induced Genitive plural on their complements, similarly to other complement-taking nouns, as well as clause-level agreement with the predicate.

(12) a. ta pjatj butylok
that-F.SG. five-F.SG. bottles- F.GEN.PL
‘those five bottles’

b. tretja pjatj prishla
third-F.SG.NOM five-F.SG.NOM came
‘The third (group of) five (people) came’

(cf. Ivanov 1983:324, Babby 1987)

Later development of the numeral system led to gradual loss of the adjectival and nominal properties of the lower and higher numerals, respectively, and their subsequent numeralization or settlement as a separate quantificational category (cf. Section 1). The numeralization of *dva* 'two' was accompanied by a number of phonological and morphological changes, among which was, in particular, neutralization of the masculine–neuter gender distinction, which affected corresponding agreement patterns. Nevertheless, numeralization of *dva* apparently was not absolute: some overt residues of the adjectival-like agreement still remain in Modern Russian, as seen in Table 2.[6]

According to the traditional accounts, two factors contributed to spreading the former dual agreement pattern to numerals *tri* 'three' and *četyre* 'four': (a) the loss of the dual number, and (b) homophony between the Genitive singular and the dual on the nouns with the numeral *dva* 'two'. Similarly to *dva*, the numerals *tri* and *četyre* themselves underwent a similar numeralization process (partially) losing the adjectival properties (see also Section 5.2). In turn, numerals greater than 4 eventually lost their nominal properties, such as gender and number morphology, as well as modification by demonstratives (cf. (12)). Genitive Case, however, remained as a marker of quantity and partitivity; its assigning property has been retained in the newly formed numeral as an important part of its internal syntax. Altogether, the loss of the dual number, morpho-phonological changes in number words themselves, loss of gender distinctions for 2 (partially), 3 and 4, and, finally, loss of typical noun distinctions for the former nouns for 5–9 were factors that, according to the traditional views, led to an eventual establishment of the category Numeral, with the corresponding lexical items showing

typical properties of this category, viz. lack of agreement, in the modern language (cf. Ivanov 1983).

It is interesting to note that although the traditional historical grammars almost uniformly trace the settlement of Genitive singular marker in Modern Russian to the former dual, they do not explain WHY it only targeted the numeral phrases with an otherwise plural agreement pattern (cf. the agreement patterns with 'three' in Table 2) but did not spread to numerals higher than 4. The question is: does the loss of the dual necessarily correlate with the spread of the Genitive singular agreement to 3 and 4? Below we sketch an account that may possibly shed light to this and related issues.

4. THEORETICAL COMPONENTS OF THE PROPOSAL

Previous research on the syntax and semantics of numeral phrases has identified three major factors that play a key role in determining an agreement pattern between the numeral and the respective noun: quantification, number and countability. Our principal guiding idea in the present work is that each of these factors represents a respective STRUCTURAL LAYER in the numeral-based QP, that is, a piece of structure responsible for the computation of the respective property in syntax and semantics. Below we discuss these structural layers in more detail paying a special attention to the intricacies underlying their featural makeup.

4.1 *The quantificational layer*

The quantificational aspect of nominal expressions in Russian and Slavic is well known in the literature on Slavic morphosyntax (Franks 1995, Brown 1999, among many others) via Genitive of quantification that shows up in various guises. The classic case is regular Genitive of negation:

(13) (a) Ja chital eti knigi.
I read these-ACC books-ACC
'I read these books.'

(b) Ja ne chital etih knig.
I not read these-GEN books-GEN
'I did not read these books.'

Genitive of quantification also shows up with (null) existential quantifiers and partitives:

(14) Ja kupil (nemnogo) xleba / knig.
I bought some bread-GEN books-GEN
'I bought some bread/books.'

(15) Ja vypil (stakan) vody.
I drank glass water-GEN
'I drank a glass of water.'

We adopt the traditional view that the Genitive in numeral-based QPs has a quantificational source (but see Pesetsky 2013 and Section 2.6). Syntactically, quantification is encoded in a functional head Q, and the QP layer hosts the numeral itself. Thus, there exists a one-to-one correspondence between the presence of the numeral and the Genitive Case (though its actual morphological realization is a different matter; see below).

4.2 *The number layer*

The bulk of the contemporary syntactic literature maintains that number is an inherent (formal) property of the noun that has a semantic interpretation. Syntactically, it participates in the computation of agreement in the DP. For instance, number agreement between an adjective and a noun is often expressed as valuation of the (unvalued) number feature on the adjective by the (intrinsically valued) number feature of the noun. An alternative view, pursued in semantically-oriented research, is that number is a property not of the noun, but of the entire DP, and, as such, it may be represented by a separate syntactic head (see Sauerland 2003, 2008 and Harbour 2011a, b for arguments and evidence). Following Sauerland (2003), we annotate this functional head as φ. This view amounts to the claim that morphological number is not an actual expression of semantic plurality of the noun. Evidence for this view comes, for instance, from examples like the following (from Chierchia 1998: fn.7, credited to Manfred Krifka; see also Krifka 1995):

(16) The average Italian family buys 0.5 cars/*car per year.

The claim is also directly corroborated by languages like Chinese that do not have the (Indo-European) system of morphological number, as well as by languages like Turkish which use singular, rather than plural, ending on the noun in numeral phrases (example from Watanabe 2010):

(17) iki öğrenci (*-ler)
two student PL
'two students'

In line with this alternative view, we take φP to be a representation of the corresponding structural layer. The number layer is responsible for the overall plural marking on the numeral-based QP. In order to account for the three-valued, singular–dual–plural number system in Old Russian, we adopt the view whereby the number category is not seen as a primitive feature, but can be decomposed into two binary features [±singular] and [±augmented] (Noyer 1997, Harbour 2011a, b, Nevins 2011, Despić 2013). Roughly, the [±augmented] sub-feature expresses the degree of satisfaction of some other feature, ranging from minimal to non-minimal (i.e. augmented). Following the view on features as semantic properties along the lines of Harbour (2011a, b), we adopt here the following definitions:

(18) (a) $[\![+\text{singular}]\!] = \lambda P.\lambda x.P(x) \wedge |x| = 1$

(b) $[\![-F]\!](P)(x) = \neg[\![+F]\!](P)(x)$

(19) $[\![+\text{augmented}]\!] = \lambda P.\lambda x.P(x) \wedge \exists y[P(y) \wedge y \sqsubset x]$

The definition in (18a) implies that the function denoted by $[\![+\text{singular}]\!]$ applies to some predicate *P* and an individual *x* to be mapped to True when that x is *P* and *x* is a singular individual. According to (19), $[\![+\text{augmented}]\!]$, too, is a property modifier: it applies to a predicate *P* and is then mapped to the set of those individuals x in the extension of *P* which are non-minimal and have sub-elements. That is, $[\![+\text{augmented}]\!]$ checks for non-atomic elements inside non-atoms. The opposite value of features is regulated by the equation in (18b). The three-way number system is then represented as follows:

(20) Singular: [+singular, –augmented] (singular individuals)

Dual: [–singular, –augmented] (plural individuals without non-singular sub-elements)

Plural: [–singular, +augmented] (plural individuals with non-singular sub-elements)

Numerals interact with the number system in non-trivial ways, adding the dimension of numerosity to atomic and non-atomic entities. This interaction is best understood via the notion of countability to which we turn next.

4.3 *The countability layer*

It is a straightforward, though sometimes overlooked, fact that using a numeral with a noun is only possible if the referent of the noun is a set of individuated and, therefore, countable elements. Conversely, we cannot attribute numerosity to something that cannot be counted. This observation seems to be robust and cross-linguistically invariant. Consequently, countability is a property of the noun that makes it possible to count the objects denoted by it. Languages reflect this property in distinguishing nouns denoting mass and count objects (e.g. Acquaviva 2017, Borer 2005). The impossibility of combining numerals with mass nouns suggests that we need to individuate a structural level on which the objects can be counted (Kratzer 1989). The semantic research on mass/count distinction suggests that what is countable is a set of atoms (see e.g. Carlson 1977, Krifka 1995, Chierchia 1998). Following Chierchia (1998), we adopt the following fundamental restriction:

(21) Only atoms can be counted.

In order for the object to become countable, therefore, a suitable atomizing criterion must be applied. Good candidates for an appropriate atomizer are classifier phrases. An atomizing function maps the denotation of a noun, i.e. a kind into a set of atoms, as argued by Chierchia (1998). In particular, both plural nouns like *dogs*, as well as mass nouns like *furniture* may be seen as different expressions of the notion KIND. That is, the denotation of each of these nouns comes from the lexicon as an entity closed under a group forming operation, and can be further atomized into a semantically singular predicate.

There are at least two major views in the literature as to when and how this atomization procedure takes place. Krifka (1995) takes the Chinese classifier system as a prototypical case for the respective structural mechanism. For Krifka, nouns come from the lexicon as kinds, unspecified for atomicity. Then an operation takes a kind and yields a predicate applying to 'specimens or subspecies of this kind', i.e. essentially, a set of atoms. For Chinese, a relevant operator may be associated with a (overt) classifier and therefore integrated into whatever syntactic structure one may assume for classifiers. For languages like English with no overt classifier system, Krifka argues that the same atomizing operator is 'built-in' the lexical meaning of the respective noun, so that English nouns come into syntax already atomized. The parametric variation (and the ensuing distinction in the use of NPs) therefore lies at the respective level of representation.

An alternative mode of deriving atomic interpretations stems from the influential view on pluralities as complete atomic join semilattices in the sense of Link (1983), adopted in Chierchia (1998) (see also Landman 1989). In this view, the interpretation of

a plural noun like *dogs* forms a join semilattice consisting of the atomic entities themselves (e.g. *Fido*, *Barky*, *Tom*, etc.) as well as various combinations thereof (e.g. plural individuals ⟦Fido⟧⊕⟦Barky⟧, ⟦Fido⟧⊕⟦Tom⟧, etc.). Thus, the specification of atomicity is hard-wired in the lexical entries. There is no specific structural mechanism that identifies individual atoms: those are already identified before forming a semilattice. Under Chierchia's parallelism view, mass nouns may involve a similar structure. Furthermore, there is a straightforward translation algorithm between properties (e.g. 'being a dog') and corresponding kinds. Languages differ in how they structure NPs in terms of reference to kinds as defined via a set of operations that are sensitive to their argument and/or predicate status.

We step on the shoulders of Chierchia (1998) and Krifka (1995) with respect to the possibility to shift a kind interpretation to a predicative interpretation. However, unlike Krifka (1995), we distinguish between the procedures for atomization and countability and do not associate these with a single structural position of the classifier. For the present purposes, in the spirit of Link (1983) and Chierchia (1998), we posit an atomizing operator ∪, as in (22), which may in principle be encoded either in the lexicon or in the syntax, along with related definitions:[7]

(22) (a) $⟦\cup⟧ = \lambda x.\lambda y.y \leq x \wedge \text{atom}(y)$

(b) $⟦a⟧ \leq ⟦b⟧$ iff $⟦a⟧ \sqcup ⟦b⟧ = ⟦b⟧$

(c) $⟦\text{atom}⟧ = \lambda x.\lambda y.\ (y \leq x) \rightarrow (y = x)$,

where ≤ is essentially Link's 'individual part relation' denoting 'part of' or 'equal to' in the case of singularities. The symbol ⊔ denotes a join operation on a semilattice. The function expressed by ∪ is to be differentiated from that of [±singular]. More

specifically, the contribution of ⟦−singular⟧ is to only verify the non-singular atomic nature of the argument to which it applies. We further assume that an atomization operator such as ∪ exists in Modern Russian, the details of which are discussed below.

In order to get a better idea about the syntactic shape of the countability layer, we will make use of two important ideas explored in Watanabe's (2010) syntactic account of Japanese numeral phrases and some related works. First, Watanabe argues that the countability structure may be universally available. Syntactically, it is based on a head category #. Numerals or measure phrases (e.g. *four liters*) are base-generated in Spec–#P. Furthermore, Watanabe proposes that numerals are licensed only when the # head is marked for [±augmented], understood along the lines of the definition in (19).[8]

Watanabe's second idea, also following Harbour's work, is that the basic features [±singular] and [±augmented] constituting the number system (cf. (20)) may be separated or 'divorced'. According to Watanabe, in Turkish and Japanese, counted nouns in numeral phrases are marked singular or not marked for number at all (cf. (17)) because the # head in these languages is not specified for [±singular], but only for [±augmented]. In contrast, in English the # head may be specified for both [±singular] and [±augmented], so the plural marker on English nouns is a result of the [–singular] specification on #.[9] For Watanabe, it is the feature [±augmented] that is associated with the classifier(-like) structure responsible for countability. Recall, however, that in order for a definition of [±augmented] like that in (19) above to work, atomization, the procedure different from countability, must take place in the lexicon and the predicate # realizing the [±augmented] function operates on already atomized units (see above). This must also be the case in Watanabe's system.

The following section lays out derivations of examples in Old and Modern Russian involving the [–singular] counterpart of φ, in which the cardinality of the counted set is greater than 1, although we return to the [+singular] setting briefly in footnote 15 below. In the course of this outline, syntactic and semantic properties of # are elaborated in greater detail. This will provide a theoretical backdrop for addressing the original puzzles Q1–Q3 in (4) above later in the discussion.

5. THE PROPOSAL AND DERIVATIONS

5.1 *Old Russian*

As noted in Section 3.2, higher numerals in Old Russian are categorically nouns, and lower numerals are adjectives. We assume the following structures for the Old Russian φP:

(23) Dual: $[_{\varphi P}\ \varphi\,[_{NP}\,[_{AP}\,(\textit{dŭv-}\textbf{\textit{a}}\,)]\ \ \textit{stol-}\textbf{\textit{a}}]]]$ φ = [–singular, –augmented]

two-DU chair-DU

'two bottles'

Plural: $[_{\varphi P}\ \varphi\,[_{NP}$ pjat' $[_{NP}$ lět-∅]]] φ = [–singular, +augmented]

five-F years-GEN.PL

'five years'

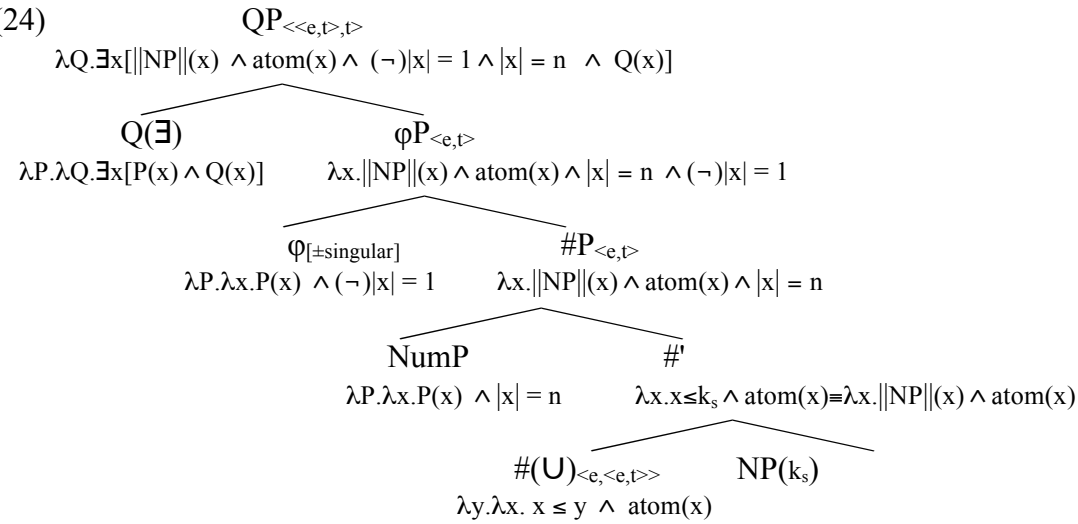
In this 'simpler' architecture, atomization of the noun, possibly along the lines of (22), must have taken place in the lexicon, or alternatively, at the NP level if Old Russian quantity words are understood as classifiers of sorts (see Sections 4.3 and 5.2). Counting takes place at the φ level. The computation of the meaning of the φP proceeds along the lines of Harbour (2011b; see his (14)): the function denoted by [–singular]

applies to the denotation of NP first, yielding non-atomic entities; the denotation of [–augmented] or [+augmented] then applies to the denotation of the respective entities, testing for sub-groups, yielding plural in the case of their presence, and dual in the case of their absence (we omit here further details for reasons of space).

5.2 *The emergence of the countability layer*

The gist of the diachronic part of our proposal is that the countability-related feature [±augmented], packaged together with [±singular] within the number layer in Old Russian, has changed its meaning from the non-atomic property checker in the sense of (19) above to a classifier-type atomization operator ∪, in modern Russian. ∪, of semantic type <e,<e,t>>, applies to a kind meaning of a noun $k_{<e>}$, and maps it to a predicative meaning of that noun whose characterized set is a join semilattice, along the lines of the semantic proposals reviewed in Section 4.2 (cf. (22)). Syntactically, this new operator is realized as a SEPARATE head category akin to Watanabe's # (cf. Section 4.3). The head of the new number layer φ continues to host the feature [±singular] only. The Q head is still responsible for the quantificational properties and Genitive Case (Section 4.1). We further assume that the Q head hosts a null determiner whose meaning is close to 'some'. Combining these theoretical components into a coherent whole results in the following structural template for Modern Russian (to be elaborated below):

(24)



In this structure, object atomization and counting takes place at the #P level.[10] The traditional singular and plural agreement endings are computed at the φ level. (Recall that, according to the definition in (18a), the predicate ⟦+singular⟧ applied to some other predicate *P* returns an atomic entity that satisfies *P*, corresponds to the traditional singular morphology. This also induces a singular clause-level (e.g. subject–verb) agreement of QP). The final representation of QP is a generalized quantifier of atomized individual(s) in the form of either a singularity or a plurality (that is, atomic or non-atomic), counted by virtue of a restrictive modifier assigning particular cardinality *n* to the atoms that have the property denoted by NP.

Thus, the two features originally constituting the number layer in the older language were split, or 'divorced', in the syntactic sense, in the new language, whereby one of these split features, now designated as #, gave rise to the new countability layer. We believe that at least part of what triggered the splitting, or divorcing, reanalysis was precisely the process of numeralization, transformation of formerly adjectival (lower) and nominal (higher) quantity words into a new functional category Numeral (A >> Num_A and N >> Num_N), in the sense of Sections 1 and 3.2. In the course of this

reorganization, the grammatical gender and number features, as well as clause-level agreement between the newly establshed numeral and the predicate, typical for the old language (cf. (12)), got lost, whereas features typical for numerals became prominent.

In other words, we propose that the former quantity words in Old Russian functioned similarly to overt numerical classifiers of the Chinese type, with numerosity lexically encoded in them. The postulated syntactic change associated with numeralization resulted in splitting these quantity words into an atomization operator, on the one hand, and the numeral proper, on the other, thus largely retaining the function typically associated with numerical classifiers. Also as part of the numeralization process, features typical for numeral categories have become prominent, in particular, the cardinality-assigning property.

Assuming that the atomization procedure is still in place, the new language needs to distinguish only singular and plural. As noted above, the feature [±augmented] used to differentiate the dual, gradually got obsolete in this new system. It is worth noting that, in the older language, the quantity word for 2 is marked [–augmented], whereas those for 3 and 4 are [+augmented]. As the dual distinction was gradually getting lost, the mismatch between the former syntactic category of the newly emerging lower (that is, adjective-sourced) numerals and their inconsistent status with respect to [±augmented] might have also contributed to abolishing the original meaning of the [±augmented] feature altogether.

In sum, as long as the featural makeup underlying the morphological exponents of dual had collapsed, so did the connection between their sound and meaning. Consequently, the morphological resources of the former dual had been reallocated to the meaning of the new head # expressing countability. Therefore, it is the

morphological expression of countability, rather than number per se, that we observe on the nouns under lower numerals in Modern Russian.

A number of authors linked the availability of classifiers in a language with the absence of morphological means for the mass/count distinction and grammatical number (Greenberg 1963, Chierchia 1998, Borer 2005, among others). Conversely, it is sometimes implied in the literature that languages with grammatical number should not have a classifier structure (cf. Ionin & Matushansky 2006). Since Modern Russian does not have overt quantity-denoting classifiers of the kind observed in East Asian languages, a potential objection may arise as to the feasibility of such classifier structure in Russian. However, we have three reasons to doubt that a strict connection of this kind exists. First, Cheng & Sybesma (1999), on the basis of more fine-grained Chinese data, show that this language does, in fact, make a mass/count distinction, despite the presence of classifiers. Second, classifier phrases may arguably exist in languages like English, as important work on pseudo-partitives shows (e.g. Selkirk 1977). Third, we side with Doetjes (1997) which suggests that both grammatical number and classifiers can be viewed as two different means serving to make semantic partitioning of the NP SYNTACTICALLY VISIBLE. Our interpretation of Doetjes' proposal is that it is the countability itself that has to be morpho-syntactically visible, and the above two means constitute values of a corresponding (morpho-)lexical parameter that serves the visibility purpose. Crucially, this parametric choice does not directly bear on the presence of syntactic structure for countability, or the countability layer.

The proposal divorcing [±singular] and [±augmented] into different morpho-syntactic entities is consistent with Watanabe's (2010) account, in which plurality is distributed over # and Q. Note that with respect to #, we essentially follow Watanabe's

syntactic proposal, as far as Old Russian is concerned, but depart from this work regarding the semantic content of this head for Modern Russian (recall that for Watanabe, atomization must take place in the lexicon, and the # head hosts the [±augmented] feature checking non-atomic entities). The proposal also recalls some previous feature-splitting accounts in the literature. In the split C system of the clausal left-periphery, Rizzi (1997) proposes that Force and Fin are often realized in a single bundle (C), if the intermediate topic/focus projections are not activated. In the spirit of Rizzi (1997), Giusti (2012) proposes that the structure of nominal expressions includes functional features Case and Num (as well as a designated left-peripheral feature projecting a phrase termed KonP; see also Giusti & Iovino 2014). Exploring a diachronic line of changes in Italo-Romance nominal expressions, Giusti argues that in Latin, Case and Num were bundled with N. In contrast, in Old and Modern Italian, Case remains separate from N, the former being realized (together with) the article. Similar proposals were made in the domain of Tense and Agreement, as well as Cause and Voice features, where respective realizations are envisioned in parametric terms (Bobaljik & Thrainsson 1998, Pylkkänen 2008). Yet another parallel may be drawn with morphological *fission* in the sense of Distributed Morphology (DM). In DM, morphological manipulations may take place over syntactic terminal nodes prior to vocabulary insertion. Fission is one such process whereby a single terminal node splits into two nodes prior to lexical insertion, allowing two or more features to be realized via separate syntactic terminals (the opposite process, fusion, effectively joins two separate terminal nodes into one). Some of the relevant cases discussed in the literature include split clitics in Catalan (Bonet 1991), split Arabic Infl morphology (Noyer 1997) and split pronominal clitics in Georgian (Halle & Marantz 1993). In DM, fission applies

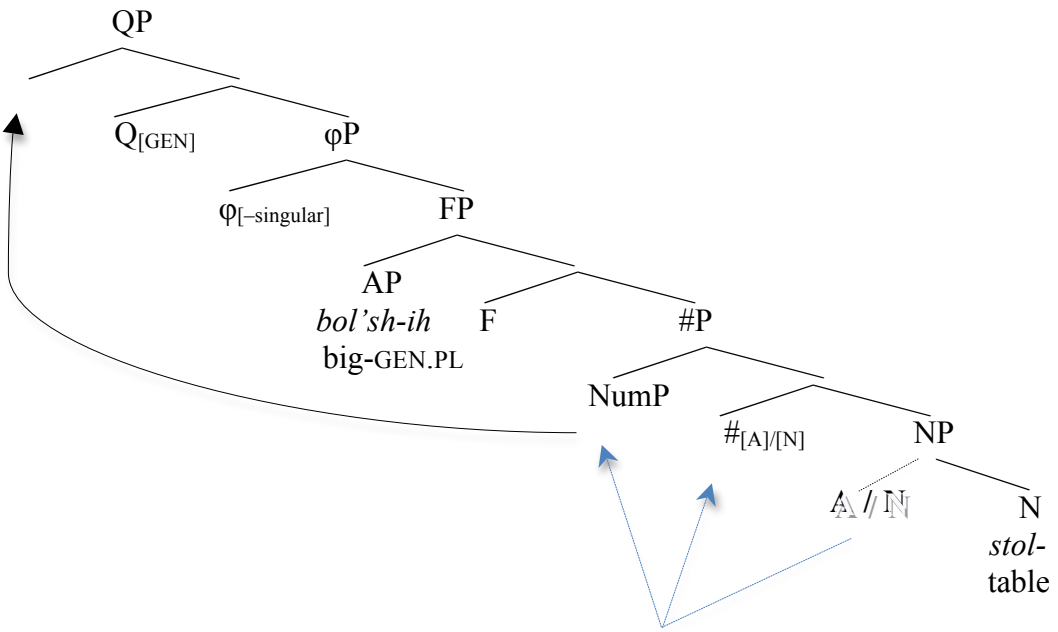
in Morphological Structure, an intermediate level between syntactic spell-out and phonological form. In our case, we might be dealing, rather, with a syntactic homologue of fission, grammaticalized in the course of the language change.[11] It remains for further research to determine whether there exist more general constraints on such separation of features in diachronic terms, and what relevant conditions at the morphology–syntax interface beyond those discussed here might be at work.

The proposed two-step scenario takes into account major historical changes documented and discussed by the traditional grammarians of Russian with respect to numerosity, but also views them in the context of a particular templatic structure of the numeral-based QP, various parts of which were argued for independently. Already at this more or less sketchy level, the scenario suggests a direction for a particular solution to the puzzle of 'broken' agreement in Russian with the lower numerals (Q2): the plural morphology on the adjective is a manifestation of the [±singular] feature, while the apparently 'Genitive singular' morphology on the noun is a morphological expression of COUNTABILITY. One question that naturally arises under this scenario is why we see the result of this process only with lower numerals. In order to shed light on this and other important questions, we now look at the splitting process in slightly greater detail.

5.3 *The resulting structural disposition*

Consider a structure of a numeral-based QP in Modern Russian, including modifying adjective(s), as in *X bol'shix stola*/*stolov* 'X(NUM) big tables'. We assume that adjectives are located in a specialized phrasal projection FP in the extended nominal domain (Grimshaw 1990, Cinque 2010). A somewhat simplified schema of the syntactic template is illustrated in (25):

(25)



The dashed arrows in (25) indicate the diachronic process of numeralization. We propose that the # head inherited the information about the categorical status of the former count word, adjectival in the case of low numerals and nominal in the case of high numerals, which for the present purposes we designate as the [A] and [N] specifications of #, respectively. The importance of this step will become clear below.[12] The NP itself is now marked for countability via assignment by #, in a head-complement configuration. We also assume that the newly emerged numeral checks its quantificational Q feature by overt movement to Spec–QP (that is, Q's respective feature has an 'EPP property'), thus deriving the surface word order.

In essence, then, we postulate that there are two historically motivated versions of the countability head # in the lexicon of Modern Russian, roughly $\#_{[A]}$ and $\#_{[N]}$. Each version is defined by the categorical type of the appropriate number word undergoing historical reanalysis as a numeral.[13]

5.4 *The computational particulars*

We assume the Minimalist approach to syntactic computation (Chomsky 2000, 2008). We further assume, without discussion, that the QP in (25) above is either a strong phase or a fragment of it, entailing that agreement-triggering heads have access to the entirety of their respective c-command domains (Abels 2003, Svenonius 2004, Chomsky 2008). Consequently, there are no Relativized Minimality type of restrictions on the respective agreement and movement processes within the QP. We also put aside the DP-level modification at this point. Following Franks (1995), we assume that both higher and lower numerals in their quantificational capacity are Caseless, that is, not subject to Nominative and Accusative Case assignment. Supporting evidence comes from adjectival quantifiers that appear to be restricted only to oblique forms, while lacking Nominative and Accusative forms, e.g. *skol'kih*/**skol'kie* 'how many-GEN/NOM', *neskol'kih*/**neskol'kie* 'several-GEN/NOM' (more discussion of oblique forms is in Section 5.7 below).[14]

In the case of higher numerals, # = $\#_{[N]}$. As noted above, the entire QP is marked for a [–singular] feature of φ, which determines its clause-level agreement. But φ also transmits this value to its complement domain via Feature assignment or via Agree, whereby nouns and adjectives marked for number within FP receive the respective valuation. In (25), both the noun and adjective are marked plural. In a similar vein, $Q_{[GEN]}$ assigns Genitive to the suitable (nominal) elements of its complement domain. Put in the Minimalist terms, both φ and Q are probes endowed with features ([–singular] and [GEN], respectively), and the noun and adjectives, but not the numeral, are goals that have respective unvalued features. The noun–adjective agreement is a morpho-syntactic

output of this mechanism. Furthermore, in the case of the higher numerals, the respective $\#_{[N]}$ head has zero morphological realization.

Consider now the lower numeral scenario, where # = $\#_{[A]}$. Each step outlined above applies here as well. In particular, in this case, the noun receives a [–singular] valuation from φ and a Genitive Case valuation from Q. The main difference is that $\#_{[A]}$ is morphologically realized as what is usually thought as Genitive singular on the complement N (itself a morphological descendant of the dual, as discussed above). The realization of the sub-features of φ and # is illustrated in Table 3.[15]

Numeral	Featural specifications	Morphological realizations
> 1	$\varphi_{[-singular]}$	–PL
2–4	$\#_{[A]}$	–COUNT (=GEN.SG)
> 4	$\#_{[N]}$	∅

Table 3

Morphological realization of φ and # in Modern Russian.

In terms of syntax–morphology mapping, and assuming the basic tenets of DM (see above), we may posit separate morphological exponents for the # head and for the φ+Q conglomerate, the former realized as in Table 3, and the latter as a regular Genitive plural marking. The φ+Q realization is not specific to numeral phrases: it is the usual marking in non-numeral quantificational phrases (e.g. *mnogo stolov* 'many tables, GEN.PL'). We suggest that the 'GEN.SG' (or, rather, COUNT) ending of $\#_{[A]}$, e.g. -*a* in the masculine form, is morphologically realized on the (counted) noun by a version of the 'affix hopping' rule in the post-syntactic component, perhaps in the form of morphological Merger or a similar operation under the condition of strict adjacency (cf. Chomsky 1957; Bobaljik 1995, 2002; Lasnik et al. 2000).[16] In contrast, with higher

numerals, since $\#_{[N]}$ is not morphologically realized, the noun only receives the GEN.PL realization of the φ+Q bundle.

Thus, in the case of the lower numerals, but not in the case of higher numerals, the noun is marked with TWO morphologically overt exponents, those corresponding to $\#_{[A]}$ and $\varphi_{[-\text{singular}]}$ + Q. However, it is known that Russian nouns generally do not allow more than one inflectional suffix, resorting to syncretism otherwise (for instance, plural inflection neutralizes gender distinction in Russian). We may capture this intuition with a DM-style impoverishment rule which deletes features from a fully specified feature matrix prior to Vocabulary insertion, yielding systematic neutralization in surface forms (cf. Bonet 1991; Nevins 2011):

(26) {–singular, [A]} → {[A]}

Alternatively, we may assume a version Pesetsky's (2013) 'one-suffix' redundancy rule that deletes all inflectional suffixes except one (Section 2.6). For Pesetsky, it is the last suffix that gets deleted by the 'one-suffix' rule. However, since in our model # is located below φ (and Q), it is the first suffix that gets retained, rather than the last. We therefore adopt a modified version of the rule as follows:

(27) Delete all suffixes on N except the first one.[17]

Thus, in the case the lower numerals, the rule in (27) will leave only the # suffix. In the case of higher numerals, the noun will be marked only for φ+Q in the first place, and this marker gets retained.

5.5 *Further evidence for #*

To reiterate, it follows from our proposal that the alleged 'Genitive singular' with lower numerals is not an expression of any particular number and/or Case at all, but, rather, a morphological exponent of # which is semantically related to atomization and countability.

This view is not entirely novel. The existence of a special 'count form' or *SCHËTNAJA FORMA*, has been previously recognized in the Russian linguistic literature (e.g. Zaliznjak 1967: 46–48; Mel'chuk 1985: 430–438) as a special 'adnumerative' form used for countable elements occurring with numerals. According to these authors, the 'count form' may reveal itself at various levels of linguistic description. The stress-changing pattern with an otherwise Genitive singular forms in masculine and feminine genders, which we have seen in (5) above, is one manifestation of it at the phonological level. The 'count form' can also be distinguished lexically, e.g. 'people': *ljudej* (GEN.PL) – *čelovek* (COUNT), or morphologically, e.g. 'flowers': *cvetov* (GEN.PL) – *cvetkov* (COUNT), 'kilogram': *kilogrammov* (GEN.PL) – *kilogram* (COUNT), and similarly for other measurement units such as *angstrem*, *mikron*, etc. The first element in each of these pairs may appear with quantifiers that do not require obligatory atomization but prefer and/or require an interpretation as a group or conglomeration. In contrast, the second element is used when the object is countable, i.e. has an atomic structure:

(28) (a) ogromnye massy ljudej / *čelovek
great masses people-GEN.PL people-COUNT
'great masses of people'

(b) desjat' čelovek / *ljudej
ten people-COUNT people-GEN.PL
'ten people'

Of course, not all Russian nouns are subject to this lexical variation. The traditional view therefore restricted the count form only to morphological and/or lexical deviations of the kind observed above. Our proposal is to generalize this view to *all* nouns occurring with the numerals. In essence, we are proposing that Russian utilizes a pseudo-classifier structure, with respective noun endings appearing as 'affixal classifiers' realizing the count form. Theoretically, this move is justified as the entire inflection paradigm for nouns comes from the separate functional head # responsible for countability. As we will see below, this generalized view on the count form is also justified from the broader Slavic perspective.

Watanabe (2010) discusses similar countability-related lexical distinctions using the distribution of 'vague' quantifiers that are compatible with either an atomic or group interpretation of the noun (note that the group interpretation does not imply that there is no atomization; rather it implies some sort of shift of focus from the atomic elements to their collective character in the interpretation procedure). For instance, *number* is only compatible with atomic/countable nouns, whereas *amount* is not:

(29) (a) John has a large number/*amount of friends.
(b) John has a large amount/*number of money.

(Kayne 2005: 145)

Watanabe encodes this contrast in terms of placement of the respective quantifiers: numerals, measure phrases and other quantificational elements expressing overt

cardinality are placed in Spec–#P, whereas 'vague' ones not directly expressing cardinality but nevertheless carrying a quantificational meaning (e.g. *many*, *a lot*, etc.) are placed in Spec–QP. We believe this step can be easily integrated in our analysis as well.

5.6 *Q1 and Q2 revisited*

The structural and (partially) diachronic perspective outlined so far allow us to approach the first two of the original three puzzles formulated in (4) above, repeated here:

Q1. How does the cardinality of the numeral affect the morphological ending on the noun, e.g. what is the relevant difference between the higher and lower numerals?

The answer to Q1 follows from our perspective on the underlying morpho-syntactic interaction of the structural layers headed by #, Q and φ. What matters is the value of the categorical feature of the # head that can be specified either as [A], for lower numerals, or as [N], for higher numerals. The former, but not the latter, has overt morphological realization (see Table 2) and it is this morphology that we see in the case of lower numerals. Thus, we claim that there are no particular semantic factors such as those related to 'paucity' or other semantically-relevant notions that affect the agreement pattern on the noun. This view is welcome on the modularity approach to grammar where (morpho-)syntactic agreement distinction is triggered by factors at the same, syntactic, level avoiding unnecessary complications involving different levels of representation.

Our solution also implies that the 'lower/higher' cut among the numerals does not correlate with the number features per se. This seems intuitively correct. The fact that the lower numerals occur with what appears to be singular morphology on the noun is an artifact of the diachronic process of restructuring of the former adjectival paradigm to the new numeral category. This view is in line also with the cross-linguistic evidence, which suggests that there is no a priori agreement with a particular number morphology in numeral phrases. If our account is correct, then the plural number morphology in numeral phrases in English, as well as the singular number morphology in Turkish, alongside with classifiers in East Asian languages, are simply different manifestations of the countability/atomization property, or, more generally, the semantic partitioning in Doetjes' (1997) sense.

We now turn to the second puzzle, repeated here from (4) above:

Q2. What is the source of the 'broken' agreement pattern or number mismatch in noun–adjective combinations?

It should be clear from our discussion above that there is no number mismatch per se. What appears to be a number mismatch between the noun and the adjective is in fact a surface reflection of different valuating mechanisms, those of #, φ, and Q, on the noun and on the adjective. As discussed in Section 5.4, the noun receives morphological marking from the φ+Q as well as from the # head, but an additional rule deletes the former marking, leaving only the latter. As for the adjective, it only receives valuation and the corresponding morphological marking from the φ+Q conglomerate and thus surfaces as Genitive plural. The 'broken' agreement pattern does not obtain with the higher numerals because [N] is morphologically null and the [–singular] plural marker

of φ surfaces unsuppressed. Russian then presents itself as a language in which each structural layer receives a separate morphological representation, thus offering a convenient test case for investigating these layers and their syntactic interaction.

5.7 *Q3: The oblique Case metamorphosis*

Consider now the third puzzle and the relevant example in (3), repeated here:

> Q3. What is the source of the structural change from the 'heterogeneous' to 'homogeneous' agreement pattern in oblique Cases?

(3) vladet'/s pjatj-ju / tremja star-**ymi** gorod-**ami**
own/with five-INS three-INS old-INS.PL town-INS.PL
'to own/with five/three old towns'

The descriptive generalization seems to be that in oblique Cases the numerals (both higher and lower) behave as modifiers, similarly to adjectives, showing full agreement, while in structural Cases the numerals behave themselves as Case-assigning heads. As discussed in Sections 2.3–2.5, this apparent duality was encoded in different terms in the literature, either in terms of relative timing of Case assignment (Babby 1987) or a dual categorical status of the numerals (Franks 1995), or the choice of a valued or an unvalued Case feature on the numeral (Rappaport 2002). These accounts are incomplete in that they do not explain what might lie behind this duality (put differently, the question is why the categorical distribution is not the reverse in each case, among other logical possibilities). Let us see how our structural proposal may shed light on this issue.

We assume that the source of an oblique Case is some external head such as P assigning a lexical or a thematically-grounded Case to the QP. Thus, in essence, phrases marked with an oblique Case are PPs (cf. Bayer, Bader & Meng 2001, Woolford 2006). It follows that the source of the (homogeneous) Case pattern is necessarily external, not pertaining to the QP-internal heads #, φ and Q themselves. In fact, the role of these three internal heads may be drastically reduced in this case, compared to the non-oblique contexts.

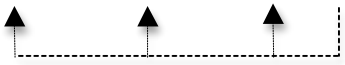
The basic idea is the old one: prepositions assign a lexical and/or thematic Case to their NP complements. A NP complement of P constitutes one of the three classical Case-licensing configurations (Chomsky 1986). Let us take this idea here quite literally, to mean that (a) the recipient of an oblique Case assigned by P must necessarily bear a full set of nominal characteristics; and (b) some kind of structural adjacency is required for lexical Case assignment (Bayer, Bader & Meng 2001, Woolford 2006). In the Minimalist formulation, the second requirement translates into a strict locality restriction on Agree with P as a probe, which needs matching with the categorical N feature. This is different from structural Case assignment, which is not subject to such a stringent locality requirement, so that a probe licensing a structural Case can search for its goal at a greater depth within the DP.[18]

This distinction provides a necessary conceptual backdrop for spelling out the homogeneous Case pattern. Both structural and oblique Case licensors require an access to a full set of nominal characteristics. A structural Case licensor such as Tense or little *v* can probe into the structure of QP to identify this set, including number, person and gender. On the other hand, an oblique Case licensor such as P is restricted in this access to the features of the (adjacent) highest projection of QP, which includes only [GEN] in

Russian, but not the N feature. This is insufficient for Case assignment, and failure to assign Case here would result in a violation of the Case filter (in whatever form we assume it to function in the grammar).

We believe that the observed homogeneous agreement pattern is a grammatical 'remedy' circumventing a Case filter violation. We envision the remedying solution in the form of effectively extending the relevant nominal domain via some version of RESTRUCTURING. Typically, QP and φP are not Case domains as they do not contain the information about the nominal, and therefore, they cannot receive Case assigned by *s* 'with'. In this case, we suggest that a series of head movements progressively incorporating heads #, F, φ and Q frees up the boundaries of the nominal domain, thus making the nominal information available 'at the top'. This is shown schematically in (30):

(30) $[_{PP}\ P\ [_{QP}\ Q_{[GEN]}\ [_{\varphi P}\ \varphi\ [_{FP}\ AP\ \ F\ [_{\#P}\ \#\ [_{NP}\ N]]]]]$



Formally, there are several ways to make this intuition more precise. In effect, the same intuition lies behind Baker's (1988) Government Transparency Corollary (GTC), whereby a head category with another head incorporated into it governs everything that the incorporated head used to govern in its original structural position. Exploring a Minimalist version of the GTC, Stepanov (2012) points out that on the assumption that head movement does not leave traces (still being part of core syntax), the Minimalist bare phrase structure theory allows for a possibility that movement of head α into head β effectively collapses αP and rearranges the target phrase βP as a composite phrase (α+β)P which includes all relevant information contained in both heads such as

respective feature matrices as well as a structural hierarchy within the head cluster resulting after movement. A radical version of this account entails that the final landing site of the series of the head movement is a conglomerate head that shares the properties of the categories through which movement took place. In our case, this means that the highest phrase resulting from these movement series will contain the features of F, the original moving head, which, by hypothesis, contains the categorical information about the NP (as well as #P). As a result of this domain extension, a restructured NP in the position of the complement of P is accessible for oblique Case assignment via Feature assignment or Agree. That is, all elements of this restructured NP projection receive Case, even the numeral itself, in accord with the empirical facts.

Feature-wise, there are no changes in the structural makeup of the extended NP projection except two: (a) when Q participates in the domain extension by hosting the # head, it no longer assigns Genitive; (b) the # head itself is no longer realized on the noun. Both restrictions are likely to result because the conditions on realization of the respective morpho-syntactic features are no longer fulfilled. Concerning the first exception, the failure to assign Genitive on the quantified noun is likely to be a result of a Case conflict or competition between Genitive and the oblique or inherent Case from an external assigner (preposition or verb). One possibility is that this conflict is morphological in the sense of the general tendency of Russian to preserve only one Case affix (Section 5.4). Another possibility is that the conflict is syntactic. To resolve a similar kind of conflict, extrinsing ordering of Case assignment was postulated previously, with oblique Case taking precedence over quantificational Genitive (cf. Section 2.4). Such extrinsic ordering is a rather unappealing solution in a computationally efficient syntactic system. Bošković (2006) offers a re-interpretation of

this Case conflict resolution in the Minimalist syntax, which avoids postulating explicit Case assignment priorities. This work derives the conflict from theta-theoretic considerations (Theta Criterion; Chomsky 1986), coupled with considerations of derivational economy understood in the sense of the Minimalist program. Consider the homogeneous Case pattern in example (3) repeated in the beginning of this section, compared with an ungrammatical Genitive-driven alternative:

(31) *vladet'/s pjat-ju star-**yx** gorod-**ov**
own/with five-INS old-GEN.PL town-GEN.PL
'to own/with five old towns'

Bošković (2006) assumes that an external oblique/inherent Case assigner will theta-mark its object iff it assigns it the inherent Case in question. The derivation in (31) then cannot converge because the external Case assigner will fail to theta-mark its object. Since derivational economy compares only convergent derivations, the derivation (31) does not actually compete with the derivation of (3): rather the derivation of (3) is forced. Thus, no extrinsic ordering of Case application is needed. From a more 'QP-internal' perspective, the situation neutralizing the Genitive-assigning ability of Q is also abstractly reminiscent of the English passives, in which the interaction of the passive morpheme *-en* and the verb destroys the configuration for Accusative Case assignment (Baker, Johnson & Roberts 1989), although, as a *JL* referee rightly points out, in our case the underlying mechanism is quite different (for the above authors, the Case assigning ability of the verb is realized on the passive morpheme *-en* which is itself a verbal argument). The conjecture that the [φ +F + #] head cluster 'absorbs' the Genitive-assigning ability of Q also recalls the proposal in Bailyn (2012) discussed in

Section 2.5 (although we depart from Bailyn in that for him it is the numeral that is generated in Q instead).

Concerning the second exception, the lack of morphological realization of # on the noun, we believe it is due to similar considerations related to head movement. It is well known that the 'affix hopping' family of rules, including morphological Merger, operate under the condition of structural adjacency (see Section 5.4. for discussion and references). Once # has moved, it is no longer adjacent to the noun, therefore, morphological Merger/affix hopping cannot proceed. A related intuition is expressed in Baker's (1988) Case Frame Preservation Principle which states, roughly, that a complex head of the category α cannot have more properties than those allowed for simple items of the same category. Feature-wise, the movement of # may trigger PF deletion of the corresponding feature [A] via another morphological impoverishment rule. Conditions for application of this rule may include the loss of structural adjacency with the noun, as outlined above, and/or potential incompatibility of $\#_{[A]}$ with other features in the resulting head cluster, in particular (oblique) Case. Some indication that a process along these lines might indeed take place comes from the fact that $\#_{[A]}$ is itself historically a (Genitive) Case marker, as we have seen above. Thus, a featural conflict may arise in the case of # as well, perhaps along the lines of a markedness hierarchy, as is often suggested (Noyer 1997, Nevins 2011, among others).

The proposed solution can be restated in the phase perspective. Recall that on our assumptions, a QP (or a DP dominating it), is possibly a strong phase (Section 5.4). One way of further streamlining the solution would be to assume that Structural Case-inducing probes such as T or little *v*, on the one hand, and oblique Case-inducing probes such as P, on the other, differ in that the latter, but not the former, require valuation of

the categorical [N] feature, in addition to the other features (e.g. *phi*-features and/or a D feature). Movement of the head #, presumably specified with the N feature, would in this case serve to make the categorical information 'visible' at the root of the phase, in much the same way phrasal movement to the edge of the phase makes the moving element visible from outside (on the assumption that the head of the phase, along with its edge, are visible to external probes, see Chomsky 2008).

Generally speaking, then, a fully spelled out formal restatement should follow from the core syntactic principles responsible for syntactic restructuring. Establishing these core principles and their interaction is itself an important theme in the recent literature, at least in the verbal domain (e.g. Wurmbrand 2001). It is important to realize that, unlike in the previous accounts, for us there is no internal 'duality' pertaining to some element in the numeral phrase; rather, the heterogeneous/homogeneous switch is determined by the 'external' properties, namely, the nature of the Case-inducing probes that target the numeral phrase. Although the proposed restructuring solution does not directly bear on the presence of the countability property per se, we believe it sheds more light on its role in the syntactic computation and its featural requirements, including interaction with other heads in the respective structures.

6. A SLAVIC PERSPECTIVE ON COUNTABILITY: BULGARIAN

Interesting (though necessarily indirect) cross-linguistic evidence for the main proposal we advance for Russian comes from another Slavic language, Bulgarian. Bulgarian is typologically remarkable within the Slavic family in that it lost the rich Case-marking morphology on nouns (though not on pronouns and clitics) typical for the rest of the Slavic languages. This fact makes Bulgarian an excellent test case for separating Case from other sources of morphological assignment internal to numeral-based QPs. If a

numeral is added into a nominal expression in Bulgarian, this triggers a special sort of marking on the noun denoting a countable quantity of objects in a set:

(32) (a) pet(-te) (dârven-**i)** prozorec-**a** / *prozorc-**i**
five-the wooden-PL window-COUNT window-PL
'(the) five (wooden) windows'

(b) (dârven-**i)** (-te) prozorc-**i** / *prozorec-**a**)
wooden-PL (the-PL) window-PL window-COUNT
'(the) (wooden) windows'

The morphologically visible count form *–a* is limited to a subset of the masculine paradigm. Standardly, it is only used with inanimate or non-person nouns. Some Bulgarian dialects, however, allow its use with animate masculine nouns as well (Pashov 1989). In non-masculine paradigms, simple plural is used instead.

Examples such as (32) are striking in two respects. First, they demonstrate a nominal morphology that is crucially different from Case and appears only in the context of a numeral. Second, numeral expressions such as (32) manifest the same 'broken' noun–adjective agreement pattern that we find in Russian. The only and important difference with Russian is that the 'broken' agreement pattern in Bulgarian is not restricted to the lower numerals, but, instead, is generalized to the entire numeral paradigm. This reinforces the view we argued for above, that whatever factor underlies the 'count' form, it does not have to be restricted to a particular subset of numbers.

If our proposed approach is on the right track, the appearance of a special morphology on the noun occurring with numerals in Bulgarian is not at all surprising. For us, this special marker is an instantiation of the countability head # whose distinct

morphological realization covers the entire numeral paradigm. The fact that it is not a Case morphology, as well as its restricted distribution, supports the view that we may not be dealing with Case morphology on the noun in Russian numeral-based QPs either. Furthermore, the 'broken' agreement pattern suggests a mechanism of morphological realization of plurality and countability similar to the one we suggested for Russian, based on the interaction of the number and countability layers, responsible for plurality and atomization. In effect, for Bulgarian we would postulate a structure close or identical to (25), modulo Genitive Case assignment by Q.

The 'count' form received a close attention in the Bulgarian traditional and typological literature. Even a cursory look at this literature reveals two major views on the origin on the 'count' form. The dominant view is that the count form is a direct descendant of the Old Church Slavonic dual number (e.g. Mirchev 1958, Pashov 1989). An alternative view sees it as a descendant of the Genitive singular form, similarly to Russian (Kharalampiev 2001, Anna-Maria Totomanova, p.c.). Our proposed model may potentially reconcile these two competing views on the origin of the count form in the Bulgarian grammatical tradition. Essentially, our model suggests that both the dominant and the alternative views are correct, but reflect different stages of grammatical change. In effect, we may postulate a trajectory of historical change underlying this 'count' form, whereby the dual form constitutes the initial (Old Church Slavonic) stage, the Genitive singular the intermediate stage, and the modern 'count' form the most recent development of the Modern Bulgarian. The major syntactic change, as we saw, happened between the first and second stages, upon emergence of the new grammatical category # (Section 4.3). Given the fact that Bulgarian lost the productive Case morphology, the modern 'count' form is in effect a 'frozen' Genitive singular marker

that lost its initial membership in the respective Case and number declensions. In fact, there is diachronic evidence reported in the literature that the development of the modern 'count' form proceeded largely in parallel with the gradual loss of Genitive Case in Bulgarian (and its replacement with prepositional forms) both of which can be traced back to the period around the 13th–14th centuries (e.g. Mirchev 1958). One may therefore tentatively hypothesize that Bulgarian represents a diachronic stage of development of the Slavic syntactic numeral system whereby the 'lower' numeral $\#_{[A]}$-marked declension has been generalized to the entire numeral paradigm.

7. COUNTABILITY: PSYCHOLINGUISTIC EVIDENCE FOR

7.1 *The experiment*

We conducted a sentence completion experiment that was aimed to test whether the morpho-syntactic mechanism of countability that we postulated for Russian may also be reflected in language use. In particular, we wanted to see whether Russian speakers are sensitive to the countability layer in the structure of numeral-based QPs, in contrast with the usual quantificational Case morphology. In the structure that we are assuming for Russian, the source of the Genitive is Q, the head of the quantificational layer (Section 4.1). Showing that speakers may distinguish the morphology induced by the countability head # and the morphology induced by the quantificational Genitive, in the same structure, may serve as additional evidence for our argument that the nominal morphology in Russian numeral phrases is in effect a morphological realization of the countability head #.

Previous psycholinguistic research on so-called 'agreement errors' or 'agreement attraction' demonstrated that the human speech production system often faces difficulties in correctly encoding grammatical information such as a particular

feature, if/when it mistakenly identifies more than one structural trigger as the relevant source of that feature. For instance, it has long been known that English speakers may make performance errors as in 'The editor of the books ARE tall' under spontaneous or experimentally induced conditions such as limiting the time frame or taxing the working memory. In examples like this, the source of the agreement feature on the auxiliary verb is the head *editor*, but another, interfering, element *books* may be seen by the production system as a competing source of agreement (Bock & Miller 1991, among others). This kind of error was argued to signal a special role of the grammatical encoding,module of the speech production system, sensitive to and/or interacting with such hardcore syntactic notions as c-command, hierarchy and even Relativized Minimality (Bock & Levelt 1994, Vigliocco & Nicol 1998, Franck, Vigliocco & Nicol 2002, Franck et al. 2006, among others). A different, though conceptually related, strand of research on 'similarity-based interference' reports interference effects in language comprehension as well as production, whereby temporary holding an element bearing a particular kind of feature in the working memory impedes recall of another element that bears the same type of feature, even if the actual value of that feature is different (e.g. Lewis & Vasishth 2005). These works demonstrate that encoding a feature in the presence of a potential competitor presents a challenge for the production system.

Taking these considerations as a point of departure, we can formulate our empirical predictions. Consider again the original Russian agreement pattern in examples (1) and (2) above. Recall that Russian allows only one inflectional suffix on the noun (there is no 'Case-stacking'). Grammatically, this restriction can be captured by a special rule, such as a version of the 'one-suffix' rule in (27)). Let us suppose, as a null hypothesis, that the *-a* ending in (2) is nothing more than a Genitive singular. The

source of that Case ending would then be the only (Genitive) Case assigning head, namely Q, heading the quantificational layer. We hypothesize that if Genitive singular and Genitive plural are 'alike' and come from the same source, they may potentially compete with each other in performance under memory-taxing conditions, inducing a similar kind of performance errors as observed in case of 'agreement errors', or difficulties caused by similarity-based interference, in language production. Specifically, we expect that (a) speakers will make performance errors in choosing the right ending; and (b) such performance errors in both directions (that is, GEN.SG instead of GEN.PL, and GEN.PL instead of GEN.SG) should be equally likely (close to a 1:1 ratio); and finally (c) the likelihood of such errors will increase proportionally to the respective difficulty in recall.

As the relevant memory-taxing conditions, we employed the linear (and structural) distance between the Q licensor and the respective noun, in terms of the number of intervening adjectives. That is, in a schematically represented situation in (33), the likelihood of errors is expected to increase progressively on the scale from (33a) to (33c).

(33) (a) Q … Adj … N

(b) Q … Adj … Adj … N

(c) Q … Adj … Adj … Ad j… N

If, on the other hand, the GEN.SG and GEN.PL are not 'alike' in the sense of not coming from the same source, then the competition may not be at issue or at least not be the only factor that affects the choice. We then expect that the ratio of observed

performance errors should not be at random. The grammatical sources of each of these endings are then likely to be different.

7.2 *Materials*

Materials consisted of 48 sentential preambles ending with a numeral phrase containing a missing noun. A lemma for the missing noun was provided separately. The following is an example of a sentential preamble:

(34) V koridore stojali pjatj starinn-yx lakirovann-yx … | [stul]
in corridor stood five old-GEN.PL lacquered-GEN.PL chair
'In the corridor there were five old lacquered …' | ['chair']

The variables manipulated were the following: (a) the type of numeral (higher, lower); (b) the number of intervening adjectives with GEN.PL (one, two or three); and (c) the gender of the noun (masculine, feminine, neuter). The items were equally represented across each of these three dimensions, and were counterbalanced using a Latin square design. The items were presented in a pseudo-randomized order to the participants, so that each participant had a unique ordering of items during the trial. Each experimental sentence was followed by a yes/no comprehension question, intended to control whether the participants were attentive to the content of the sentence. In addition, four practice sentences were added to the experimental items.

7.3 *Participants*

Sixty-two self-reported monolingual Russian adult native speakers participated in the experiment voluntarily and anonymously for no material compensation (age 24–46 years, median age = 31 years). All participants were recruited via email and social networking forums and were located in the greater areas of the cities of Moscow,

Novgorod and Perm. They had normal or corrected to normal vision. The participants were naïve as to the real purposes of the study.

7.4 *Procedure*

The task was a cloze-type sentence completion task. First, an incomplete sentence was presented to a participant word by word in an auto-paced mode: each word was presented for exactly 300 ms followed by a short pause of 100 ms between words. Words appeared exactly in the center of the computer screen replacing one another, so that in order to assess the meaning of the entire sentence the participants could only rely on their working memory, excluding other potential cues, such as relative 'topological' positions of the words on the screen. As (34) exemplifies, the numeral-based QP was always sentence-final, usually the (postverbal) subject of the sentence. After the lastly presented item in the sentence, the screen showed the lemma for the missing noun and a type-in window was provided immediately below.

The participants were asked to read each sentence and fill in the missing word form in the window provided, on the basis of their first intuitions and without dwelling on the answer. When a participant typed in the required input, the latter was recorded and the screen showed a comprehension question, followed by two possible answers to it, 'yes' or 'no', immediately below. The participants were then instructed to either click on the respective answer with the mouse or press the key corresponding to the number of that answer. The participants were also allowed to take a short break before reading the next item, if needed. The experiment was programmed using the Ibex web-based software (by Alex Drummond, http://spellout.net/ibexfarm/).

7.5 *Results*

The data from five participants were removed from the count: four participants did not understand the task and gave irrelevant responses, and one person's score on the correctness of the comprehension questions was at chance. This left the data from 57 participants to be included in the analysis. The results were manually coded for the correctness of the word form provided by the participants to the respective sentential preamble. The correctness in this case means choosing either a GEN.SG or a GEN.PL form on the final noun, depending on the preceding numeral.

We found that participants indeed made performance errors under the induced experimental conditions, in the form of substituting Genitive singular for Genitive plural and vice versa. A total of 231 errors was found, of which 174 (75%) were false GEN.SG, 42 (18%) GEN.PL, and 15 (6%) miscellaneous errors (e.g. repetition of the lemma). The overall ratio of errors was about 8%, which is comparable to a typical amount of spontaneous production errors in subject–verb agreement reported in the literature (e.g. Franck et al. 2006). The distribution of these performance errors was thus heavily biased towards Genitive singular with a ratio of about 4:1 (Figure 1).

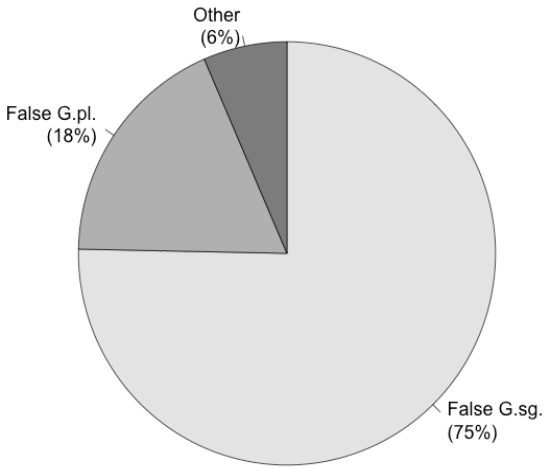


Figure 1

Distribution of performance errors in assigning the correct word form after the numeral, in total.

Our statistical models revealed a robust main effect of the numeral type ($\chi^2(1) = 14.20$, $p = .00016$).[19] There was no main effect of the number of adjectives accompanying the noun ($\chi^2(2) = 3.67$, $p = .159$), and there was a marginal effect of gender ($\chi^2(2) = 5.00$, $p = .0817$), with masculine and neuter nouns being slightly more likely to induce errors than feminine nouns. In other words, in our study the numeral type emerged as the only significant predictor that affected the error ratio.

7.6 *Discussion*

The experiment was intended to determine whether the Genitive singular and Genitive plural may compete with each other as two possible values of (Q-induced) Genitive Case under memory-taxing conditions potentially complicating recall of their respective licensor (the quantificational Q head). If the two values competed with each other, then we could possibly see their distribution more or less at random. The observed overwhelming dominance of Genitive singular errors strongly suggests that Russian speakers do not manipulate the two morphological values on an equal basis. This, in turn, is consistent with the hypothesis that the sources for each value are different. This interpretation is further strengthened in light of our finding that in making those errors, speakers are insensitive to the purely performance-based conditions such as the linear/structural distance between the alleged Q licensor and the noun, which would have affected their likelihood if the distance were measured from the same, single licensor.

Under our proposed architecture of the QP, whereby the sources of Genitive plural and Genitive singular are the quantificational (Q) and the countability (#) heads,

respectively, the results are not surprising. When lacking full access to the respective feature licensor, as evidenced by the errors, speakers choose to encode something that is present *in every item*, irrespectively of the numeral type, namely, countability. Put differently, the observed response and error patterns suggest that, when supplying the respective noun forms, the speakers may have not necessarily remembered what the preceding numeral was, but they were always certain that there was a numeral, and consequently, an atomized plurality in the input. Again, this is consistent with the invariant presence of the syntactic licensor of countability in the syntactic structure, consulted by the speakers' parsing routines.

8. Concluding remarks

This work started with the three well-known agreement puzzles in the internal syntax of Russian numeral phrases, which concern unusual or unexpected patterns of agreement on the noun and accompanying adjective in the presence of a numeral. In the present work we sketched a theory of numeral-based QPs that offers a unified account of these three puzzles. Our proposed account stems from the premise that the structure of a numeral-based QP comprises (at least) three structural layers: the quantificational layer headed by Q, the number layer headed by φ, and the countability layer headed by #. Each of these heads consists of a particular set of features. Some of those features, such as those of #, are subject to diachronic processes. We have shown that the puzzling agreement patterns in Russian QP are better understood if one takes into account the semantic notion of countability as well as the specifics of the diachronic change within the Russian numeral phrase, at the same time avoiding major overhauls of the syntactic (or semantic) theory. We see the main contribution of the present study as mainly integrative, bringing together different numerosity-related components and spelling out

the details of their interaction. We believe our proposals are corroborated by independent cross-linguistic and psycholinguistic considerations (Sections 6 and 7). In view of these results, Russian emerges as a revealing test case for a unified structure of numeral-based QPs along the proposed lines.

Some questions remain open for further study. For instance, the loss of Genitive-assigning ability of Q in the oblique Case environment needs to be specified more precisely, exploring other possible structural conditions under which it may occur. We also deliberately limited ourselves with simple (one-word) numerals, leaving aside complex cardinals such as 'twenty three' or 'one thousand one hundred and twelve', although, given the current understanding of the syntax and semantics of these complex cardinals, we expect that the account proposed here can naturally be extended (see Ionin & Matushansky 2006 for relevant discussion). It also seems promising to explore the consequences for Slavic languages beyond Bulgarian (Section 6), as well as for a larger cross-linguistic sample.

REFERENCES

Abels, Klaus. 2003. *Successive cyclicity, anti-locality and adposition stranding*. Ph.D. dissertation, University of Connecticut.

Acquaviva, Paolo. 2017. *Categorization as noun construction. Gender, number, and entity types.* Ms., University Colledge Dublin.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59.4, 390–412.

Babby, Leonard. 1987. Case, prequantifiers, and discontinous agreement in Russian. *Natural Language & Linguistic Theory* 5, 91–138.

Bailyn, John F. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press.

Bailyn, John F. & Andrew Ira Nevins. 2008. Russian Genitive plurals are impostors. In Asaf Bachrach & Andrew Ira Nevins (eds.), *Inflectional identity*, 237–270. Oxford: Oxford University Press.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function*. Chicago, IL: University of Chicago Press.

Baker, Mark, Kyle Johnson & Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20, 219–251.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67.1, 1–48.

Bayer, Josef, Markus Bader & Michael Meng. 2001. Morphological underspecification meets oblique Case: Syntactic and processing effects in German. *Lingua* 111, 465–514.

Bobaljik, Jonathan David. 1995. *Morphosyntax: The syntax of verbal inflection*. Ph.D. dissertation, MIT.

Bobaljik, Jonathan David. 2002. Realizing germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6.2/3, 129–167.

Bobaljik, Jonathan David & Höskuldur Thrainsson. 1998. Two heads aren't always better than one. *Syntax* 1(1). 37–71.

Bock, Kathryn & Willem J. M. Levelt. 1994. Language production: Grammatical encoding. In Morton A. Gernsbacher (ed.), *Handbook of psycholinguistics*, 945–984. San Diego, CA: Academic Press.

Bock, Kathryn & Carol A. Miller. 1991. Broken agreement. *Cognitive Psychology* 23, 45–93.

Bonet, Eulàlia. 1991. *Morphology after syntax: Pronominal clitics in Romance.* Ph.D. dissertation, MIT.

Borer, Hagit. 2005. *Structuring sense*, vol.1: *In name only*. Oxford: Oxford University Press.

Bošković, Željko. 2006. Case and agreement with genitive of quantification in Russian, In Cedric Boeckx (ed.), *Agreement systems* , 99-121. John Benjamins, Amsterdam.]

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45.1, 27–89.

Brown, Sue. 1999. *The syntax of negation in Russian: A Minimalist approach*. Stanford, CA: Center for the Study of Language and Information.

Carlson, Gregory N. 1977, *Reference to Kinds in English*, Ph.D. dissertation, University of Massachusetts at Amherst. Published in 1980 by Garland, New York.

Cheng, Lisa Lai-Shen & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30.4, 509–542.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6.4, 339–405.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays in Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarretta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambrdige, MA: MIT Press.

Cinque, Guglielmo. 2010. *The syntax of adjectives: A comparative study*. Cambridge, MA: MIT Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Dékány, Éva. 2015. Review of Pesetsky (2013). *Lingua* 159, 18–26.

Despić, Miloje. 2013. A note on paucal, agreement and Case. *Formal Approaches to Slavic Linguistics #20: The Second MIT Meeting*, 57–72.

Doetjes, Jenny Sandra. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English* (HIL Dissertations 32). The Hague: Holland Acad.

Franck, Julie, Glenda Lassi, Ulrich H. Frauenfelder & Luigi Rizzi. 2006. Agreement and movement: A syntactic analysis of attraction. *Cognition* 101, 173–216.

Franck, Julie, Gabriella Vigliocco & Janet Nicol. 2002. Subject–verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes* 17, 371–404.

Franks, Steven. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language & Linguistic Theory* 12, 570–649.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. New York: Oxford University press.

Giusti, Giuliana. 2012. On Force and Fin, Case and Num. In Valentina Bianchi & Cristano Chesi (eds.), *Enjoy linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday*, 205–217. Siena: CISCL.

Giusti, Giuliana & Rossella Iovino. 2014. A split-DP hypothesis for Latin and Italo-Romance. *Olomouc Linguistics Colloquium 2014*, 127–143.

Greenberg, Joseph. 1963. *Universals of language*. Cambridge, MA: MIT Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Halle, Morris. 1990. An approach to morphology. *North East Linguistic Society 20* (NELS 20), 150–184.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2011a. Descriptive and explanatory markedness. *Morphology* 21.2, 223–245.

Harbour, Daniel. 2011b. Valence and atomic number. *Linguistic Inquiry* 42.4, 561–594.

Ionin, Tania & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23, 315–360.

Ivanov, V. V. 1983. *Istoricheskaja grammatika russkogo jazyka* [A historical grammar of Russian]. Moscow: Prosveschenie.

Jakop, Tjaša. 2008. *The dual in Slovene dialects*. Bochum: Brockmeyer.

Kayne, Richard S. 2005. *Movement and silence*. Oxford: Oxford University Press.

Kharalampiev, Ivan. 2001. *Istoricheska gramatika na bŭlgarskiia ezik* [A historical grammar of Bulgarian]. Veliko Tŭrnovo: Faber.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12.5, 607–653.

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 398–411. Chicago: The University of Chicago Press.

Landman, Fred. 1989. Groups, part I. *Linguistics and Philosophy* 12, 559–605.

Lasnik, Howard. 1999. *Minimalist analysis*. Malden, MA.: Blackwell.

Lasnik, Howard, Marcela Depiante & Arthur Stepanov. 2000. Syntactic Structures *revisited: Contemporary lectures on classic transformational theory*. Cambridge, MA.: MIT Press.

Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press.

Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29.3, 375–419.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice theoretical approach. In Rainer Bauerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 302–323. Berlin: Walter de Gruyter.

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–418.

Mel'chuk, Igor A. 1985. *Poverxnostnyj sintaksis russkix čislovyx vyraženij* [Surface syntax of Russian numeral expressions] (Wiener Slawischer Almanach Sonderband 16). Vienna: Institut für Slawistik der Universität Wien.

Mirchev, Kiril. 1958. *Istoricheska gramatika na bŭlgarskiija ezik* [A historical grammar of Bulgarian]. Sofia: Nauka i izkustvo.

Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42.3, 413–444.

Noyer, Rolf. 1997. Features, positions and affixes in Autonomous Morphological Structure. New York: Garland Press.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In Jereon Groenendijk & Martin Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris.

Pashov, Peter. 1989. *Practical grammar of Bulgarian (Prakticheska bâlgarska grammatika)*. Sofia: Narodna prosveta.

Pereltsvaig, Asya. 2010. As easy as two, three, four? *Formal Approaches to Slavic Linguistics 18: The Cornell Meeting*, 417–434.

Pesetsky, David Michael. 2013. *Russian Case morphology and the syntactic categories*. Cambridge, MA: MIT Press.

Peškovskij, Aleksandr Matveevič. 1956. *Russkij sintaksis v naučnom osveščenii* [A scientific approach to Russian syntax]. Moscow: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo Ministerstva prosveščenija RSFSR.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.

R Core Team. 2014. R: A language and environment for statistical computing. R Vienna: Foundation for Statistical Computing. http://www.R-project.org/.

Rappaport, Gilbert. 2002. Numeral phrases in Russian: A Minimalist approach. *Journal of Slavic Linguistics* 10, 329–342.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. London: Longman.

Sauerland, Uli. 2003. A new semantics for number. *Semantics and Linguistic Theory 13* (SALT 13).

Sauerland, Uli. 2008. On the semantic markedness of *phi*-features. In Daniel Harbour, David Adger & Susana Bejar (eds.), Phi-*features*, 57–82. Oxford: Oxford University Press.

Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 413–434. Berlin: Mouton de Gruyter.

Selkirk, Elisabeth. 1977. Some remarks on noun phrase structure. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 285–316. New York: Academic Press.

Shahmatov, Alexander A. 1957. *Istoricheskaja morfologija russkogo jazyka* [The historical morphology of Russian]. Moscow: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo Ministerstva prosveščenija RSFSR.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 243–281. Palgrave: McMillan.

Stepanov, Arthur. 2012. Voiding island effects via head movement. *Linguistic Inquiry* 43.4, 680–693.

Svenonius, Peter. 2004. On the edge. In David Adger, Cecile de Cat & George Tsoulas (eds.), *Peripheries: Syntactic edges and their effects*, 259–287. Dordrecht: Kluwer.

Vigliocco, Gabriella & Janet Nicol. 1998. Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition* 68, B13–B29.

Watanabe, Akira. 2010. Vague quantity, numerals, and natural numbers. *Syntax* 13.1, 37–77.

Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37, 111–130.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

Yip, Moira, Joan Maling & Ray Jackendoff. 1987. Case in tiers. *Language* 63, 217–250.

Zaliznjak, Andrey A. 1967. *Russkoe imennoe slovoizmenenie* [Russian nominal inflection]. Moscow: Nauka.

Authors' address: *Center for Cogntive Science of Language, University of Nova Gorica, Vipavska 13, SI-5000 Nova Gorica, Slovenia*
arthur.stepanov@ung.si, penka.stateva@ung.si

FOOTNOTES

[1] Portions of this article were presented at a colloquium at the Department of Linguistics of the University of Konstanz. We thank two anonymous *JL* referees for their careful reading and thoughtful comments on an earlier draft. This work was financially supported by research project J7-5569 from the Slovenian Research Agency (ARRS).

[2] The same is true for *poltora* 'one and a half', *pol* 'half' and some other fractional quantifiers (Mel'chuk 1985) which we put aside in this work.

[3] The GEN.SG marking on these nouns is provisional. In the course of the discussion that follows, we argue for a different source of the relevant ending. Correspondingly, we will switch to a different marking at some point. See Section 5.2. for details.

[4] Q2 was recognized in the Russian descriptive tradition. For instance, for Peškovskij (1956), a sentence like (2) 'constitutes a typical example of the extent to which grammatical and logical thinking may diverge' (cited as translated in Pesetsky 2013: xiii).

[5] From now onwards, Case in glosses is assumed to be nominative, unless indicated otherwise.

[6] If Modern Russian had retained the dual, the adjectival agreement pattern would still obtain with the lower numerals. One may in fact observe this agreement pattern almost in its entirety in contemporary Slovenian which indeed kept the dual from the Old Church Slavonic stage (although many dialectal varieties appear to be losing it nowadays; cf. e.g. Jakop 2008):

(i)	(a)	(dv-a)	krog-a	(b)	tri-je	krog-i	(c)	štir-je	krog-i
		two-M.	circle-M.DU		three-M.	circle-M.PL.		four-M.	circle-M.PL
		'two circles'			'three circles'			'four circles'	

The only relevant difference between Old Russian (and Old Church Slavonic) and Modern Slovenian in our terms, is that in the latter, but not in the former, higher numerals underwent the numeralization process (cf. footnote 13). In both languages, though, the quantificational layer is a locus of the Genitive assignment on the nouns.

[7] Chierchia (1998) assumes that the conversion from e-type kinds *d* to <e,t>-type predicates of individuals (i.e. atomization) results from a type-shifting operation $^{\cup}(d)$ akin to the type-shifting operation *Id* proposed by Partee (1986). We depart from these works only in assuming that there is a structurally represented operator responsible for that change which gets us to the targed type-raising in a compositional manner. We also follow Chierchia (1998) in treating plurals and mass nouns in a similar vein. Thus, we assume both plurals and mass nouns need to be atomized at some point, although we put mass nouns aside for the time being.

[8] For Watanabe, the feature [±augmented] realizes the successor function. This is made possible using Harbour's (2011a, b) particular feature calculus which allows proliferation of different values of the same feature in the same featural specification. For instance, the number 3 corresponds to [–singular, +augmented, –augmented], etc. See these works for details.

[9] Relevant in the present context, Nevins (2011: 423–424, see also the references therein) discusses so-called 'constructed duals' in Hopi, in which [±singular] and [±augmented] appear on distinct syntactic elements. The difference is that for Nevins, the [±singular] feature is a property of (pro)nouns, whereas [±augmented] is a feature that distinguishes verbs. For us, both features remain within the numeral phrase constituent, even after the splitting.

[10] The suggested lexical entry for the numerals under NumP is inspired by Ionin & Matushansky (2006).

[11] An anonymous *JL* referee correctly points out that under the fission approach, a priori one would probably expect the fissioned heads to be adjacent, whereas in our proposed configuration this is not so. Working out the fission approach in more detail would therefore require a more principled answer to this potential adjacency concern.

[12] It might have been more informative to label the respective varieties of # as something like [low] and [high], instead of [A] and [N]. The suggested specifications are intended to underscore the historical origins of # in Russian. How could this categorical distinction have affected the emergence of the new countability head #? Adjectives, including the numeral ('low') ones in the old language, agree with their

nouns. This ‘agreeing propensity’ is part of their inherent core. Furthermore, adjectivals, including the old numeral ones, usually have a specific syntactic position relative to the noun, that of a modifier, viz. adjunct or a specifier. On the other hand, nouns, including the numeral (‘high’) ones, do not agree with their complements. Furthermore, nominal numerals have a different syntactic position relative to their counted noun: they are usually heads, taking the counted noun as its complement. It is therefore possible that when the proposed ‘numeralization’ took place, the ‘lower’ numeral change A >> Num_A was accompanying by transferring the ‘agreeing propensity’ of the former adjectival, which might or might not have correlated also with its syntactic modifier status, to the new countability head $\#_A$. In contrast, the ‘higher’ numeral change N >> Num_N was not accompanied with this kind of bias, and was based on a slightly different syntactic relation (head). Therefore, the distinction of two versions of # is essentially lexico-syntactic, not semantic, and could be stated in terms of features that encode the respective parts of the diachronic change, perhaps in subcategorizational terms.

[13] Further indirect evidence for two versions of # comes from Modern Slovenian, a language which retained the dual number as well as the adjectival agreement with lower numerals (cf. footnote 6). In present terms, this suggests that Modern Slovenian has $\#_{[N],}$ but not $\#_{[A]}$ in the lexicon, the lower numerals thus retaining their non-quantificational status from Old Church Slavonic.

[14] A few exceptional numerals like *tysjača* 'thousand' still have the Accusative form, perhaps retaining their old categorical status of a noun and the semantic one of a classifier (see Section 5.1).

[15] Concerning the [+singular] setting of φ, and numeral *odin* ‘one’, we exploit the fact that this numeral actually retained its adjectival status in the modern language, even more so than the lower numerals (it agrees with the noun in all genders and numbers), in that the ‘agreeing propensity’ in the sense of footnote 12 above is fully retained on the numeral itself, rather than transferred to the respective value of # (call it $\#_{[A+]}$). Thus, we expect the morphological realization of $\#_{[A+]}$ to be different from $\#_{[A].}$ For the present purposes, we assume the morphological marker of $\#_{[A+]}$ to be null, similarly to $\#_{[N].}$ We also need a special proviso that Genitive is not assigned by Q when the latter selects φ

with the value [+singular]; this can be viewed on a par with similar known configurations, e.g. assigning Accusative Case by *v* in combination with a transitive V in many Minimalist accounts.

[16] As an alternative to the DM view, we may assume a 'mixed' system in which functional nodes can be either featural (whose morphological realization may be determined by postsyntactic rules), or AFFIXAL already in syntax. An example of such a system is a 'hybrid' account of English verbal morphology assuming a 'dual' status of Infl (Lasnik 1999: Chapter 5). Under this alternative view, both φ and Q heads are featural, as before, but $\#_{[A]}$ is affixal. Again, in accord with the general 'affix hopping' intuition, $\#_{[A]}$ must be realized on the complement noun under the condition of strict adjacency. We will not decide among these alternatives here: for the purposes of discussion, we will broadly assume a DM-type approach.

[17] An anonymous *JL* referee suggests that the 'delete first affix' strategy may be less computationally efficient compared to the 'delete last' strategy, as it would seem to involve a more proliferate searching algorithm which, in order to determine whether α is the first affix, would first need to go one level down to see whether β, immediately dominated by α, is an affix or not, whereas in the 'delete last' strategy the last element to be deleted is identified straightforwardly being visible by the spell-out procedure. There is, however, an alternative rationale based on the assumption that the deletion mechanism is 'post-syntactic' enough so as not to refer to syntactic dominance. Rather, the relevant affixation tier (e.g. in the sense of McCarthy 1981 or Yip, Maling & Jackendoff 1987) may be organized computationally as a linear STACK or a QUEUE. The former operates via the 'last-in-first-out' strategy, which would be equivalent to 'delete last' in our case, whereas the latter operates via the 'first-in-first-out' strategy, targeting, instead, the first element. Both deletion algorithms are comparable in terms of computational complexity and may be employed in different empirical parts of the syntax–morphology interface. We leave a further empirical evaluation of this interesting implementational aspect for future research.

[18] This distinction could possibly be made to follow from the contextual/dynamic approach to phases whereby the status of X with respect to phasehood changes depending on the syntactic context in which X occurs (Bošković 2014).

[19] For the statistical analyses, we used linear logistic mixed effects models (Baayen, Davidson & Bates 2008). The numeral type, number of adjectives and gender were used as fixed factors, whereas participants and items were treated as random factors with an intercept. We report *p* values based on the likelihood-ratio test. The models were fitted using the *glmer* function of the *lme4* library (Bates et al. 2015) in the R programming environment (R Core Team 2014). Main effects, including *p*-values, were estimated by using the *drop1* function of the main 'stats' library.