

APPROXIMATION IN RUSSIAN
AND THE SINGLE-WORD CONSTRAINT

Loren Allen Billings

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE SLAVIC LANGUAGES DEPARTMENT

November 1995

[Corrections added as of June 1999/L.A.B.]

© Copyright by Loren Allen Billings, 1995. All rights reserved.

Russian quantifiers are known for their complexity. This dissertation investigates expressions of indefinite quantity—specifically, accusative-assigning *s* ‘about’ of approximate measure.

This preposition has undergone a somewhat unique diachronic change which now requires that its complement consist of only a single word. I chronicle the advent of the single-word restriction (LONE-WD), showing historical data with multi-word complements of *s*. Adjective-noun and numeral-noun complements were once attested; Russian now requires only one word after *s*.

This study investigates various apparent exceptions to LONE-WD, which are violated only under very specific circumstances. These exceptions clarify the morphosyntax of

- paucal numerals (‘two’ through ‘four’ and the fractions *pol* ‘half’ and *čtvrt* ‘quarter’),
- “prequantifier” adjectives,
- syntactic compounds (adjective-noun sequences which inflect separately but are treated by the syntax as a single word), and
- large-quantity numbers (*tysjača* ‘thousand’ and greater).

Distributions of special genitive-singular and -plural forms, assigned only by quantifiers, are shown to be distinct: Only paucal numerals in morphological-nominative case assign “ADPAUCAL” genitive-singular forms (such as end-stressed *čas* ‘hours’); a number of elements, not just numerals, trigger “COUNT” genitive-plural forms (*človek* ‘people’). Other constructions discussed include *okolo* ‘approximately’, approximative inversion, *ětak* ‘about’, and *neskol’ko* ‘several’:

Quantification is not a syntactic category but a semantic feature for which *okolo* is unmarked; *okolo* is quantificational only if its sister is a quantifier. Otherwise *okolo* is merely proximative: ‘near’. Tests confirm that quantificational *okolo* heads a prepositional phrase within the noun phrase. While most prepositional quantifiers

have this structure, accusative-assigning *s* is the relativized head of a hybrid phrase due to featural deficiencies.

Numeral-noun complements of *s* undergo approximative inversion—the noun moving to specifier position—to circumvent LONE-WD. Approximative inversion is likewise subject to a variant of LONE-WD, which requires a single **prosodic** word in the quantified constituent. When inversion is impossible a pleonastic count noun is inserted instead.

An Optimality-theoretic model is proposed, formalizing LONE-WD and constraints requiring prosodic contiguity and exceptions to LONE-WD caused by words expressing more closely defined measure.

Acknowledgments

Thanks to J. Allen, T. Avgustinova, L. Babby, S. Blackwell, L. Bogoslaw, I. Boguslavsky, W. Browne, P. Chaput, C. Chvany, L. DePerno, L. Dorfman, E. Dumanis, G. Fowler, S. Flank, S. Franks, G. Gerhart, F. Gladney, K. Goeringer, J. Jacobson, Ju. Kadukov, H. Kadukova, I. Kadukova, M. Kenstowicz, A. Kofman, N. Kondrashova, K. Křivinková, J. Lavine, A. Lebedev, D. Livingston, H. Loos, A. Marantz, D. Matthews, M. McShane, I. Mel'čuk, O. Mladenova, H. Olmsted, R. Orr, S. Paperno, L. Parrott, R. Patterson, I. Raitsimring, D. Rooker, Z. Saloni, A. Solovyova, A. Statiev, W. Sullivan, G. Thobe, K. Wertz, I. West, E. Williams, M. Yadroff and O. Yokoyama, as well as audiences at Harvard University and AATSEEL (cf. Billings 1993a-b) for helpful comments, discussion, and suggestions. Any errors that appear are, of course, my own doing.

This dissertation would not have been possible were it not for the following who taught me Slavic languages and linguistics: W. Smith, A. Holodiloff, N. Balint, R. Valentine, G. Rappaport, S. Yen, I. Thompson, J.-R. Vergnaud, D. Davidson, G. Webelhuth, R. Brecht, C. Martin, J. Uriagereka, J. Stonham, N. Hornstein, O. Rassudova, C. Townsend, E. Williams, G. Volodina, A. Prince, J. Holub, K. von Kunes, O. Yokoyama, P. Chaput, S. Kuno, M. Kenstowicz, N. Chomsky, and J. Grimshaw. To the memory of P.M. Krichevsky, my very first Russian teacher. Most of all, I am indebted to Leonard Babby, my adviser and mentor, who directed this dissertation.

Thanks to Princeton University, for continued funding, including the Howard T. Behrman Fellowship Fund, and to Rutgers University, for employment during the past academic year and for providing a linguistics department to supplement my Slavic-specific endeavors. Thanks also to the people of New Jersey for expediting the completion of this degree in ways unimaginable. Progress toward this dissertation

during the past two summers was also supported in part by National Science Foundation grant No. SBR-9223725 to Brandeis University; I thank Joan Maling, primary investigator, who employed me as research assistant. Thanks especially to Florida State University, for the upcoming employment that expedited this dissertation.

The aforementioned institutions and people were the result of my own decisions to pursue a Slavic-linguistics career. The following people, however, made even more fundamental decisions long before I was able to choose such a path. Were it not for the intellectual values of my father David Elliott Billings and his mother Ina Eunice Elliott Billings, two people of humble logging and farming roots, I would not have been motivated to seek a higher degree. A letter from her encouraging him **not** to attend law school was also vital to what I am today. Likewise, were it not for three generations of missionary forbears in the provinces of China and the Philippines, and the expressed desire of my mother, Ruth Louise Allen Billings, that I not attend linguistically isolated boarding schools early on, I would not have had the international and bilingual background that has enabled me to analyze human language as an adult. Of these, only one person survives for me to thank personally. Thanks mom.

Finally, thanks to my best friend, Irina Kadukova, to whom this work is gratefully dedicated.

TABLE OF CONTENTS

	Abstract	iii
	Acknowledgments	v
	Table of contents	vii
	Introduction	1
Chapter 1	The diachronic change that restricted <i>s</i> +ACC	9
Chapter 2	A similar-looking yet distinct construction (<i>s</i> +GEN)	15
Chapter 3	Properties shared with other prepositions	18
3.1	Against pluralized ACC complements of <i>s</i>	19
3.2	The <i>v</i> +ACC-of-identity construction	24
3.3	The animate ACC with paucal numerals and prepositional quantifiers	28
Chapter 4	Ruling out multi-word complements of <i>s</i>	35
4.1	Ruling out <i>s</i> + [noun + prepositional phrase]	35
4.2	Accounting for <i>s</i> + adjective + noun _{ACC.SG}	36
4.2.1	Prequantifiers	37
4.2.2	Syntactic compounds	40
4.2.3	Adjectives which specifically delimit a noun's measure	43
4.2.4	Calcified examples	49
4.3	Against <i>s</i> + numeral + noun	53
4.3.1	A brief background of the Russian numerals	54
4.3.2	The behavior of <i>četvert'</i> 'quarter'	59
4.3.3	Measure nouns	68
4.3.4	The large numbers: 'thousand', 'million', etc.	72
4.3.5	The numeral <i>pol</i> 'half'	80
4.4	Complements of <i>s</i> with adnominal-GEN structures	90

4.5	Defining “single word”	91
4.6	Other constructions with a single-word restriction	96
4.6.1	The <i>učit'sja na</i> +ACC construction	97
4.6.2	Single word in the complement of <i>pol</i> ‘half’	99
4.6.3	The so-called second-GEN case	101
4.6.4	Single-word limitations on ADPAUC/COUNT forms	105
4.6.5	Possible single-syllable restrictions	132
Chapter 5	Other approximate-measure constructions	137
5.1	On approximate-measure constructions with <i>okolo</i>	138
5.2	Approximative inversion	162
5.3	Regarding <i>ètak</i> ‘about/approximately’	189
5.4	Regarding <i>neskol'ko</i> ‘several’	191
Chapter 6	Optimality-theoretic treatment of <i>s</i> +ACC	199
6.1	A summary of the crucial data	199
6.2	A brief introduction to Optimality Theory (as applied to syntax)	202
6.3	Formalizing Russian approximative inversion	209
6.4	The treatment of <i>s</i> +ACC proper	219
6.4.1	Formalizing the single-word constraint	219
6.4.2	Formalizing non-numerical exceptions to the single-word constraint	220
6.4.3	Modeling numerical <i>s</i> +ACC complements	223
6.5	The universal viability of the constraints proposed	233
	Conclusion	236
	References	238

In this dissertation I study the issue of approximation in Russian. I specifically investigate a relatively rare construction in modern Russian which is apparently subject to a single-word restriction. It is the ACC-assigning preposition *s*, which ascribes a meaning of approximate measure to its complement:

- (1a) Prošlo *s* **nedelju**. ‘About (a) **week** passed.’
 passed about week
 (V)PAST.NEUT.SG (N.FEM)ACC.SG [= ex. 16a in Babby (1985:100)]
- (1a) *Prošlo *s* **odnu nedelju**. ‘About **one week** passed.’

This preposition is unique in that it tends to require complements which consist of just a single word. For example, the additional word *odnu* ‘one_{(ADJ)FEM.ACC.SG}’, in (1b), is not allowed.

In addition to any theory-driven reasons for studying *s*+ACC (as I will hereafter call this construction) there is the startling fact that this preposition has not, to my knowledge, been treated in any study for more than a couple of paragraphs of commentary or a few examples. Some historical grammars mention *s*+ACC in passing. There are also comparative-historical articles on *s*+ACC overall in Slavic, but they fail to deal with the modern-Russian facts. Even studies of approximative-quantificational or related morphosyntactic phenomena usually merely list *s*+ACC along with other so-called prepositional quantifiers. Dictionaries usually list the ACC-assigning uses of *s* along with the more frequent INST- and GEN-assigning uses of *s* (meaning ‘with’ and ‘off of’, respectively). None, however, has attempted to gather and analyze all the data on *s*+ACC systematically. This dissertation, in addition to its analytical contribution, is therefore also intended to be a repository of empirical data on *s*+ACC. In the course of my exposition I also correct errors in the literature wherever they are observed. For this reason, and because the data will surely outlive any theory, I

present the data in the most theory-neutrally manner possible, relegating the analysis to the final chapter.

There are, in addition to any descriptive goals, several theoretical reasons for investigating *s+ACC*. The data I discuss are of interest to three different schools of linguistics: historical linguistics, Optimality Theory, and Slavic morphosyntax.

First, the single-word restriction illustrates how language changes incrementally, which is of interest to historical linguists. Language, instead of changing abruptly, as any linguist knows, usually undergoes step-by-step change. The single-word restriction, in this case, appears to be one of the steps taken by *s+ACC* in gradual transition from being a fully productive construction to being a marginalized one, perhaps headed toward eventual extinction. Whereas this construction is far from being extinct, it shows indications of dying out. I also show other phenomena in Russian which are subject to a single-word (or -syllable) restriction, many of them also apparently in diachronic transition. This study sheds light on the latest stages of the development of the numeral as a distinct part of speech in Russian.

Next, the single-word restriction must have a mechanism in the grammar to generate the data as attested. The generative-linguistic school, until recently, has not had an adequate mechanism to deal with apparent restrictions on the size of a constituent. I employ a theoretical framework that makes use of output constraints to account for this restriction. This framework is known as Optimality Theory (Prince & Smolensky 1993, as applied to syntax in Grimshaw 1993;1995). Crucial to the Optimality approach is the notion that constraints are **violable**—a particular constraint A can be violated if a more highly ranked constraint B makes a conflicting requirement on the output of the grammar. The construction I study here is of interest to Optimality Theory because the single-word constraint does not apply categorically—it is overridden in certain specific circumstances. That is, whereas the ACC complement

of *s* is usually a single word, this constituent can consist of more than one word, under certain circumstances. The exceptions to the single-word generalization are, at first glance, numerous and unrelated. I present prosodic, morphological, syntactic and semantic constraints, all quite distinct from each other, which interact with the requirement that the ACC complement of *s* be a single word. I conclude that a violable-constraint theory such as Optimality is the only way to account for such seemingly diverse “exceptions” as are observed with *s*+ACC.

Finally, *s*+ACC sheds light on a number of inter-related phenomena often referred to by the blanket term Slavic morphosyntax. The *s*+ACC construction has been referred to as a prepositional quantifier, one of a number of prepositions denoting quantity which have, according to Babby (1980; 1985; 1987), become reanalyzed categorially as “quantifiers”. There are certain plurality and animacy-marking restrictions shared by *s*+ACC and other prepositional quantifiers. That is, whereas ordinary prepositions can have a pluralized complement, quantificational ones cannot. Likewise, quantificational prepositions are unable to take the so-called animate accusative when the complement is quantified by a paucal numeral.¹ Unlike the other prepositional quantifiers, however, *s*+ACC requires a single-word complement, usually an unmodified noun. I show, however, that several types of adjectives **are** allowed between *s* and its ACC-case complement. Some are only apparent exceptions to the single-word restriction; others represent an actual violation of this constraint. When the complement of *s* involves a numeral, generally only the numeral occurs after *s*, with the quantified noun **required** to undergo approximative inversion: instead of **s pjat' nedel'* ‘about five weeks’, the order, if there is a numeral, must be *nedel' s pjat'*. There is one numeral, however, *pol* ‘half’, with unique morphological properties,

¹ The term “paucal”, pertaining to a **few** items, is the term in Slavic linguistics used to refer to the numbers ‘two’ through ‘four’, which have distinct syntactic properties, as I show throughout the study.

which **does** allow overt *s* + numeral + noun order; this special numeral constitutes one of the significant exceptions to the restriction against more than one syntactic word in the complement of *s*.

I also compare *s*+ACC with two other quite common and extensively studied means of expressing approximate measure in Russian: the preposition *okolo* ‘around/near/about/approximately’ and approximative inversion, in which a numeral and the noun it quantifies are juxtaposed to express approximation. New evidence is provided to support a rather controversial proposal about the phrase structure of *okolo*. I make concrete proposals about the syntactic and prosodic structure of approximative inversion as well.

This dissertation has been written with two distinct audiences in mind: For the specialist in Russian and Slavic this study has ample data and explains any terminology or theory new to the Slavic field. The material is also presented in such a way that allows the non-Slavic linguist to follow the argumentation, glossing all data and explaining certain terms common only in the Slavic-linguistic literature. In order to accomplish this twofold aim, it is necessary to provide numerous footnotes and references.²

This study has the following organization: In the first chapter I outline the diachronic change with regard to the more limited distribution of *s*+ACC—to just single-word complements—providing examples of older forms and showing which ones are no longer acceptable to modern speakers. Then, in chapter 2, I briefly dis-

² I transliterate all Cyrillic text using the system (**not** equal to the Library of Congress system) in the back of any issue of *Slavic and East European Journal*. I also render diacritics as they are shown in the sources I quote. I also render the now-archaic orthography as it is shown. For example, Russian no longer has the Cyrillic letter *ѣ* but I gloss it as *ĭ* to distinguish it from the Cyrillic letter *і*, transliterated as *i*. Generally speaking, at about the time of the Russian revolution of 1917 there were orthographic changes, doing away with word-final " and (*inter alia*) merging *ě* with *e*. Certain works published outside of Russia, including Aleksandrov" (1925), continued the pre-revolutionary convention.

cuss a construction that looks deceptively similar to *s+ACC*, one that has been mistakenly interpreted as *s+ACC* on numerous occasions in the literature. I then discuss the properties which *s+ACC* has in common with other prepositional quantifiers, discussing the plurality and animacy restrictions which these prepositions share (chapter 3). In chapter 4, I look specifically at the feature that separates *s+ACC* from other prepositional quantifiers: the single-word restriction. There I investigate three types of violations of this restriction involving prepositional-phrase adjuncts (§4.1), adjectives (§4.2), numeral-noun sequences (§4.3), and adnominal-GEN structures (§4.4); I also determine that the single-word restriction specifically requires a single **syntactic** word (§4.5). I conclude the chapter with a survey of other single-word (and -syllable) phenomena (§4.6). In chapter 5, other common ways of expressing approximation are discussed, including *okolo* ‘approximately’ and approximative inversion. Chapter 6 accounts for the data in the preceding chapters using an Optimality-theoretic constraint hierarchy. There it is argued that *s+ACC*’s single-word requirement is ranked below constraints that require multiple word complements due to semantic reasons and the special morphological properties of *pol* ‘half’.

Before dealing with the particulars I should mention that I rely heavily on the publications and comments of three people, Leonard Babby, Steven Franks and Igor’ Mel’čuk, who have all extensively investigated both numerical expressions and interactions of syntactic and morphological case-assignment in Russian. None of these, however, has proposed a detailed analysis of *s+ACC*. In order to clarify the properties of this construction, I have found it necessary to conduct in-depth analyses of related constructions (which these three researchers and others **have** investigated extensively) that bear on this construction. The reason for this necessity will become clearer though the course of the study.

The seed idea for this project was a comment by Leonard Babby at a history-of-Russian seminar early in 1993, stating that *s*+ACC, unlike other quantificational prepositions, requires a single-word complement—the idea which I pursue the details of in chapter 4 and formalize in chapter 6. This was followed by a misguided attempt to limit the complement of *s* using a prosodic-word criterion (Billings 1993a; 1993b). The essence, however, of the solution to this problem arose from a brief but informative discussion of this construction with William Sullivan in early 1995.

A few comments on phrase-structure notation are in order as well: My approach is generative—i.e., with variously as Government and Binding, Principles and Parameters, or other more recent labels. It is necessary in such a framework to make explicit the phrase structures of the data. With regard to the structure of the Russian noun phrase (NP), there are two general structures that are specific enough to discuss here. I have decided to use as the starting point the specific but somewhat controversial model of the Russian NP in Babby (1987). It not only differs from other generative-syntactic models in being far more complex, with five X-bar levels instead of the customary two, but also differs from most existing models in being very precise, using only a single maximal projection to include all numerals and some prepositions which quantify the noun. Another model in the literature which is sufficiently articulated is the one in Franks (1994; 1995), which essentially translates the distinction between Pesetsky's (1982) NP and quantifier phrase (QP), by using recent advances in functional categories, into a determiner phrase (DP) and QP, respectively. Babby proposes in one five-level projection what Franks does primarily in three slimmer projections: NP within QP within DP. I have determined that Babby's (1987) model is sufficient and even the preferable one of the two. In the course of the dissertation, however, I mention ways in which my approach might be translated into that of Franks (1994; 1995).

I present here, without critical commentary, the essentials of Babby’s (1987) NP phrase structure. I do this in order to distinguish his proposals from any of mine in the course of the dissertation.

Words denoting numbers, many of which were nouns originally, became historically reanalyzed as quantifiers. That is, several hundred years ago (Old Russian) the subject noun phrase of *ta pjat’ butylok prišla*, literally, ‘that_{FEM.NOM.SG} five_{(NOUN.FEM)NOM.SG} bottles_{(NOUN.FEM)GEN.PL} arrived_{(V)PAST.FEM.SG}’ had the structure [*ta pjat’ [butylok]_{NP}*] _{NP}, in which one NP is within another NP. The head of the matrix NP was *pjat’* ‘five’, as evidenced by agreement on both the determiner *ta* ‘that_{FEM.SG}’ and the clausal predicate *prišla* ‘arrived_{(V)PAST.FEM.SG}’. In the modern language the agreement patterns of the same words are drastically different: In *te pjat’ butylok prišli*, literally, ‘those_{NOM.PL} five_{(QUANTIFIER)NOM} bottles_{(NOUN.FEM)GEN.PL} arrived_{(V)PAST.PL}’ the subject NP has the following phrase structure: [*te pjat’ butylok*] _{NP}, in which there is only a single NP, headed by *butylok* ‘bottles_{(NOUN.FEM)GEN.PL}’.

Within this modern-Russian NP are **five** X-bar projections based on five distinct kinds of non-head daughters at each of these levels: NP, the maximal projection, has N^{''''} and a determiner as its two possible daughters. N^{''''} has as its two possible daughters N^{'''} and an adjective phrase (AP) which is **not** within the scope of quantification. N^{'''} can have more than two daughters: Aside from N^{''}, the possible daughters of N^{'''} are a QP and a “prequantifier” AP. N^{'''} constitutes the scope of negation—i.e., the constituent any daughter of which the QP c-commands and to which QP assigns the GEN-of-quantification. The two possible daughters of N^{''} are an AP, an element within the scope of quantification, and N[']. Finally, the two possible daughters of N['] are the N[°] and its complement. The N[°] is the head of this NP. Its complement is usually another NP, which is assigned adnominal GEN case.

If the NP bears syntactic NOM or ACC case—the so-called “direct” cases—and there is a QP under N^{'''}, then only the determiner, the AP under N^{''''} and the Q can bear this direct case. All other constituents, being within the scope of quantification, including the head N[°], are assigned the GEN of quantification. This means that it is possible for the NP to be syntactically assigned one case and for its head to bear a different case.

Two primary kinds of structures can fill the QP, these being a prepositional QP or a numerical QP. I discuss these two structures in detail below (in §4.3 and §5.1). Prequantifiers are adjectives which do not modify the noun or projection thereof but rather express the speaker’s opinion about the quantity being expressed. I also discuss prequantifiers below (in §4.2.1). The detailed discussion of case assignment within the quantified NP I defer as well until the beginning of the section on numerals (in §4.3.1).

As in any detailed linguistic study, however, certain findings require me to depart from Babby (1987). Primary among these is his conception of Q[quantifier] as a part of speech. I argue that Q is instead a feature, possessed by several different syntactic categories. In the spirit of Babby’s quantifier-as-category model, I propose the category called Numeral (Num), containing a very specific kind of quantifier, which is invariably [+ Q].

This concludes the introductory remarks. I now proceed to the diachronic development which makes the modern *s*+ACC construction so unique.

Chapter 1 The diachronic change that restricted *s*+ACC:

Older examples of the *s*+ACC construction—some as late as last century—have nominal complements of *s* followed by numerals or modifiers. The modern standard language³ generally restricts the overt *s*+ACC complement to a single word. In this chapter I show the exact extent to which the *s*+ACC complement has become restricted. I present historical-textual examples and show which ones are still acceptable—in the relevant respects (namely, the bold-faced parts)—to modern-day speakers.⁴

Historical grammars of Russian generally agree that the following change in the language took place since approximately the mid 19th century:

“Combining the preposition *s* [+ACC ...] with numerals in the modern language has been lost, and an approximative meaning is [now] expressed by moving the noun to the front, for example:

³ I specify the **standard** language not because I rely only on literary or normative sources, but rather because there are indications that various modern, non-standard dialects have developed differently: Potapova (1987:80) reports that Perm’ dialects no longer have any ACC-assigning constructions with *s*; Staniševa (1966:134), citing sources from as late as the 1920s, reports instances of the construction *s ètu storonu* ‘on this side_{ACC.SG}’ in (other) Siberian and N. Russian dialect studies. I do not consider this latter use (but Ivšić 1950 does). Demidova (1978:96) lists dialect examples of *s* with an approximative meaning. Matveeva (1954), a work I was not able to consult, appears from its title to address this issue. Suffice it to say that in standard Russian the only use of *s*+ACC is that of approximate measure. Staniševa (1966:134-35) and Dal’ (1991:373) both comment that constructions like *s ... storonu* ‘on ... side_{ACC.SG}’ have since become *s*+GEN in Russian. Potebnja (1941)—originally penned in the 1880s (according to p. 5)—strangely does **not** comment on the archaicism of the following example:

Zajti ... so levuju ... so pravuju storonušku.
go(V)INFIN TO(P) left(ADJ)FEM.ACC.SG to(P) right(ADJ)FEM.ACC.SG side(DIMINUTIVE(N.FEM)ACC.SG
[Potebnja (1941:272), citing Barsov" (1872:74)]

In any event, I do not ignore data considered substandard or colloquial. In a study such as this, with diachronic phenomena, it is necessary to concentrate on one dialect. The data are most plentiful in the standard one, so I use it. None of the crucial data below are either prescriptive or socially stigmatized

⁴ I use an informal corpus of about 100 examples—both old and new—supplemented occasionally by elicited data from informants, from which I determine the current state of affairs.

[2] Časa tam tri xodil. ”
 hour_{N,MASC} there three walked-around_V
 GEN.SG (ADV) ACC PAST.MASC.SG

‘(He) walked around there for about three hours’.

[Bukatevič (1958:146), quoting ex. from Krylov (1946:95); my translation/LAB⁵]

The following characterization refers to just time-expression uses of *s*+ACC:

“The preposition *s* in Contemporary Standard Russian is used to express approximation of a particular period of time **only with nouns not accompanied by numerals**. All of them [i.e., the nouns] enter this construction **without modifier words and only in the singular ...**”

[Lomtev (1956:350); my translation, bold-facing added/LAB⁶]

Mel’čuk (1985:371), stipulates without further explanation that if approximation is expressed using *s*+ACC, then the complement is limited to nouns (including measure nouns like *litr-* ‘liter’).

The following are older examples of *s*+ACC that are no longer acceptable in modern Russian. Note the use of multi-word complements.⁷ My own informants’ judgments regarding the acceptability of the bold-faced portions are shown in parentheses following each example:⁸

⁵ This quote in the original: „Sočettanie predloga *s* takogo značenijsa s čislitel’nymi v sovremennom jazyke poterjano, i značenie priblizitel’nosti vyražetsja vydviženiem na pervoe mesto suščestvitel’nogo, napr.: [followed by ex. (2)]”

⁶ This quote in the original: „V sovremennom russkom literaturnom jazyke predlog *s* upotrebljaetsja dlja oboznačenijsa priblizitel’nosti opredelennogo sroka tol’ko s imenami suščestvitel’nymi bez čislitel’nyx. Vse oni vxodjat v dannuju konstrukciju bez opredeljajuščego slova i tol’ko v edinstvennom čisle ...”

⁷ In example (49) I also show that a multi-word complement which does not involve numbers has remained as a fixed expression. Presumably, such expressions were also productive at one point.

⁸ An asterisk (*) represents unacceptability, while a check mark (✓) represents full acceptability. Question marks represent varying degrees of unacceptability, where one (°) is slightly marginal and more are increasingly worse. In a few instances I use the degree symbol (°), which means something like “I wouldn’t say it myself, but it sounds like something I might hear others say.”

- (3) a inii **s trista ix"** vbegoša vo P"skov". (*)
 other about 300 them rushed to Pskov
 ACC GEN.PL
 ‘and another **about 300 of them** rushed into Pskov.’
 [= ex. 3 in Bukatevič (1958:145), quoting *Ermolinskaja* (1910:86)]
- (4) Est' u tebjā sily **s dvu-menja**, [...] (*)
 about two me
 ACC GEN.SG
 ‘you have the strength **of about two of me**’
 [Širokova (1963:36), quoting Rybnikov (1861:90)⁹]
- (5) A budet" tēx" strjapčix" **s" vosm' sot" čelověk"**. (*)
 about eight hundred people
 ACC GEN? GEN.PL
 ‘And of these lawyers there will be **about eight hundred**.’
 [Bukatevič (1958:145), quoting Kotošixin (1906:25)—originally published in 1840]
- (6) M"gla stojala po rjadu **s" dva mēsjača** (??)
 about two month
 ACC GEN.SG
 ‘(The) gloom hung around **for about two months**.’
 [Lomtev (1956:350), quoting *Suzdal'skaja* (1928:534)]
- (7) Poxodiv, po krajnej mere, **s tri časa** [...] (??)
 about three hour
 (P) ACC GEN.SG
 ‘Having walked for **about three hours** at least ...’
 [*Sintaksis* (1960:183), quoting Čexov's *Moja žizn'* (no cit.)]

Consider, however, the following pre-20th-century examples with numeral complements of *s*, all of which my informants accept. Each of these has the noun to the left

⁹ In example (4) I label *dvu* ‘two’ as “ACC”. Fryšćák (1969:13) reports that the morphological gen of this word was originally *d'voju* (homophonous with the locative case). Due to “interference from the nominal declension” the form became *d(°)vu* and later *dvux*. The form *dvu* was attested until the mid 1700s; see ex. (105d) below. Since it quantifies an animate pronoun, then the animate-ACC form (= the morphological gen) appears in this example. In a footnote following example (14) below I discuss why some monosyllabic numerals are not as unacceptable to modern speakers as polysyllabic numerals. This monosyllabic numeral is in the animate ACC (morphological gen) and no such examples are acceptable in the modern language. See also §3.3 for my discussion of this issue.

of the numeral. (I discuss approximative and emphatic-thematic inversion in §5.2. below)

- (8) **časa s tri** plačjuči [!] u brega stojali. (√ with ČAsa stress; cf. §4.3)
hour about three
GEN.SG ACC
‘(They) stood by the riverbank crying **for about three hours.**’
[Bukatevič (1958:145), quoting Avvakum (1934:67); also in Avvakum (1960:56)]
- (9) **osetrof [sic.] s sorok** (√)
sturgeons about forty
GEN.PL ACC
‘about forty sturgeons’ [Popova (1969:149), quoting Avvakum (1960:17)]
- (10) i stojali dlja radi buri **časov" s" pjat´.** (√)
hours about five
GEN.PL ACC
‘and (they/we) stood **for about five hours** because of (?) the storm’
[Bukatevič (1958:145), quoting Peter (1887:34)]
- (11) Prišlo **čelovek" s" desjat´, so sto.** (First one √; the second ??)
people about ten about hundred
GEN.PL ACC ACC
‘**About ten, about 100 people** arrived.’ [Vostokov" (1831:285; 1839:289)]
- (12) **Pušek" polkovyx" u vas" budet" s" dvadcat´.** (√)
regimental cannons about twenty
GEN.PL ACC
‘(As for) **regimental cannons**, you will have **about twenty.**’
[Bukatevič (1958:145), quoting Peter (1893:50)]
- (13) [...] kupil" igumen" Xristofor" tomu **lēt" s" šestdesjat".** (√)
years about sixty
GEN.PL ACC
‘Abbott Christopher bought ... **about sixty years** ago (?).’
[Lomtev (1956:350), citing Sreznevskij (1906:638), quoting *Pravaja gram. Kirillovu mon.*]
- (14) [...] a morem k Venecěi šli **verst" s" pjat´**
and by-sea to Venice went versts five
INST.SG DAT.SG 3.PL GEN.PL ACC
‘... and (they) travelled by sea for about five versts [1 verst ≈ 1 km.] to Venice.’
[= ex. 4 in Bukatevič (1958:145), quoting „Snošenĵa ...” (1951:875)]

In each of (8) through (14) the order of the bold-faced elements is noun + *s* + numeral.¹⁰

It would appear that under certain circumstances a lone numeral after *s* without a noun **is** acceptable in the modern language despite the preceding excerpts by Bukatevič and Lomtev.¹¹ Note the contrast between the no longer acceptable examples with numerals—examples (3) through (7) above—and these in (8) through (14). The examples in (8) through (14) each have only a single word following *s*. The ones in (3) through (7) each involve multi-word complements of *s*. This in itself suggests that a category-blind constraint restricting the sheer size of the complement is necessary (in addition to possible part-of-speech/plurality restrictions on the complement).

It is not entirely clear why the second part of ex. (11) is unacceptable. I have attempted to test the two parts separately, and yielded the following results: \checkmark *Prišlo čelovek s dvadcat*’, $??$ *Prišlo čelovek so sto*. Example (15a) below, in order to rhyme both prosodically (with penultimate stress on each syntagma) and segmentally (with underlying /o/ reducing to [a] in non-stressed position: [*PROsta* ... *SOsta*]), must have the stress on the preposition. In (15b) the stress is actually marked this way and the phonetic [a] (of underlying final /o/) is spelled as such:

- (15a) *Živi* *prosto, proživeš’* **let** **so** **sto.** (??)
 live simply live-out years about hundred
 (V)IMPERATIVE (ADV) (V)FUT.2.SG GEN.PL (P) ACC

‘Live simply (and) you’ll live-out **about a hundred years.**’
 [Zolotova (1988:222), quoting Dal’ who (1991:373) actually uses the verb *vyživeš’*.]

¹⁰ In (12) there is actually a different order, so-called emphatic-thematic inversion, discussed in § 5.2 below. Nevertheless, there is only a single numeral after the *s* preposition.

¹¹ The informal corpus that I collected includes three examples of apparently modern usage of *s* + numeral_{ACC} + noun_{GEN}, shown below in (110c-e). None was fully acceptable to my informants.

Chapter 2 A similar-looking yet distinct construction:

In this brief chapter I show that one type of example, often considered along with *s*+ACC, which does not assign ACC, but instead the GEN case. These are, in many cases, hard to distinguish from the real ACC-assigning *s* because animacy factors in morphological-case selection.

The following examples are mistakenly labelled as ACC-assigning in some of the handbooks. (Instead of glossing these examples word-for-word, I show the construction in bold-faced type in both the example and the gloss.)

- (16a) Počítali malost', i budet **s vas**.
'You've read a little, and that'll do **for you**_{GEN}' [Ušakov (1940:15), quoting Čexov (no cit.)¹²]
- (16b) Dovol'no **s vas**, raby bezumnye.
'That's enough **for you**_{GEN.PL}, (you) crazy slaves.' [Ušakov (1940:15), quoting Puškin (no cit.)]
- (16c) **s vas** xvátit
'that's enough for you' [Isačenko (1962:576)]
- (17) Xvatit ètogo **s tebjja**.
'That's enough of this **for you**_{GEN.SG}' [Ušakov (1940:15), also quoted in Stang (1956:515)]
- (18a) Skol'ko **s menja**?
'How much do **I**_{GEN.SG} owe?' [LAB]
- (18b) **s" menja** ètogo budet"
'it is enough **for me**_{GEN.SG}' [Aleksandrov" (1923:625)]

¹² Here *vas* is morphologically PL but can be the polite-SG form. I cannot determine from this example the number of addressees. Chapter 3 shows that *s*+ACC cannot take a pluralized complement.

(18c) [...] Ja zastavil tebj_a vystrelit' po mne, **s menja** dovol'no.

'... I forced you to fire upon me, that's enough **for me**_{GEN.SG}.'
 [Slovar' (1962:20), quoting Puškin (1942:15); transliteration modified]

(19) **So staruxi** i ètogo xvatit.

'This should be enough **for the old lady**_{GEN.SG}.' [Ušakov (1940:15)]

This is due to the fact that all declensional classes of nouns except the so-called *-a* class and all personal pronouns exhibit morphological syncretism between the ACC and GEN cases in animate nouns. Furthermore, this construction invariably has animate complements.¹³ Ušakov (1940:15) even lists a singular (SG) noun of the *-a* declension; *staruxi* 'old woman' in example (19) is unmistakably GEN.SG and **not** ACC.SG (which would end in *-u*). Modern Russian remains unchanged with respect to this use of *s*: Only the GEN is allowed (i.e., the ACC, where morphologically distinct, is unacceptable in (19): **So staruxu*_{ACC} *i ètogo xvatit*). This oversight has led to yet other treatments of the *s*+ACC construction, where Isačenko (1962:576) and Stang (1956:515) attempt to fit *dovol'no s vas*-type sentences into the explanation of the truly uniform case-assignment of *s*+ACC data in Russian.¹⁴ It is clearly difficult to distinguish some *s*+GEN and *s*+ACC examples, as is shown in (20).

(20) Rublej **s" pjatok"** izderžal", budet s" menja.
 roubles fiver spent will-be me
 GEN PL ACC (= NOM) MASC SG 3 SG GEN (= ACC)

'He spent **about a fiver** (and) that'll about do it for me.' [Dal' (1991:373)]

¹³ See §3.3 regarding the "animate" (morphological-gen) ACC case.

¹⁴ A less glaring mistake is exhibited by Ušakov's example in (16b). I show in chapter 3 that *s Vas* (= (25) below)—namely, the truly ACC-assigning construction meaning 'about as big as you'—can only refer to one person. The following words *raby bezumnye* 'crazy slaves' show that the speaker is addressing more than one person, thus indicating that this is not a modern *s*+ACC construction. From the fragment of text in (16a) it is impossible to determine how many addressees there are. Isačenko's German gloss of (16c) shows the same type of mistake: *es reicht für euch* 'that's enough for you_{PL}'.

The first, bold-faced PP (*s'' pjabatok''*) is the ACC -assigning approximate-measure construction, while the second, underlined one (*s'' menja*) is the similar, yet deceptive, s+GEN construction.

In this brief chapter I have shown that one group of data, often grouped together with s+ACC, are not really ACC-assigning. The data in (17) through (19) have nothing to do with approximation. I mention them merely because these two similar-looking constructions have been confused in past studies, obscuring any uniform analysis of s+ACC.

Chapter 3 Properties shared with other prepositions:

I show in this chapter that some of the seemingly unique properties possessed by *s*+ACC are shared with a number of other ACC-assigning prepositions which also express quantity. I confirm the validity of the requirement that the complement of *s* must bear SG number (cf. Lomtev quote above in chapter 1). I begin with a detailed treatment of ACC complements of *s*, showing that, for the most part, this restriction is accurate—and unlike the change described in chapter 1—this restriction seems not to be a recent one. I go on to show that this particular restriction is not limited to just *s*+ACC, but applies as well to other so-called prepositional quantifiers, a group of prepositions which quantify their complements and function syntactically as NPs. Specifically, I investigate a construction with semantics quite similar to those of *s*+ACC. I then discuss a related characteristic common to these prepositional-quantifier constructions: the lack of the “animate ACC” when the paucal numbers are involved.¹⁵ I show that *s*+ACC is unique among prepositional quantifiers as it cannot be followed by a numeral-plus-noun complement.¹⁶

¹⁵ The antonym of *paucal* is *multiple* (at least according to Hockett 1958:234, where he discusses a “multiple” number, used for **many** items, in Fijian. Because of the **multiple** meanings of *multiple* I will use “non-paucal” to refer to numerals that do not assign paucal forms to the nouns they quantify.

¹⁶ In §5.2 I show that *s*+ACC phrases, unlike other prepositional-quantifier phrases, cannot be conjoined with (non-prepositionally) numerically quantified NPs.

3.1 Against pluralized ACC complements of *s*

It appears that *s*+ACC does restrict against pluralized complements.¹⁷ The following examples show various types of approximate-measure comparisons using *s*+ACC. In each of these examples the underlined NPs are being compared to the **bold-faced** NPs (i.e., the ACC complements of *s*). It is striking that none of the *s*+ACC complements in (21a-g) is in the plural (PL). Not considering the examples with numerals in the complement of *s*, none of the examples so far in this study has involved a PL-noun complement of *s*. I might add that there appears to be no diachronic contrast with regard to this restriction; even the older examples I collected, as in (3) through (6) above, conform to this characterization.¹⁸

- (21a) Ivolgi, krasivye [...] pticy, veličinoj **s** **golubja,** [...]

Orioles, pretty birds, size dove

(FEM)NOM.PL (ADJ)NOM.PL (FEM)NOM.PL (FEM)INST.SG (MASC)ACC.SG

'Orioles, pretty ... birds, **about the size of a dove,** ...'

[Vlaxov & Muckov (1974:35), quoting V. Arsen'ev (no cit.)]

¹⁷ I have encountered one example which cannot be accounted for in this section. It is from a translation of a work originally published in English. My own informants recoil instantly at this example, but then come to tolerate it somewhat with repeated exposure. I suspect that whoever translated this example spoke Russian natively and was influenced by the PL in the original:

- kolos'ja jačmenja vyšinoj s derev'ja (*)

ear_{(N.MASC)NOM/ACC} barley_{(N.MASC)GEN.SG} height_{(N.FEM)INST.SG} about_(P) trees_{(N.NEUT)ACC.PL}

'ears of barley the height of trees'

[Pete (1984:74), quoting "translation of Swift's *Gulliver's travels*" (no cit.); my glosses/LAB]

¹⁸ Examples (21c-d) have a multiple-word complement of *s*. I address this problem in §4.4 below.

- (21b) Vo sne on videl cvetuščie višni i bol'šix, **s" vorob'ja** monastyrskix mux.
 large sparrow monastery flies
 ACC.SG ACC.SG (ADJ)GEN.PL GEN.PL
 'Asleep, he saw blooming cherry trees and large abbey flies **about-the-size-of (a) sparrow**.'
 [Zolotova (1988:224), quoting Vs. Ivanov (no cit.)¹⁹]
- (21c) Tel'ce u gornostaja gibkoe, dlinnoe, a nožki korotkie, odnako [...] provorny,
 'The ermine's body is supple (and) long, but its legs (are) short, however [...] agile,
černye glaza **so šljapku** **sapožnogo** **gvozdika**.
 black eyes cap shoe nail
 NOM.PL (MASC)NOM.PL (N.FEM)ACC.SG (ADJ)MASC.GEN.SG (MASC)GEN.SG
 (its) black eyes **about-the-size-of the head of a cobbler's nail**.'
 [Zolotova (1988:222), quoting V. Bočarnikov (no cit.)]
- (21d) Mašina, poxožaja na zubovračebnoe kreslo, vybrasyvala s legkim groxotom
 'A machine, similar to a dentist's chair, tossed out with a mild din
ottiski razmerom **s list** **pisčej** **bumagi**.
 prints size leaf writing paper
 (MASC)ACC.PL (MASC)INST.SG (MASC)ACC.SG (ADJ)FEM.GEN.SG (FEM)GEN.SG
prints **about-the size of a sheet of writing paper**.
 [Zolotova (1988:223), quoting Paustovskij (no cit.)]
- (21e) [...] tel'jata byli **s mužickuju** **korovu**. [...]
 calves were peasant's cow
 NOM.PL PL (ADJ)FEM.ACC.SG (FEM)ACC.SG
 '... (the) calves were **about-the-size-of (a) peasant's cow**.'
 [Sajkiev (1955:61), quoting L. Tolstoj (no cit.)]
- (21f) Èti [...] častički veličinoj **s bulavočnuju** **golovku**.
 these particles size pin head
 NOM.PL (FEM)NOM.PL (FEM)INST.PL (ADJ)FEM.ACC.SG (FEM)DIM.ACC.SG
 'These ... particles **about the size of a pin head**.'
 [Babov (1968:171), no citation]
- (21g) Odni rybki **s veršok** i bol'še, drugie ne dlinnee nogtja.
 some fishes veršok
 (ADJ)NOM.PL (FEM)DIM.NOM.PL (MASC)ACC.SG
 'Some fishes (were) **about a veršok** and larger; others (were) no longer than a fingernail.'
 [Zolotova (1988:222), quoting Čexov (no cit.); 1 veršok ≈ 1.75 inches.]

¹⁹ L. Babby and A. Lebedev inform me that 'monastery flies' in (21b) are/were known for being fat since they have no predators. Apparently the monks did not kill them. This detail is pertinent because, though not in the complement of *s*, it pertains to the size of the item being measured. See §4.2.3 below. A. Lebedev points out that a 'peasant's cow' is smaller than the 'landowner's cow', thus it is a yardstick for a "smallish" head of cattle. Pete (1984:74) lists a similar example, but with the adjective *porodistuju* 'pedigreed_{(ADJ)FEM.ACC.SG}', indicating that this is a **big** cow.

(21h)	<u>celye</u>	okoroka	— s	greckij orex
	whole	legs-of-ham		walnut
	(ADJ)NOM/ACC.PL	(N.MASC)NOM/ACC.PL		(MASC)ACC.SG

‘Whole legs of ham (are) about the size of a walnut.’ (cf. also ex. (40d) re *greckij*)
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]

In none of these examples is the ACC-case complement of *s* that is allowed to have morphological-PL inflection. There is no reason, *prima facie*, for at least some of these *s*+ACC complements to be in the PL. In (21a), for example, **each** oriole (*ivolg-*) is being compared to a sparrow (*golub*´).²⁰ Other periphrastic expressions do allow the PL: *ivolgi razmerom podobny golubjam* ‘orioles_{(N.FEM)NOM.PL} size_{(N.FEM)INST.SG} similar_{PL} pigeons_{(N.MASC)DAT.PL}’.

I should explain the INST-case nouns in (21a, d, f) and in many of the examples to come: I am assuming that these words are adverbial in function and adjoined to the prepositional phrase which is headed by *s*. Such adverbs further specify what type of approximate measure is being applied. In all of these examples either *veliĉinoj* or *razmerom* is used, both of which mean ‘size’. In some of the examples to come a more specific word is used—for example, *rostom* ‘height’ in (25), (26) and (27b)—to further specify the dimension measured. I have even found one example with two such words conjoined, shown in (44b) below. If there is no such word, then the default dimension is mass or size in general.²¹

²⁰ There is so-called vowel-zero alternation in some stems which end in more than one consonant when the declensional ending is not vowel-initial. That is, if the declensional ending is either $-\emptyset$ (or consonant initial, i.e., /-ju/, the INST.SG for stems of the *-i* declension) then a vowel—usually /o/, but also /e/ or /i/—appears before the last consonant (cf. Levin 1978:33-36). Since this particular phenomenon does not affect any of the data presented here, I merely show the stems without any notation of the so-called “zerovowel”.

²¹ In time expressions it is also possible in time expressions to have the GEN.SG word *vremeni* ‘time’ along with the *s*+ACC phrase, which, I assume, is adjoined to the PP headed by *s*. See one example in the footnote immediately before ex. (120) below. More rarely, the GEN.SG word *mesta* ‘space/place’ is used, showing a spatial term for the corresponding temporal concept:

Footnote continued on next page

- (24) Vozle doma postavili takuju reklamu, s vorota prjamo!
 near house erected such-a billboard gate straight
 (PP) (V)PAST.3.PL.PAST ACC.SG (N.FEM)ACC.SG (PL)ACC ADV

‘Next to (our) building they put up such a (big) billboard; the size of a gate (it was)!’
 [grat. Ju. Kadukov; my glosses/LAB]

Another type of *pluralia tantum* of sorts is the polite-form of the second-person personal pronoun that is used to address a single person: *Vy*. This word also appears with morphological-PL number. Note that the use of *Vas* (the ACC-case form of *Vy*) here can only have the reading of a single person (thanks to H. Olmsted for this observation).

- (25) On rostom s Vas. ‘He is about your height.’
 he height about you
 NOM (N.MASC)INST.SG (P) (PL)ACC [Aleksandrov" (1923:625)]

Examples like (23) through (25) show that there is no limitation against having PL *s*+ACC complements as such. The only limitation is against pluralizing a noun which is can otherwise exhibit morphological-SG marking. For example, lexical items like *vorota* (23)-(24) or *Vas* (25) have lexically idiosyncratic morphological-PL properties.²⁴

I have shown in this section that the *s*+ACC complement indeed has a restriction against being pluralized. Only *pluralia tantum* elements—i.e., elements that

²⁴ The idiom *about the size of* in English usually takes a SG noun as well (with an indefinite article):

- (i) Tiny artichokes, about the size of a baby’s **fist**, were often brought to my family’s kitchen [...]
 (ii) You wrap these [...] sandwiches—about the size of a baby’s **fist**—in paper towels and [...]
 [Jakobsmeier (1994:21), quoting A. Allegra, M. Jacobson, *Los Angeles Times*, 2.5.93, 9.4.92, resp.]

In these examples the underlined NP is PL. Grat. K. Křivinková for bringing these data to my attention.

lexically require PL morphology—are allowed to appear with ACC.PL inflection.²⁵ I will further consider other so-called prepositional quantifiers which assign the ACC case in the following two sections, beginning with *v*+ACC, which has a meaning quite close to *s*+ACC and, as such, is worth investigating for that reason as well.

3.2 The *v*+ACC-of-identity construction

One preposition appears to mean nearly the same thing as *s*+ACC, which is *v*+ACC in one of its functions. Since there are several other uses of *v* that assign ACC case in modern Russian, I will refer to this function as “*v*+ACC of identity” (following Peškovskij 1956:306). The *v*+ACC-of-identity and *s*+ACC constructions differ in a few ways: Širokova (1963:33), reports that *v*+ACC is used when the meaning is “resemblance” (*podobie*), while *s*+ACC is used to mean “equality” (*ravenstvo*). She quotes the following example:

- (26) Agrippa **licom** **v** **mater´**, **rostop** **s** **otca**
 Agrippa face mother height father
 (FEM)NOM.SG (NEUT)INST.SG (FEM)ACC.SG (MASC)INST.SG (MASC)ACC.SG
 ‘Agrippa has her mother’s face and her father’s height.’
 [Lomonosov (1952:571-72; 1755/1975:204)²⁶]

Peškovskij (1956:306) adds, however, that although *v*+ACC of identity can mean resemblance in general, *s*+ACC is used to compare two items with regard to only a single dimension or characteristic (*priznak*). He supplies the following near-minimal pair:

²⁵ S. Franks has indicated to me that the restriction may be against PL referentiality, a semantic and not a syntactic restriction. I was unable to test this. It seems, *prima facie*, to be a valid line of inquiry.

²⁶ As Širokova (1963:33) points out, there is an apparent misprint in Lomonosov (1952:571), incorrectly listing *po* instead of *v(o)*. This misprint was in the original version, as evidenced in a recent photoreprint of it (Lomonosov 1755/1975:204).

(27a) Aleksej byl v batjušku.
 Aleksej was dad
 (MASC)NOM.SG (V)MASC.SG (MASC)ACC.SG
 ‘Aleksej took after (his) dad.’

(27b) ... s batjušku rostom
 dad height
 (MASC)ACC.SG (MASC)INST.SG
 ‘... about (his) dad’s height’

[Peškovskij (1956:306); also quoted in Gladney (1986)]

The *s+ACC* example in (27b) is restricted to the dimension of height using the *INST*-case word; if there is no such *INST* word, as in the part of (26) after the comma (the *s+ACC* part), then the default dimension is size or mass. Širokova (1963) supplies the following example, apparently to make the same point.

(28) Brovi v nitku, černee saži
 eyebrows thread blacker soot
 (FEM)NOM.PL (FEM)ACC.SG COMPAR (FEM)GEN.SG
 ‘eyebrows like thread, blacker than soot’ [Širokova (1963:35), citing Radiščev (1961:41)]

She specifies that *brovi v nitku* means ‘brows like a thread’ (*brovi kak nitka*), presumably meaning that this construction does **not** mean ‘as thin as thread’. My informants, however, glossed this phrase as ‘eyebrows thin as thread ...’, thus suggesting that the meanings of *s+ACC* and *v+ACC* of identity are often distinguishable. Širokova (1963:36) adds that *s+ACC* has been used throughout the history of Russian for quantificational comparisons. My own corpus of examples of the two constructions confirms this: *s+ACC* is generally used when there is some measurable similarity between the item compared and the “yardstick” item, whether it be length/height/distance, time, volume, or some other easily measurable scale; *v+ACC*

of identity, a construction that compares the resemblance of two heights, weights, etc., can also be used to show the resemblance of one item/person to another.²⁷

These two constructions differ crucially, however, in the way they take complements which include numerals (cf. Bukatevič 1958:143). Recall from chapter 1 that if *s*+ACC takes a complement consisting of a numeral and a noun, then the order must be noun + *s* + numeral (which I discuss at length in §5.2 below); *v*+ACC of identity allows such constituent orders if it has a numeral in its complement, as the examples in (29a-b) show, but does not require it, as in examples (30a-d):

- (29a) *tolščinoju palca [sic.] v dva* ‘about two fingers in thickness’
 thickness_{FEM} finger_{MASC} two
 INST.SG GEN.SG ACC [Širokova (1963:35), citing Peter (1948:304)]
- (29b) *Vošla ženščina let v tridcat’, prijetnaja licom.*
 entered_{PERF} woman_{FEM} years thirty pleasant face_{NEUT}
 FEM.SG NOM.SG GEN.PL ACC (ADJ)_{FEM.NOM.SG} INST.SG
 ‘A woman came in, about thirty years old, with a pleasant face.’
 [Širokova (1963:39), quoting Karamzin’s *Pis’ma ruskogo puteščestvennika* (no cit.)]
- (30a) *Kuplju lentu v tri aršina* ‘I’ll buy a string three aršins in length’
 buy_{PERF} string_{FEM} three aršin [NB: 1 aršin = 0.711 meter]
 1.SG ACC.SG ACC GEN.SG [Bukatevič (1958:143), citing Bardin (1940:254)]
- (30b) *Sejčas utrennik, moroz v tri gradusa*
 now morning-frost_{MASC} freezing_{MASC} three degree_{MASC}
 ADV NOM.SG INST.SG ACC GEN.SG
 ‘there’s a morning frost; it’s three degrees below freezing’
 [Bukatevič (1958:143), citing Čexov (1935:404)]
- (30c) *tok v tri ampera* ‘a three-ampere current’
 current_{MASC} three ampere_{MASC}
 NOM/ACC.SG ACC GEN.SG [Bukatevič (1958:143), no citation]

²⁷ The resemblance of one person’s face to another person’s face in (26) is a classic example: Whereas with computers it is possible to digitally measure the similarity of one person’s face to some other person’s face, the folk understanding is that faces resemble each other in some way that is not quantifiable. Each work in the literature that treats *v*+ACC of identity usually includes a set of examples, like (26), with the meaning of some child ‘taking after’ (i.e., resembling) a parent or other older relative.

- (30d) Kiber-dvornik [...] Siloj rovno v tri medvedja
 cyber-gardener strength_{FEM} equal three bear_{MASC}
 (N.MASC)NOM.SG INST.SG ACC GEN.SG
- ‘(a robot-gardener) equal to three-bear-power in strength’
 [≈ ex. 2a in both Mel’čuk (1981:117; 1985:438), quoting Strugackij & Strugackij (1975:118)]

It would appear that the meaning of *v*+ACC of identity includes not just the similarity (*podobie*) meaning discussed in Širokova, but also a ‘comparison-to-set-measure’ meaning. Whether or not this construction and *s*+ACC crucially differ in their meaning, it is clear from that only *s*+ACC requires the noun-*s*-numeral order.²⁸

Recall from the preceding section that there is a restriction on *s*+ACC which requires that a plurality of items be compared to a morphologically SG *s*+ACC object. Example (28)—and certainly the other *v*+ACC-of-identity examples in Širokova (1963:33-36) and Bukatevič (1958:132, 142)—seem to confirm that a similar restriction holds of the *v*+ACC-of-identity construction as well. To restate this point, while only *v*+ACC of identity can have *s*-numeral-noun order, neither construction appears to be able to take a pluralized-noun complement (without a numeral). I did try to elicit examples of the type *Ona licom v ego brat’ev-bliznecov* ‘She has the face of her brothers (who are) twins.’ Such examples appeared strange, but not

²⁸ While I have not been too specific about the actual differences in the semantics of these two constructions, it is clear that there is not total overlap. For example, in (30a-b) *s*+ACC cannot be substituted for *v* (regardless of whether the noun precedes or follows the preposition and numeral). Additionally, it is possible to combine these two constructions in the same example:

Koridor	byl	v širinu	s polmetra,	da ešče,	požaluj,	i djujma	četyre	sverx	togo.
hallway	was	width	half-meter						
(N.MASC)	(V)PAST	(N.FEM)	ACC						
NOM.SG	MASC.SG	ACC.SG							

‘The hallway was about half a meter wide, and still, perhaps, about another four inches beyond that.’
 [Zolotova (1988:222), quoting A. Grin (no cit.)]

See also example (46a) for another, albeit quite archaic example with both constructions. I assume that the PP headed by *v* is adjoined (adverbial) to the PP headed by *s*, just as INST-case nouns are. I should also point out that I have been referring to these two phenomena as “constructions” not in some formal sense, but merely for convenience. It is perhaps more accurate to refer to the difference between the two as properties of the lexical items *v* and *s*.

ungrammatical. The question, therefore, remains open as to whether these two constructions share an anti-pluralizing restriction for the same reason.

It would also be interesting to test whether other prepositional quantifiers (such as the ones to be discussed in the next section) can have pluralized complements without numbers. It appears from preliminary work with informants that they too cannot have pluralized complements (without being quantified by a numeral).²⁹ Such an excursus, unfortunately, must await future research.³⁰

I have shown in this section that the *v+ACC-of-identity* construction is like *s+ACC* in that it requires its complements to be in the singular. I have also shown that while the two constructions are similar in meaning, their semantic components differ slightly. I will show in the following section that these two constructions, as well as other ACC-assigning prepositional quantifiers, also share the property of disallowing the animate accusative with paucal numerals.

3.3 The animate ACC with paucal numerals and prepositional quantifiers

One property which *s+ACC* clearly, but not obviously, shares with other so-called prepositional quantifiers is the morphological case-marking of an animate ACC-case noun quantified by a paucal numeral. Both restrict against morphological-gen ACC.

Babby (1985) uses the term “prepositional quantifier” to refer to a small group of prepositions which have quantificational force, discussing primarily *po+ACC/DAT*

²⁹ The only rationale I could find in the literature for this is the following one: DePerno (1991:ch. 4:9) suggests without further explanation that a “Q[uantifier] cannot combine with plural noun”. I have not pursued this suggestion further, but it seems to be a valid approach.

³⁰ It is far more difficult to deny the existence of pluralized (non-numerical) quantifiers of the other quantificational prepositions for several reasons: First, whereas the only meaning of *s+ACC* in modern Russian is the one being discussed here, each of *v*, *na*, *za*, *čerez*, and *po* have other uses which assign the ACC case, making it difficult to consult reference grammars and dictionaries. Also, it is not entirely clear as to where some of these prepositions’ semantics and subcategorization overlap. Therefore, it is a much larger task to track down other prepositional quantifiers with anti-pluralization requirements.

‘apiece/each’ and *okolo*+GEN ‘approximately/about’. I do not discuss *po* in detail, cf. (32e), (34e) and (11c) below for examples. I discuss the quantificational uses of *okolo* at length (in §5.1) below. Babby (1985) also lists several other prepositions of this type, including example (1a) above, in which the prepositional phrase has been re-analyzed historically such that this PP is a QP (cf. Babby’s 1987 NP structure outlined immediately before chapter 1 above). All of these prepositions share the property of somehow quantifying their complements.

Generally, when an ACC-assigning word, such as a verb or a preposition, has a complement consisting of a numerically quantified nominal expression in which the numeral is paucal (i.e., a noun quantified by the number ‘two’, ‘three’ or ‘four’) **and** the noun is animate, then the numeral and noun can take one of two case-marking strategies:

- (31a) Ja vižu **četyre** **studenta.** ‘I see four students.’
 I see four students
 NOM.SG 1.SG ACC.**nom** GEN.**SG**
- (31b) Ja vižu **četyrëx** **studentov.** ‘I see four students.’
 I see four students
 NOM.SG 1.SG ACC.**gen** GEN.**PL**

The two forms of ‘four’ in (31a-b), therefore, are the **morphological** -nom and -gen forms of this stem, but both function in the **syntactic** ACC. Where necessary, I will show the morphological realization of a particular case in lower-case and continue to use the small-caps abbreviation to mean syntactic case.

As I explain below (in §4.3.1), if a numerically quantified nominal expression is syntactically assigned a direct case, then the numeral appears in that direct case and the noun appears in the morphological-gen case. If the numerically quantified nominal expression is syntactically assigned oblique case (i.e., not NOM or ACC), then both the numeral and the noun appear in that oblique case. Structures like (31) require

the numeral to select either the morphological-nom or the morphological-gen form to express the syntactic ACC. The literary norm in Russian is for such paucal-numerical structures to select the morphological-gen form if the noun is animate; the spoken language, however, is moving in the direction of using the morphological-nom case in such structures regardless of whether the noun is animate.³¹ It is necessary to note in this discussion that such an animacy distinction is **possible** in the language (even though it may be on the wane).

Though I use a verb in (31), most prepositions also have the same animacy split if they assign the ACC.³² There is, however, a group of prepositions, each of which assigns the ACC case and has a quantificational meaning but does not allow the morphological-gen numeral: *v* (of identity, discussed in §3.2), *na*, *za*, *čerez*, *po*, and *nazad*.³³ I refrain from discussing the internal structure of such prepositional constructions until chapter 5, where I compare two approximative prepositions, *s*+ACC and *okolo*+GEN ‘approximately’; I conclude there (§5.1) that *s*+ACC heads a matrix PP while others head a PP within the NP. Many of these—all except *nazad* (which is actually postpositional)—have non-quantificational ACC-assigning uses in the

³¹ The details are quite complicated but discussed in detail in Grannes (1984; 1986) and Mel’čuk (1980; 1981; 1985:438-52). Babby (1987:111) supplies examples analogous to these, marking as ungrammatical the one analogous to (31a), his ex. 40a. He now admits that this reflects an increasingly prescriptive judgment. Nonetheless, for Babby’s arguments and mine here, it is important only that there **can** be the morphological-gen forms in such structures and not whether it is required.

³² There is one phenomenon, in which verbs of transference trigger the morphological-**nom.PL** of animate profession names—e.g., *idti v soldaty* ‘to-go_{(V)INFIN} to_(P) soldiers_{(N,MASC)nom.PL}’ (= ‘become a soldier’)—is discussed in Ickovič (1980:84-85), Mel’čuk (1985:461-88) and Tolbert (1974:113).

³³ This is not the entire list of prepositional quantifiers, only those which exclusively assign the ACC case. The other two examples of prepositional quantifiers listed in Babby (1985:101) are the following: *Emu bylo let pod sorok*, literally: ‘him_{DAT} was_{(V)NEUT.SG} years_{GEN.PL} toward_(P) forty_{(NUM)ACC}’ (= ‘He was pushing forty.’/LAB) and *Pri areste konfiskovano na 200 tysjač dollarov kokaina i geroina*, ‘at_(P) arrest_{(N)PREP.SG} confiscated_{(PRT)NEUT.SG} worth_(P) 200,000_{(NUM)ACC} dollars_{(N)GEN.PL} cocaine_{(N)GEN.PL} and heroin_{(N)GEN.PL}’ [= his exx. (16b-c, respectively)]. Babby (1991) also lists the following example, in which a numerical range is expressed by the *ot ... do* prepositional pair: [*Ot dvux do trex millionov čelovek*]_{NP:NOM} *poseščajut zoopark* ‘from_(P) two_{(NUM)GEN} to_(P) three_{(NUM)GEN} million_{(N,MASC)GEN.PL} people_{(N,MASC)GEN.PL.COUNT} visit_{(V)PRES.3.PL} (the) zoo_{(N,MASC)ACC.SG} (each year).’ [= ex. 5 in Babby (1991:17)]. (Both *ot* and *do* assign the GEN case to their complements.) Cf. also (35) and (59b).

language, so I will refer to these as “quantificational” *v*, *na*, etc. Each P assigns the syntactic ACC but the morphological nom; I have added the corresponding ungrammatical counterparts—with ACC/gen complements—in brackets.³⁴

- (32a) (siloj rovno) **v** tri medvedja [***v** trex medvedej]
 three bears three bears
nom (MASC)GEN.SG **gen** (MASC)GEN.PL
 ‘(with the power of exactly) three bears’ (i.e., ‘3-bear-power ...’) [ex. (30d) above]
- (32b) (bol’še) **na** dva mal’čika [***na** dvux mal’čikov]
 two boys two boys
nom (MASC)GEN.SG **gen** (MASC)GEN.PL
 ‘two boys (more)’
- (32c) (apel’siny končiliš’) **za** četyre človeka (do menja) [***za** četyrex človek]
 four persons four people
nom (MASC)GEN.SG **gen** GEN.PL
 ‘(the oranges ran out) four people (ahead of me [in line])’
- (32d) (on stojal v očeredi) **čerez** četyre človeka (ot menja) [***čerez** četyrex človek]
 four persons four people
nom (MASC)GEN.SG **gen** GEN.PL
 ‘(he stood in line) four people away (from me).’
- (32e) **po** troe bol’nyx (v palatu) [***po** troix bol’nyx]
 three patients three patients
nom (ADJ)GEN.PL **gen** GEN.PL
 ‘three patients (into each ward)’³⁵
- (32f) dve ženy tomu **nazad** [***dvux** žen (tomu) nazad]
 two wives ago two wives
nom GEN.SG **gen** GEN.PL
 ‘two wives ago’ [quoting Kurt Vonnegut (no cit.)³⁶
 [≈ exx. 2a-f in Mel’čuk (1981:117); also exx. 2a, b, v, g, d, e in Mel’čuk (1985:438)]

³⁴ Since the prepositions themselves are quite difficult to render word-for-word into English, I show the entire phrase (using the glosses in the 1981 English-language version translated by Steven Franks).

³⁵ The numeral form in (32e) is a so-called collective form required when the noun is morphologically adjectival, as is the case here. See the source of this example, Mel’čuk (1981:117; 1985:438) for more discussion of such forms.

³⁶ I have found one example of *nazad* which postposes an *s*+ACC phrase: *s god tomu nazad* ‘about a year ago’. It is unclear whether the ACC case in (32f) is assigned by *tomu nazad* ‘ago’ or by *s*. See also (13) above, where *tomu* ‘ago’ precedes the *s*+ACC phrase. Regardless of the source of ACC case in (32f), the same animacy restriction is observed.

For each of these prepositions, the use of the morphological-gen numeral—and the corresponding GEN.PL noun—is ungrammatical **with these meanings**.

Not surprisingly, the *s*+ACC construction patterns the same way, with one caveat: The only acceptable constituent order in the modern language is noun_{GEN} + *s* +numeral_{ACC}, as in (8)-(14) above, none of these examples showing both a paucal numeral and an animate noun.³⁷ I did find the example in (33a); I have supplied its ungrammatical counterpart in (33b), with the numeral in the morphological-gen case (and the noun in the GEN.PL):

(33a) √ Pošlo **čeloveka** **s** **tri** ‘About three people set out (on foot).’
 set-out persons about three
 (V)PAST.NEUT.SG (MASC)GEN.SG (P) **nom** [Elenskij (1977:51), no citation]

(33b) * Pošlo **čelovek** **s** **trěx**
 set-out persons about three
 (V)PAST.NEUT.SG (MASC)GEN.PL (P) **gen**

Examples (33a-b) show, therefore, that *s*+ACC is likewise restricted from expressing the animate (morphological-gen) ACC with paucal numbers.

The other prepositional-quantifier constructions in (32a-f) can also have such numeral-first order (which has an added approximative meaning, which I discuss in detail in §5.2 below).³⁸

(34a) (siloj) medvedja v tri ‘**about** three-bear-power’
 power bears three
 (FEM)INST.SG (MASC)GEN.SG nom

³⁷ One of the unacceptable examples, (4), has the archaic form of a morphologically gen paucal number, *dvu* ‘two’, but then has a GEN **pronoun**, which makes this example difficult to assess, since personal pronouns **invariably** show ACC/gen regardless even of animacy.

³⁸ I show in §5.2 below that it is also possible to invert just the numeral and the noun, with the preposition in front of both. That order is characteristic of the colloquial register. I also treat such data in §6.4.3 below.

- (34b) (bol'še) mal'čika na dva '(larger [in number] by **about** two boys'
 boys three
 (MASC)GEN.SG nom
- (34c) (apel'siny končiliš') čeloveka za četyre (do menja)
 persons four
 (MASC)GEN.SG nom
 '(the oranges ran out) **about** four people (ahead of me)'
- (34d) (On stojal v očeredi) čeloveka čerez četyre (ot menja)
 persons four
 (MASC)GEN.SG nom
 '(Jura stood in line) **about** three people away (from me).'
- (34e) bol'nyx po tri (v palatu) '**about** three patients (per hospital room)'³⁹
 patients three
 (ADJ)GEN.PL nom
- (34f) ženy četyre nazad '**about** four wives ago'⁴⁰
 wives four ago
 GEN.SG nom

In all the good examples in this section—(32a-f), (33a) and (34a-f)—the numeral is morphologically **nom**; in the bad examples—(38a-f) and (33b)—the numeral is morphologically **gen**. This shows quite clearly that *s+ACC* patterns exactly like the other quantificational prepositions; *s+ACC* is different from the other constructions only in that it cannot overtly take both the numeral and the noun.⁴¹

³⁹ Instead of the so-called collective numeral form *troje* 'three' in (32e), my informants prefer the non-collective form *tri* in the approximative-inversion example. Mel'čuk (1985:147, 149) does not mention morphologically adjectival nouns like the one in (32e) and (34e) in this regard. Mel'čuk points out, however, that collective numerals can be used in approximative inversion only with *pluralia tantum* nouns. In any event, the adjectival noun remains in the morphological-gen case regardless of inversion and the numeral is morphologically **nom** in (32e) and morphologically **nom** in (34e). In fact, some of my informants prefer *tri*—the non-collective form of 'three'—in both examples.

⁴⁰ In (34f) I have changed the number to 'four' for pragmatic reasons: It is easier to envision the approximation when it is a greater number of wives ago. Additionally, *tomu*, a word that often accompanies *nazad* and more characteristic of formal Russian, has been omitted. Cf. the apparently prepositional *tomu* in (13) above. See also the footnote referred to in the citation of example (32f).

⁴¹ One of the prepositional quantifiers listed here, *po* 'apiece/each'—cf. (32e), (34e), (59c), (100c), and (157a)—has unique case-assignment properties. When its complement does not include a numeral, then it obligatorily assigns the DAT case. When there is a numeral, *po* can optionally (although it is extremely archaic) assign the DAT to the numeral (but not to the quantified noun). See Babby (1985) and Franks (1995:139-57) for details. As Franks (1995:144) shows, when the numeral is 'two' through 'four', the DAT is not an option. This restriction is possibly related to the restriction against the animate

In this chapter I have accomplished the following: I began by showing that *s+ACC* cannot have a pluralized-noun complement (§3.1). I also showed that a semantically similar construction, *v+ACC* of identity, likewise has the anti-PL restriction, but differs from *s+ACC* by not restricting against overt complements that consist of a numeral plus a noun (§3.2). Finally, I have shown that all ACC-assigning prepositional quantifiers share a restriction against the option of the so-called “animate” (morphological-gen) ACC with paucal numbers (§3.3). In summation, in this chapter I have isolated those properties which *s+ACC* shares with other prepositional quantifiers. This shows that *s+ACC* is not entirely unique in Russian. In the following chapter I will discuss the single-word restriction, a property which sets *s+ACC* apart.

morphological-gen case. All prepositional quantifiers which **can** assign the ACC appear to be restricted from expressing its complement morphologically with an oblique case if there is a paucal numeral.

Chapter 4 Ruling out multi-word complements of *s*:

In the preceding chapter I showed that some of the apparently idiosyncratic properties of *s*+ACC are in fact shared with other ACC-assigning quantificational prepositions. I showed briefly in the course of that discussion that *s*+ACC is unique among such constructions in not allowing the following surface order: *[*s* numeral noun]. In this chapter I show that this is part of an overall restriction against multi-word complements of *s*. Several example types are considered. I show in each example that the *s*+ACC construction is limited to single-word environments. A few exceptions will remain, those which I will re-assess in the final chapter.

4.1 Ruling out *s* + [noun + prepositional phrase]

One type of *s*+ACC example which appears to violate the single-word restriction is when there are N-plus-PP complements of *s*. The following example is acceptable to my informants only when there is a pause after *sosnu*:⁴²

- (35) rostom s sosnu, ot stanovogo kornja do makuški
height_{FEM} about pine-tree_{FEM} from main_{ADJ} root_{MASC} up-to crown_{FEM}
INST.SG (P) ACC.SG (P) MASC.GEN.SG GEN.SG (P) GEN.SG
- ‘about the height of a pine tree from the main (i.e., deepest) root to the crown’
[Širokova (1963:36), quoting Sadovnikov (1959:167)]

⁴² There is impossible to affirm that a pause existed when Sadovnikov first published this example (in 1876). The comma after *sosnu* suggests that there was a pause even during that period. Note that Širokova failed to quote the comma after *rostom*, which is unfortunate, considering the required intonational break (in modern Russian at least). This comes from a collection of riddles and appears in full in Sadovnikov (1959:167) as follows:

Rostom s sosnu, ot stanovogo kornja do makuški, a ot zemli ne vidat´.
but from ground not to-see

‘About as tall as a pine tree from the apex to the main root, but can’t be seen from the ground.’

Answer: *serdcevina* ‘pith (i.e., the innermost core of a tree trunk)’.

This example apparently requires the pause in modern Russian to avoid a multi-word *s*+ACC complement. That is, I posit the following (simplified) phrase structure:⁴³

(36a) \checkmark [**s** [**sosnu**]_{NP} [*ot stanovogo kornja do makuški*]_{PP}]_{PP}

(36b) * [**s** [[**sosnu**]_{NP} [**ot stanovogo kornja do makuški**]_{PP}]_{NP}]_{PP}

The prepositional phrase(s) beginning with *ot* is either adjoined to the prepositional phrase headed by *s* or attached to a projection of P higher than P°.⁴⁴ Crucially it is not in the complement of *s*. In (36b) the *ot*-PP is adjoined to the noun phrase *sosnu*.⁴⁵ The pause apparently makes explicit the phrase structure in (36a).

I have shown in this brief section that adjunct prepositional phrases are apparently disallowed as part of the *s*+ACC complement. Due to the semantic intricacies of adjunction, I have not attempted to elicit more examples of this type. In the next three sections I look at better understood structures: ACC complements of *s* with either adjectives, numerals, or adnominal-GEN complements.⁴⁶

4.2 Accounting for *s* + adjective + noun_{ACC.SG}

There are various data of the type *s* + adjective + noun_{ACC.SG} that correspond to radically different structures. In this section I assess each of these structures

⁴³ For clarity, *s* and its complement are shown in bold face; I've also enlarged the brackets corresponding to the complement of *s*.

⁴⁴ I refer to the *ot do* phrase here as a PP. Actually, it appears to be two PPs in apposition. Cf. example (59b) for a quantificational use of this pair.

⁴⁵ Again, the structure in (36b) may actually have the PP inside the NP headed by *sosnu*. Crucially, the PP in this unacceptable structure is within the complement of *s*.

⁴⁶ In §5.1 I conclude that an *s*+ACC phrase has a relativized head, with features percolating upward from both *s* and its NP complement. These relativized features result in the equivalent of an NP. The *ot ... do ...* phrase conjoins to that hybrid-NP node.

The adjectives *cel-*, *xoroš-* and *dobr-*, if they are functioning as ordinary (modifier) adjectives, mean ‘entire’, ‘good’ and ‘kind’, respectively. Babby (1987) formalizes a distinction between these and other adjectives in Russian numerical expressions as follows:

(38a) Ja vypil $\sqrt{\text{dobryx/ *dobrye}}$ pjat' butylok vina.
 I drank good(ly) five bottles wine
 NOM.SG MASC.SG.PAST GEN.PL/ACC.PL ACC GEN.SG GEN.SG

‘I drank a good five bottles of wine.’

(38b) Ja vypil $\sqrt{\text{*poslednix/ \sqrt{poslednie}}}$ pjat' butylok vina.
 I drank last five bottles wine
 NOM.SG MASC.SG.PAST GEN.PL/ACC.PL ACC GEN.SG GEN.SG

‘I drank the last five bottles of wine.’

[≈ exx. 55, 54 (resp.) in Babby (1987:118)]

A supplementary explanation is necessary of Babby’s (1987) structure of the quantified noun phrase: As I show immediately before chapter 1 above, Babby argues for five X-bar levels:

(39a) [[[[**dobryx**]_{AP} [[pjat']_{QP} [[[butylok]_{N°} vina_{NP}]_{N'}]_{N''}]_{N'''}]_{N''''}]_{NP}
 good(ly) five bottles wine
 GEN.PL ACC GEN.SG GEN.SG

[≈ exx. 63 in Babby (1987:123)]

(39b) [[[**poslednie**]_{AP} [[pjat']_{QP} [[[butylok]_{N°} vina_{NP}]_{N'}]_{N''}]_{N'''}]_{N''''}]_{NP}
 last five bottles wine
 ACC.PL ACC GEN.SG GEN.SG

[≈ (the phrase structure of) ex. 54 in Babby (1987:118, based on his ex. 79 (p. 134)]

Babby’s system is essentially equivalent to the conventional Government/Binding framework, with N° being the head, with N' being the level at which the adnominal complement is added, and with NP the maximal projection—the level at which a determiner is located. He adds the N'', N''', and N'''' levels based on where other adjectival and quantifier elements attach to the structure: N'' is where an adjective modifier attaches within the scope of quantification; N''' is the scope of quantification; and N'''' is another level at which adjective modifiers attach outside

the scope of quantification. The NP and N' levels are not necessary for the examples in (39a-b); I merely show them to accurately render Babby's model. The structures in (39a-b) are crucial for showing that the adjective *dobryx* in (39a) is immediate daughter to N''', while in (39b) the adjective *poslednie* is the immediate daughter of N'''. This means that only the GEN.PL adjective in (39a) is within the scope of quantification. Moreover, Babby argues for a triple-branching structure in (39a), with the adjective, numeral, and N'' as the three daughters of N'''. In (39b) there is no triple-branching structure: the adjective's only sister is N'''. Babby argues that anything within the scope of the numeral (i.e., c-commanded by the numeral) gets assigned GEN case, except the numeral itself.⁴⁸ Following Babby (1987), I posit the distinctive triple-branching structure in (39a) for prequantifier adjectives.

Apparently prequantifiers are not limited to just numerals: numerals take GEN.PL prequantifiers⁴⁹ but nouns take agreeing prequantifiers, which accounts for the ACC.SG case on *celuju*, *dobroe* and *xorošij* in (37a-c), in which there are no numerals. The GEN.PL on *dobryx* in (37d) is due to the fact that *pol* 'half' is a numeral (cf. §4.3.5

⁴⁸ Babby uses the notation Q[quantifier]P but admits that it is a preliminary notation. This label translates to Num[eral]P[hrase] in my formulation below.

⁴⁹ Franks (1994:610, n. 15) mentions that he has elicited preferences for NOM/ACC.PL prequantifiers when the numeral is paucal.

Vsego-to ostalos' žit' kakie-nibud' polgoda
 in-all_(ADV) remained_{(V.PERFECTIVE)PAST.NEUT.SG} live_{(V)INFIN} only_{(ADJ)ACC.PL} [half-a-year]_{ACC}

'In all, there was only some half a year left to live.' [Chey (1967:105), citing Suprun (1964:98)]

Blažev (1962), a study of adjectives which precede forms with *pol* 'half', consistently reports prequantifier-type adjectives in the GEN.PL and modifier-type adjectives in the NOM/ACC.SG. Tolbert (1974:39) reports that either *celyx polgoda* ('whole_{(ADJ)GEN.PL} half_{(NUM)NOM/ACC} year_{(N.MASC)GEN.SG}') or *celye polgoda* ('whole_{(ADJ)NOM/ACC.PL} half_{(NUM)NOM/ACC} year_{(N.MASC)GEN.SG}') can mean 'a whole half year'. It is unclear, however, whether both these forms have a prequantifier interpretation. Cf. also exx. 109e, h in Crockett (1976:398). Elsewhere, Tolbert (1974:20) lists *celye polčasa*, which he glosses as either 'a whole half-hour' or 'whole half-hours' (p. 20); here he is clearly discussing prequantifiers. My informants reject the last two NOM/ACC.PL prequantifiers with *pol*.

below, where I show that *pol* is a numeral).⁵⁰ It is not necessary at this point to discuss the internal structure of *polversty* ‘half-verst’; the important factor is that the syntactic case assigned to this compound constituent is ACC. I use prequantifiers as a test of various other quantificational phenomena in the rest of the present study. Suffice it to say that prequantifiers are a distinct type of adjective in Russian. The three bold-faced constituents in (37a-d), following Babby’s model, are sisters in a triple-branching structure, thus making *s* and the final noun in each sisters, even though these two sisters are not adjacent, but separated by the prequantifier adjective. Note also that this triple-branching structure entails a structure in which no one sister of *s* is more than a single word in size. Thus, prequantifier adjectives do not constitute a real exception to the restriction against *s* + numeral_{ACC} + noun_{GEN}.

I have shown in this subsection that so-called prequantifier adjectives are not an actual exception to the restriction that the ACC-case complement of *s* consist of no more than one word. The remainder of this section deals with other types of adjectives that also appear to violate the single-word generalization.

4.2.2 Syntactic compounds: In this subsection I look at another somewhat specialized use of an adjectival stem in Russian. Certain combinations of an adjective and a noun

⁵⁰ Non-quantificational nouns (like *koleso* ‘wheel’ and *ogurec* ‘cucumber’ in (37b-c) above) require an agreeing prequantifier adjective. Measure nouns, discussed in §4.3.3, which I propose to be non-numerals but nonetheless quantificational, take either GEN.PL or NOM/ACC.PL prequantifiers. For example, in (37a) the measure word *sotnju* ‘unit-of-hundred’ is in the ACC.SG (and is apparently not allowed to be in the GEN.PL according to my informants: **Pozdravlenij — s celyx sotnju.*). I have found other examples of measure words with GEN.PL prequantifiers (cf. also the preceding footnote):

(za) kakix-nibud’ paru časov. ‘(within_(P)) only_{(ADJ)GEN.PL} a-few_{(FEM)ACC.SG} hours_{(MASC)GEN.PL}’
[DePerno (1991:ch.8:6), a slight modification of ex. 77 in Babby (1987:134)]

The overall nominal expression here is assigned ACC case by the (quantificational) preposition *za*; the measure word is *paru* ‘pair/couple/a-few’; the prequantifier is *kakix-nibud’*, which means ‘only’, cf. Babby (1987:121), Crockett (1976:346) and (Pesetsky 1982:221, n. 32) for other examples of this prequantifier, which these authors gloss as either ‘about’ or ‘some’. Cf. also Crockett (1976:389, incl. fn. 29) regarding agreeing and GEN.PL prequantifiers in non-numerical quantifier expressions.

appear to behave as single syntactic words. I call these phrases “syntactic compounds”. When the *s* has such a constituent as its complement, then there is no violation of the single-word restriction, so long as “word” is interpreted here as a “syntactic word”.

Syntactic compounds (sometimes referred to in the Russian-language linguistic literature as *slovosocetanie*, literally ‘word group’), an individual lexical item that consists of an adjective and a noun, constitute another apparent exception to the restriction against *s* + adjective + noun. For example, *čajnuju ložku* ‘teaspoon_{ACC.SG}’ in (40a) consists of *čajnuju* ‘tea_{(ADJ)FEM.ACC.SG}’ and *ložku* ‘spoon_{(N.FEM)ACC.SG}’. As the adjective-noun order of these elements is fixed, no other **syntactic** elements cannot intrude between the two parts.⁵¹ They are treated by the syntax as atomic (i.e., indivisible) but are inflected as separate words and do not appear to involve any prosodic subordination (i.e., are separate matrix prosodic words). This means that the prosody and morphology treat these groups as separate words while the syntax and lexicon treat them as indivisible units. Such Russian forms readily gloss into English (and other Germanic languages) as morphological compounds—a grouping of prosodically subordinated word-stems that have a morphological rather than syntactic internal phrase structure. Why such structures do not form ordinary morphological compounds in Russian is not entirely clear.⁵²

⁵¹ It would, however, appear that discourse particles can intrude between the two constituent parts of a syntactic compound, as in *Čajnaja že ložka ...* or *Čajnaja ved’ ložka ...*, where *že* and *ved’* are discourse particles that, when exercising sentential scope, usually encliticize to the initial prosodic word of the clause. Cf. Parrott (1992) for further details on discourse clitics in Russian. I have also shown that the yes/no interrogative clitic *li* must follow the first prosodic word (Billings 1994b).

⁵² An interesting example is *železnaja doroga* ‘railway’ (literally: iron_{ADJ} road_N). The noun form is a syntactic compound while the corresponding adjective is a morphological compound: *železnodorožnyj*. The noun forms a syntactic compound while the adjective forms a morphological compound: one reason for this is that a constraint interaction is at play with regard to borrowing such forms (I assume that this is a calque from a Germanic morphological compound such as *Eisenbahn* (German, lit. ‘iron road’). Such a structure may be borrowed as a syntactic compound or as a morphological compound (i.e., the traditional notion of “compound”). The former requires an adjective to be produced but does not require complex morphological or prosodic structure. The latter requires complex morphological

Footnote continued on next page

The following are examples of syntactic compounds as complements of *s*.

- (40a) **s čajnuju ložku**
 tea spoon
 (ADJ)FEM.ACC.SG (FEM)ACC.SG
 ‘about a teaspoon(full)’ [L. Babby (lectures)]
- (40b) **knižka so spičičnyj korobok**
 match box
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘(a) book about the size of a matchbook’ [*Sintaksis* (1980:448)]
- (40c) **Bukaška s bulavočnuju golovku.**
 pin head
 (ADJ)FEM.ACC.SG (FEM)DIM.ACC.SG
 ‘(The) bug is about the size of a pin’s head’ [Ušakov (1940:15); cf. also (21f), (22) above]
- (40d) **Kogda vskryli grudnuju kletku, uvideli oskolok veličinnoj s greckij orex.**
 “Greek”_{ADJ} nut_{N,MASC}
 MASC.ACC.SG ACC.SG
 ‘When the chest cavity was opened up a fragment the size of a walnut was found.’
 [Zolotova (1988:222-23), quoting *Pravda*, June, 1983; cf. also ex. (21h) above]
- (40e) [...] **golova s pivnoj kotel**
 beer vat
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘... (his) head is **about the size of a beer vat.**’
 [Bukatevič (1958:132), quoting Belinskij (1948a:26); also in Babov (1968:172)]
- (40f) **jabloko s pivnoj bočonok**
 apple beer keg
 (N.NEUT)NOM/ACC.SG (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘(an) apple **about the size of a beer keg**’
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]

The examples in (40a-d) each exhibit properties of syntactic compounds. Namely, the constituent *čajnuju*, just as *tea-* in English, does not directly predict that one is

and prosodic structure but does not require the derivation of a new adjective. Optimality-theoretic constraints (cf. chapter 6) could be fashioned to account for this “choice”: NOCOMPLEXWORDS » NOCOMPLEXSYNTAX (the negative approach) or DERIVESYNTAX » DERIVEWORDS (the positive approach). Whichever of these turns out to be the case, it would appear that the Germanic languages (including non-Latinate English) use the opposite constraint rankings, thus preferring morphological and prosodic complexity to syntactic complexity.

speaking about a small spoon. Likewise, in (40b) the constituent *spičičnyj*, only from pragmatic knowledge, reveals that a *matchbox* is of a particular (small) size. In (40c), there is also meaning unrecoverable from the sum of the two parts. The adjective portion *greckij* in (40d) probably originally from ‘Greek’ (cf. the modern-Russian adjective *grečeskij* ‘Greek’), has been lost, leaving behind a few such syntactic compounds; *greckij orex* means a specific variety of nut, and thus a specific size of spheroid. In (40e-f) ‘beer vat’ and ‘beer keg’ specify the size of container; vats and barrels come in various sizes and would be insufficient. Determining whether such a combination is a syntactic compound can also be tested using the adverb *očen’* ‘very’. This adverb cannot modify the first portion of a syntactic compound: **očen’ greckij orex* ‘*a very walnut’. These two-word combinations are clearly single lexemes and are mapped into the syntax as single X° constituents.⁵³ It appears, therefore, that whereas these syntactic compounds consist of two morphological (and prosodic) words, the syntax nonetheless treats these pairs as simplex entities.⁵⁴

I have shown in this subsection that the first constituent in so-called syntactic compounds, like prequantifiers, does not constitute actual violations of the single-word restriction as long as this restriction is interpreted in syntactic terms. In (40a-d) there is no multiple-**syntactic**-word complement of *s*.

4.2.3 Adjectives which specifically delimit a noun’s measure: The one kind of real, modifier adjective that violates the single-word restriction is one which further delimits the measure of the noun complement of *s*. In this subsection I show examples

⁵³ Cf. another Princeton dissertation, in progress as of this writing, by Daniel Rooker on so-called relational adjectives. Rooker informs me that there is evidence to show that despite the separate morphological adjective and noun, the two nonetheless occupy the position of a noun (N°) in the syntax.

⁵⁴ I show additional evidence to support this claim in my discussion of *pol* ‘half’ in §4.3.5 below.

- (43b) Dynja — s **detskuju** **golovu.**
 child's head
 (ADJ)NEUT.ACC.SG (NEUT)ACC.SG
 ‘(The) cantaloupe is **about the size of a child’s head.**’ [Sintaksis (1980:301)]
- (43c) lošadi s novorožděnnogo kotěnka
 horses newborn kitten
 (N.FEM)NOM.PL (ADJ)MASC.ACC.SG (N.MASC)ACC.SG
 ‘horses **about the size of newborn kittens**’ (literally: ‘... a newborn kitten; cf. §3.1)
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (44a) Mne nado byt’ s **goroxovyj** **stručok** rostrom [...].
 pea pod
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘I should be **about the height of a pea pod ...**’
 [Zolotova (1988:222), quoting Šoloxov (no cit.)]
- (44b) [...] razvernul svitok dlinoj i širinoj s **berězovoj** [sic.] **listok**
 unfurled scroll length and width birch leaf
 (V) (N.MASC) (N.FEM) (N.FEM) (ADJ) (N.MASC)
 PAST.MASC.SG ACC.SG INST.SG INST.SG MASC.ACC.SG ACC.SG
 ‘(He) unfurled (a) scroll **about the size of a birch leaf** in length and width.’
 [Pete (1984:73), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (44c) spelye zěrna veličinoj s **krupnuju** **elovuju** **šisku**
 ripe grains size large fir/spruce cone
 (ADJ) (N.NEUT) (N.FEM) (ADJ) (ADJ) (N.FEM)
 NOM/ACC.PL NOM/ACC.PL INST.SG FEM.INST.SG FEM.INST.SG ACC.SG
 ‘ripe (cereal) grains about the size of a {whopping/large}pinecone’
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (45a) Každij korotyška byl rostrom s **nebol’šoj** **ogurec**
 small cucumber
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘Each munchkin was **about the height of a small cucumber.**’ [Nosov (1987:3)]
- (45b) kameški veličinoj s **nebol’šuju** **rybač’ju** **xižinu**
 stones size small fisherman’s hut
 (N.MASC) (N.FEM) (ADJ) (ADJ) (N.FEM)
 NOM/ACC.PL INST.SG FEM.ACC.SG FEM.ACC.SG ACC.SG
 ‘stones **about the size of small fishing huts**’ (literally: ‘... of a small ... hut’; cf. §3.1)
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]

- (45c) V [...] lagere mačta dlja flaga byla vysotoj s **dvuxètažnyj dom.**
 two-storey building
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘At the ... camp the flagpole was **about the height of a two-storey building.**’
 [Babov (1968:171)]
- (45c) stol byl očen´ vysokij — [...] s **dvuxètažnyj dom**
 table was very tall two-storey building
 (N.MASC) (V) (ADV) (ADJ) (ADJ) (N.MASC)
 NOM.SG PAST.MASC.SG MASC.NOM.SG MASC.ACC.SG ACC.SG
 ‘the table was very tall — ...about as high as a two-storey building’
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (45d) brěvna byli tolščinoju s **obyknovennuju trostočku**
 logs were thickness ordinary walking-stick
 (N.NEUT)NOM.PL (V)PAST.PL (N.FEM)INST.SG (ADJ)FEM.ACC.SG (N.FEM)ACC.SG
 ‘(the) logs were **about as big around as an ordinary walking stick.**’
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (45e) krysa veličinoj s **bol’šuju dvornjagu**
 rat size big mongrel
 (N.FEM)NOM.SG (N.FEM)INST.SG (ADJ)FEM.ACC.SG (N.FEM)ACC.SG
 ‘a rat about the size of a large mongrel (dog)’
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (45f) okuni veličinoj s **bol’šuju akulu**
 perches size big shark
 (N.FEM)NOM/ACC.PL (N.FEM)INST.SG (ADJ)FEM.ACC.SG (N.FEM)ACC.SG
 ‘perches (kind of fish) about the size of sharks’ (literally ‘... a shark’; cf. §3.1 above)
 [Pete (1984:74), quoting “translation of Swift’s *Gulliver’s travels*” (no cit.); my glosses/LAB]
- (46a) iskopasta mi jamu nogty svoimi v" glubînu s" **muža stojšče**
 man standing
 (MASC)ACC.SG (PRT)MASC.ACC.SG
 ‘... they dug me a hole with their fingernails to a depth of **about a man standing.**’
 [Staniševa (1966:135-6), quoting Tixonravov” (1863:1973:76)]
- (46b) vysotoju s **čelovečeskij rost**
 person’s stature
 (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘about a man’s stature in height’ [Sinaksis (1980:72)]

The examples in (41) through (46) likewise each involve complements that consist of more than a single word in size. Moreover, it is highly unlikely that any of these words constitute single lexical or syntactic units (i.e., they are unlikely to be examples

meaning of it anymore).⁵⁶ The use of the adjective apparently specifies the specific person being referred to and therefore falls under the same category as the other adjective examples in this subsection. (See also example (51) below for another tall-person-as-yardstick example.) Likewise, the use of the possessive *našu* ‘our’ in (47b) likewise clarifies which kind of string, and therefore a particular (albeit approximate) thickness.

Nevertheless, it is not likely that these are atomic lexical or syntactic constituents, as in the preceding syntactic-compound subsection, and must be accounted for somehow. It appears that modifiers can be added as long as they delimit the meaning of the noun, further specifying the measurement being expressed by the s+ACC construction. Example (48) expresses the extreme of using such adjectives:

- (48) Èta kružka—s tvoju sinjuju farforovuju kitajskuju vazu.
 thy_{ADJ} blue_{ADJ} porcelain_{ADJ} Chinese_{ADJ} vase_N
 FEM.ACC.SG FEM.ACC.SG FEM.ACC.SG FEM.ACC.SG (FEM)ACC.SG
- ‘This mug is **about the size of your blue porcelain Chinese vase.**’
 [grat. O. Yokoyama for coming up with this example]

⁵⁶ *Slovar’* lists a definition of *tambur[-]mažor*: “The main regimental drummer in the French Army of the 17th and 18th centuries and of the Russian Army of the 19th century.” One example listed there refers to a particular height:

Vperedi bėžali [...] mal’čiški, i **vysokŷjtambur"-mažor"** šagal", otmaxivaja takt" bol’šim" žezlom".

‘Out in front ran ... boys, and a **tall drum major** strode, keeping time with a big staff.’
 [*Slovar’* (1963:91), quoting Korolenko (1914:85); shown in original spelling.]

An example of *tambur-mažor* in Gercen (1955:375) adds that such a person is usually tall (Gercen 1955:511). Drum majors were known for their height and thus are a reasonable yardstick to be used as the complement of *s*. *Orfoèpičeskij* (1989:560) lists *tambùrmažór* (presumably to indicate secondary stress on the first part of the compound word, consistent with Russian stress rules), while *Orfografičeskij* (1980:416) lists only the main word stress on the very last syllable; it is not altogether unreasonable to expect the boundary between these two parts to disappear with time (especially if one of the two parts, *mažór* is not a part of the modern language. Nor, practically speaking, is *tambUR* ‘drum’ a word in the modern language, although there is a word *TAMbur*, but it means the vestibule of a passenger railcar and has initial stress). Meta-linguistic evidence for the boundary loss is that the hyphen between *tambur* and *mažor* is, according to *Slovar’* (1963:511), now considered archaic.

Imagine two collectors of fine china at a shop; each is familiar with the other's extensive collections of such vessels. Since both have so many items of various colors, media, origins and types, it is necessary to actually say 'your', 'blue', 'porcelain', 'Chinese' and 'vase' just to rule out any other item known to both speakers. This example, as well as those in (41) through (47), shows that functional considerations override the single-word restriction on the ACC-case complement of *s*.⁵⁷ This does **not**, however, mean that there is no single-word restriction on the *s*+ACC complement. I explain in chapter 6 that the fundamental distinction of Optimality Theory is that constraints are **violated** in order to conform to certain more highly ranked ones. Constraints are no longer the absolutes that typify (and bedevil) previous generative-linguistic theories.

In this subsection I have shown the first actual data which violate the single-word constraint on the complement of *s*. The noun complement of *s* can be modified by an adjective that further specifies the noun's size. Before leaving the issue of adjectives, I investigate one more type of adjective in the complement of *s*.

4.2.4 Calcified examples: There is one final type of adjective that apparently violates the single-word restriction. In this subsection I investigate several calcified expressions—or frozen lexical units—in the modern language that are the etymological result of *s*+ACC, but which no longer have a compositional meaning.

- (49a) ... sena-to na zimu **s gul'kin nos!**
 '... the hay for this winter is **very little** [Zolotova (1988:222), quoting A. Galiev (no cit.)]
- (49b) Deneg u nego — **s gul'kin nos!**
 'The money he has is **very little.**' [= ex. 118 in House (1982:66); gloss modified/LAB]

⁵⁷ Cf. Billings & Rudin (1994) regarding a similar phenomenon in Bulgarian, that of functional considerations overriding syntax.

(50a) Napisat' **s tri koroba** rukovodjaščix statej.

'To write **a lot** of governing rules/articles'

[Ušakov (1935:1471), quoting Saltykov-Ščedrin (no cit.)]

(50b) Emu **s tri koroba** navrěš'

'You can lie to him **a lot**'

[from the film *Priključenija Buratino*; grat. I. Kadukova]

(50c) nagovorit' **s tri koroba**

[transliteration modified/LAB]

'to spin a long yarn, talk the hind leg off a donkey'

[Grosberg (1957:175-176)]

Mel'čuk (1985:28) lists *s gul'kin nos* in (49a-b) as a "non-numerical ('qualitative') characteristic of quantity". The meaning here is clearly non-compositional.⁵⁸ The same is true of *s tri koroba* in (50a-c), which, although etymologically consisting of *s* + numeral + noun (literally: 'about_(P) three_(NUM)ACC baskets_{(N.MASC)GEN.SG}'), is now just a fixed expression that means, according to Ožegov (1983:263), 'to end up saying too much'. This expression, too, is thus non-compositional and does not constitute a productive violation of the restriction against *s* + numeral + noun in the modern language. As frozen expressions, *s gul'kin nos* and *s tri koroba*, do not constitute complements of *s* that consist of more than one word, because there is no longer any productive use of *s*+ACC in such examples.⁵⁹

Likewise, *Petra Velikogo* 'Peter the Great' in (51) is a well known personality. This czar is known to most Russian speakers (even today) to have stood a phenomenal two meters or so tall, thus a distinct yardstick by which to describe a very tall person.

⁵⁸ Apparently the etymology is 'about the size of a dove's beak', where *gul'kin* is the pronominal adjective formed from *gul'ka*, the diminutive of a children's term for 'dove' and *nos* 'nose', which can also mean 'beak'. House (1982:177) mistakenly translates the expression's etymology as 'from a booming nose'. This is clearly not the case. The 'from' meaning of *s* is expressed only if the complement is in the GEN case. I cannot see at all how 'booming' relates to *gul'kin*.

⁵⁹ Lubensky (1995:306, 414) lists these two phrases as well as *s vorob'inyj nos* (same gloss as (49)).

- (51) [...] seržant Mišin, rostom **s Petra Velikogo**... uže posmatrival na časy.
Peter Great
(M)ACC.SG (ADJ)MASC.ACC.SG
- ‘... sergeant Mišin, **about the height of Peter the Great** ...already began to check the clock.’
[Sajkiv (1955:61), quoting A. Fadeev (no cit.)]

The adjective *velikogo* ‘great_{MASC.ANIM.ACC.SG}’ is not a separate syntactic word. Rather, the two-morphological-word combination is a syntactic atom similar to the adjective-noun examples discussed above (in §4.2.2), except that here the non-canonical order of noun plus adjective is attested.

Yet another type of calcified use of *s* is the word *skol’k-* ‘how many/much’, as in example (18a) above: *Skol’ko s menja?* ‘How much do I owe?’⁶⁰ *Spojky* (1980:361) and other etymological references list this item as originally consisting of *s* and the common-Slavic *k*-initial interrogative (*wh*) word meaning ‘how much/how many’. East Slavic is distinct in having *s*-initial forms for this word.⁶¹ Russian no longer has a stand-alone form of this root, but does have etymologically related larger words with the /kol’-/ root (as in *količestvo* ‘quantity’).⁶²

This subsection has shown lexically calcified instances of multi-word complements of *s*, those which do not constitute actual violations of the single-word restriction. Recall that this restriction has come into force in the language during

⁶⁰ Cf. also *stol’ko* ‘so-/as-many’, derived from *s* + the demonstrative-quantity stem.

⁶¹ S. and W. Slavic languages that use this root do not have the initial *s*; for example, Czech *kolik* ‘how-much/-many’. The other E. Slavic languages, Ukrainian and Belarusian, use combinations of *k*- and *sk*- forms with these meanings. In §5.4 I discuss *neskol’ko* ‘several’ and its current part of speech.

⁶² One other calcified use of *s*+ACC is the expression *nebo s ovčin(k)u pokazalos’* (cf., e.g., Isengalieva 1959:142, citing Puškin’s *Kapitanskaja dočka*), which literally means ‘the sky appeared to be about the size of a sheepskin’. Ušakov (1938:745) define this expression as losing the ability to see or discern something due to a strong shock (primarily from fear or pain)”, also listing a similar example by Dostoevskij. Cf. also Dal’ (1989b:641), Lubensky (1995:394) and Zolotova (1988:222). This calcified expression does not constitute a violation of the single-word restriction in any way.

approximately the past century. Such expressions were presumably productive before the single-word restriction went into effect. While I have only found one calcified expression with an adjective-noun complement of *s*, it would seem reasonable to suggest that such expressions were entirely productive in the past. It is thus not coincidental that the loss of numeral-noun and adjective-noun complements happened at the same time; the same single-word restriction applied to both.

This section, which discussed *s* + adjective + noun, has illustrated several facts: Only certain adjective-plus-noun combinations can follow *s*. The most distinct of these is the group of so-called prequantifiers, which, I suggest, actually enter into a triple-branching structure—in which *s*, the adjectival prequantifier, and the noun are all sisters. No one sister of *s* exceeds a word in size (§4.2.1). In addition, I presented so-called syntactic (adjective-noun) compounds in which compelling evidence indicates that the two prosodic/morphological words are a single lexical entry and are mapped into the syntax as an indivisible unit (§4.2.2). I also show that certain actual adjectives can modify the complement of *s* when the adjective serves to further delimit the approximate measure being expressed (§4.2.3). Finally, I list another distinct group of examples in which one or more of the three parts—*s* + adjective + noun—is a fixed expression. In each of these examples there is no **productive** multi-word complement of *s* (§4.2.4) occurs.

The only problematic data are those in which some mechanism is needed to license an “override” of the one-word restriction because the lone-noun complement is not sufficiently specific to delimit the “yardstick” item to which some object or person is compared (§4.2.3). I return to this and other theoretical proposals in the last chapter.

In the remaining sections of this chapter I assess other potential exceptions to the single-word restriction: complements with both a numeral and a noun. I also look at a few other constructions in Russian with single-word restrictions of their own.

4.3 Against *s* + numeral + noun

In addition to the adjective data discussed in the preceding section, there is one other structure which constitutes an apparent exception to the single-word restriction: *s* + numeral_{ACC} + noun_{GEN}.

Comparison of (3)-(8) with (9)-(14) shows quite clearly that the modern language no longer tolerates sequences of this type. Specifically, this **surface** order—e.g., *s dva mesjaca* ‘about two months [\approx (6) above]—is not allowed. Inverting the order—to noun + *s* + numeral—**is** allowed, as (8)-(14) attest: *časov s pjat´*, literally ‘hours about five’ [\approx (10) above] (which I discuss below in §5.2).⁶³ There are, however, several apparent exceptions to the single-word restriction without such inversion. These are examples in which the constituent after *s* and before the noun_{GEN} is one of a limited set of words that are either currently making the diachronic transition to being a numeral or—as well established numerals—are morphologically distinct. The first, *četvert´* ‘quarter’, is historically a FEM noun of the *-i* declensional class. I also discuss certain nouns—which I call “measure nouns”—that are similar to but not identical to numerals. The second, *tysjača* ‘thousand’, is likewise historically a FEM noun but from the *-a* declensional class. The morphologically distinct numeral is *pol* ‘half’ that must form a special kind of morphological compound with its complement noun. Before launching into the particulars, however, I should summarize the diachrony and current state of the system of numerals in Russian.

⁶³ Example (12)—*Pušek polkovyx u vas budet s dvadcat´* ‘(As for) regimental cannons, you will have about twenty’—is not an instance of approximative inversion but rather emphatic-thematic inversion with a predicate filled by an *s*+ACC phrase (with the noun quantified by *dvadcat´* elided).

4.3.1 *A brief background of the Russian numerals:* Russian originally had three morphological numbers: singular (SG), dual (DL) and plural (PL). The words that have come to act as a distinct part of speech, which I will call numerals, were once either adjectives or nouns. Specifically, the stems for ‘one’, ‘two’, ‘three’ and ‘four’ were adjectives, while the stems for ‘five’ through ‘ten’, ‘forty’⁶⁴, ‘hundred’ and ‘thousand’ were nouns. Integers for ‘five’ and greater were built from combinations of these stems. The adjectives for ‘one’ through ‘four’ (including numerical compounds ending in these stems) agreed with the nouns they quantified. Fryšćák (1969:12-20) provides a detailed chronology of the changes in ‘two’ through ‘four’ in Russian.

The numerals for ‘five’ and greater, as nouns, were the heads of their NPs.⁶⁵ As such, they governed the adnominal GEN in the nouns they quantified. Since all nominal number stems were non-SG and non-DL, the nominal numbers triggered specifically the GEN.PL in the nouns they quantified.⁶⁶

During the past millennium the morphological DL was lost, leaving only SG and PL. In a large number of the nominal declensional classes the NOM/ ACC.DL forms⁶⁷ happened to be homophonous with the GEN.SG of the same declensional class. For example, *sestry* was both the GEN.SG and the NOM/ACC.DL of ‘sister’, a noun of the *-a* declension; *roda* was the same two forms for ‘lineage/family/stock’, of the *-Ø* (or *-ǔ*) declension.⁶⁸ Since the nouns for ‘five’ and greater triggered the GEN.PL and

⁶⁴ The word for ‘forty’, *sorok*, was a noun used to refer to a common quantity that later replaced the compound ‘four-ten’ word as the numeral in modern Russian. This may be an instance of a measure noun that came to replace the number word as the numeral. Cf. Schütz (1986), which shows archaic examples of *sorok* in the plural, which is a good indicator that this was a noun.

⁶⁵ I largely ignore compound numerals like *dvadcat’ tri* ‘twenty-three’. Cf., however, Mayer (1974).

⁶⁶ See Babby (1985: §5.2; 1987, fns. 13-14) for a syntactic explanation of the category changes (nouns and prepositions becoming quantifiers) to which I refer frequently in this subsection.

⁶⁷ The NOM.DL and ACC.DL were homophonous in all declensional classes.

⁶⁸ Cf. Ivanov (1990:246-49) for declensional-class tables.

‘two’ **appeared** to trigger the GEN.SG, the erstwhile NOM/ACC.DL (i.e., the adjective for ‘two’ plus the apparent GEN.SG of the noun) was re-analyzed as the paucal counterpart of the GEN.PL-assigning larger numbers. For various reasons, the words for ‘three’ and ‘four’ also began to trigger the GEN.SG in the nouns they quantified. The morphologically adjectival stems for ‘two’ through ‘four’, often referred to as the “paucal” numbers, continued to agree with the nouns they quantified in the non-NOM or -ACC (or “oblique”) cases, but triggered the GEN.SG in the quantified nouns if the entire nominal expression was syntactically assigned the NOM or ACC (the “direct” cases).

The erstwhile nominal stems for ‘five’ and greater likewise began to trigger the GEN.PL in the nouns they quantified when the entire nominal expression was assigned a direct case. However, they began to agree in case with the noun when the entire expression was assigned an oblique case. This entailed the loss of a distinct morphological-PL paradigm in the nouns for ‘five’ and greater. For example, whereas *pjat* ‘five’ had a full PL paradigm when it was a noun; it now has no morphological PL in its own declensional paradigm as a numeral. The result is that all numerals agree in case with the noun they quantify if the entire nominal expression is assigned an oblique case, and they assign GEN to the same noun (and themselves appear in the ambiguous NOM/ACC) if the overall expression is assigned a direct case. The morphological number of the GEN-case noun is **SG** when the direct-case numeral is ‘two’ through ‘four’ and is GEN.**PL** when the direct-case numeral is ‘five’ or greater.⁶⁹

⁶⁹ Any adjectives modifying the quantified noun are in the morphological PL even if the noun itself is in the GEN.SG (i.e., following a direct-case numeral for ‘two’ through ‘four’). The morphological case of the adjective is not easy to describe and, in any event, not pertinent to this discussion. Naylor (1977) argues convincingly that although Russian nouns quantified by a paucal numeral appear to be in the GEN.SG, this is not really what is taking place. I use his arguments, bolstered by additional polemics of my own, to propose a similar account for these data. Although I agree with Naylor’s reasoning, which can be traced back to Zaliznjak’s (1967:46-48) and Isačenko’s (1962:529-30) proposed count “case”, I continue to use the labels GEN.SG and GEN.PL to refer to the forms governed by numerals. See below in this section (in §4.3.4), as well as §4.6.4 below.

To date, all integers except for the very highest (i.e., *tysjača* ‘thousand’ and larger) are syntactically numerals. (I discuss *tysjača*, which is currently transitional in this respect, in §4.3.4 below.)

The word for ‘one’ has remained an adjective in the modern language, agreeing with the noun which it accompanies in case and gender. It is not a numeral. For example, ‘one’ never causes the noun it modifies to take a different case.

Certain words denoting fractions have become numerals: *pol*, one way to say ‘half’, is invariably a numeral in the modern language, and *čtvrt* ‘quarter’ appears to be in transition from noun-hood to numeral-hood (see §4.3.5 and §4.3.2 below, respectively).

The end result is that Russian has a distinct class of words which function syntactically in a cohesive way, which I call numerals. The numerals for ‘five’ and greater tend to act as a subgroup, while the words meaning ‘two’, ‘three’, and ‘four’ (and, to some extent, the fraction numerals) act as another subgroup, which I will call the “paucal” numerals. Numerals differ morphologically from other nominal categories (i.e., pronouns, pronominal adjectives, nouns or adjectives) in that their inflectional paradigms no longer have separate SG and PL morphological numbers; numerals have a form for each morphological case but not separate SG and PL forms for each case. Numerals differ syntactically from other nominal categories in that they all show an asymmetry between direct (NOM and ACC) and oblique (all other) cases. When a direct case is assigned to the overall nominal expression containing a numeral, then the numeral bears that direct case and triggers the GEN in the noun which it quantifies; if the overall expression is assigned an oblique case, then both the numeral and the noun take that oblique case. I use these morphological and syntactic properties as diagnostics of numeral-hood of various words below.

It is also important to mention that there are special GEN.PL and GEN.SG forms used only after numerals or other quantifiers. For example, the MASC noun *čas* ‘hour’ ordinarily has the GEN.SG form *ČAsa*, with initial stress.⁷⁰ When this noun is quantified by a paucal numeral—i.e., one of the numerals that assigns the GEN.SG—a special form is used, with final-syllable stress: *čaSA*.⁷¹

Franks (1995:52) observes yet another instance of a distinct adpaucal form, this time in FEM surnames which end in *-ina* or *-ova*. Such surnames have a mixed paradigm, declining like nouns of the *-a* declension in the direct (NOM and ACC) cases and like adjectives in the oblique (i.e., all other) cases. Thus, the GEN.SG of the surname *Puškina* is usually *Puškinoj*, but the ADPAUC is *Puškiny*, which is the form expected of any **noun** of the *-a* declension. Franks uses this fact to support his proposal that the GEN case assigned through quantification is a direct case, as opposed to the oblique case. The direct/oblique distinction is usually between NOM and ACC on the one hand and the remaining cases, including GEN, on the other (respectively). The same facts are also in Tolbert (1974:29).⁷²

⁷⁰ I show main word stress by placing the entire stressed syllable in upper case and secondary word stress with small caps. I deviate from the traditional custom of using acute and grave stress (respectively), because one Cyrillic vowel letter, è, is transliterated with a grave accent mark, thus making it difficult to use the traditional stress notation clearly; I will need to show secondary stress quite frequently. For example: *POLčaSA* ‘half an hour’.

⁷¹ The same is reportedly true of *šag* ‘pace/step’, which is end-stressed with *dva*, *tri*, *četyre* and *pol*. (grat. C. Chvany, J. Lavine, and O. Yokoyama, who each separately pointed this out to me; see also Žalznjak 1987:147, 231 and *Orfoèpičeskij* 1989:623.) In addition to these, Fowler (1988:41; 57, n. 28; 59, n. 46) lists three more nouns that behave similarly: *rjad* ‘row’, *sled* ‘trace’, and *šar* with the specific meaning of ‘billiard ball’. Additionally, C. Chvany informs me that *raz-* ‘time/instance’ also behaves like these other nouns. Mel’čuk (1985:323) reports that only the non-ADPAUC GEN.SG forms of *šag-*, *rjad-* and *šar-* can be used if preceded by *četvert’*, which he uses as evidence against the numeral-hood of *četvert’* (he does not include *sled-* in this list). Unfortunately for these purposes, the combination *četvert’ šaga* ‘a quarter step’ is pragmatically odd and therefore somewhat useless to this investigation.

⁷² Tolbert also mentions that MASC surnames, like *Puškin* quantified by a paucal numeral in an NP syntactically assigned NOM case can be either *tri Puškina* ‘three_(NUM)NOM Puškin_{(N.MASC)GEN.SG}’ or *tri Puškinyx* ‘three_(NUM)NOM Puškins_{(ADJ)GEN.PL}’. When the overall NP is assigned ACC case, Tolbert adds, then the form is *trex Puškinyx* ‘three_{(NUM)gen} Puškins_{(ADJ)GEN.PL}’; my informants also allow *tri Puškinyx* ‘three_(NUM)NOM Puškins_{(ADJ)GEN.PL}’. See §3.3 regarding the so-called animate ACC.

Other nouns have special GEN.PL forms when they are quantified: The most common example is the noun which means ‘person’ or ‘people’. The NOM.SG is *čelovek*; the rest of the SG paradigm consists of adding monosyllabic case endings to the stem *čelovek-*. Just as in English, where ‘person’ is usually the SG stem and ‘people’ is the PL stem, there is suppletion in this Russian word. The PL paradigm is generally formed from the stem /ljudj-/; specifically, the GEN.PL is *ljudej*. The distinct GEN.PL form is *čelovek* (which happens to be homophonous with the NOM.SG of this word), and is used only after certain quantificational elements.⁷³

Such special GEN-case forms in the literature are referred to by various names: *adnumerative*, *Count I and II*, *quantification form*, *numeral form*, *paucal*, etc. I use the following terms (see §4.3.2, §4.3.4, and §4.6.4): “ADPAUC” refers to the special form that segmentally resembles the GEN.SG of a noun but has final-syllable stress and is quantified by a paucal numeral in the morphological-nom case, while “COUNT” refers to the special inflectional form that replaces the GEN.PL (e.g., *čelovek* instead of *ljudej* mentioned above). These separate terms are preferable to a single term, like *adnumerative*, because these two forms have quite distinct distributions—except for a single-word restriction they share (discussed in §4.6.4)—which I show below in this section.

I now return to a discussion of the four primary types of exception to the generalization that modern Russian does not tolerate the overt sequence *s* + numeral +

⁷³ Strangely, as Mel’čuk (1985:430) points out, no single noun appears to have both (GEN.SG) ADPAUC and (GEN.PL) COUNT forms. Since so few words have these forms, this may just be coincidence. He also mentions more extensive phenomena of this type in two other Slavic languages: In Ukrainian nouns usually bear NOM.PL when quantified by a paucal numeral; a very large group of MASC nouns has a special ADPAUC form with the segments of the NOM.PL ending (-y) but the prosody of the GEN.SG of that same word. An apparent analogue of the Russian COUNT is observed in Bulgarian, in which nouns no longer have overt case marking. Thus, nouns generally have one SG and one PL form each. When a noun is quantified by a numeral it shows the PL (as in English). There are nouns that have a special “second” PL used only if quantified. What is especially telling about Bulgarian (which Mel’čuk fails to observe) is the following: Because the language has no overt case-marking on nouns, it is unlikely that the “second PL” is a case.

noun: *čtvrť* ‘quarter’ (§4.3.2), measure nouns (§4.3.3), large numerals like *tysjača* ‘thousand’ (§4.3.4), and finally *pol* ‘half’ (§4.3.5).

4.3.2 *The behavior of čtvrť ‘quarter’*: The word *čtvrť* is noteworthy because of the interaction of the diachronic phenomenon discussed in chapter 1 with another recent diachronic phenomenon wherein *čtvrť*, formerly just a noun, has begun to take on properties of a numeral. As is mentioned briefly in the preceding section, one test for numerals is the assignment of special (ADPAUC) GEN.SG forms to certain nouns. For those words that are exclusively numerals in the modern language, such as *tri* ‘three’, only the ADPAUC form is allowed. Only the following form is allowed: *tri čaSA* ‘three_{NOM/ACC} hours_{GEN.SG(ADPAUC)}’, (not **tri ČAsa*). In the case of *čtvrť*, both stresses are allowed on the quantified noun, suggesting that *čtvrť* functions either as a noun (triggering non-ADPAUC stem stress on *ČAsa*) or a numeral (triggering end-stressed *čaSA*):⁷⁴

(52a)	<i>čtvrť</i>	ČAsa	(52b)	<i>čtvrť</i>	čaSA
	quarter	hour		quarter	hour
	(NOUN) _{NOM/ACC}	(NOUN) _{GEN.SG}		(NUM) _{NOM/ACC}	(NOUN) _{GEN.SG.ADPAUC}

The fractions are especially helpful to this investigation since they either have distinct noun and numeral variants or are undergoing a part-of-speech change: *polovina* ‘half’ was (and still is) a noun, while the etymologically related shorter form *pol*, which also means ‘half’, acts only as a numeral; hence the stress distinction *polovina ČAsa* vs.

⁷⁴ Gladney (1986:141) assesses *čtvrť časa* without considering the stress or part-of-speech facts. This is unfortunate, because it is unlikely that end-stressed *čaSA* receives its case through the “adnominal genitive rule”, as he suggests; the GEN assigned by nouns is invariably the non-ADPAUC form (i.e., *ČAsa*). Chey (1967:39, 42-43, 63) considers *{ěti /každye}čtvrť časa* ‘{these/each}_{PL} quarter hour’ and likewise without considering the stress on *časa*. Chey considers this PL agreement to be “exceptions”, “the use of the plural adjective with a noun quantifier expressed in the singular”. “This syntactic peculiarity is shared by other quantifiers, namely, numerals.” If the stress of *časa* is taken into consideration, then there is little doubt that *čtvrť* is indeed a numeral in such examples.

polčaSA which both mean ‘half hour’. (I discuss *pol* in detail in §4.3.5 below.) Similarly, *čtvrt’* ‘quarter’ is historically a noun but has almost completely transformed into a quantifier for most modern speakers.⁷⁵ As the following table shows, *čtvrt’*, as a **noun**, has its predictable FEM.SG agreement features, as manifested on both modifier (*èta* ‘this_{FEM.SG}’) and predicate (*byla* ‘was_{FEM.SG}’/*budet* ‘will-be_{3.SG}’). When acting as a numeral, the nominal expression which includes *čtvrt’* triggers PL agreement (*èti* ‘these’ and *byli* ‘were_{PL}’/*budut* ‘will-be_{PL}’):

(53a)	*	<i>èta</i> this _{NOM.F.SG}	<i>čtvrt’</i> quarter _{(NOUN.FEM)NOM.SG}	<i>čaSA</i> hour _{GEN.SG(ADPAUC)}	<i>byla/budet ...</i> was _{FEM.SG} /will-be _{3.SG}
(53b)	√	<i>èta</i> this _{NOM.F.SG}	<i>čtvrt’</i> quarter _{(NOUN.FEM)NOM.SG}	<i>ČAsa</i> hour _{GEN.SG(NON-ADPAUC)}	<i>byla/budet ...</i> was _{FEM.SG} /will-be _{3.SG}
(53c)	√	<i>èti</i> these _{NOM.PL}	<i>čtvrt’</i> quarter _{(NUMERAL)NOM}	<i>čaSA</i> hour _{GEN.SG(ADPAUC)}	<i>byli_{PL}/budut_{3.PL} ...</i> were _{PL} /will-be _{3.PL}
(53d)	*	<i>èti</i> these _{NOM.PL}	<i>čtvrt’</i> quarter _{(NUMERAL)NOM}	<i>ČAsa</i> hour _{GEN.SG(NON-ADPAUC)}	<i>byli_{PL}/budut_{3.PL} ...</i> were _{PL} /will-be _{3.PL}

‘This quarter of an hour was ...’

This set of examples shows that *čtvrt’* can function as either noun or numeral. As a noun, *čtvrt’* heads the nominal expression and triggers agreement in case (shown here in the NOM), in number (here in the SG) and in gender (this noun is always FEM). As the sentential subject, as is shown grammatically in (53b), the predicate shows either FEM.SG or 3.SG agreement (depending on the tense). However, with *čtvrt’* as a numeral, shown grammatically in (53c), the demonstrative pronoun, if present, takes

⁷⁵ Zaliznjak (1987:543) and *Orfoèpičeskij* (1989:617) both list only *čtvrt’ čaSA* (without mentioning, or ruling out, *čtvrt’ ČAsa*). Leonard Babby informs me that he has also encountered examples of *tret’ čaSA* ‘(a) third (of an) hour’, suggesting that this fraction also has at least begun the switch from noun to numeral. These inflectional-paradigm dictionaries only list *dva*, *tri*, *četyre*, *pol*, and *čtvrt’* as those numerals which trigger the end-stressed GEN.SG form *čaSA*. Mel’čuk (1985:322-25) convincingly argues (contrary to Worth 1959:119) that *tret’* is not a numeral in any way. I tend to agree (modulo Babby’s one example of *tret’ čaSA*, the source of which cannot be recovered for further analysis). It would not be surprising, however, if *tret’*, too, began to undergo a gradual change from noun to numeral.

the PL; likewise, the entire nominal expression, if the sentential subject, triggers PL agreement.⁷⁶ As a numeral in (53c), *četvert'* triggers in its syntactic sister noun the end-stressed ADPAUC form *čaSA*; as a noun—in (53b)—it triggers the stem-stressed form *ČAsa*.

I should point out that (53b-c)—i.e., the two acceptable examples in (53)—mean different things. As shown below (at the end of §5.1) in my finalized definitions of “numeral”, “measure noun” and “simple noun”, *četvert'* can function as a numeral only if the noun it quantifies is non-referential. Mel'čuk (1985:322-25), who considers *četvert'* to be a noun despite the fact that it possesses the inherent qualities of a numeral, reports that if *četvert'* has a GEN.PL prequantifier, while suggesting it is functioning as a numeral (cf. discussion, including fn. after ex. (38) above), then its complement must be non-referential, and is usually a measure noun:

⁷⁶ In (53) I have only shown data with demonstratives, which must appear either in the {3/FEM}.SG or PL, depending upon whether *četvert'* is a noun or numeral (respectively). When subject NPs contain a demonstrative, then the predicate agreement must match the demonstrative's agreement. If, however, there is no demonstrative, and *četvert'* is a numeral, then the predicate agreement can be {NEUT/3}.SG—i.e., non-agreement, or “neutral” agreement in Corbett's (1978b, 1980, 1986, 1988, 1991) framework (specifically either NEUT.SG in the past tense or 3.SG in the non-past)—as shown:

Mne	nužno	četvert'	čaSA.	'I need a quarter of an hour.'
me	is-necessary	quarter	hour	
DAT	NEUT.SG	(NUM)NOM	(N.MASC)GEN.SG.ADPAUC	[≈ ex. 98a in Crockett (1976:387)]

In this example I further elicited the final stress on *čaSA*, which further led me to modify Crockett's glosses and part-of-speech labels (specifically, changing the label of *četvert'* form “nom sg fem” to “(NUM)NOM”). It appears that PL agreement is also possible with end-stressed *čaSA*; *if the stress is ČAsa*, then the predicative agreement must be FEM.SG. In any event, I am merely showing the general possibilities of Russian numerical agreement. Cf. chapters 4-5 of Franks (1995) for a detailed treatment of the Russian facts (compared to other Slavic languages as well). (See also ex. 98b in Crockett 1976:387, which shows NEUT.SG agreement with *tret'* ‘one third’.) See also Chey (1967:72, 84).

(54a) Kole vydelili dlja raboty {dobryx/celyx} četvert´ komnaty.
 good/entire quarter room
 (ADJ)GEN.PL ACC (N.FEM)ACC.SG

‘Kolja was allocated **a goodly/whopping quarter of a room** for (his) work.’

(54b) Kole vydelili dlja raboty {dobruju/celuju} četvert´ našej komnaty.
 good/entire quarter our room
 (ADJ)FEM.ACC.SG ACC FEM.GEN.SG (N.FEM)ACC.SG

‘Kolja was allocated **a goodly/whopping quarter of our room** for (his) work.’

Mel’čuk adds to (54a) the following: “that is, *the area the size of about a quarter of a room*; the [speaker] does not have any particular room in mind.” Russian does not have articles to express referentiality. Mel’čuk in (54b) uses the word našej ‘our’ to force the referential interpretation.

To summarize briefly, *četvert´* ‘quarter’ can be either a noun or a numeral. As a noun it triggers the non-ADPAUC GEN.SG. If *četvert´* is a numeral, then it triggers the special ADPAUC form.

Recall from examples (3) through (14) above that sequences of *s* + numeral_{ACC} + noun_{GEN} are not grammatical in the modern language. This predicts that only (52a) and not (52b) should be acceptable as the complement of *s* in modern Russian. I have collected four such examples, shown in (55a-d):

(55a) ... proexav ešče **s četvert´ časa**, my ne vidali ni odnogo verstovogo stolbca.
 quarter hour
 ACC(SG?) GEN.SG

‘... having travelled **for about a quarter of an hour** more, we didn’t see a single milepost.’
 [Sajkiev (1955:52-53), quoting L. Tolstoj (no cit.)]

(55b) Car´ postojal **s četvert´ časa** i opjat´ zadremal.

‘The czar ... stood **for about a quarter hour** then dozed off again.’ [Afanas’ev (1940:91)]

(55c) Car´ postojal ešče **s četvert´ časa**; [...] svalilsja on ná pol i zasnul.

‘The czar stood **for yet about a quarter hour**; ... he collapsed to the floor and fell asleep.’
 [Afanas’ev (1940:91), also quoted in Bukatevič (1958:146)]

(55d) **S četvert´ časa** deržal on obeimi rukami ruki Čičikova.

‘**For about a quarter hour** he held Čičikov’s hands with both of his own hands.’
 [Koka (1955:111) quoting Gogol’ (no cit.); = (116a) below]

Whereas it is impossible to determine the stress on *časa* was at the time these examples were first produced (two of them were written by 19th-century authors), it is clear that for them to be acceptable in the modern language the stress must be *ČAsa*. My informants' responses indicate a slight but consistently decisive preference for (56b) over (56a):

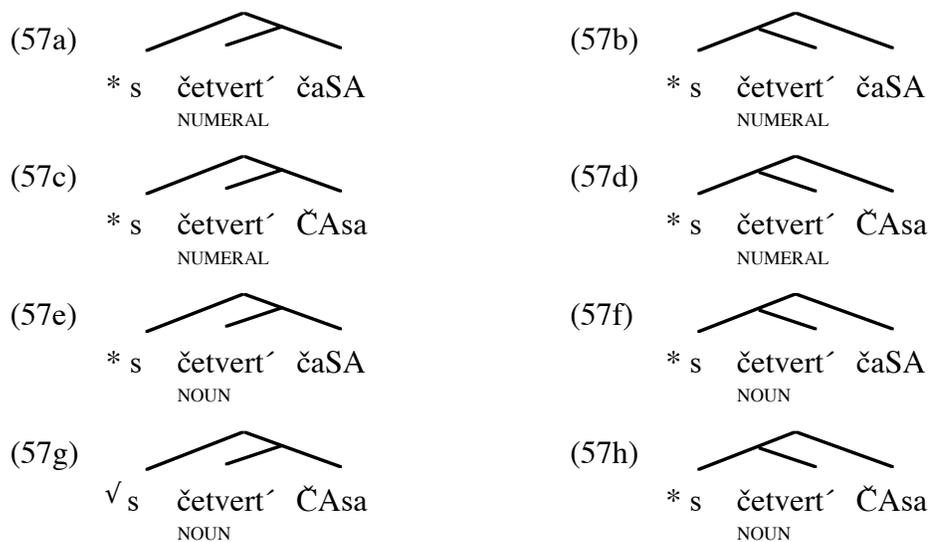
(56a) ? ... s *četvert' čaSA*, ...

(56b) ✓ ... s *četvert' ČAsa*, ...

Note that in (56a) the end-stressed form *čaSA*, which I call the ADPAUC form, suggests that this word is the sister of the paucal **numeral** *četvert'*, further meaning that *s* then has a complement consisting of both numeral and noun.⁷⁷ Recall also that Russian no longer tolerates numeral-noun combinations in the complement of *s*.

What then is the phrase structure of (56b)? It will be necessary to consider eight basic structures in all because of the following factors: (i) *četvert'* can be either a noun or a numeral; (ii) *časa* can have two stresses, thus meaning that each should be considered for every structure proposed; and (iii) there can be one of two structures (assuming that branching is not ternary): [s [*četvert' časa*]] or [[s *četvert'*] *časa*]. The binary variables in (i) through (iii) predict eight (2 x 2 x 2) permutations, shown in (57a-h):

⁷⁷ In the earlier versions of this work (Billings 1993a; 1993b), I argued that it was the twofold part-of-speech status of *četvert'* which determines the stress of *časa*: The numeral variant triggers *čaSA* while the noun variant triggers *ČAsa*. Here I argue that the more stringent structural relationship of **sisterhood** to a numeral is the determining factor for this noun to be end-stressed.



Each of these structures, except for (57g), is preceded by an asterisk. Note that I use symbols which have a slightly different meaning from that generally do in the literature: An asterisk here simply means that this is not the preferred form either because this is not the attested order of these three constituents, because this is not the attested stress on *časa*, or because this particular tree and part of speech of *četvert´* couldn't possibly correspond to the stress and constituent order shown. I likewise mark with an asterisk impossible structures ascribed to attested utterances. In any event, justification is provided for ruling out each of the starred figures in (57), as well as why I do not rule out (57g).

Examples (57a-d) are all ruled out on empirical grounds: With the exception of *pol* 'half', in the modern language *s* never precedes a numeral **and** the noun which that numeral quantifies, regardless of the stress on *časa*. As proof, I use another GEN.SG-assigning "paucal" numeral, 'two': **s tri čaSA*/**s tri ČAsa* 'about three hours'. Additionally, (57a-b, e-f) are ruled out because this is not the preferred stress on *časa*.

I should explain the caveat "with the exception of *pol*" in the preceding paragraph: *četvert´* is unlike *pol* and like all other numerals in this respect. It might

be argued that *čtvrt'* ‘quarter’ and *pol* ‘half’ act as a group because these are the only two fractions that act as numerals. These two differ, however, in several ways: First, as shown below, *pol* doesn’t and *čtvrt'* does undergo approximative inversion (see §4.3.5 for the primary discussion of *pol* and §5.2 on approximative inversion). Second (as I also show below in §4.3.5), *pol* possesses morphological properties that are possessed by no other numeral (*pol* is always a numeral, with a different form, *polovina* as a different form always acting as the corresponding noun meaning ‘half’); *čtvrt'* is the homophonous form of the noun and the numeral that both mean ‘quarter’. Finally, these two fractions trigger differing stresses on *časa*: $\sqrt{pol}čaSA$, * $pol\check{C}A$ sa ‘half (an) hour’, but $\sqrt{čtvrt'}čaSA$, $\sqrt{čtvrt'}\check{C}A$ sa ‘quarter (of an) hour’; $\sqrt{s} polčaSA$, * $s pol\check{C}A$ sa ‘about half an hour’, but $?s čtvrt'čaSA$, $\sqrt{s} čtvrt'\check{C}A$ sa ‘about a quarter of an hour’ (= 18a-b).

There is one apparent way in which *pol* ‘half’ and *čtvrt'* ‘quarter’ behave alike: To the exclusion of all other numerals that assign ADPAUC (or even COUNT, for that matter), the morphological number of an adjective modifying the quantified noun is never PL. Whereas the modifier of a noun quantified by a paucal integer invariably exhibits the morphological PL, an adjective modifying the noun quantified by either *pol* or *čtvrt'* will always be in the morphological SG.⁷⁸ It is unfortunate for these purposes that the distinct ADPAUC form is restricted from appearing when there is such a modifier because (as I show in §4.6.4) the ADPAUC is restricted to a single-word restriction, so the use of the case of adjectives that modify a paucally quantified noun—as Fowler (1988:44-45) does—is not a reliable exercise. In addition, *čtvrt'* has homophonous noun and numeral forms; what determines (since the actual ADPAUC stress is not attested when there is such a modifier) that the adjectives used are not in a

⁷⁸ Note the use of the term “modifier”. So-called prequantifier adjectives, discussed above in §4.2.1 and §4.3.5, can be in the GEN.PL; cf. also the distinction between examples (91a-b) in this respect.

structure with the **noun** version of *čtvrt'*? Furthermore, due to the prosodic peculiarities of *pol*, it is very rare that one finds examples of *pol* + adjective_{GEN.SG} + noun_{GEN.SG}. What **is** a reliable test is quantifying a de-adjectival noun like *stolovaja* ‘dining room’ with the fractions, as Crockett (1976:399, fn. 32) does:⁷⁹

- (58a) *dve* $\sqrt{\text{stolovyx}}/*\text{stolovoj}$ ‘two dining rooms’
 two dining-room(s)
 NOM/ACC GEN.PL/GEN.SG
- (58b) *pol* $*\text{stolovyx}/\sqrt{\text{stolovoj}}$ ‘half a dining room’
 half dining-room(s)
 NOM/ACC GEN.PL/GEN.SG
- (58c) *čtvrt'* $*\text{stolovyx}/\sqrt{\text{stolovoj}}$ ‘a quarter of a dining room’
 quarter dining-room(s)
 NOM/ACC GEN.PL/GEN.SG

Examples (58b-c), as opposed to a paucal integer in (58a), prove that there exists an inherent difference between the paucal fraction numerals and the paucal integers. L. Babby has suggested to me that there is a semantic feature of [\pm PLURAL] which the integer numerals (including non-paucals) possess to the exclusion of the fraction numerals, *nul'/nul'* ‘zero’, and forms of ‘one’ (as an adjective, not a numeral); he suggests that this feature is optionally applied when the ACC case is assigned to a paucal numeral that quantifies an animate noun: If applied, then both words show morphological gen (with the noun in the morphological PL); if not, then the numeral shows morphological nom and the noun shows the GEN.SG (and ADPAUC if that noun

⁷⁹ The sequence *čtvrt' stolovyx* is not outright ungrammatical, it just does not mean ‘a quarter of a dining room’, but rather ‘a quarter of (the) dining rooms’. Tolbert (1974:34) lists a pair similar to the two possibilities in (58c). The noun *čtvrt'* may take either a PL or a SG complement; the numeral *čtvrt'* may take only the SG one. Note also that in (58b) *pol*, which can only be a numeral, cannot take a PL complement in any meaning. Chey (1967:66-67, 70) hopelessly confuses prequantifiers (as defined in §2.2.2 above) with modifier adjectives: “The obligatory use of plural in the adjective applies also to the cases where the numeral quantifiers (*dva, tri, čtyre*, [‘two’, ‘three’, ‘four’ ...] *pol-* [‘half’ ...]) are followed by the genitive singular of the quantified nouns. [...] *dva novyx žurnala* [‘two_(NUM)NOM/ACC new_(ADJ)GEN.PL magazine_{(N.MASC)GEN.SG}’ ...] *celyx poltora časa* [‘a-whole_(ADJ)GEN.PL one-and-a-half_(NUM)NOM/ACC hour_{(N.MASC)GEN.SG}’ ...]” The last example has a prequantifier adjective.

has the special form; cf. §3.3 above). It is fortunate for these purposes that this semantic feature is not crucial to this study.

A decision remains to be made between (57g-h) to be decided between: The initial-syllable stress on *ČAsa* is consistent with both structures. This indicates that a morphological-nom paucal numeral sister of this noun exists in neither of the structures. One argument in favor of (57h) is the structure represented by the other prepositional quantifiers *okolo* ‘about/approximately’ and *po* ‘each/apiece’, as argued in Babby (1985). I must reject (57h), however, because *s+ACC* phrases cannot be conjoined with numerically quantified NPs while phrases with other prepositional quantifiers can be conjoined with such NPs:⁸⁰

- (59a) [[Vosem´QP.NOM [krepostnyx sten]N´.GEN]NP.NOM i [okolo desjatka]QP.NOM
 eight fortified walls and about unit-of-ten
 [nebol´šix fortov]N´.GEN]NP.NOM [zaščiščajutV.PRES.3PL gorodNP.ACC]VP
 small forts defend city
 ‘Eight fortified walls and about a dozen small forts defend the city.’
 [= exx. 10, 12 in Babby (1985:96-97), citing *Izvestija*.]
- (59b) Vosem´ krepostnyx sten i ot pjati do semi fortov zaščiščajut každyj gorod.
 eight fortified walls & from five to seven forts defend each city
 NOM GEN.PL (P) GEN (P) GEN GEN.PL(V)3.PL (ADJ) ACC
 ‘Eight fortified walls and from five to seven forts defend each city.’
- (59c) Vosem´ istrebitelej i po desjat´ tjaželyx tankov zaščiščajut polkovye štaby.
 eight interceptors & each ten heavy tanks defend regimental HQs
 NOM (N)GEN.PL (P) ACC (ADJ) GEN.PL (V)3.PL (ADJ) ACC.PL
 ‘[Eight interceptors] and [ten heavy tanks each] defend the regimental headquarters_{PL}.’
- (59d) * Vosem´ krepostnyx sten i s desjatok nebol´šix fortov zaščiščajut gorod.
 eight fortified walls & about unit-of-ten small forts defend city
 NOM GEN.PL (P) (N)ACC.SG (ADJ) GEN.PL(V)3.PL ACC.SG

⁸⁰ In (59b) I added *každyj* ‘each’ to ease the acceptability of the range of numbers. In (59c) I made sure that the understanding was that there existed a total of eight interceptors, but ten heavy tanks per regimental headquarters. The use of aviation, usually a flexible defense asset, allowed for this reading.

Below (in §5.1) I show that quantificational *okolo* ‘approximately’ and *s* are synonymous. The reason that replacing *okolo* with *s* (and, correspondingly, marking the noun *desjatk-* with ACC case) is not permitted is because *s* is the head, with *čtvrt’ ČAsa* as its complement.⁸¹ A nominal expression with *s*+ACC is a PP and not an NP.⁸²

I conclude, therefore, that (57g) is the only licit phrase structure, prosody (i.e., stress on *ČAsa*) and labeling for this sequence of words. That is to say, **in the overt string** *s čtvrt’ časa*, in (55a-d), the part of speech of *čtvrt’* must be noun (and not numeral). Furthermore, *čtvrt’* and *časa* must be immediate sisters, and *časa* must have non-ADPAUC stem-stress (i.e., the stress must be *ČAsa*).

I have shown in this subsection that the fraction word *čtvrt’* ‘quarter’ can be either a noun or a numeral. In the string *s čtvrt’ ČAsa* ‘about a quarter of an hour’ *čtvrt’* is a noun. There is, then, no violation here of the generalization that *s* cannot take a numeral-plus-noun complement.

4.3.3 Measure nouns: The question is then raised about the part of speech of *čtvrt’* in the licit structure *s čtvrt’ ČAsa* ‘about a quarter hour’? I consider it to be a member of a group of nouns which I call measure nouns, which are categorially nouns, but with quantificational semantics similar to that of numerals. In this group are various unit-of-measure nouns (e.g., *djužina* ‘dozen’, *metr* ‘meter’, and *para* ‘pair/couple’; cf. Vinogradov 1979), container-size words (e.g., *čáška* ‘cup’), etc.⁸³

⁸¹ Another possible reason to reject (57h) is that whereas *okolo* is **not** inherently quantificational, as I also show below (in §5.1), *s* need not have such a structure to be quantificational since it is inherently endowed with “approximate measure” as part of its lexically encoded meaning.

⁸² I modify this claim significantly in §5.1 below. This distinction, however, is essentially accurate.

⁸³ Crockett (1976:385-86) has devised a different taxonomy of various numeral-like quantifiers in Russian. She groups together those number words with not-exactly numerical function (*nol’* ‘zero’ and *tysjača* ‘thousand’ and greater), which she calls “digital numeral-like expressions”; stems like *djužina* ‘dozen’, *desjatok* ‘unit-of-ten’, and *sotnja* ‘unit-of-hundred’, which she calls “non-digital numeral-like expressions”; “adverbial-like quantifiers” (which are hardly discussed here); and “fractional quantifiers”, which she divides into “definite” (for the fraction words themselves: *čtvrt’* ‘quarter’,

The most interesting are the nouns derived from numeral stems to refer to specific counted quantities of items. For example, whereas *desjat´* is the numeral ‘ten’, there is also *desjatok*, the noun used to refer to a unit of ten items. In fact, *desjatok* is far more common than *djužina* ‘dozen’ in modern Russian, since most items in everyday life are dispensed in decimal quantities. I have more than once heard professional interpreters use *desjatok* for English *dozen*.⁸⁴ Other numeral-noun pairs referring to the same quantity are *pjat´* ‘five’/*pjatok* ‘unit-of-five’ and *sto* ‘hundred’/*sotnja* ‘unit-of-hundred’. I propose (in §5.1) that numerals and measure nouns share a semantic feature of quantification which ordinary nouns do not have. It is clear from the following nouns, all of which consist of *s* + measure noun + quantified ordinary noun, that there is no restriction against this sequence. To save on space in the examples, I will stop glossing ACC-assigning *s* itself.

- (60) [...] on znaet **s djužinu nauk** da s poldjužiny drevnix i novyx jazykov.
dozen sciences
(N)ACC.SG (N)GEN.PL
‘he knows **about a dozen disciplines** and about half a dozen ancient and modern languages.’
[Ušakov (1940:15); Zolotova (1988:223); all quoting Gončarov [1965:13]⁸⁵]

pol/polovina ‘half’, *tret´* ‘one third’) and “indefinite” (for words like *bolšinstvo* ‘majority/most’). While Crockett is right in separating a group of quantifiers from the numerals proper, she is mistaken in putting certain fraction words here: *pol* is always a numeral and *čtvrt´* and *tysjača* often are numerals. Furthermore, I see no reason to divide the “numeral-like nominal quantifiers” into “digital” and “non-digital”; both groups trigger the COUNT GEN.PL in the nouns they quantify, for example.

⁸⁴ In sentential glosses of examples with *desjatok* I will often use ‘dozen’, which if not accurate numerically, is more accurate as to function. The word *desjatok* is often used in the plural (or with numbers) to mean some rough figure as in the following example; here *desjatok* glosses best as ‘dozen’:

- [...] potom vysypalo poltora **desjatka** soldat ...
one-and-a-half unit-of-ten soldiers
(NUM)NOM (N.MASC)GEN.SG (N.MASC)GEN.PL
‘... then a **dozen** and a half soldiers emptied out (of the airplane) ...’
[Skoblikova (1959:105-06), quoting Simonov’s *Tovarišči po oružiju*, chapter 33.]

⁸⁵ I discuss the second half of this example in §4.3.5 below.

- (61a) Nautro v atel´e govorjat, čto on ne uspel k vam ili nazývajúť
 ešče **s desjatok podobnyx „uvažitel´nyx”pričin**
 unit-of-ten similar legitimate reasons
 (N)ACC.SG GEN.PL GEN.PL (N)GEN.PL
- ‘The next day at the shop they tell you that he wasn’t able to get to your place or give you **about a dozen more similar “legitimate” reasons.**’
 [Zolotova (1988:223), quoting *Pravda*, May, 1983]
- (61b) kupit´ **s desjatok otkrytok**
 unit-of-ten postcards
 (N)ACC.SG (N)GEN.PL
- ‘to buy about a dozen postcards’ [Sintaksis (1980:448)]
- (61c) **S desjatok učeníkov** bol´ny.
 unit-of-ten pupils sick
 (N)ACC.SG (N)GEN.PL (ADJ.SHORT-FORM)PL
- ‘**About a dozen schoolchildren** are sick.’
 [Sintaksis (1980:448); ex. 102d in Crockett (1976:391)⁸⁶]
- (61d) Na stapeljax stajalo **s desjatok korpusov nedokončennyx lodok.**
 unit-of-ten hulls unfinished boats
 (N)ACC.SG (N)GEN.PL GEN.PL (N)GEN.PL
- ‘On the stocks stood **about a dozen hulls of unfinished boats.**’
 [Sintaksis (1980:243), quoting K. Simonov (no cit.)]
- (61e) protjanul [...] **s desjatok otdel´nyx kuskov nitok** [...]
 unit-of-ten separate pieces thread
 (MASC)ACC.SG (ADJ)GEN.PL (MASC)GEN.PL (FEM)GEN.PL
- ‘(he) held out ... **about ten separate pieces of threads** [sic.] ...’
 [DePerno (1991:ch.7:7), quoting Solženicyñ (1969:487)]
- (62a) polučil **s sotnju pozdravlenij**
 unit-of-hundred invitations
 (N)ACC.SG (N)GEN.PL
- ‘(he) received **about a hundred invitations.**’ [Sintaksis (1980:448)]
- (62b) Rasstojanie — **s sotnju metrov.**
 unit-of-hundred meters
 (N)ACC.SG (N)GEN.PL
- ‘The distance is **about a hundred meters.**’ [Sintaksis (1980:448)]
- (62c) Polučeno **s sotnju zajavok.**
 unit-of-hundred requisitions
 (N)ACC.SG (N)GEN.PL
- ‘**About a hundred requisitions** have been received.’ [Sintaksis (1980:448)]

⁸⁶ See this example, repeated below with (122), and footnotes, regarding PL agreement with s+ACC.

The following minimal pair supports the argument that these measure nouns are actually nouns:

- | | | | |
|-------|--------------------------------|-------|--------------------------------------|
| (63a) | Desjatok zajcev plyl. | (63b) | S desjatok zajcev plylo. |
| | unit-of-ten hares swam | | about unit-of-ten hares swam |
| | (N)NOM.SG (N)GEN.PL (V)MASC.SG | | (P) (N)ACC.SG (N)GEN.PL(V)NEUT.SG |
| | ‘A dozen hares swam.’ | | ‘About a dozen hares swam.’ |
| | | | [Zolotova (1988:223) ⁸⁷] |

Without *s* in (63a), the predicate takes MASC.SG agreement, indicating that *desjatok* is the syntactic head of the subject nominal expression (*desjatok* happens to have homophonous NOM.SG and ACC.SG forms). My informants also accept PL verbal agreement in (63a), but accept only NEUT.SG agreement in (63b). This is due to the fact that with *s*, as in (63b), the prepositional phrase which includes the same noun phrase is the clausal subject, triggering default NEUT.SG agreement.

In each of the examples in (61)-(62) and (63b) it is perfectly grammatical to replace *s desjatok* with the numeral *desjat* ‘ten’ or replace *s sotnju* with *sto* ‘hundred’. Another grammatical option is to remove *s* and lose the ‘approximately’ meaning. If the remaining NP is the sentential subject, as in (61c-d) and (62c), then the nouns would appear as nominative *desjatok* and *sotnja*, also having corresponding predicate agreement. However, it is **not** possible to replace just *desjatok* or *sotnju* with the corresponding numeral: **s desjat* ‘podobnyx’ „uvažitel’nyx” pričin, **s desjat* ‘otkrytok, **S desjat* ‘učenikov, **s desjat* ‘korpusev, **s(o) sto* pozdravlenij, **s(o) sto* metrov, **s(o) sto* zajavok, **S desjat* ‘zajcev. This is due to the fact that only as

⁸⁷ I have further elicited the following acceptable predicate agreements for these two examples: In (63a) either *plylo*_{NEUT.SG} or *plyli*_{PL} is acceptable; in (63b), however, only the NEUT.SG is acceptable (not the PL or the MASC.SG). In general I stay away from the complications of predicative agreement in this study. One thing that **is** worth mentioning here is that aside from so-called semantic-PL or default (NEUT.SG/3.SG) predicative agreement, clauses with prepositional-phrase subjects such as (63b) cannot have agreement with the noun inside the prepositional phrase. (See, however, a crucial correction in the footnote immediately before ex. (120) below.)

measure nouns, as in (21b) and (60), can the forms in (61a-d), (62a-c), and (63b) avoid violating the restriction against an ACC-case numeral and the noun which that numeral quantifies following *s*.

In this subsection I have shown that the noun version of *čtvrt'* 'quarter' is a measure noun. As such, it is the head of the NP and can take an adnominal complement and still be in the complement of *s*. I return to such structures (in 4.4) below, explaining why there can be a complement of *s* consisting of a noun and its own NP complement, but not a complement of *s* consisting of a numeral and noun.

4.3.4 The large numbers 'thousand', 'million', etc.: I show in this subsection that apparent instances of *s* + large numeral + quantified noun are not real examples of *s* + numeral + noun. Similar to *čtvrt'* 'quarter', these large-number words are nouns.

As is mentioned briefly above, certain words for large quantities (e.g., *tysjača* 'thousand'), like *čtvrt'* 'quarter', are still making the transition to numeral-hood. Historically a noun of the *-a* declensional class, *tysjača* 'thousand' has come to take on certain numerical properties. Recall (from §4.3.1) that numerals have no separate morphological-PL paradigms. Nouns that have fully made the transition to numeral-hood do not have distinct SG and PL forms of the same morphological case. The case-marking of *tysjača* 'thousand' is especially revealing in this respect. It shows an apparent paradigm split in the INST(SG).⁸⁸ The form *tysjačej* is the canonical INST.SG form, the form expected of a noun of the *-a* declensional class; it is used in the modern language when *tysjača* is a noun. When it is used as a numeral, however, the newer form *tysjač'ju* has emerged. Due to most of the numeral stems (formerly nouns) coming from the so-called *-ŕ* declensional class, which has the INST.SG form in *-ju* (cf.

⁸⁸ See also Mel'čuk 1985:289, *Orfoèpičeskij* 1989:578, Strelkov 1950:35, Zaliznjak 1987:68, 225.

pjat'ju ‘five_{INST-SG}’), this stem, by analogy to those numerals, developed the INST form *tysjač'ju*, but only for the **numeral** uses of this word. The following table shows the distribution of this word’s forms as a noun in the INST.SG and as a numeral in the INST:

(64a)	*	vladet' ètoj _{(DET)INST.FEM.SG}	tysjač'ju _{(NUMERAL)INST}	rublej _{(N)GEN.PL}
(64b)	°	vladet' ètoj _{(DET)INST.FEM.SG}	tysjačej _{(NOUN.FEM)INST.SG}	rublej _{(N)GEN.PL}
(64c)	√	vladet' ètimi _{(DET)INST.PL}	tysjač'ju _{(NUMERAL)INST}	rubljami _{(N)INST.PL}
(64d)	*	vladet' ètimi _{(DET)INST.PL}	tysjačej _{(NOUN.FEM)INST.SG}	rubljami _{(N)INST.PL}
(64e)	*	vladet' ètoj _{(DET)INST.FEM.SG}	tysjač'ju _{(NUMERAL)INST}	rubljami _{(N)INST.PL}
(64f)	*	vladet' ètoj _{(DET)INST.FEM.SG}	tysjačej _{(NOUN.FEM)INST.SG}	rubljami _{(N)INST.PL}
(64g)	*	vladet' ètimi _{(DET)INST.PL}	tysjač'ju _{(NUMERAL)INST}	rublej _{(N)GEN.PL}
(64h)	*	vladet' ètimi _{(DET)INST.PL}	tysjačej _{(NOUN.FEM)INST.SG}	rublej _{(N)GEN.PL}

(literally) ‘to-possess this/these thousand roubles’

As mentioned above (in §4.3.1), if a NP including a numeral and noun is assigned either NOM or ACC case, then the numeral appears in that case morphologically, but the head noun which it quantifies appears in the GEN case. Numbers for ‘five’ and greater assign the GEN.PL to the nouns they quantify. There are nouns with special COUNT GEN.PL forms (cf. §4.3.1 above): The word for ‘person/people’ is the commonest of these nouns: *ljudej* is the ordinary GEN.PL form, while *čelovek* is the COUNT GEN .PL form. Consider the following example: *tysjača tysjač'√ljudej'?'čelovek*, literally: ‘thousand_{NOM.SG} thousands_{GEN.PL} people_{GEN.PL(NON-COUNT/COUNT)}’. The use of a morphological PL form *tysjač* strongly suggests that the second word is acting as a noun, not as a numeral. Due to this, the third word

containing the non-COUNT form *ljudej* is therefore noticeably preferable to *čelovek*.⁸⁹

One example from the informal corpus that was collected bears on this issue:

- (65) cisterna s tysjaču litrov
 about thousand liters
 (P) (N.FEM)ACC.SG (N.MASC)GEN.PL

‘(a) tank **about a thousand liters** (in size)’ [Sintaksis (1980:448); cf. (147a-b) below]

(Note that this is *s+ACC* and the noun is the ACC form *tysjaču*, not the orthographically similar INST form *tysjač’ju*.) As I argue throughout this study, *s* cannot be followed by an ACC-case numeral and the GEN noun which that numeral quantifies. It is therefore consistent with my analysis that *tysjaču* in this example functions as a noun.

Mel’čuk (1986) argues convincingly that *tysjač-* is both a numeral and a noun in modern Russian. This would predict that only the latter is allowed from the following examples:

- | | | | | | | | |
|-------|-------|----------|----------------|-------|-------|-----------|------------------|
| (66a) | s | tysjaču | čelovek | (66b) | s | tysjaču | ljudej |
| | about | thousand | people | | about | thousand | people |
| | (P) | (NUM)ACC | (N)GEN.PL | | (P) | (N)ACC.SG | (N)GEN.PL |
| | | | COUNT | | | | NON-COUNT |

‘about a thousand people’

‘about a thousand people’

There appear, then, to be differences in the distributions of the (GEN.SG) ADPAUC and (GEN.PL) COUNT forms, which in turn make it difficult to use structures like (66a-b) as a test. Before going into these differences, I should mention that both of (66a-b) are acceptable, perhaps even with a preference for (66a). I explain the reason for this so following an excursus comparing ADPAUCs and COUNTs.

⁸⁹ Note also how the two plural forms for the English word have a similar distribution: \surd *three people*, \surd *three persons*; \surd *a thousand people*, \surd *a thousand persons*; \surd *a thousand thousand people*, \surd *a thousand thousand persons*. These are my judgments, intended merely to show a rough tendency.

Above (in §4.3.2), I point out that the ADPAUC is attested only when the quantified noun is the sister of a paucal numeral; measure nouns, for example, do not trigger it. This appears **not** to be the case with GEN.PL COUNT forms. Mel'čuk (1985:430-31) reports that certain large-number words which otherwise fail to behave as numerals nonetheless trigger the COUNT. For example, whereas *tysjača* 'thousand' functions as a numeral in many ways, the next-larger stem *million* 'million' does not function as a numeral, except for triggering COUNT. Consider the following examples:

- | | |
|---|---|
| <p>(67a) √ v tysjače knig
 in thousand books
 (P) (N)PREP.SG (N)GEN.PL
 ' in a thousand books'</p> | <p>(67b) √ v tysjače knigax
 in thousand books
 (P) (NUM)PREP (N)PREP.PL
 ' in a thousand books'</p> |
| <p>(68a) √ v milione knig
 in million books
 (P) (N)PREP.SG (N)GEN.PL
 ' in a million books'</p> | <p>(68b) * v milione knigax
 in million books
 (P) (NUM)PREP (N)PREP.PL
 ' in a million books'
 [≈ exx. 118a-b in Franks (1995:175)]</p> |

As noted above (in §4.3.1), numerals in modern Russian assign the GEN.PL to the noun they quantify only if the overall nominal expression is not assigned an oblique case. In these examples the preposition *v*, in its locative meaning in these examples, assigns the prepositional (PREP), one of the oblique cases. The left-hand examples, (67a) and (68a), show these number words as nouns, while the right-hand examples show the same words as numerals. As nouns, both words can assign the adnominal GEN.PL to their noun complements. As numerals, these words are predicted to agree in morphological case with the quantified head noun. Only *tysjača* can function as either noun or numeral in this respect; *million* is apparently restricted to functioning as noun.

Compare also the following pair of parallel examples, which shows the same distinction between paucals and other numerals:⁹⁰

- (69a) dva s lišnim {[√]čaSA / [√]ČAAs} 'just over two hours'
two_{NOM/ACC} with excess_{(ADJ)MASC.INST.SG} hour_{GEN.SG} {ADPAUC / NON-ADPAUC}
- (69b) pjat´desjat s lišnim {[√]čelovek / *ljudelj} 'just over fifty people'
fifty_{NOM/ACC} with excess_{(ADJ)MASC.INST.SG} hour_{GEN.SG} {ADPAUC / NON-ADPAUC}
[≈ exx. 13, 16, in Mel'čuk (1985:433)]

The ADPAUC form is optional in (69a), suggesting perhaps that the complex constituent [*dva s lišnim*] 'just over two' can optionally be analyzed as a numeral. In (69b) the constituent which consists of *pjat´desjat s lišnim* 'just over fifty', need not be a numeral, since not just numerals assign the GEN.PL COUNT.

Nonetheless, both *tysjača* and *million* can trigger the COUNT GEN.PL form in the following example:

- (70a) tysjača [√]čelovek / [√]ljudelj (70b) million [√]čelovek / ?ljudelj
thousand people million people
NOM (N)GEN.PL (N)NOM/ACC (N)GEN.PL
COUNT/NON-COUNT COUNT/NON-COUNT
[Elicited from informants based on exx. 2a-g in Mel'čuk (1985:430-31)]

Generally speaking, the words which mean 'million' and greater, despite not being able to function as numerals as in (68b) above, nonetheless **require** the COUNT form, as in (70b). Why then does the word *tysjača* 'thousand'—which, I've argued, can function as either noun or numeral—apparently allow the **non-COUNT** form *ljudelj* 'people_{GEN.PL}', as in (66b) above? It appears that *tysjača ljudelj* does not mean 'a

⁹⁰ Cf. also exx. (94a-c), including fn., in §4.6.4 below, regarding slightly different judgments from my informants, etc. It may seem extreme to suggest that *dva s lišnim/nebol'šim* 'just over two' is a numeral. Cf., however, exx. (94a-c) below.

thousand people’, but rather ‘a lot of people’.⁹¹ To say ‘a thousand people’ one must use the COUNT form: *tysjača čelovek*. It appears, however, that *million* does not have this twofold meaning. Mel’čuk (1985:430) controls for the ‘a-lot’ reading, apparently, by making a complex number in (71b) [= his ex. 2a]; I have added the elicited example in (71a) for comparison:

(71a) tri tysjači √čelovek / ?ljudej
 three thousand
 NOM/ACC (N)GEN.SG
 ‘three thousand people’

(71b) dva milliona √čelovek / ?ljudej
 two million
 NOM/ACC (N)GEN.SG
 ‘two thousand people’

Thus, leaving aside the reading of ‘a lot of people’, both *tysjača* and *million* have a strong preference for the GEN.PL COUNT form. Patton (1969:108), quoting *Pravda*, lists several examples of COUNT *čelovek* triggered by various GEN-assigners.

Another difference between ADPAUC and COUNT is that the latter is attested even when the numeral and noun are in the syntactic GEN case, as shown in the following example:

(72) Do **trexsot** **čelovek** togda francuzov bylo pogubleno [...]
 up-to 300 people then Frenchmen
 (P) (NUM)GEN GEN.PL.COUNT (ADV) (N.MASC)GEN.PL
 ‘Up to three hundred Frenchmen were slaughtered then ...’
 [Škoblikova (1959:113-14), quoting Sergeev-Censkij’s *Sevastopol’skaja strada*, part 4, ch. 5.]

Example (72) also involves additional semantics discussed at length below in the subsection on ADPAUCs and COUNT forms’ single-word restriction (§4.6.4).

⁹¹ Roughly the same thing happens with *skol’ko*. When this word means ‘how many’ it requires the COUNT form: *skol’ko čelovek?* ‘how many people?’. When it means ‘what a lot!’, then the COUNT form is prohibited: *skol’ko ljudej!* ‘what a lot of people’. Cf. also §5.4 below.

Not surprisingly, the group of words called measure nouns (cf. §4.3.2 above) also apparently require the COUNT form *čelovek*:

(73a)	desjatok	čelovek	(73b)	sotnja	čelovek	(73c)	para	čelovek
	unit-of-ten	people		unit-of-hundred	people		pair	people
	(N)NOM.SG			(N)NOM.SG			(N)NOM.SG	
	'a dozen people'			'a hundred people'			'a couple of people'	

The first word in (73a-b) is a measure noun derived from the corresponding numeral (cf. *desjat* 'ten', *sto* 'hundred'); *para* in (73c) is a non-numerically-derived measure noun; in each example the COUNT form is preferred.⁹²

It can thus be stated, then, that the form *čelovek* in the preceding examples, which has been called the “adnumerative” (e.g., Mel’čuk 1985) or “numeral form”⁹³ (e.g., Naylor 1977), is better referred to as COUNT based on its apparent distribution, following any element which denotes a countable quantity and not just numerals. I have shown here that both numerals and measure nouns also trigger this form. Hence my somewhat clumsy (albeit more precise) terms ADPAUC and COUNT. I should add that the COUNT is actually possible with paucal numerals, as long as the numeral is in the GEN case. See example (75b) below.

To summarize briefly, when a numerically quantified nominal expression is assigned syntactic ACC case, because none of the paucal numerals have a distinct morphological-acc form, two options result: (i) If the noun is inanimate, then the

⁹² Cf. the following sources for extensive examples of the COUNT with non-numeral quantifiers: Crockett (1976:325, ex. 22c; 350, ex. 45j; 351, ex. 46; 378, ex. 88a; 410, ex. 119i) and Skoblikova (1959:95 (twice); 96; 104).

⁹³ Strangely, Naylor (1977:92, n. 1) uses the term “numeral form” to refer to what I call COUNT, and “quantification form” to refer to what I call ADPAUC. Specifically, he uses the former to refer to phenomenon in Bulgarian and (to a lesser extent) Macedonian, in which numbers of any quantity (not just ‘four’ or less) trigger a separate inflectional form (often called the “second plural”; cf. Bernard 1954, Koneski 1954:33, Mayer 1973 and Mel’čuk 1985:437). Amazingly, Naylor does not at all mention any GEN.PL-related phenomena in Russian (such as *čelovek*). His ADPAUC data come from Russian and Serbo-Croatian.

numeral must in turn take the morphological nom, one of the direct cases, which requires the noun to have GEN.SG/ADPAUC inflection, as in (74a) below; (ii) if the noun is animate,⁹⁴ then the numeral can either take the morphological-nom (and the noun is the GEN.SG/ADPAUC), as in (74b), **or** take the morphological-gen. As one of the oblique cases, the morphological-gen requires the quantified noun to then take the PL number of that same oblique case—i.e., the GEN.PL—and that GEN.PL is the COUNT, as shown in (75a). I have added two more examples of the morphological-gen case, this time resulting from syntactic GEN, in (75b-c). (Examples in (74a-b) and (75a-c) are elicited.)

(74a) Oni videli **četyre** **gruzovika**
 they saw four trucks
 NOM.PL (V)PAST.PL (NUM)NOM (N.MASC.INANIMATE)GEN.SG
 ‘They saw four trucks.’

(74b) Oni videli **četyre** **čeloveka**
 they saw four people
 NOM.PL (V)PAST.PL (NUM)NOM (N.MASC.ANIMATE)GEN.SG
 ‘They saw four people.’

(75a) Oni videli **četyřex** **čelovek.**
 they saw four people
 NOM.PL (V)PAST.PL (NUM)GEN (N.MASC.ANIMATE)GEN.PL
 ‘They saw four people.’

(75b) Ne xvataet **četyřex** **čelovek.**
 not is-enough four people
 (NEG) (V)PRES.3.SG (NUM)GEN (N.MASC.ANIMATE)GEN.PL
 ‘Four people are missing.’

(75c) Pogiblo okolo **četyřex** **čelovek.**
 perished about four people
 (V)PAST.PL (P) (NUM)GEN (N.MASC.ANIMATE)GEN.PL
 ‘About four men died.’

⁹⁴ See §3.3 and the sources cited there for the definition of “animate”.

In each of examples (75a-c) it is possible to get the COUNT on the noun even if it is quantified by a paucal numeral. The term “ADPAUC” thus must be quantified with the proviso that it be assigned by a morphologically **nom** paucal numeral.

I conclude this subsection on the large numerals (‘thousand’ and greater) by repeating the main points: Examples of *s* + *tysjača* ‘thousand’ (or greater) + noun_{COUNT} do not represent exceptions to the generalization that *s* cannot be followed by numeral plus noun because it is argued that this is the **noun** version of *tysjača*. Number words greater than ‘1000’, although they trigger the COUNT, are always merely nouns. I show, therefore, that some non-numerals can trigger the COUNT form *človek*. The COUNT GEN.PL forms are not a sufficient test for determining whether the quantifier is specifically a numeral. None of the examples here constitutes an exception to the generalization that *s* cannot take a numeral-plus-noun complement.

4.3.5 *The numeral pol ‘half’*: In this final subsection I investigate the properties of *pol* ‘half’. Whereas none of the preceding data in this section has so far constituted a real exception to the generalization that cannot exist after *s*-numeral-noun sequences, the data in this subsection in fact **do** violate this generalization.

The numeral *pol* ‘half’ is a unique numeral in that it invariably must immediately precede the element which it quantifies.⁹⁵ That is, the quantified noun (more specifically, N'' à la Babby 1987) must be phonetically overt, there can be no intervening material of any kind, and the two cannot undergo approximative inversion (see §5.2 below). Mel'čuk (1983), a comprehensive treatment of this numeral, concludes that *pol* is (i) categorially a cardinal numeral, (ii) a “separate word-form”, and (iii) a separate lexeme. I will deal with each of these criteria in turn:

⁹⁵ See the second part of (60) above; cf. also Fryščák (1969:124) and Pete (1984:73) for examples.

Is pol a numeral? There appears to be no argument on this point. One test for this is the two possible stresses of *časa* ‘hour_{GEN.SG}’, discussed above (in §4.3.1). Final-syllable stress (*čaSA*) is the only acceptable form after *pol*. The main word stress is only attested on the final syllable of *POLčaSA* ‘half-an-hour’; the small-caps on *POL* indicate secondary word stress. Like the other paucal numerals (listed above in §4.3.2), *pol* assigns the GEN.SG to the noun it quantifies. If a noun has an ADPAUC form distinct from the GEN.SG, then *pol* must be followed by that ADPAUC form.

If the noun quantified by *pol* is modified by an adjective which appears between the numeral and noun, which happens very rarely with *pol*, then the adjective is in the GEN .SG (unlike the integer numerals, which generally take GEN.PL adjectives).⁹⁶

(76)	<i>pol</i> half NOM/ACC	<i>žutkix</i> terrible (ADJ)GEN.PL	<i>časa</i> hour (N.MASC)GEN.SG	‘a terrible half an hour’ [Mel’čuk (1983:52)]
(77)	<i>pol</i> half NOM/ACC	<i>žutkogo</i> terrible (ADJ)MASC.GEN.SG	<i>časa</i> hour (N.MASC)GEN.SG	‘half of a terrible hour’ [Mel’čuk (1983:52)]
(78a)	<i>pol</i> half NOM/ACC	<i>apel’ sinovogo</i> orange (ADJ)MASC.GEN.G	<i>piroga</i> pie (N.MASC)GEN.SG	‘half an orange pie’ [Mel’čuk (1985:37)]
(78b)	<i>pol</i> half NOM/ACC	<i>našej</i> our (ADJ)FEM.GEN.SG	<i>gruppy</i> group (N.FEM)GEN.SG	‘half of our group’ [Mel’čuk (1985:37)]

The adjective in (76) is a prequantifier, as defined above (in §4.2.1), and as such does not **modify** the noun. This order alternates, in the case of *pol*, with one in which the

⁹⁶ I return to the data in (76)-(77), comparing phrase structures and stress, in (93 a-b).

prequantifier is first: *žutkix polčaSA* [same gloss and reference].⁹⁷ Besides prequantifiers, however, the adjective appears in the GEN.SG, as in (77) and (78 a-b), which is **not** the case with the paucal integer numerals. Note that the adjective in (77) is formed from the same stem as the stem in (76) but does not function as a prequantifier (cf. also the differing glosses). The SG-PL distinction between the adjectives of fractions and of paucal integers is not conclusive, because paucal numerals always assign the morphological PL in the adjective of their complement while *pol* and *čtvrt´* (including when it is a noun) invariably require any **modifier** adjectives in their complements to have morphological-SG number. The fact that *pol* takes a GEN.SG adjective in (78a-b) is not problematic. I show in my comparison of (58a) with (58b-c) above that the fraction numerals differ from the integers in the case of the adjective modifying the noun which they quantify.

The one conclusive diagnostic test for numeral-hood is the ADPAUC: If a noun that can have an ADPAUC form (distinct from the ordinary GEN.SG) is the lone word⁹⁸ in the constituent quantified by some constituent and must appear in the ADPAUC form, then that “some constituent” must be a numeral. Of course, not all numbers can take the ADPAUC, only the paucal ones. There is no single test for non-paucal numerals.

Is pol a “separate word-form”? Matters are not as clear. While Mel´čuk considers *pol* a separate word, other researchers have referred to it as it a particle or a prefix (Buslaev 1875/1977:210 and Vostokov 1839:81, respectively, both cited in Kačevskaja 1969:325, fn. 8). The fact that *pol*, unlike the other paucal numerals, requires the immediately following noun it quantifies to be overt suggests that *pol* is a

⁹⁷ See the expanded discussion of examples (76)-(77) in §4.6.4, and of (77)-(78) in §4.6.2 below.

⁹⁸ In §4.6.4 I show that the ADPAUC (and COUNT) forms are restricted to a lone-word restriction.

type of “bound morpheme”, this being a prefix, a clitic, or the first part of a compound word.

The prosody indicates that *pol* is certainly not a prefix or clitic since it bears secondary stress. *Orfoèpičeskij* (1989:403) reports that *pol* “is pronounced with weak stress.” A reliable test of stress on such prosodically light constituents in Russian is the case for loss of lip-rounding with /o/. Stressed /o/ is pronounced with rounding while unstressed /o/ has no rounding. The /o/ in *pol* is pronounced with rounding because it is stressed.

It is necessary at this point to present a brief discussion of secondary stress in Russian: Most Russian words exhibit only primary word stress. Morphological compounds are the only elements that exhibit secondary stress: Primary stress is placed on the second part of a two-part compound, while secondary stress is placed on the first part. It is, however, unfortunate that both primary and secondary stress both maintain rounding. For this reason, it is impossible to only use a test of roundness to determine whether *pol* is an independent prosodic word.⁹⁹ *Orfoèpičeskij* (1989:409), however, mentions that in every example in which *pol* precedes an *l*-initial word—all examples that begin with palatalized /lʲ/ (e.g., *pol-litra* ‘half-liter’)—the two liquid consonants are in separate syllables and there is no palatalization assimilation as in [POL|lʲ|lʲ|tr «] (syllable breaks shown with a vertical line). *Orfoèpičeskij* (1989:11) also reports that this separate syllabification is standard for words that begin with syllables that exhibit secondary stress as with *MEŽAtomnyj* [MʲEŠ|Alt «mlnyj] ‘inter-atomic’ and *ZAVlaboraTOriej* [ZAF|l «lb «lralTO|rʲ|lʲ|jij] ‘laboratory director’ (note the unstressed /o/ without rounding in the third syllable of both examples). These last two examples are

⁹⁹ Most other phonetic tests are likewise unavailable: /l/ as a sonorant, undergoes neither word-final obstruent-devoicing nor obstruent-cluster voicing assimilation between words in connected speech. These are two tests of phonological word-hood. Cf. Rappaport (1988), a detailed study of Polish words, in which he teases apart the notions **prosodic** word, **phonological** word, and **syntactic** word.

so-called stump compounds (cf. Comrie & Stone 1978:99-101), in Russian *složnosokraščennoe slovo*.¹⁰⁰ Ward (1965:156-63) specifically isolates one type of stump compound, in which the first word is usually reduced to a single syllable and the second word is rendered in full. The term stump compound is sometimes used in order to refer to a word composed of two or more parts that are stumps. I restrict the meaning of this term in this study to a compound consisting of a stump followed by an unabbreviated word.

The preceding phonological evidence also rules out a clitic interpretation. All clitics in Russian fail to have /o/ rounding and fail to undergo word-final devoicing.¹⁰¹ Prepositions are a good example. All monosyllabic or lighter prepositions are clitics, while most disyllabic and all trisyllabic or larger prepositions are separate prosodic words. Some disyllabic prepositions can be optionally proclitic. For the preposition *pered* ‘before/in-front-of’, it is proclitic when it does not have its own word stress¹⁰² and undergoes final-obstruent devoicing as in *pered domom* [pʲirʲidDOM «m] ‘in-front-of (the) house’; when *pered* is a stand-alone prosodic word it has independent word stress and word-final devoicing: [PʲErʲit DOM «m]. Suffice it to say that *pol* does not behave like a clitic.

An argument **against** *pol*’s status as bound morpheme, however, is its ability to assign case—a property generally restricted to syntactic words, an argument raised by Worth (1959:129) and cited in Mel’čuk (1983:57-58). In this sense forms beginning with *pol* are similar to *ZAVlaboraTOriej* ‘lab director’ (literally: ‘director_(TRUNCATED)

¹⁰⁰ Note also the final devoicing of these stump morphemes’ final obstruents ; cf. also preceding fn.

¹⁰¹ I assume that *no* ‘but’, which never loses rounding (i.e., always phonetically [no]), is not a clitic.

¹⁰² There are, unfortunately, no disyllabic prepositions of this kind in which the /o/-roundness test, discussed above, can be tested. The two examples I know of, *pered* and *čerez* ‘through’ (and other meanings), both have only the vowel /e/. (In the meaning of ‘every other’ *čerez* is usually not a clitic.)

laboratory_{(N.FEM)INST.SG}'), listed two paragraphs above, which assigns INST case to the second part of the compound, because *ZAV* is the stump form of the de-participial noun *zavedujuščij* 'director', which idiosyncratically assigns INST case (as does the verb, *zavedovat'* 'manage', from which it is successively derived). Ward's (1965:159:61) otherwise quite detailed discussion of stump compounds does not include any examples like *POL-Litra* or *ZAVlaboraTOriej*, in which the stump of the stump portion governs a particular case in the second part. Nonetheless, it appears that *pol* is precisely this kind of morpheme, required to be the stump part of a stump-plus-full compound.

Using comparative data from other Slavic languages (e.g., Czech *půl* 'half'¹⁰³ < *pōl*), it is clear that the existence of a monosyllable like *pol* goes back at least to the common-Slavic period. It remains a mystery, then, how Russian *pol* came to be interpreted as a stump. Much of the discussion of such compounds and acronyms, called *abbreviatury* in Russian (cf. Ickovič 1971) is in the sociolinguistic literature, because these forms' appearance coincided with the 1917 revolution. In actuality, they were attested shortly before the Revolution,¹⁰⁴ but greatly expanded with the advent of new organizational names and other "newspeak" mentality. It is interesting, therefore, to conjecture about the chronology of *pol*'s having become a stump. I have, unfortunately, not found any reference to *pol* forms in the literature on stumps.¹⁰⁵ Maksimov (1973:53), a brief historical sketch of *pol* and *polovina*, mentions that

¹⁰³ Kačevskaja (1969:327) provides some Czech data, lacking diacritics. One of these, *půl měsíce* 'half a month', contrasts with expressions of 'integer and a half', such as *měsíc a půl* 'month and a half' (literally 'month and half'). That is, whereas Russian uses *pol* + noun to express 'half', and noun + *s polovinoj* (literally 'noun with half') to express 'and a half', Czech uses the same lexeme, *půl*, for both.

¹⁰⁴ Cf. Čukovskij (1914:110-13, referred to on the cover pages of Zalucky 1991), Jakobson (1921:10ff), Karcevskij (1923:46-47), Mazon (1920:3-12), and Seliščev (1924:169).

¹⁰⁵ I did find the following entries in Zalucky (1991:563-64): "**pol.**—*polovina* — half", "**polbánka / polbánki**(sl[ang])—*pollitrovaja butylka vodki* — half-liter bottle of vodka", "**politróvka** (coll)—*pol-litrovaja butylka (vodki)* — half-liter bottle (of vodka)". I return to these data below.

whereas constructions like (79), with *polovina*, were prevalent starting in the late 1700s, the ones in (80), with *pol*, took hold since then:¹⁰⁶

(79a) **V" polovine sed´mago časa razbudili detej [...]**
 at half seventh hour awakened children
 (P) (N.FEM)PREP.SG (ADJ)MASC.GEN.SG (MASC)GEN.SG (V)PAST.PL ACC.PL

‘At half past six [i.e., 6:30] (they) woke the children ...’

[Maksimov (1973:53), citing Karamzin’s *Derevenskie večera*.]

(79b) **Priходите v polovine dvenadcatogo**
 Arrive at half twelfth
 (V)IMPERATIVE (P) (N.FEM)PREP.SG (ADJ)MASC.GEN.SG

‘Drop by at half past eleven [i.e., 11:30].’

[Maksimov (1973:53), citing Puškin’s *Pikovaja dama*.]

(80a) [...] dlja slesarej, dlja plotnikov, vstavavšix **v pol šestogo [...]**
 at half sixth
 (NUM)ACC?.SG (ADJ)MASC.GEN.SG

‘... for the locksmiths, for the carpenters, who got up at half past five [i.e., 5:30] ...’

[Maksimov (1973:53), citing Sluckij’s *Škola dlja vzroslyx*.]

(80b) **Zavtra razbudit´ v pol vos´mogo!**
 tomorrow awaken at half eighth
 (ADV) (V)INFIN (P) (NUM)ACC?.SG (ADJ)MASC.GEN.SG

‘Wake (me?) up at half past seven [i.e., 7:30]!’

[Maksimov (1973:53), citing Majakovskij’s *Letajuščij proletarij*.]

¹⁰⁶ In (80) *polšestogo* and *polvos´mogo* are each written as single words. I break them apart only in order to gloss each part separately. Maksimov adds that the forms in (80a-b) are referred to as “conversational” in modern handbooks of Russian. Mel’čuk (1983) specifies that it is the uses of forms in *pol* which are assigned an oblique (non-NOM or -ACC) case that are problematic (but lists numerous examples from even the literary language). In (80a-b) I have labeled *pol* as “ACC?” because I am reasonably certain this is the ACC case but do not have morphological evidence to this effect. Note, then, that I label *polovina* in (79a-b) and *pol* in (80a-b) as having different cases. My rationale for this is supported by the fact that **all** other expressions for times during a particular hour (or days of the week) use *v* which assigns ACC to a numeral (e.g., *v sem´časov* ‘at 7 (o’clock)’ [at_(P) [seven_{(NUM)ACC} (hours_{(N.MASC)GEN.PL)]]]; *v pjat´ minut vos´mogo* ‘at 7:05’ [at_(P) [[five_{(NUM)ACC} minutes_{(N.FEM)GEN.PL}] eighth_{(ADJ)MASC.GEN.SG}]]; *v polvos´mogo* ‘at 7:30’ [at_(P) [half_{(NUM)ACC} eighth_{(ADJ)MASC.GEN.SG}]]; cf. (80a-b); and (v) *bez desjati vosem´ (časov)* ‘at 7:50’ [at_(P) [[without_(P) ten_{(NUM)ACC}] eight_{(NUM)ACC} (hours_{(N.MASC)GEN.PL)]]]. Only *polovina* was assigned a non-ACC case. This is somewhat similar to distributive *po* ‘apiece’, which assigns ACC to numerals-plus-noun complements and DAT to nouns complements (cf. Franks 1995:139-54).}}

Maksimov then adds the following:

“The element *pol* reminds one of a truncated form of [...] *polovina* ([just as] *zavedujuščij* [‘director’ ...]), the usage of which is influenced by other forms which begin with *pol* [...]. Thus such constructions arose a century and a half after the first constructions with *polovina* were established.” [p.53]

It seems plausible, therefore, that the existence of the forms *pol* and *polovina*, both meaning ‘half’ during the advent of the stump compound earlier this century allowed for a reanalysis of *pol* as a stump abbreviation of *polovina* even though this was not etymologically what happened.¹⁰⁷

Another historical motivation for *pol*’s having changed into a bound morpheme is that it was one of a handful of nouns in the so-called *-ŭ* (or short-*u*) declension (Preobražinskij 1914/1959:821, cited in Kačevskaja 1969:325, fn. 8). Other stems in this class—*dom* ‘house’, *měd* ‘honey’, *vol* ‘ox’, and *verx* ‘top’, shown here in their modern-Russian forms—have all either merged with another declension. Being in an unstable inflectional class often provides the impetus for categorial reanalysis.

Synchronically, then, *pol* is a hybrid entity: a syntactic word which assigns its own case, but morphologically a stump. There is a general tendency in human language for syntactic words to be morphological ones, and vice versa (cf. Prince & Smolensky’s 1993:43 constraint $LX \approx PR$, which I discuss in §6.4.3 below). I have found some indications in colloquial Russian to suggest that the case-assigning properties of *pol* are being lost. The form *polbanki* ‘half-liter bottle of vodka’, which literally means ‘half_(NUM)NOM/ACC jar_{(N.FEM)GEN.SG}’ has the colloquial variant *polbanka* (Zalucky 1991:563). The latter form appears to be a mere concatenation of

¹⁰⁷ Crockett (1976:388-89) presents a similar analysis of *pol* and *polovina*, although apparently based only on her impressions of the two forms. Cf. also Fryšćák (1969:120-22), which shows that *pol* historically behaved like modern numerals in the sense that the entire numerical expression showed oblique case, but “in the two direct cases the construction did not show grammatical agreement, since *pol*” as a quantifier required Gen. sg. of [the noun which it quantified].” [p. 120]

pol and *banka* ‘jar_{(N.FEM)NOM.SG}’, and is inflected as though *pol* were absent (Tolbert 1974:32, Borrás & Christian 1971:386-88). That is, there appears to be no syntactic subordination.¹⁰⁸ Kačevskaja (1969:327-27) discusses these as well. I cannot expound on such forms at length here, except to say that they are considered colloquial and may be an indication of where such stump compounds are heading: toward true compounds in which no syntactic relation (i.e., case-assignment by *pol*) is involved.

Based on secondary stress, non-gemination or -palatalization, and words which behave analogously, it is possible to conclude that *pol* is not an affix or clitic, but rather the first part of a morphological (stump-plus-full-word) compound.¹⁰⁹ Based on its case-assigning properties, *pol* is a separate syntactic word. This numeral is, therefore, a hybrid entity: a syntactic and prosodic word but a morphologically subordinated form.

Is pol a separate lexeme? Mel’čuk (1983:55) argues convincingly that nearly any noun can be quantified by *pol*. That is, whereas only countable nouns can be quantified by integer numerals, almost anything can have a ‘half’. It would therefore be ridiculous to list each *pol*-initial form in the lexicon. Indeed, *pol* is a separate lexeme.

What, then, is *pol*? I conclude that it is a **syntactic** word because of the case it assigns (and specifically the special ADPAUC stress it triggers in *POLčaSA*, not **POLČA*sa). The prosody, however, suggests strongly that *pol* forms a morphological

¹⁰⁸ There is still the prosodic subordination: [POL BAN ka].

¹⁰⁹ There is no direct inflectional marking on *pol* itself, which is generally limited to being either in the NOM case (i.e., as the subject or predicate nominal) or in the ACC (e.g., as the direct object of a verb, or a preposition’s object). Mel’čuk (1983) does list data with *pol* phrases in oblique cases as well. The same article also disambiguates *pol* from two forms of *polu-*, which, although etymologically related, are separate morphemes (one means ‘semi-’; the other, ‘half’ with only an oblique case distribution and, crucially, not productive.)

(stump-plus-full-word) compound with the noun it quantifies. As such a “bound morpheme” *pol* cannot be separated from this noun. Finally, *pol* can quantify virtually any noun in the language, suggesting that this morpheme is separately stored in the lexicon. This numeral is therefore a hybrid entity: syntactically a word, but morphologically a stump (bound morpheme).

I conclude this section on the apparent exceptions to the generalization *[s [numeral noun] by summarizing which data are true exceptions and those which are merely apparent ones: *čtvrt'* ‘quarter’ can be either a noun or a numeral, with only the noun version able to take a multi-word complement of *s* (§4.3.2); other measure nouns are also readily attested in multi-word *s*+ACC complements (§4.3.3); the data with the larger numerals, like *tysjača* ‘thousand’, do not show any conclusive violations, presumably because *tysjača* is a noun in the apparent violations (§4.3.4); finally, I conclude that *pol* ‘half’ is a numeral and as such **does** constitute a violation to the generalization that *s* cannot take a numeral-plus-noun complement (§4.3.5). Thus, there is this one numeral, which has distinct prosodic properties, that stands alone as the only violation of the restriction.

What then of the other, more general restriction against multi-word complements of *s*? Whereas the preceding section does not show any violations of the restriction against *s* + numeral + noun (except for *s* + *pol* + noun), the measure nouns readily violate the restriction against multi-word complements of *s*. It is this issue—complements of *s* which themselves have adnominal-NP complements—that I take up in the next section.

4.4 Complements of *s* with adnominal-GEN structures

In this section I show that under certain circumstances complements of *s*+ACC consisting of a noun and an adnominal noun phrase in the GEN case are allowed despite the single-word restriction.

I showed in the preceding section that, except for *pol* ‘half’, Russian does not tolerate complements of *s* which consist of a numeral and the noun. Examples in which the numeral is substituted with a measure noun **are** acceptable. For example, **s desjat’ otkrytok* ‘about(P) ten(NUM)ACC postcards(N.FEM)GEN.PL’, with the numeral *desjat’* ‘ten’ is unacceptable, while the following form is fully acceptable: *s desjatok otkrytok* ‘about(P) unit-of-ten(N.MASC)ACC.SG postcards(N.FEM)GEN.PL’ [≈ (61b) above]. The primary difference between these two examples is the part of speech of the number word. The *s*+ACC complement cannot contain a numeral and noun but can contain a noun with another noun (phrase). In fact, as examples (61a, d-e) and others show, the complement of *s* can include a noun with a multi-word adnominal-NP complement of its own: [*s [desjatok [korpusov [nedokončennyx lodok]_{NP}]_{NP}]_{NP}]_{PP} ‘about(P) unit-of-ten(N.MASC)ACC.SG hulls(N.MASC)GEN.PL unfinished(PRT)GEN.PL boats(N.FEM)GEN.PL’ [≈ (61e) above].*

Other examples of adnominal complements within the *s*+ACC complement include the following: [*so [šljapku [sopožnogo gvozdika]_{NP}]_{NP}]_{PP} ‘about-the-size-of (the) head (of a) cobbler’s nail’ and [*s [list [pisčej bumagi]_{NP}]_{NP}]_{PP} ‘about-the-size-of (a) sheet (of) writing paper’ [≈ (21c) and (21d), resp.].**

My explanation for this distinction is the following: When the complement of *s* includes a numeral, then there exists a means by which such a complement can exist without violating the single-word restriction. Namely, such structures undergo approximative inversion, which I have yet to discuss in detail (in §5.2) below.

Adnominal structures, however, do not have the option of inverting to express approximation. They can only appear in the order shown.

It is nevertheless necessary to add to a single-noun complement in order to properly specify the measure being expressed. Just as certain adjectives are needed to further specify the measure of a noun (cf. §4.2.3 above), it is also necessary to specify what kind of item, and therefore its measure, is being compared to. Consider once more the examples repeated in the preceding paragraphs. The phrase *s list* means ‘about the size of a leaf/page’; *so šljapku*, because of the dual meanings of the noun, could mean ‘about the size of a cap/small-hat’. It appears that additional words in such structures are allowed so long as they further delimit the measure of the head noun in the complement of *s*. I return to this kind of example in the final chapter, where I show how such exceptions can be explained systematically.

In this section I have shown that certain structures of the type [*s* [noun [...]_{NP}]_{NP}]_{PP} are allowed. Crucially, the word after *s* must be a noun and the additional NP must contribute to that noun’s measure somehow. This concludes the *s*+ACC data having to do with the single-word restriction. In the remaining two sections of this chapter I specify the exact kind of “word” needed in the single-word restriction (§4.5), then show other phenomena in Russian which appear to be subject to the same restriction (§4.6).

4.5 Defining “single word”

So far in this chapter I have shown that *s* quite consistently requires a single-word ACC-case complement “if at all possible”. I have shown (for example in §4.2) that the complement of *s* disprefers modifiers; and those modifiers which are tolerated must be crucial to the measure semantics somehow. I show in my discussion of *s*+ACC with numerals (§4.3) that the only real exceptions to the restriction against *s*-numeral-noun

sequences are when the numeral is *pol* ‘half’, which cannot undergo inversion because it is a bound morpheme. All this suggests that the complement of *s* is limited to a single word in size, but what kind of “word”? There are several ways to define word: for example, prosodic, morphological and syntactic. A prosodic word (PrWd) in Russian is, simply (and somewhat circularly) speaking, a constituent with a single word stress. A morphological word (MrWd) in Russian is a constituent with a root, possibly with derivational affixes, and with a single inflectional ending.¹¹⁰ A syntactic word (SnWd) is an X° constituent, a single item mapped into the syntax from the lexicon.¹¹¹ Not all of the examples I have shown so far that are fully acceptable in modern Russian conform to each of (81a-c). In this section I assess all of the data in this chapter in terms of the following three criteria:¹¹²

- (81a) **Prosodic approach:**
Limit the complement in the *s*+ACC construction to a single PrWd.
- (81b) **Morphological approach:**
Limit the complement in the *s*+ACC construction to a single MrWd.
- (81c) **Syntactic approach:**
Limit the complement in the *s*+ACC construction to a single SnWd.

I proceed through the data in the same order as it is presented above:

In chapter 3 I showed that *s*+ACC is not alone in restricting its complements from either appearing in either the morphological plural (if not quantified by a

¹¹⁰ Obviously, *s* must take a complement that is capable of exhibiting ACC case **morphologically**, which further limits this approach to nominal (i.e., declinable) entities. Thus “inflectional” might as well be “declensional”.

¹¹¹ Again, in order to be **assigned** ACC case **syntactically**, the complement of *s* must be nominal in some sense. See, e.g., (125) where the single word is not a noun.

¹¹² Rappaport (1988) uses an additional definition, that of phonological word distinct from prosodic word in his investigation of Slavic (mostly Polish) clitics. He does not, however, use the MrWd. I do not require separate prosodic and phonological words here. Since the phonological criteria I assess here (with the exception of my treatment of stumps in §4.3.5 above) deal fully with accentuation, I use the term prosodic word.

numeral) or showing the so-called “animate” (or morphological-gen) ACC when a syntactically ACC nominal expression contains a paucal numeral and an animate noun. This might lead one to believe that there is a restriction against adding certain morphological features to the *s*+ACC complement, in support of the morphological approach in (81b). That is, any morphological operation, including feature-addition and affixation, is disallowed. In that chapter I also show, however, that not only *s*+ACC, but also several other ACC-assigning quantificational prepositions have this restriction; I show specifically that *v*+ACC of identity, one of these other constructions, regularly takes multi-word complements. I conclude there that any restrictions on pluralizing (non-numerical) complements and expressing animate ACC in paucal-number constructions is a separate feature from the single-word phenomenon exhibited by *s*+ACC. That is, *s*+ACC is one of a number of prepositional quantifiers that is subject to anti-pluralizing and animate-GEN restrictions, but only *s*+ACC of these constructions appears to impose a single-word restriction.

The data in this chapter include the following: The chapter begins (in §4.1) with an example in which a prepositional phrase following the noun after *s* must not be interpreted as part of the complement of *s*. The structure [[*s* noun] PP] is preferable to [*s* [noun PP]], in which the complement of *s* consists of more than one word. That example, unfortunately, does not clarify the choice between any of (81a-c).

I then discussed various examples in which the complement noun of *s* appears to be modified by an adjective (§4.2). That is, there appear to be two words in the complement of *s*: the adjective and the noun. The first set of data, the so-called prequantifier adjectives (§4.2.1), are not problematic with regard to any of (81a-c), based on my assumption that such adjectives are not part of the complement of *s*; the only item in the ACC-case complement of *s* is a single word, which is at the same time

a single PrWd, MrWd and SnWd. In my discussion of syntactic compounds (§4.2.2) I show that such examples indeed consist of two PrWds/MrWds in the complement of *s*, but are a single syntactic atom. Thus, the syntactic-compound data support the syntactic approach in (81c) over the prosodic or morphological ones in (81a-b). I conclude in another set of data that some adjectives (in §4.2.3), which constitute separate PrWds, MrWds and even SnWds, are a problem for any of the three approaches in (81a-c); these data each involve an adjective that contributes to the measure semantics somehow. Finally, certain frozen lexical expressions represent no challenge to the syntactic approach in (81c), while in some cases challenging the prosodic and morphological approaches in (81a-b). In all, the data on adjectives show repeated violations of both of (81a-b) but only one actual violation of (81c)—namely, those adjectives (in §4.2.3) which serve to qualify the size, length, etc. of the yardstick noun. Such adjectives appear to override the single-word restriction on the complement of *s*. I discuss that “override” again in my Optimality-theoretical treatment of the crucial data in chapter 6 below.

The next section, dealing with the an apparent restriction against overt *s*-numeral-noun sequences in the modern language (§4.3), shows that there is only one actual violator of (81c), the numeral *pol* ‘half’ (discussed specifically in §4.3.5). This numeral, unlike any other in the modern language, is morphologically bound. That is, *pol* must form the first part of a stump-plus-full-stem compound with the noun it quantifies. This stump compound requires the two parts to be overtly adjacent; the noun must not be elided or be uttered other than immediately after *pol*. All other numerals, as I show, have the option of undergoing approximative inversion, which I discuss in detail below (in §5.2). The *pol* data with *s*+ACC have the following prosodic, morphological and syntactic structures:

(82a) Prosodic structure: $[s [pol_{PrWd} [noun]_{PrWd}]_{PrWd}]_{PrWd}$

(82b) Morphological structure: $s [pol_{STUMP} [noun]_{MrWd}]_{MrWd}$

(82c) Syntactic structure: $[s_{P^{\circ}} [pol_{NUM^{\circ}} [noun_{N^{\circ}}]_{N''}]_{NP}]_{PP}$

In (82a) the noun and *pol* are both PrWds, which in turn combine to form a larger PrWd; *s* then adjoins to that PrWd forming an even larger one.¹¹³ In (82b) only the noun is an ordinary MrWd of its own (with internal structure consisting of prefixes, the root, derivational affixes, and declensional suffix—that are not shown); *pol* is morphologically marked as a stump morpheme, which requires it to form the first part of a compound MrWd along with the noun. According to recent theory on clitics, it is possible to assume that *s* is a clitic (CL) which affixes itself at some phrasal level of morphology to the rest of the structure in (82b). The notation in (82c) is more straightforward: the three constituents are P° , N° , and Num° —in order of vocal appearance—their bracketings then form the constituents N'' , NP, and PP. In each example I have enlarged the brackets corresponding to the complement of *s* and bold-faced the labels thereof.

It is now possible to assess *pol* in light of the approaches in (81a-c). In (82a) the complement of *s* consists of a single, albeit complex PrWd. In (82b) the complement of *s* is a complex MrWd.¹¹⁴ In (82c), on the other hand, the complement of *s* consists of more than one SnWd. Unlike the *s*-adjective-noun section (§4.2), in which there are several types of data with multiple-PrWd and -MrWd complements of *s*, the *s*-numeral-noun data (in §4.3) show that the only violators of any of (81a-c) are

¹¹³ Recall that *pol* is a full PrWd. The test is /o/-rounding discussed in §4.3.5 above.

¹¹⁴ In other data the overt complement of *s* consists of a complex MrWd. Cf. the clearly compound numerals in (13) and (141). This is even more reason to rule out morphological complexity as the criterion for excluding certain multi-word complements of *s*. Below I present data with **syntactically** complex numeral phrases—cf. (94) and (134), as well as the footnotes there regarding why such numerals are illicit with *s*+ACC.

those which include *pol*, in which there is a single-PrWd and -MrWd complement of *s* but a multi-SnWd complement.

Recall also that there is one set of multi-SnWd complements among the *s*-adjective-noun data, specifically those adjectives (in §4.2.3) which are required to further delimit the measure of the yardstick noun. Finally, the data in the preceding section show that adnominal NPs are likewise allowed if they further specify the measure semantics of the noun complement of *s*. Based on these data, in my Optimality-theoretic treatment in chapter 6 below, I conclude that the criterion for judging the acceptability of *s*+ACC complements is syntactic.

To conclude this section on the definition of “word”, I have shown that whereas non-syntactic factors are helpful indicators, it appears that the prosodic and morphological definitions of word—namely, (81a-b) above—are **not** the mechanism used by the grammar to express this single-word restriction on the complement of *s*.

4.6 Other constructions with a single-word restriction in Russian

Before concluding this chapter on the single-word restriction I show several other phenomena in Russian which also appear to be restricted in some way to a “single word”. One is the construction *učit'sja na*+ACC meaning ‘study to be a’ plus some profession name; like *s*+ACC, this construction requires a single-word complement of *na*. I also discuss the additional property of *pol* ‘half’, which requires a single-word quantified element. There is also a special form of the GEN case which appears to be restricted to environments in which the word with that special marking is the lone word in its NP. Next, I look again at ADPAUC and COUNT forms, showing that the ordinary (i.e., non-ADPAUC/non-COUNT) GEN case is used if the noun is not the only word in the N'' sister of a quantificational element. These data show that *s*+ACC is not the only construction that requires a single-word constituent. Lastly, I investigate

certain possible phenomena, in which a monosyllabic constituent is exceptional with regard to case-marking; this suggests that there may be a single-**syllable** restriction at work as well in the language.

When possible, I show whether the size limitation being discussed is the initial stage of marginalization of the construction. In light of the preceding section, I also attempt to show whether the single-word restriction is syntactic.

4.6.1 The *učít'sja na+ACC construction*: In this subsection I show another construction aside from *s+ACC* which includes a preposition that appears to require a single-word complement.

The following excerpt might suggest that there is a similar restriction in the *učít'sja na+ACC* construction (thanks to W. Browne for bringing this to my attention):¹¹⁵

“If [using a ‘study’ verb] you specify the name of the future profession, only *učít'sja na* + Acc. can be used.

[83a] Pétja učít'sja na vračá (inženéra, advokáta, šofěra).
 Pete is-studying for physician (engineer, lawyer, driver)
 (MASC)NOM.SG (V)PRES.3.SG (P) (MASC)ACC.SG

‘Pete is studying to be a doctor (engineer, lawyer, driver).’

“Only the names of practical professions can be used in this construction; *učít'sja na filósofa* ‘study to be a philosopher’ sounds ironic. Also, **no adjective can precede the name of the profession** in this construction, unlike English *chemical engineer, nuclear physicist*, etc. If you cannot find any other way out, use an appositive construction:

[83b] Pétja učít'sja na inženéra xímika [...]”
 Pete is-studying for engineer- chemist
 (MASC)NOM.SG (V)PRES.3.SG (P) (MASC)ACC.SG (MASC)ACC.SG

‘Pete is studying to be [lit.] an engineer-chemist (i.e., a chemical engineer).’

[Nakhimovsky & Leed (1980:7); glosses, bold-facing, glosses added/LAB]

¹¹⁵ The complement of *na* must show “animate” (morphological-gen) ACC case; cf. §3.3 for details.

That is, **Petja učitsja na ximičeskogo inženera* ‘... chemical_{(ADJ)MASC.ACC.SG} engineer_{(N.MASC)ACC.SG}’ is not allowed.

This “appositive construction” in Russian is morphological compounding, more accurately shown with **secondary** stress on the initial stem in (83b). Whereas Nakhimovsky and Leed use the same acute stress mark on both parts of the appositive pair, the prosodic prominence of the two parts is not equal. The second part, is more prosodically prominent than the first, most likely indicating secondary stress on the first member of each: *inžeNERa-XImika*. The resulting forms still allow this preposition to take a single—albeit morphologically compound—syntactic word as its complement.

Nakhimovsky & Leed (1980) are correct in their generalization that no adjective can precede the profession-name noun. I interpret their commentary as a pedagogical one: names of professions are unlikely to consist of adjective + noun in Russian as they do in English.¹¹⁶ If one wants to specify the type of engineer, the accepted option is to use the apposition of two profession names (as in *inženera-ximika*), regardless of this construction. There are, however, exceptions such as before *gornogo inženera* ‘mining engineer’. If there is a single-word restriction on the complement of *na* similar to the complement of *s*, exceptions that are analogous to the syntactic compounds discussed above (in §4.2.2) are also possible under this restriction:

- (84) Petja učitsja na gornogo inženera
 Pete is-studying on mountain engineer
 (MASC)NOM.SG (V)PRES.3.SG (P) (ADJ)MASC.ACC.SG (MASC)ACC.SG
 ‘Pete is studying to be a mining engineer’ [Grat. O. Yokoyama for pointing out this example]

¹¹⁶ The constituent *chemical* in *chemical engineer* is not technically an adjective (as Leed and Nakhimovsky imply above). These, and *mining* in the gloss of (84) are the first part of morphological compounds in English (i.e., subject to the compound stress rule, don’t allow *very*, etc.).

Note that the two parts of this profession name cannot truly be semantically decomposed; it is not possible to fully recover the meaning of ‘mining’ from *gornogo* ‘mountain_{(ADJ)ACC.SG}’. This suggests that *gornogo inženera* is a syntactic compound just like *greckij orex* ‘walnut’. It might well be that the *učit'sja na+ACC* construction has a limitation on the size of its complement, a limitation similar to that of *s+ACC*. If true have such a limitation, then it is predicted that the same types of apparent exceptions appear: syntactic compounds (as defined in §4.2.2 above).

In this subsection I have shown that there may be a single-word restriction applying to the *učit'sja na* ‘study to be a ...’ construction. Preliminary indications suggest that this construction requires a single **syntactic** word as the complement of *na*.¹¹⁷ Since this construction has such a specialized use, it is impossible to determine conclusively whether the single word must be defined in terms of syntactic criteria. Nor have I been able to determine whether the single-word restriction of this construction is a stage in some sort of gradual extinction of this construction.

4.6.2 Single word in the complement of pol ‘half’: I return in the this subsection to the morphologically unique numeral *pol* and show that the constituent which it quantifies must be a single syntactic word.

Like the prepositions *na* in the preceding subsection and *s+ACC*, *pol* also generally requires a single-word complement. This requirement holds only of the formal language, however. Several examples with *pol* + adjective + noun are repeated here as (85a-d); the only other examples of this type I’ve found are in (85d-f):

¹¹⁷ Other exceptions analogous to §4.4 are also possible: I have elicited *Petja učitsja na voditelja tramvaja* ‘Pete is studying to be a streetcar driver’, in which the complement of *na* is the noun *voditelja* ‘driver_{(N.MASC)ACC.SG}’ with its adnominal complement NP *tramvaja* ‘streetcar_{(N.MASC)GEN.SG}’. (Thanks to Steve Franks for suggesting that I try such data.) I was not able to elicit examples analogous to §4.2.3, with the structure [*na* [adjective [noun]], perhaps because profession names do not readily appear in this form (except for syntactic compounds like the one in example (87)).

(85a)	pol half NOM/ACC	žutkix terrible (ADJ)GEN.PL	časa hour (N.MASC)GEN.SG	‘a terrible half an hour’ [= (76) above]
(85b)	pol half NOM/ACC	žutkogo terrible (ADJ)MASC.GEN.SG	časa hour (N.MASC)GEN.SG	‘half of a terrible hour’ [= (77) above]
(85c)	pol half NOM/ACC	apel´sinovogo orange (ADJ)MASC.GEN.SG	piroga pie (N.MASC)GEN.SG	‘half an orange pie’ [= (78a) above]
(85d)	pol half NOM/ACC	našej our (ADJ)FEM.GEN.SG	gruppy group (N.FEM)GEN.SG	‘half of our group’ [= (78b) above]
(85e)	pol half NOM/ACC	stolovoj table (ADJ)FEM.GEN.SG	ložki spoon (N.FEM)GEN.SG	‘half a tablespoon’ [Orfoèpičeskij (1989:403)]
(85f)	pol half NOM/ACC	čajnoj tea (ADJ)FEM.GEN.SG	ložki spoon (N.FEM)GEN.SG	‘half a teaspoon’ [Rozentál´ (1974:159; 1977:141)]

Example (85a) has a prequantifier adjective and is not at all problematic (cf. §4.2.1 above).¹¹⁸ Only the prequantifiers among the adjectives in (85) are morphologically-PL; the remaining adjectives are in the GEN.SG and agree in gender with the noun.¹¹⁹ Based on the triple-branching structure proposed so far for prequantifiers, it is sufficient to account for (85a), as long as the single-word restriction is interpreted as applying to each of *pol*’s sisters, not to all of its sisters combined.

Mel´čuk (1985:37) points out that examples (85c-d), are somewhat “conversational”-sounding. Presumably he would make a similar statement about (85b), as my informants do. Examples (85e-f) have syntactic compounds, as defined

¹¹⁸ As I mention in a footnote following example (95), *pol* is the only numeral that can take a prequantifier adjective after it. It is odd indeed for an element, which must for a specialized stump compound with its complement (cf. §4.3.5), to be the only numeral to allow the third sister to intervene between it and its complement.

¹¹⁹ Whereas with paucal **integer** numerals there can be **ACC.PL** adjectives, the **ACC.SG** is not an option with the paucal fraction numerals under any circumstances.

above (in §4.2.2): adjective-noun combinations treated by the syntax as single words. If *pol* has a one-word-complement restriction, then it is precisely this type of complement that will be the apparent exception in the handbooks' treatments of *pol*. This example does **not** have the same marked conversational tenor that (85b-d) do.¹²⁰

Thus, it is only the standard, literary language that imposes a single-word restriction on the complement of *pol*. Because of the less-than-rigorous term “colloquial”, nothing in my arguments about *s+ACC* hinges on this apparent restriction on *pol*'s complement. The only crucial distinction about *pol* (to my study of *s*) is that *pol* is inseparable from its complement (either by ellipsis or inversion).

In this subsection I have shown that, like the prepositions *na* (in §4.6.1) and *s+ACC*, the numeral *pol* ‘half’ requires that its complement in the formal register be a single word. Examples of syntactic compounds show that this restriction, as in the case of the previous subsection and *s+ACC*, must be worded in terms of a single **syntactic** word.

In the next two subsections I show that the single-word restriction is not limited to the complement of some case-assigner, but rather to the distribution of specialized inflectional forms, both of the GEN case.

4.6.3 The so-called second-GEN case: In this subsection I show that one specialized form of the GEN case is also restricted to single-word environments. Unlike *s+ACC* or *učit'sja na* ‘study to be a ...’ (in §4.6.1), this single-word phenomenon is not linked to

¹²⁰ Rozental' (1974:159; 1977:141) refers to (85f) as “colloquial”. (Cf. also ex. (40a).) My informants disagree: It would seem that Rozental' is being prescriptive. From my experience, Rozental''s works, while oriented towards stylistics, are generally descriptive. *Orfoèpičeskij*, on the other hand, is usually quite prescriptive. The fact that the latter lists (85f) without further comment suggests that there is nothing problematic about either of (85e-f).

a specific preposition or other case-assigning word. Instead, the restriction limits the distribution of a specific inflectional affix.

There is one very specialized inflectional ending placed only on certain nouns of a particular declensional class and gender to express a partitive meaning. Such an ending, called either “partitive” or “second” GEN, has been called a separate case in some of the linguistic literature, an issue reviewed in Fowler (1988:75-87).

The GEN-2, as I will call it, is peculiar in many ways; I summarize only a few of these peculiarities here: It has a single ending, *-u*, used only in the SG. As Fowler (1988:78-79) points out, all of the uses of this special inflectional suffix are linked to quantification somehow, and the nouns that take it—within the MASC \emptyset declensional class—are also a semantically definable set. It is nonetheless **not** entirely predictable whether or not a noun will take GEN-2, suggesting that individual nouns must lexically specify whether they take this special form. The verb governing the noun must also allow GEN-2 (Babby 1980:79-83, Pesetsky 1982:201-02). In addition, Fowler reports that the GEN-2 is clearly being phased out of the language, with fewer and fewer words taking this form.

The reason I consider GEN-2 here is that this special inflection is attested only when the word bearing this form is alone in its noun phrase.¹²¹ That is, the distribution of nouns with GEN-2 is limited to a single-word environment:

- | | |
|---|--|
| <p>(86a) √ Xoču čaju.
 want_{1.SG} tea_{(MASC)GEN-2.SG}
 ‘I want some tea.’</p> | <p>(86b) * Xoču tureckogo čaju.
 want_{1.SG} Turkish_{GEN.SG} tea_{(MASC)GEN-2.SG}
 ‘I want some Turkish tea.’</p> |
|---|--|

¹²¹ This is not one of Fowler’s (1988:85) nine criteria for considering the GEN-2 and PREP-2 (see ex. (86) and (87) below). Fowler (1988:82) points out that GEN-2 (and PREP-2) have no special adjectival form. Of course GEN-2 has no adjectival form; the nominal form cannot be used if there is an adjective with it in the NP. I cannot, however, hold this against Fowler, since I have him to thank for suggesting to me that the GEN-2 may be another instance of a single-word restriction in Russian.

Example (86a) shows a single-noun NP in the GEN-2. If the same noun is modified, as in (86b), then the noun cannot show GEN-2 but instead shows ordinary GEN(-1).

- (87a) √ Xoču čaja.
want_{1.SG} tea_{(MASC)GEN-1.SG}
 ‘I want some tea.’
- (87b) √ Xoču tureckogo čaja.
want_{1.SG} Turkish_{GEN.SG} tea_{(MASC)GEN-1.SG}
 ‘I want some Turkish tea.’

The GEN-2 form is optional except in a few very limited, possibly lexified, phrases, as shown in (87a-b) with the ordinary GEN-1 allowed in both modified and unmodified structures corresponding to (86a-b).

One other phenomenon often discussed along with the GEN-2 is the so-called “second prepositional” or PREP-2 case, which shares many properties with GEN-2. The single-word restriction does not apply, however, to PREP-2. That is, unlike GEN-2, the distribution of PREP-2 is **not** affected by the presence or absence of another word in the noun phrase:

- (88a) √ v sneGU.
in snow_{(MASC)GEN(P2).SG}
 ‘in (the) snow’
- (88b) √ v černom sneGU.
in black_{PREP.SG} snow_{(MASC)GEN(P2).SG}
 ‘in (the) black snow’

Also unlike GEN-1, PREP-2 is not optional, as (89a-b) show.¹²²

- (89a) * v SNEge.
in snow_{(MASC)GEN(P1).SG}
 ‘in (the) snow’
- (89b) * v černom SNEge.
in black_{PREP.SG} snow_{(MASC)GEN(P1).SG}
 ‘in (the) black snow’

Another difference between the two is that PREP-2 forms are attested with two different forms in more than one declensional class. On MASC -Ø nouns the ending is stressed-*u*, while on MASC and FEM -*i* stems the inflectional suffix is stressed-*i*. If

¹²² Fowler (1988:67-75, 79-87) and Jakobson (1958:147ff) give various contexts in which either PREP-1 or PREP-2 is preferable or required.

there were a single-word restriction with PREP-2, it would possibly have to be written into both variants' lexical codes.¹²³ With GEN-2 there the a single inflectional suffix *-u*, which lexically requires that the word to which it is suffixed be alone in its noun phrase.

It seems reasonable, therefore, to propose that GEN-2 is subject to a single-word restriction akin to the one that apparently restricts *s+ACC*. Why the PREP-2 does not pattern identically is not entirely clear. The fact that only GEN-2 and not PREP-2 is subject to this restriction is not surprising if one assumes that the restriction is encoded lexically.

The preceding subsection has shown yet another phenomenon, GEN-2, which is likewise restricted to a single-word restriction. A comparison of GEN-2 and PREP-2 demonstrates that the single-word restriction is quite arbitrarily assigned. I propose that this peculiarity is specified in the lexical code of the GEN-2 suffix, *-u*, which requires itself to be in a noun phrase consisting of a single word. It appears that the single-word restriction in GEN-2 is one step toward eventual extinction, which also seems to be the case with *s+ACC*. I have not been able to determine whether this particular restriction specifically calls for a syntactic word, but the data in this subsection are not inconsistent with such a specification.

¹²³ The GEN-2 and PREP-2 distributions share many properties with ADPAUC and COUNT forms, discussed above in §4.3 and again in the next subsection (where I discuss its single-word distribution in detail). I tabulate these three forms, properties in (i) through (vii):

	GEN-2	PREP-2	ADPAUC	COUNT
(i) Only lexically specified nouns exhibit this form:	yes	yes	yes	yes
(ii) Attested only in the morphological SG:	yes	yes	yes	no
(iii) Attested only in MASC nouns of the $-\emptyset$ declension:	yes	no	yes	yes
(iv) Only a single form of this suffix:	yes	no	yes	yes
(v) Limited to single-word environments (cf. also §4.6.4 below):	yes	no	yes	yes
(vi) Optional (cf. Jakobson 1958:147ff) and Fowler 1988:67-87:	yes	no	yes	no
(vii) Final-syllable stress in all forms (cf. exx. (88a-b) and §4.3.2):	no	yes	yes	no

4.6.4 Single-word limitations on ADPAUC/COUNT forms: In this subsection I show that ADPAUC and COUNT forms that are discussed above in this chapter at length, are both also subject to a single-word restriction.

In the earlier discussion of these two forms (in §4.3) I showed that the respective distributions of ADPAUC and COUNT were distinct; the ADPAUC is attested only if that word is the sister of a paucal number in the morphological-nom case, while the COUNT is assigned not just by numerals but by any of a number of elements which designate a countable quantity. For the purposes of this subsection these two forms have identical distributions: both are attested only when they are the immediate sister of the quantifier. That is, while the ADPAUC is licensed by a very specific kind of quantifier (a paucal numeral in the morphological-nom case) and the COUNT by a slightly different kind of quantifier expressing countable quantity, their distributions with regard to the single-word restriction is identical. I will, however, continue to refer to the two using the separate terms.

Neither the ADPAUC form nor the COUNT form is attested when some other word aside from the head noun is in the quantified constituent. In terms of the model of the noun phrase in Babby (1987), this is the N'' constituent. That is, the ADPAUC/COUNT form is attested only when the N'' consists of a single word.

Two very frequently quantified nouns, both of which are used extensively throughout this study, are /čas-/ 'hour_{MASC}', and /čelovek-, ljudj-/ 'person/people_{MASC}'.¹²⁴ The former exhibits an ADPAUC form distinct from its regular GEN.SG form, while the latter has COUNT form distinct from its regular GEN.PL form:

¹²⁴ I do not come close to exhausting the intricacies of the ADPAUC/COUNT in Russian. I merely attempt to accomplish two things here: First, I show that these forms are subject to a single-word restriction. Next, I mention details needed for my exposition below. For almost any other detail on this complicated phenomenon, see Mel'čuk's (1985:430-37) excursus on "adnumeratives" or Fowler's (1988:41-59) discussion of "Count I (paucal)" and "Count II", and the references cited therein.

The table in (90) shows the paradigms of /čas-/ ‘hour’; note the bold-faced ADPAUC form:

(90) Paradigm of /čas-/ ‘hour’

Singular:

NOM	ACC	GEN-1	GEN-2	DAT	PREP1	PREP-2	INST
ČAS	ČAS	ČA SA / čaSA (NON-ADPAUC/ADPAUC)	ČAsu	ČAsu	ČAse	čaSU	ČAsom

Plural:

NOM	ACC	GEN	—	DAT	PREP	—	INST
čaSY	čaSY	čaSOV	—	čaSAM	čaSAX	—	čaSAmi

I discuss these variants briefly above (in §4.3.5). The various inflectional-paradigm dictionaries describe *čaSA* as appearing with the (GEN .SG-assigning) numeral quantifiers *četyre* ‘four’, *tri* ‘three’, *dve/dva* ‘two’, *poltory/poltora* ‘one-and-a-half’ *pol-* ‘half’ and *četvert’* ‘quarter’.¹²⁵ The stem-stressed form *ČA**SA*** (or the GEN-2 *ČAsu*) is attested elsewhere. Zaliznjak (1967:46-48) discusses these peculiarities in the paradigm of *čas-*, calling the ADPAUC forms a separate morphological case.¹²⁶ Mel’čuk (1985:430-37) adds the following details about GEN.PL COUNT forms:¹²⁷

¹²⁵ Disappointingly, both Zaliznjak’s (1987) and *Orfografičeskij*’s entries for *čas* fail to mention that *poltora* ‘one and a half’ also invariably triggers ADPAUC stress on this word.

¹²⁶ This idea was apparently expressed in print even earlier by Isačenko (1962:530). Zaliznjak (1967) cites Isačenko (1962) in his bibliography, but does not actually cite Isačenko in the pages of his book dealing with this phenomenon.

¹²⁷ Even earlier treatments of COUNT are in Bider *et al.* (1978) and Plotnikova & Krasil’nikova (1983:194), who discuss the possible productivity of such forms.

(91) Paradigm of /čelovek-, ljudj-/ ‘person/people’¹²⁸

Singular:

NOM	ACC	GEN	DAT	PREP	INST
čelovek	čeloveka	čeloveka	čeloveku	čeloveke	čelovekom

Plural:

NOM	ACC	GEN (NON-COUNT/COUNT)	DAT	PREP	INST
ljudi	ljudej	ljudej/ čelovek	ljudjam	ljudjax	ljud´mi

The primary complicating factor in this paradigm is the suppletion between the stems /čelovek-/ and /ljudj-/ which roughly correspond to the SG and PL, respectively.¹²⁹ It is the GEN.PL where there is suppletion analogous to the *časA* ~ *ČAsa* phenomenon in the GEN.SG of (90). Here the COUNT (GEN.PL) form is homophonous with the NOM.SG form (*čelovek*). That is, the COUNT form *čelovek* is used (i) when a numeral is either in the NOM/ACC.SG, which in turn requires the noun to take the GEN.PL; (ii) when the numeral is in the GEN case, thus requiring both numeral and noun to exhibit GEN morphology;¹³⁰ or (iii) if a non-numeral quantifies this word.

Defining exactly the kind of quantifier that can trigger *čelovek* is a complicated issue by itself. Numerals and measure nouns both trigger the COUNT. Henriksen (1993) and Xajzer (1976) discuss the uses of *čelovek* vs. *ljudej* following the “indeterminate quantifiers” (*ne)malo* ‘(not a) few’, (*ne)mного* ‘(not) a lot’, *stol´ko* ‘so/as many’, *neskol´ko* ‘several’ and *skol´ko*, which means either ‘how many?’ or ‘what a lot!’. The generalization is that all but *neskol´ko* and *skol´ko* must take non-

¹²⁸ The GEN-2 and PREP-2 are shown in (90) simply because /čas-/ also happens to have these forms; the paradigm in (91) does not. Nor is stress a crucial factor in distinguishing the COUNT form: All forms from the /čelovek-/ stem have stress on the third syllable; /ljudj-/ forms are stressed on the last syllable only when the inflectional ending is not vowel-initial.

¹²⁹ Mel´čuk (1985:430) lists various other pairs, including *cvetkov/cvetov* ‘flowers_{(N.MASC)GEN.PL(COUNT/NON-COUNT)}’, and a possibly productive set of measure nouns in which there is a tendency for the COUNT GEN.PL to be in *-Ø* and the non-COUNT GEN.PL is in *-ov*: *kilogramm/kilogrammov* ‘kilogram_{(N.MASC)GEN.PL(COUNT/NON-COUNT)}’. See also Fowler (1988:47-48).

¹³⁰ I show below in (92) that the other oblique cases also have COUNT forms.

COUNT *ljudej* and *neskol'ko* take *čelovek* almost without exception in the modern language; and *skol'ko* takes *čelovek* whenever it means ‘how many?’ and not when it means ‘what a lot!’. (Cf. my treatment of *neskol'ko* in §5.4 below.) The distribution of *čelovek* has apparently expanded during the last century. Crockett (1976:319) reports that in nineteenth-century Russian only numerals (probably not even collective ones¹³¹) triggered *čelovek*. If true, then *čelovek* might be more accurately referred to as “adnumerative” during that period, and not COUNT. I show in the next chapter (§5.1) that a certain preposition-plus-numeral constituents can also trigger the COUNT.

The reason for going into ADPAUC/ COUNT in such detail is an apparent additional restriction: If a noun has such a distinctive form, then this form is attested only when no adjective modifies that noun (cf., e.g., Crockett 1976:319, fn. 1 and Mel'čuk 1985:432). I interpret this restriction to mean that the distinct (bold-faced) forms in (90) and (91) are restricted to environments when that word is the only word in the complement of the numeral. In terms of the phrase structure proposed in Babby (1987), this constituent is N''.

¹³¹ It is not clear from Crockett’s explanation whether collective numerals triggered *čelovek*: “In the nineteenth century [*čelovek*] was used only when the noun was modified [sic.] by numerals (Ščerbakov 1969, 12-13). In current usage, *čelovek* is also the preferred form when the noun is modified by a ‘collective’ numeral (Ščerbakov, *ibid.*) or by *skol'ko* ‘how many’ or *neskol'ko* ‘several’ [...]” Furthermore, I haven’t been able to consult Ščerbakov (1969). Mel'čuk (1985:188-89) admits that the current situation is quite unclear, concluding rather vaguely that collective numerals (which he calls “personal-quantificational”) with *čelovek* is “awkward/difficult” (*zatrudnitel'no*) in the modern language. Cf. also the following example from a novel apparently published in the 1950s:

... dvoe **ljudej**, vylezšix iz raskalenoj stal'noj korobki, ležali [...] vozle bronevička ...
 two people
 (NUM.COLL)NOM NON-ADNUM

‘... **two people** who’d crawled out of (the) burning steel box, lay ... near the armored vehicle ...’
 [Skoblikova (1959:103), quoting Simonov’s *Tovarišči po oružiju*, chapter 24]

This noun, of course, is modified by the participial phrase headed by *vylezšix* ‘who had crawled out’, and as such does not satisfy the single-word restriction discussed here. See also example (122), with a measure noun, *desjatok* ‘unit-of-ten’, takes non-COUNT *ljudej*. It is decidedly archaic, however.

I might add that COUNT forms are apparently not just restricted to the GEN case. It is possible in colloquial Russian for the other oblique-case forms in (91) to have apparent COUNT forms in the PL. Deviations from the standard register are underlined and all COUNT forms are bold-faced:¹³²

(92) ‘persons/people’ **Plural:**

NOM	ACC	GEN		DAT		PREP		INST
ljudi	ljudej	ljudej	čelovek	ljudjam	čelovekam	ljudjax	čelovekax	ljud’mi
								čelovekami

Mel’čuk (1985:431) and Zaliznjak (1987:441) specify that these forms, too, are only attested after numerals. I have further confirmed that these COUNT forms, like the bold-faced ADPAUC/COUNT forms in (90) and (91), are restricted to the same single-word environment.¹³³

An opportune test of the single-word restriction comes from comparing the following examples, repeated here with the stress indicated on the last word, as they appear originally in Mel’čuk (1983:52), and with my proposed phrase-structure bracketings added:

- (93a) [[pol]_{NumP} [žutkix]_{AP} [**čaSA**]_{N''}]_{NP} ‘a terrible half an hour’
half terrible_{GEN.PL} hour_{(MASC)GEN.SG(ADPAUC)} [≈ ex. (76) above]
- (93b) [[pol]_{NumP} [žutkogo]_{AP} [**ČAsa**]_{N''}]_{NP} ‘half of a terrible hour’
half terrible_{GEN.SG.MASC} hour_{(MASC)GEN.SG(NON-ADPAUC)} [≈ ex. (77) above]

¹³² Mel’čuk does not fully agree that the bold-faced forms in (92) are COUNT forms. He suggests instead that there are two lexemes *čelovek*, the gist of which I repeat below following example (101).

¹³³ At least one author apparently does not consider these forms to be markedly colloquial. Rozenal’ (1974:152; 1987:170) reports that *pjati čelovek* ‘five_{GEN} people_{GEN.PL}’, *pjati čelovekam* ‘five_{DAT} people_{DAT.PL}’, and *s pjat’ju čelovekami* ‘with five_{INST} people_{INST.PL}’ are required, that the corresponding forms with the stem /ljudj-/ are ungrammatical: **pjati ljudej*, **pjati ljudjam*, **s pjat’ju ljud’mi*. The first two editions of this book do not include this section. A different edition (1977:135) words this section slightly differently: “*pjati čelovek*, *pjati čelovekam* (*pjat’ neznakomyx ljudej* [‘five_{NOM/ACC} unfamiliar_{(ADJ)GEN.PL} people_{GEN.PL.NON-COUNT}’] is also possible).”

The bracketing in these examples is intended to show that *pol* has two sisters in (93a) and only one sister in (93b). In (93a) the adjective *žutkix* is a prequantifier (as defined in §4.2.1 above) and does not prevent the ADPAUC from being realized. This is due to the fact that *čaSA* is only one of the **two** sisters of *pol*. In (93b) *pol*'s only sister is [*žutkogo ČAsa*]_{N'}—i.e., more than one word—and thus requires the **non**-ADPAUC form *ČAsa*.

There are numerous examples of complex numerals that trigger ADPAUC forms in the nouns they quantify, much of it seemingly problematic to the single-word restriction being pursued in this chapter:¹³⁴

(94a) dva s lišnim {√čaSA / √ČAsa}
 two_{NOM/ACC} with excess_{INST.SG} hour_{GEN.SG} {ADPAUC / NON-ADPAUC}
 ‘just over two hours’

(94b) dva s polovinoj {√čaSA / *ČAsa}
 two_{NOM/ACC} with half_{(NOUN.FEM)INST.SG} hour_{GEN.SG} {ADPAUC / NON-ADPAUC}
 ‘two and a half hours’

(94c) dva s četvert'ju {√/?/*čaSA / √ČAsa}
 two_{NOM/ACC} with quarter_{(NOUN?.FEM)INST.SG} hour_{GEN.SG} {ADPAUC / NON-ADPAUC}
 ‘two and a quarter hours’

[≈ exx. 13-15 in Mel'čuk (1985:433)]

NB: *s* here is the **INST**-case assigning preposition meaning ‘with’, **not** *s*+ACC!

Mel'čuk (1985:35-36) suggests that the first three constituents of (94b-c) are each syntactically a single numeral. Some other points are also clear: *polovinoj* ‘half_{INST.SG}’ in (94b) is unmistakably a noun (cf. §4.3.5 above). If *časa* were the

¹³⁴ My own consultations with speakers yield similar judgments. Mel'čuk, in addition to (94a) as shown, also reports *dva s nebol'sim časa* (same overall gloss; *nebol'sim* means ‘a-little_{(ADJ)INST.SG}’) as acceptable with either stress on *časa*. In that example my informants consistently prefer *ČAsa*. I have no explanation for this. The same is true if *dva* is substituted with *četyre* ‘four’ in any of these examples (to control for possible single-syllable effects, cf. §4.6.5). See additional examples like (94b-c) in (134a-b) below. In (94c) and (134b) I assume, non-crucially, that *četvert'ju* is the INST.SG form of the **noun**, not the INST form of the numeral. The two are homophonous (cf. §4.3.2 above). My only reason for assuming this is that the parallel word in (94b) and (134a) is the noun *polovinoj*.

complement of *polovinoj*, then only the non-ADPAUC form would be acceptable: *[s [polovinoj [ČAsa]_{NP}]_{NP}]_{PP} ‘with_(P) half_{[N.FEM]INST.SG} (an) hour_{(N.MASC)GEN.SG}(NON-ADPAUC)’. Thus, there is no interference from the surface-consecutive order of *polovin-* and *čas-*. If anything, one would expect **dva s polovinoj ČAsa* if such interference did exist.¹³⁵

In (94c), on the other hand, if *čtvrt’(ju)* ‘quarter’ **can** be a numeral in the modern language, as I argue above (in §4.3.2), then why is the non-ADPAUC form *ČAsa* apparently preferred? The phrase structures for all three examples in (94) appear to be the same: [[numeral]_{Num°} [‘with’ ...]_{PP}]_{NumP} noun]_{NP}. I have a possible explanation: Recall (from §4.6.4) that the word *poltora* ‘one and a half’—etymologically *pol* + *vtora* ‘half_{NOM/ACC} second_{(ADJ.SF)GEN.SG}’—also assigns the ADPAUC form: *poltora čaSA* ‘one and a half hours’ (**poltora ČAsa*).¹³⁶ It is plausible that *dva s polovinoj* ‘two and a half’—the only way to avoid the equivalent of *two point five* in Russian—is also a numeral. Furthermore, it is possible that *dva s čtvrt’ju* ‘two and a quarter’ is not interpreted as a numeral constituent as in the case of *dva s polovinoj*

¹³⁵ I have uncovered the following example of such interference:

V 1954 godu	svyšē	5,5	milliona	detej	otdyxali	v pionerskix lagerjax [...]
	over		million	children	vacationed	
	(ADV)		(N.MASC)GEN.SG	(N)GEN.PL	(V)PAST.PL	

‘In 1954 over 5.5 million children vacationed at Young Pioneer camps ...’
[Skoblikova (1959:113), quoting an unspecified newspaper]

The etymologically comparative quantifier *svyšē* assigns GEN case, most likely to the entire quantified NP. The digits 5,5 can be pronounced a number of ways: *pjati i pjat’* ‘five_{GEN} point five’, *pjati i pjat’ desjatyx* ‘five_{GEN} and five tenths’ or *pjati s polovinoj* ‘five and a half’ (literally: ‘five_{GEN} with_(P) half_{(N.FEM)INST.SG}’). Some of these, with certain speakers, according to Mel’čuk (1985:225-34; 250, n. 9) allow the GEN.SG in the following quantified noun. I do not pursue this issue further in this study.

¹³⁶ Zaliznjak (1987:66), citing Zaliznjak (1967), mentions that in conversational Russian there is also the form *POLtoro* or *POLtora* used to describe one and a half of a *pluralia tantum* noun. I will not be using this particular word in this study. There is sufficient evidence that forms in *poltor-* are (paucal) numerals: They trigger the ADPAUC (*poltora čaSA* ‘one and a half hours’) and the adjective modifying the quantified noun is invariably in the GEN.PL: *poltora bol’six arbuza* ‘one and a half big watermelons’. See also Butorin (1968) for historical and contemporary data on this unique numeral.

in parentheses after (95b), considering only the non-prequantifier interpretation as a possible construction.¹³⁸

Mel'čuk (1983:52) continues, stating that cardinal numerals other than *pol* 'half', behave differently, thus providing the following pair:

(96a) pjat' žutkix časov
 five terrible hours
 NOM/ACC GEN.PL GEN.PL

'five (of possibly many) terrible hours'

(96b) žutkix pjat' časov
 terrible five hours
 GEN.PL NOM/ACC GEN.PL

'five terrible hours of many (not necessarily terrible) hours'

[both from Mel'čuk (1983:52); glosses sic.]

The differing glosses of these two examples show that only (96b) has a prequantifier interpretation. If only *pol* 'half' allows post-numeric prequantifiers, then *žutkix* in (95b) cannot be a prequantifier. Example (95b), therefore, remains problematic to the one-word restriction being pursued here.

Along with (95b) Mel'čuk lists an example of COUNT -overgeneration, analogous to the ADPAUC-overgeneration in (95b):

(97) Peredo mnoj stojalo [[četyre puški] i [dvadcat' pjat' vzroslyx čelovek]]. (?)
 before me stood four cannons andtwenty five grown-up people
 (P) INST NEUT.SG NOM NOM NOM GEN.PL COUNT

'There stood before me four cannon and twenty-five adults.'

[= ex. 12b in Mel'čuk (1985:433), citing Vinokurov (1964:8); brackets added/LAB]

(95b) and (97) are the only examples I have found of a non-prequantifier adjective appearing between a numeral and a quantified ADPAUC/COUNT noun. Unlike (95b), which my informants outright reject, (97) appears to be decidedly strange, but not

¹³⁸ I have, accordingly, glossed (95b) using as a template the gloss of (96a), which fortunately comes from an article that **is** translated into English (by Paul Gorgen) and therefore has glossed examples.

completely so.¹³⁹ Generally the ADPAUC/COUNT form is not attested when there is a modifier, as shown in (98):

- (98) V našej kvartire žívět sem´ odinokix **ljudej** [...]
 in our apartment lives seven single people
 (P) FEM.PREP.SG (FEM)PREP.SG (V)3.SG (NUM)NOM (ADJ)GEN.PL (MASC)GEN.PL
 ‘In our apartment building (there) live seven single persons ...’
 [≈ ex. 45g in Crockett (1976:350)]

I consider two possible solutions to the recalcitrant examples in (95b) and (97): One is prosodic explanation suggested by Mel’čuk, which I bolster with additional accentuational evidence. The other solution is based on a suggestion elsewhere in Mel’čuk (1985) which is not actually used to account for these data. It may seem odd that I belabor these two examples, since my own informants consider them both to be less than acceptable. A closer investigation, specifically of (97), will also shed more light on the single-word restriction with regard to ADPAUC and COUNT forms. I begin with Mel’čuk’s intended explanation first:

Trochaic adjectives can be exceptions: Based on (95b) and (97), Mel’čuk (1985:433) suggests that prosody might well be a factor. Specifically, in each example the problematic adjective is trochaic (i.e., a disyllable with initial stress: *ŽUTkix* and *VZROSlyx*, respectively).¹⁴⁰ I cannot add to this observation aside from pointing out a very general tendency of Russian adjectival accentuation: If both the

¹³⁹ I. Mel’čuk, whom I contacted personally about ex. (97), agrees: “The expression *25 vzroslyx čelovek* IS odd; I agree with your informants. BUT it is not ungrammatical: Vinokurov masters Russian quite well; it is rather a calculated artistic effect.” While I neglected to ask him about example (95a), of which he does not list the source, I presume that Mel’čuk considers it to likewise be extralinguistic word play of some sort. Immediately prior to these two examples, however, he writes that the ADPAUC/COUNT forms are possible “in both the spoken language and in written text[s]” [p. 432].

¹⁴⁰ Mel’čuk (1985) mentions another possible prosodic effect, which I discuss in §4.6.5 below.

stem and the inflectional affix are unaccented, then there will be initial stress.¹⁴¹ Most inflectional affixes (i.e., declensional desinences) on adjectives are unaccented. One affix ending which **is** accented is the (short-form) FEM.SG ending *-a*. According to *Orfoèpičeskij* (1980:113), *Orfografičeskij* (1989:144) and Zaliznjak (1987:317), the adjective stem *žutk-* ‘terrible’ has initial-syllable stress in all forms except the short-form FEM.SG, which has ending-stress: *žutKA*. This, then, is the classic example of an inherently unaccented adjectival stem.

As for *vzrosl-* ‘mature/grown-up/adult’, the adjective in (97), all three aforementioned dictionaries show fixed initial stress in all long forms. As for the short forms, the data are sketchy: Zaliznjak (1985:355) cautions that the FEM.SG and MASC.SG forms are difficult to produce, while *Orfografičeskij* (1989:65) lists fixed initial stress throughout the short-form subparadigm. Thus, the only data, from *Orfografičeskij*, suggest that the */vzrosl-/* stem may be inherently **accented** on the initial syllable.

Outside of inflection proper there is one other affix that is inherently accented—the first syllable of the (productive) disyllabic comparative ending *-ee*. Unfortunately for these purposes, the comparative of *žutk-*, according to Zaliznjak, is likewise difficult to get; *Orfografičeskij* agrees that this adjective’s comparative is “not freely” derived, but nonetheless lists *žutče*, reflecting a non-productive affix (different from *-ee*), which in turn tells us nothing about the accentuation of the stem. As for the *vzrosl-* stem, *Orfografičeskij* and Zaliznjak both list *vzrosLEe*, consistent with the statement that this adjective is inherently unaccented. (*Orfoèpičeskij* lists no short forms or comparatives for either adjective).

¹⁴¹ This generalization is widely known among Slavists as the Basic Accentuation Principle. The adjectival data are summarized by Levin (1978: chapter 4), without referring to this principle directly.

Suffice it to say that the adjectives in (95b) and (97) might not only be trochaic specifically in their GEN.PL inflectional forms but also may also be inherently unaccented. That said, if a specific stress type does reveal an exceptional trait, it would not be surprising if the **least marked** accentual type were the class to constitute that exception. These two adjectives appear to belong to the type that is least marked. To fully corroborate Mel'čuk's suggestion about prosody being involved, however, extensive research would be required. So, it is not possible to confirm Mel'čuk's suggestion, but only bolster it with these accentuational facts.

My own reanalysis of (95b) and (97): The other solution that accounts for (95b) and (97), to my knowledge, is somewhat more complicated: In this approach I do not use the same explanation for both (95b) and (97). Instead, I entertain a prosodic approach to (95b), based on other phenomena in which monosyllabic numerals can be exceptional. I consider a different phrase structure in (97), one in which the problematic adjective *vzroslyx* does not function syntactically as an adjective.

I start with (97), repeated here as (99):

- (99) Peredo mnoj stojalo [[četyre puški] i [dvadcat' pjat' **vzroslyx** čelovek]]. (?)
 before me stood four cannons andtwenty five grown-up people
 (P) INST NEUT.SG NOM NOM NOM GEN.PL COUNT
 'There stood before me four cannon and twenty-five adults.'

The word *vzroslyx*, although morphologically an adjective, is probably syntactically a noun, even in (99) where this adjective appears to modify the noun *čelovek* 'people'. Other examples of such "adjectival nouns" (as such words are commonly called, also "substantivized adjectives") are shown in (100a-c):

- (100a) V SŠA nasčityvaetsja svyšē desjati millionov **negramotnyx**
 number over ten million illiterates
 (V)PRE.3.SG COMPAR (NUM)GEN (N.MASC)GEN.PL (ADJ)GEN.PL

'In the USA illiterates number over ten million.'
 [Skoblikova (1959:96), quoting *Pravda*, 19.11.1953]

(100b) [...] prinjalo učastie svyšē 2300 **trudjaščixsja** Kazaxstana.
 over 2300 workers Kazakhstan
 COMPAR (NUM)[GEN] (ADJ)GEN.PL (MASC)GEN.SG

‘... over 2300 workers of Kazakhstan took part.’
 [Skoblikova (1959:96-97), quoting *Pravda*, 6.12.1955]

(100c) po troe **bol’nyx** ‘three patients apiece’
 three patients_{(ADJ)GEN.PL} [= (32e) above]

In (100c), the word *bol’nyx*, if functioning syntactically and semantically as a modifier, means ‘sick’. But this word can also stand in for a noun, as it does in (32e), and mean ‘patient’ (i.e., ‘sick one’). Another common example is the stem *russk-*, which means either the adjective ‘Russian’ (as in *russkij jazyk* ‘**Russian** language’) or the adjectival noun (as in *On russkij*. ‘He is a **Russian**.’). Without listing all the peculiarities of adjectival nouns here (but cf. Fowler 1988:43-46), I can say the following: First, there are certain peculiarities in the behavior of adjectival nouns when they are quantified by numerals. As (100c) shows (cf. also Mel’čuk 1985:390-91), adjectival nouns in some registers are preferably quantified by special collective forms of numerals (i.e., *troe* instead of the ordinary *tri*; both mean ‘three’). Second, related to the first, when adjectival nouns are quantified by numerals, they often have pleonastic count nouns inserted, as in (101):

(101a) Komanda sobralas’ pëstraja: neskol’ko grekov, dvoe ital’jancev, dva turka,

‘The crew that formed was a motley one: some Greeks, two Italians, two Turks,

negr, i pjatnadcat’ **čelovek russkix**.
 Negro and fifteen people Russians
 NOM.SG NOM GEN.PL (ADJ)GEN.PL

a Negro, and fifteen Russians.’

[Mel’čuk (1985:209, n. 2)]

- (101b) Ja nanjal" neskol'ko **čelověk" rabočix"**.
 I hired several people workers
 NOM (V)PAST.MASC.SG (Q)ACC GEN.PL (ADJ)GEN.PL
 'I hired several workers.' [Aleksandrov" (1923:701); my glosses/LAB¹⁴²]
- (101c) U zabora tolpilos' **čelovek dvadcat' štatskix**.
 By fence crowded people twenty civilians
 (P) (N.MASC)GEN.SG (V)PAST.PL GEN.PL (NUM)NOM (ADJ)GEN.PL
 'About twenty civilians crowded by the fence.'
 [Pete (1984:76), quoting Simonov (no cit.); my glosses/LAB]

Common pleonastic nouns are *čelovek* and *duš* 'soul(N.FEM)GEN.PL' used for humans, and *štuk* 'item(N.FEM)GEN.PL', used (mostly) for non-humans.¹⁴³ See Sussex (1976) for a detailed treatment of such words. (Example (101c) shows how the same structure expresses approximation, cf. also (106)-(107) below.)

I assume that adjectival nouns occupy a noun's position in the syntax even if they inflect like adjectives. That said, *vzroslyx* in (97) and (99) is most likely not a modifier, despite the appearance of modifying *čelovek*.

It is then important to determine the exact nature of *čelovek* in (97)/(99), and even in (101). Fryšćák (1969:211-12) writes that the combination of words like *štuk* and *čelovek*, along with an ordinary cardinal number as in (102b), has largely replaced one of the uses of the collective numeral, as in (102a):

¹⁴² The slight differences in spelling merely reflect an older orthographic convention, not important to the discussion here. I discuss the part-of-speech status of *neskol'ko* in §5.4 below.

¹⁴³ Two words—*golova* 'head', used specifically for livestock, and *mesto* 'place', to individuate pieces of luggage (DePerno 1990, citing W. Browne p.c.; also Chey 1967:39)—are slightly different, because they allow non-count words like *skot* 'livestock' and *bagaž* 'baggage' to be countable. In this respect the former has both functions; *golova* can be used with countable nouns like *korova* 'cow'.

(102a) *pjatero* *detej*
 five children
 (NUM.COLL)NOM/ACC (N)GEN.PL

(102b) *pjat'* *čelovek* *detej*
 five people children
 (NUM)NOM/ACC (N.MASC)GEN.PL (N)GEN.PL

'(group of) five children'

[Fryščák (1969:211-12)]

Presumably both structures render the 'group' meaning.¹⁴⁴ Chey (1967) specifies that pairs like (102a-b) arose because the decline of the collective numerals:

"[...] Since the use of the collective numerals *pjatero* ['five(some)_{NOM}'] and up is rare, the paraphrased type *devjat' duš detej* ['nine_{(NUM)NOM} souls_{(N.FEM)GEN.PL} children_{(N)GEN.PL}'] or *šest' čelovek mužikov* ['six_{(NUM)NOM} people_{(N.MASC)GEN.PL} peasants_{(N.MASC)GEN.PL}'] is preferred to the construction *d[e]vjatero detej* ['nine(some)_{NOM} children_{(N)GEN.PL}'] or *šestero [mužikov* 'six(some)_{NOM} peasants_{(N.MASC)GEN.PL}']."

[Chey (1967:56-57) citing, *inter alia*, Unbegaun (1957/1960:145)]

For reasons not directly pertinent to this study, the viability of certain collective numerals became limited.¹⁴⁵ Now the apparent way to express this "group" or "collective" meaning with larger numbers is by placing a pleonastic count noun after a non-collective numeral.

Mel'čuk (1985:209, n. 2) lists (101a) as an example of a separate "numeralive" lexeme, which he labels "*čelovek-2*", which differs from "*čelovek-1*"—the form I use

¹⁴⁴ Strangely, Fryščák mentions this in his conclusion, but does not appear to cover this issue in the body of the dissertation. The other example he provides is of a *pluralia tantum* noun: *dvoe perčatok* (literally: 'two_{(NUM.COLL)NOM/ACC} gloves_{(N)GEN.PL}') being replaced by *dve pary perčatok* (literally: 'two_{(NUM)FEM.NOM/ACC} pairs_{(N.FEM)GEN.SG} gloves_{(N)GEN.PL}'); both mean 'two pairs of gloves' [p. 211]. I have found another such pair:

<i>šest'</i>	<i>čelovek" tatar"</i>	(ii)	<i>šestero</i>	<i>tatar"</i>
six	people Tatars		six	Tatars
(NUM)NOM/ACC	GEN.PL (N.MASC)GEN.PL		(NUM.COLL)NOM.ACC	(N.MASC)GEN.PL

'six Tatars' (same gloss for both)

[Unbegaun (1935:311), also in DePerno (1990:2)]

Unbegaun (1935:311) lists these as examples having been synonymous even in the fifteenth century.

¹⁴⁵ Unbegaun (1957/1960:145) writes that the use of collectives for 'five' and greater is on the wane. It is not surprising that the distinction has been drawn between 'four' or less on the one hand and 'five' or greater on the other—the paucal/non-paucal distinction. See also Tolbert (1974:12).

throughout this study, which has only the one (GEN.PL) COUNT form—in the following ways: First, only *čelovek-1* has a singular paradigm, repeated here as (103):

(103) Paradigm of /*čelovek-*, *ljudj-* ‘person’ (Mel’čuk’s *čelovek-1*)

Singular:

NOM	ACC	GEN	DAT	PREP	INST
čelovek	čeloveka	čeloveka	čeloveku	čeloveke	čelovekom

Plural:

NOM	ACC	GEN (NON-COUNT/COUNT)	DAT	PREP	INST
ljudi	ljudej	ljudej/ čelovek	ljudjam	ljudjax	ljud’mi

Second, *čelovek-2* has no suppletion with the /*ljudj-*/ stem. The entire paradigm for *čelovek-2* is shown in (104):

(104) Paradigm of *čelovek-2* [cf. (92) above]

Plural:

NOM	ACC	GEN	DAT	PREP	INST
—	—	čelovek	čelovekam	čelovekax	čelovekami

Third, as (104) also shows, *čelovek-2* has no direct-case (NOM or ACC) forms. This is because the distribution of *čelovek-2* is limited only to numerical expressions. As a result, only the oblique cases are represented since NOM or ACC numerical elements require the nouns they quantify to be in the GEN case. Examples (105a-e), which include one example of the inanimate count noun *štuk*, come from various other sources (all numerals have been spelled out):

(105a) Trista **čelovek** interbrigadovcev podošli k xolmu [...]

three-hundred people *interbrigada*-ists approached to hill

(NUM)NOM (MASC)GEN.PL (MASC)GEN.PL (V)PAST.PL (P) (MASC)DAT.SG

‘Three hundred *interbrigada*-ists approached the hill.’

[Skoblikova (1959:101), quoting Simonov’s *Tovarišči po oružiju*, chapter 23.]

(105b) nas bylo devjat’ **čelovek** detej

us was nine people children

GEN (V)PAST.NEUT.SG (NUM)NOM GEN.PL (N.NEUT)GEN.PL

‘There were nine of us kids (in the family).’

[= ex. 1a in Sussex (1976:145), quoting Puškin’s *Kapitanskaja dočka*; glosses mine/LAB]

- (105c) [...] Zimovalo dvadcat' štuk ovec. [...]
wintere-d-over twenty items sheep
(V)PAST.NEUT.SG (NUM)NOM (N.FEM)GEN.PL (N.FEM)GEN.PL
'... Twenty (head of) sheep wintered over. ...'
[= ex. 2d in Sussex (1976:145), quoting Bunin's *Derevnja*; glosses mine/LAB]
- (105d) ... Poslat' [...] dva [sic.] čelovek saldat [sic.] s ynstrukcieju [sic.]
send two people soldiers with instruction
(V)INFIN (NUM)ACC GEN.PL (N.MASC)GEN.PL (P) (N.FEM)INST.SG
'... to send ... two soldiers with instruction[s]'
[Fryščák (1969:13), citing Bulaxovskij (1958:196), quoting *Ukaz Manufaktur-kollegii 1752 g.*]
- (105e) V komnate bylo tol'ko dvadcat' čelovek oficerov.
in room was only twenty people officers
(P) PREP.SG (V)NEUT.SG (ADV) (NUM)NOM (N.MASC)GEN.PL (N.MASC)GEN.PL
'There were only twenty officers in the room.'
[I. Mel'čuk (personal communication); my glosses/LAB]

A few comments on these examples are in order:¹⁴⁶ First, none of (105a-e) expresses approximation as such, but see (106a-e). Second, there is an added emphasis on the quantity. For example, in (105b) nine children is a lot for one family to have, thus making it likely that the sheer quantity is what is being expressed. Third, the canonical order of *čelovek-2* (when there is no approximative inversion) is after the numeral and before the quantified noun. DePerno (1990; 1991: chapter 9) refers to this use of nouns as “postquantifiers” (following Babby's 1985 term “prequantifiers” discussed in §4.2.1 above). Finally, this use of pleonastic count nouns is frequently attested with adjectival nouns, as in (101a-b) above.

¹⁴⁶ The following example does not actually possess all the criteria for the construction being discussed here.

Vposledstvii on spas žizn' vsem nam, tridcati čelovekam komandy.
subsequently he saved life all us thirty people crew
(ADJ) NOM MASC.SG ACC.SG DAT.PL DAT.PL DAT DAT.PL GEN.SG

'Subsequently he saved the lives of all of us, a crew of thirty.' [Mel'čuk (1985:209, n. 2)]

Instead of a countable, GEN.PL noun following either a form of /čelovek-/ or /štuk-/, as in (105a-e), this example has the GEN.SG noun *komandy* 'crew/team'. Mel'čuk specifies that *ljudjam* (i.e., the ordinary DAT.PL counterpart of *čelovekam*) is unacceptable here. My informants allow *členam* 'members_{DAT.PL}', but this version loses the emphasis on quantity which the use of *čelovekam* entails.

It is also possible to express approximation in conjunction with this construction by ordering the pleonastic noun before the numeral:

(106a) **čelovek** dvadcat´ partizan ležali vokrug kostra
 people twenty partisans lay around campfire
 GEN.PL (NUM)NOM (N.MASC)GEN.PL (V)PAST.PL (P) (MASC)GEN.SG
 ‘**about twenty partisans** lay around (the/a) campfire’
 [= ex. 1c in Sussex (1976:145), quoting Fadeev’s *Razgrom*; glosses mine/LAB]

(106b) Nikolaj... sgreb **štuk** desjat´ suxarej
 Nikolaj gathered items ten croutons
 (MASC)NOM.SG (V)PAST.MASC.SG (N.FEM)GEN.PL (NUM)ACC (N.MASC)GEN.PL
 ‘Nikolaj ... gathered **about ten croutons** together’
 [= ex. 2b in Sussex (1976:145), quoting Turgenev’s *Nakanune*; glosses mine/LAB]

(106c) [...] soedinjaet **štuk** pjat´, **štuk** desjat´ anekdotov [...]
 items five items ten anecdotes
 GEN.PL (NUM)ACC GEN.PL (NUM)ACC (N.MASC)GEN.PL
 ‘... connects **about five** (or) **ten anecdotes** ...’
 [= ex. 2e in Sussex (1976:145), quoting G. Uspenskij’s *Peterburgskie pis´ma*; my glosses/LAB]

(106d) V komnate tolpilos´ **čelovek** desjat´ mužikov.
 In room crowded people ten peasants
 (P) (FEM)PREP.SG (V)MIDDLE.PAST.NEUT.SG (MASC)GEN.PL (NUM) (MASC)GEN.PL
 ‘The room was crowded with **about ten peasants**.’ [Mel’čuk (1985:209, n. 2)]

(106e) V temnote **čelovek** dvadcat´ ljudej okružilo P´era.
 In darkness people twenty people surrounded Pierre
 (P) (FEM)PREP.SG (MASC)GEN.PL (NUM) (MASC)GEN.PL (V)PAST.PL ACC.SG
 ‘In the darkness **about twenty people** surrounded Pierre.’
 [Pete (1984:74), quoting Tolstoj (no cit.)]

The order in (106) is related to approximative inversion to be discussed in the next chapter (§5.2). Omitting either *čelovek* or *štuk* in (106a-e) removes the approximative interpretation as well as the emphasis on the quantity. Example (107) shows this quite clearly:

(107) Ne odna ved´. **Čelovek** dvadcat´ ix sobralos´.
 not alone after-all people twenty them gathered
 (NEG) (ADJ)FEM.NOM.SG (CL) GEN.PL (NUM)NOM GEN.PL (V)PAST.NEUT.SG
 ‘After all, (she) is not alone. There are **about twenty of them** who have gathered.’
 [= ex. 30c in Crockett (1976:333)]

The implication in the second clause of (107) is that ‘about twenty people’ must be more than enough people to keep anyone from being alone. Again, the emphasis is on the sheer quantity.

Before proceeding further a brief comment on (106e) is in order. In that sentence the numeral is preceded by the pleonastic count noun *človek* and followed by the quantified noun *ljudej*. The latter, which is alone in the N'' and should be in the COUNT form. Apparently the pleonastic noun blocks the COUNT form in the lexical noun. I propose in the next chapter that approximative inversion is the movement of the lexical noun to Spec-of-NP (or -PP) position. When there is more than one word in N'', then instead of moving the noun a pleonastic noun appears in that same Spec position. The data in (105) through (107) are special in that there is emphasis on the quantity, which I discuss more below. The examples in (105) do not have approximation, while those in (106)-(107) do express approximation. I propose, quite tentatively, that the pleonastic noun in these emphasis-on-quantity examples is in the complement of Num°, where there is no approximation, and in the Spec of NumP when expressing approximation, in (106)-(107). Furthermore, a numeral discharges the special COUNT form once, to its sister. When there is a pleonastic noun within NumP, then the numeral discharges the COUNT inflection on that NumP-internal noun and the noun-head of the matrix NP does not get the COUNT form. This particular proposal, the position of the pleonastic noun in emphasis-on-quantity constructions, is far from proven; I merely provide a possible way to account for this use of *človek* consistent with other data which I formalize more precisely.¹⁴⁷

¹⁴⁷ I further assume that the pleonastic noun is in the complement of Num° in (101a-b) and in Spec of NumP in (101c). Adjectival nouns are deficient somehow and the insertion of a pleonastic noun allows the numeral to discharge certain quantificational features which adjectival stem cannot bear.

This construction is obscured by yet another complication: In my view the construction in (105)-(107) is distinct from yet another use of *čelovek* and *štuk*:¹⁴⁸

(108a) odnaždy **čelovek** desjat´ našix oficerov obedali u Sil´vio
 once people ten our officers dined at Silvio's
 GEN.PL (NUM)NOM GEN.PL (N.MASC)GEN.PL (V)PL (PP)

‘Once **about ten of our officers** dined at Silvio’s’

[= ex. 1b in Sussex (1976:145); Franks (1994:661, n. 73), Pete (1984:76), quoting Puškin’s *Vystrel*]

(108b) (Dymov) ... proiznes **štuk** pjat´ nexorošix slov
 items five bad words
 GEN.PL (NUM)ACC (ADJ)GEN.PL (N.NEUT)GEN.PL

‘Dymov ... uttered **about five obscene words**.’

[= ex. 2c in Sussex (1976:145), quoting Čexov’s *Step*; glosses mine/LAB]

(108c) [...] vpolzali **štuk** desjat´ malen´kix devoček s knižkami
 items ten little girls
 GEN.PL (NUM)NOM (ADJ)GEN.PL (N.FEM)GEN.PL

‘... **about ten little girls** with books would creep into (the gates of her house).’

[= ex. 2f in Sussex (1976:145), quoting G. Uspenskij’s *Iz činovnič´ego byta*; my glosses /LAB]

(108d) On kupil **štuk** desjat´ starinnyx knig.
 he bought items ten antique books
 NOM.SG (V)PAST.MASC.SG GEN.PL (NUM)ACC (ADJ)GEN.PL (N.FEM)GEN.PL

‘He bought **about ten antique books**’

[= (135c) in §5.2 below]

(108e) [...] sidelo **čelovek** sem´desjat slučajnyx posetitelej [...]
 sat people seventy chance spectators
 (V)PAST.NEUT.SG GEN.PL (NUM)NOM (ADJ)GEN.PL (N.MASC)GEN.PL

‘... there sat **about seventy chance visitors** ...’

[Skoblikova (1959:101), quoting Sajanov’s *Nebo i zemlja*, part 3, chapter 1]

(108f) [...] **čelovek** poltorasta anglijskix soldat ostalis´ [...]
 people 150 English soldiers remained
 GEN.PL (NUM)NOM (ADJ)GEN.PL (N.MASC)GEN.PL (V)PAST.NEUT.SG

‘... **about 150 English soldiers** remained ...’

[Skoblikova (1959:113), quoting Sergeev-Censkij’s *Sevastopol´skaja strada*, part 3, ch. 6.]

(108g) [...] stojalo **čelovek** pjat´ skromno odetyx ljudej.
 stood people five modestly dressed people
 (V)PAST.NEUT.SG GEN.PL (NUM)NOM (ADV) (ADJ)GEN.PL (N.MASC)GEN.PL

‘... there stood **about five modestly dressed people**.’

[Chey (1967:59), quoting Il’f & Petrov (1961:283); my glosses/LAB]

¹⁴⁸ Because of the additional material in the N’’ constituent in examples (108g-i), the non-COUNT, GEN.PL form *ljudej* is used. See example (141), as well as the footnote preceding that example.

(108h) V nebol'šoj komnate prisjažnyx bylo **čelovek** desjat' raznogo sorta ljudej.
 jurors was people ten of-various-kinds people
 (ADJ) (V) (N.MASC) (NUM) (NP)GEN.SG (N.MASC)
 GEN.PL PAST. GEN.PL. NOM
 NEUT COUNT
 .SG .NON-COUNT

'There were **aboutten jurorsofvariouskinds** in the small room.'

[*Sintaksis* (1980:331), quoting L. Tolstoj (no cit.)]

(108i) Okolo nego tolpilos' **čelovek**pjat' dvorovyx ljudej
 around him crowded people five court people
 (P) ACC.SG (V)PAST.PL GEN.PL (NUM)NOM(ADJ)GEN.PL (N.MASC)GEN.PL

'About twenty court servants crowded around him.'

[Pete (1984:74), quoting Turgenev (no cit.); my glosses/LAB]

As in (106), placing the pleonastic count noun before the numeral renders an approximative meaning. My primary discussion of approximative inversion is in the next chapter (§5.2). All that can be said here is that such inversion is usually a juxtaposition of the numeral and the noun which it quantifies. Such a juxtaposition is impossible if the constituent quantified by the numeral consists of more than one word. For example, if a numeral quantifies a noun modified by an adjective, then approximative inversion is not allowed: **posetitelej sem' desjat slučajnyx*, **slučajnyx sem' desjat posetitelej* are both illicit as ways of expressing example (108e). The additional material—usually an adjective phrase, italicized in (108a-g, i), but possibly an adnominal NP, as in the underlined words in (108h)—disallows approximative inversion. In such environments *čelovek* or *štuk* is uttered immediately before the numeral in order to achieve an approximative interpretation. The reason I claim that the examples in (108) are distinct from the preceding ones is due to their semantics: (108a-h) do not carry the nuance of emphasized quantity which (105) through (107) possess. It would appear, furthermore, that this use of a pleonastic count noun is related to the uses of such as word in (101a-b), neither of which carries this emphasis-

on-quantity interpretation.¹⁴⁹ I delay a structural analysis of these forms until after the primary discussion of approximative inversion (in §5.2).

I summarize the non-COUNT uses of *čelovek* (and *štuk*) briefly: (i) Adjectival nouns quantified by a numeral often have a pleonastic count noun with no added semantics, as shown in (101a-b), but not obligatorily, as attested by (100a-c). (ii) When a numeral modifies a multi-word constituent, as in (108a-i), then a pleonastic count noun precedes the numeral to express approximation, again with no added emphasis on quantity. (iii) When the numeral quantifies just a lone noun, as in (106a-e), then it is possible to insert *čelovek* or *štuk* between the numeral and noun to arrive at an interpretation of emphasized quantity, but no approximation. (iv) Such structures (described in the preceding sentence) can also place *čelovek* or *štuk* **before** both the numeral and noun and achieve an approximative interpretation and emphasis on the amount, as in (29a-d) and (107).

What, then, is example (99), repeated here as (109)?

- (109) Peredo mnoj stojalo [[četyre puški] i [dvadcat' pjat' **vzroslyx** čelovek]]. (?)
 before me stood four cannons and twenty five grown-up people
 (P) INST NEUT.SG NOM NOM NOM GEN.PL COUNT
 'There stood before me four cannon and twenty-five adults.'

This example unfortunately does not match the word order of any of the preceding emphasized-quantity uses of *čelovek*, which are either numeral + *čelovek* + quantified noun, as in (105) **or** *čelovek* + numeral + quantified noun, in (106a-e) and (107); or

¹⁴⁹ Unfortunately for his study, Sussex (1976) intersperses approximative uses of *čelovek* and *štuk* (in (108a-c)) with the other uses of these words, thereby totally obscuring the semantics of either of these constructions.

even the non-semantic examples of numeral + *čelovek* + adjectival noun, as in (101a-b).¹⁵⁰

Example (109) **does** appear to conform to the emphasized-quantity **semantics** of *čelovek* in (105), however. This can be demonstrated by the two paragraphs of text which precede this example, which shows quite clearly that ‘twenty five grown-ups’ is being emphasized:

“[...] Not having finished the tenth grade, the day right after welcoming the new year in 1943 I left for artillery officer’s school. [...] The two-year course of study had been crammed without abridgment into nine months.

“In the fall of that same year I took command of an artillery platoon. I had not even turned eighteen yet; before me stood four cannons and **twenty-five grown-ups**. ...”

[Vinokurov (1964:7-8); my translation/LAB]

In other words, the author is emphasizing the sheer number of people under his command. Additionally the author is unlikely to have approximated the number of soldiers, since he was their commander and would probably wish to express precisely the number of men. My informants, who find this example strange, prefer either *dvadcat’ pjat’ čelovek vzroslyx* (with the same emphasis on quantity) or just *dvadcat’ pjat’ vzroslyx* (but without such semantic overlay).

The upshot of this analysis of example (109) is that there are uses of *čelovek* that are related to what I call COUNT, but also semantically marked. It is clear from

¹⁵⁰ Sussex (1976) also provides examples of the following constituent orders: noun + numeral + pleonastic noun and noun + pleonastic noun + numeral. I have uncovered one example of the latter order which also has s:

Da	detej	štuk	s	pjatero
and	children _{GEN.PL}	items _{GEN.PL}	about(p)	five _{(COLL.NUM)ACC}

‘And there are **about five** children.’

[DePerno (1991:ch.1:11), citing Suprun (1964:68), quoting Gogol’s *Majskaja noč*.]

I suspect that these orders represent so-called genitive-initial sentences of House (1982), also called genitive themes by Franks & House (1982), which I discuss in §5.2 below using Mel’čuk’s term “emphatic-thematic inversion”. In such sentences the GEN noun need not actually be clause-initial, as example (137b) below shows. Example (97)/(99)/(109) is **not** this type of sentence.

the semantics of (109) that the adjective does not modify *človek* but is rather in N° position. I leave undecided the structural position of *človek* in amount-emphasis constructs, assuming that it is not within N'', the constituent quantified by the numeral. I show in chapter 6 that pre-numeric, approximative *štok* and *človek* occupy spec-of-NP (or -PP) position.

In order not to resort to Mel'čuk's trochaic-adjective explanation outlined above, it is necessary to also account as well for example (95b)—*dva žutkix čaSA* 'two terrible hours_{ADPAUC}'. Unlike the somewhat strange example (109), example (95b) is judged by my informants to be outright ungrammatical. Still, assuming that there are those who accept (95b), I propose the following brief observation:

Note that the numeral in (95b), *dva* 'two', is monosyllabic. I have encountered a similar set of judgments in the first few examples of this study above. Specifically, in my list of older examples which involve *s*-numeral-noun sequences in (8) through (14), all of which are unacceptable in modern Russian, two sentences in particular, repeated here in (110a-b), rather consistently garner a less-than-completely-bad judgment from my informants. That is, whereas the other examples with *s*-numeral-noun sequences are completely unacceptable, these only receive a double question mark. I list three more such examples in (110c-e) which come from recent studies and presumably quote modern sources. As in chapter 1, my informants' judgments about these examples are shown in parentheses following each example.

(110a) M"glá stojála po rjadu **s"** **dva** **měsjaca** (??)
 about two month
 ACC GEN.SG

'(The) gloom hung around **for about two months.**' [= (6) in chapter 1 above]

(110b) Poxodiv, po krajnej mere, **s** **tri** **časa** [...] (??)
 having-walked at least about three hour
 (P) ACC GEN.SG

'Having walked at least about three hours ...' [= (7) in chapter 1 above]

are both required for a numeral to be less than fully unacceptable in the structures in (110a-e). Returning to (95b), it may well be that there is such an exception in this case. I will return to both (110a-e) and (95b) in my discussion of monosyllabicity effects discussed in the next subsection.

To conclude the discussion, then, of the two solutions for the examples in (95b) and (97), I have pursued some possible solutions. Namely, either trochaic adjectives are exceptions to the generalization that adjectives cannot intervene between a numeral and an ADPAUC/COUNT form **or** the structure of (97) involves a non-COUNT use of *čelovek* and the monosyllabic, paucal numeral in (95b) is exceptional. Clearly none of these solutions is conclusive. Nonetheless, pursuing them has yielded other insights: in particular, an understanding of pleonastic count nouns with a meaning of unexpectedly high/low number.

Incidentally, Mel'čuk (1985:432) views what I call the single-word restriction in terms of obligatory contact between the numeral and noun. That is, the numeral and noun must be consecutive for the noun to exhibit the ADPAUC or COUNT form. I have not found any crucial evidence to decide between Mel'čuk's obligatory-contact observation and my single-word proposal. I merely pursue the merits of my single-word proposal here since my overall intention is to show that other constructions aside from *s+ACC* are subject to a single-word restriction.¹⁵³

Yet a third approach, aside from Mel'čuk's obligatory-contact proposal and my single-word restriction, to why a noun with ADPAUC/COUNT form is attested only in

¹⁵³ In a footnote above n §4.2.2 I mention that second-position clitics—either discourse particles or the YES/NO interrogative *li*—can break up a syntactic compound. I have not been able to elicit discourse particles like *že* or *ved'* between numeral and ADPAUC/COUNT noun, but I did elicit ..., *četyre li {čaSA/*ČAsa}*... 'four_{(NUM)NOM/ACC} YES/NO hour_{(N.MASC)GEN.SG{ADPAUC/*NON-ADPAUC}}' and ..., *pjat' li {čelovek/*ljudej}*... 'five_{(NUM)NOM/ACC} YES/NO people_{(N.MASC)GEN.PL{COUNT/*NON-COUNT}}'. Cf. Billings (1994b) and Parrott (1992). Such clitics, however, do not prove that the numeral and noun are syntactically separated. Still, such data would have to be accounted for by Mel'čuk's adjacency theory.

single-word environments is the following: It may well be that the ADPAUC/COUNT feature is not the property of just the noun but of the entire constituent quantified by the numeral or quantifier. Such a constituent, in the X-bar framework of Babby (1987), is N^{''}. Following Naylor (1977)—who convincingly argues that adjectival inflection must also be taken into consideration in ADPAUC phenomena, and using an ADPAUC phenomenon in Serbo-Croatian, in which only modifiers of paucally - quantified nouns show a distinct form—it is plausible that the ADPAUC is realized as a special form **resembling** the GEN.SG on **nouns** (in all but a few stems) and GEN.PL on **adjectives**. Unfortunately, however, Naylor fails to mention that the distribution of end-stressed ADPAUC nouns is limited to unmodified environments.¹⁵⁴ Moreover, this ADPAUC (or COUNT) feature must be realized morphologically **no more than once** in the quantified N^{''} constituent. That is, either the adjective exhibits GEN.PL inflection (more precisely, ADPAUC/ COUNT inflection) and the noun shows ordinary-GEN.SG (i.e., **non-ADPAUC**) or -GEN .PL (**non-COUNT**) inflection, **or**—when there is no adjective in the N^{''}—the noun exhibits distinctive ADPAUC or COUNT inflection. This idea might not be valid if the fraction numerals are considered: *pol* and *četvert´* never trigger morphological PL, but only the SG, in the nouns **or adjectives** they quantify (cf. Crockett 1976:399, fn 32). Thus, adpaucity, as far as Russian is concerned, is only exhibited by the noun, not by the N^{''} constituent as a whole.

¹⁵⁴ Even more unfortunately, Naylor lists as evidence the following example, which my informants **reject**, accepting only non-ADPAUC stress.

èti	tri	interesnyx	čaSA	‘those three interesting hours’
these/those	three	interesting	hour	[stress notation modified]
(DET)NOM.PL	(NUM)NOM	(ADJ)MASC.GEN.PL	(MASC)GEN.SG.ADPAUC	[Naylor (1977:91)]

This mistake is especially surprising considering that elsewhere in the article Naylor refers to “phonological limitations” on the distribution of the GEN-2 and PREP-2 cases. I assume he is referring to the apparent single-word restriction on the GEN-2, which I discuss in §4.6.3 above. Naylor also makes no mention of GEN.PL COUNT phenomena whatsoever.

In this subsection I have shown that ADPAUC and COUNT forms are, like *s+ACC*, subject to single-word restriction. These special forms of the GEN.SG and GEN.PL are used only if the noun is the **sister** of the paucal numeral or quantifier (respectively). I cannot conclude specifically that it is a syntactic word, as in *s+ACC*, but the data here are not inconsistent with such a specification. This concludes the four phenomena, aside from *s+ACC*, in which I have identified a single-word restriction. Before concluding this section, however, I investigate some possible single-syllable phenomena in Russian.

4.6.5 *Possible single-syllable restrictions:* The phenomena discussed so far in this section all deal with single-**word** restrictions of one kind or another; the following is a single-**syllable** restriction:¹⁵⁵

Mel'čuk (1985:223) reports that in certain two-part compound numerals (i.e., numerals consisting of two numeral stems, as in *dvadcat' dva* 'twenty-two') the first stem can apparently optionally fail to show inflection if the second stem is either 'two' or 'three'. Specifically, if the compound numeral is part of an overall nominal expression assigned GEN, PREP or DAT case, and the latter of the two numeral stems is 'two' or 'three', then the first stem of the compound numeral can optionally **not** show the GEN, PREP or DAT case but instead be homophonous with the NOM/ACC form:

(111a) V tridcat' **dvux** redakcijax [...]
 in thirty_{NOM/ACC} two_{PREP} editorial-offices_{PREP.PL}

'In the editorial offices of thirty-two newspapers ...'
 [= exx. 9a in Mel'čuk (1985:223), citing A.N. Tolstoj's *Giperboloid inženera Garina*.¹⁵⁶]

¹⁵⁵ Mel'čuk (1985:433) refers to yet another prosodic effect with trochaic adjectives which I mention in §4.6.4 above. At the end of that same subsection I consider another single-syllable effect.

¹⁵⁶ Thanks to A. Lebedev and A. Rakityanskaya for assistance with the glosses of this example.

- (111b) k *dvadcat'* **trěm** *studentam*
to *twenty*_{NOM/ACC} *three*_{DAT} *students*_{DAT.PL}
'towards twenty-three students' (or 'to twenty-three students' homes')
[= exx. 9b in Mel'čuk (1985:223); no citation.]

In prescriptive Russian both numeral stems must show inflection: *tridcati* 'thirty_{PREP}' in (111a) and *dvadcati* 'twenty_{DAT}' in (111b). If the second numeral is *dvux* 'two_{GEN/PREP}', *dvum* 'two_{DAT}', *trěx* 'three_{GEN/PREP}', or *trěm* 'three_{DAT}', then the first numeral can fail to show respective GEN, PREP or DAT agreement. Note that this characterization excludes the INST-case forms of 'two' and 'three'—*dvumja* and *tremja*, shown in (112a-b), respectively—and all forms of 'four', as shown in (113):

- (112a) s {**tridcat'/*tridcati/√tridcat'ju*} *dvumja* *studentami*
with {*thirty*_{NOM/ACC}/*thirty*_{GEN/DAT/PREP}/*thirty*_{INST}} *two*_{INST} *students*_{INST.PL}
- (112b) s {**dvadcat'/*dvadcati/√dvadcat'ju*} *tremja* *studentami*
with {*twenty*_{NOM/ACC}/*twenty*_{GEN/DAT/PREP}/*twenty*_{INST}} *three*_{INST} *students*_{INST.PL}
[= exx. 9v-g in Mel'čuk (1985:233)]
- (113) v {**tridcat'√tridcati*} *četyrěx* *redakcijax*
in {*thirty*_{NOM/ACC}/*thirty*_{GEN/DAT/PREP}} *four*_{GEN/PREP} *editorial-offices*_{PREP.PL} [= his exx. 9d]

It is peculiar indeed for the oblique cases (GEN, PREP, DAT, INST) not to behave alike.¹⁵⁷ It is likewise odd that the paucal integers fail to act uniformly (i.e., 'four' is not allowed to optionally undergo this phenomenon as 'two' and 'three' do).¹⁵⁸

¹⁵⁷ There is one other way in which the INST case differs from the other oblique cases in Russian. The closest other distinction between INST and the other oblique cases that I am aware of is found in Ukrainian and W. Slavic languages: In those languages third-person personal pronouns (like Russian) have *n*-initial forms regardless of whether that pronoun is the object of a preposition. Ferrell (1958) reports that in Ukrainian, Slovak, Polish and Upper Sorbian an INST third-person personal pronoun (unlike Russian) is always *n*-initial, regardless of whether it is the object of a preposition. Townsend (1990:65; 66, n. 6) reports something similar in spoken Prague Czech (SPC), at least for the INST.PL third-person personal pronoun *jima* versus *nima*. Whereas he lists both pronouns in his table (p. 65), he mentions in a note that the former is not used much in SPC because in many cases the use of the prepositionless INST has been replaced by an *s*+INST preposition phrase. He also tells me personally that the same is true of the INST.SG third-person personal pronouns. It appears, then, that this Ukrainian and W. Slavic phenomenon has to do with the infrequency of non-prepositionally governed pronouns and not with any special syntactic properties of the INST case *per se*.

¹⁵⁸ The closest other phenomenon that comes to mind with regard to 'four' acting differently from either 'two' or 'three' in Russian is the forms for 'forty' (*sorok*) being of an entirely different stem

Mel'čuk (1985:247) conjectures that the crucial factor distinguishing *dvux*, *dvum*, *trëx*, and *trëm* is that these are the only monosyllabic numerals in an oblique case; the INST forms of 'two' and 'three' are disyllabic; all case forms of 'four' in Russian are trisyllabic.

I would modify Mel'čuk's observation as follows: First, I agree that prosody **does** appear to be the factor. I would, however, not refer to this as "non-declension" only in the oblique cases.¹⁵⁹ Instead, this phenomenon extends to **all** six cases (including the NOM and ACC). Whenever the second part of a complex numeral is monosyllabic, regardless of case, the first part optionally fails to show inflection. Examples (114a-b) show the declensional paradigms of the numerals 'twenty' and 'three' (the endings of 'thirty' and 'two' are analogous to these, respectively):

	NOM	ACC ¹⁶⁰	GEN	DAT	PREP	INST
(114a) 'twenty'	dvadcat´	dvadcat´	dvadcati	dvadcati	dvadcati	dvadcat´ju
(114b) 'three'	tri	tri / trëx	trëx	trëm	trëx	tremja

I wish to make the following point: If the second part of the compound numeral (in this case the 'three' in '23') is monosyllabic, **regardless of its case**, then the preceding part of the complex numeral—i.e., the first digit in '23'—can optionally have no inflectional ending: /dvadcatj + Ø/, which is spelled *dvadcat´*, homophonous with

while *dvadcat´* and *tridcat´* are etymologically related to *dva* and *tri*. This is, I should emphasize, not a productive derivation (and the factor responsible for this diversion was most likely a very prevalent measure noun, like *dozen*, meaning 'unit-of-forty'), as though *dozen* had replaced the word *twelve* in English, for example. This may be an example of a noun diachronically becoming a numeral. Synchronically, in modern Russian *sorok* is only a numeral. Cf. also Schütz (1986).

¹⁵⁹ Mayer (1967:308-09) concludes that Slavic cardinal numerals are heading away from their nominal roots to being "simply a marker of quantity and nothing else." This may be a step toward it.

¹⁶⁰ The two ACC forms for 'three' correspond to the inanimate and animate forms (though the latter is becoming less obligatory in the spoken language with time; cf. Blažev 1966 for recent data and Franks 1995:156-57, including n. 47 on p. 214, for a structural distinction between the two). Cf. §3.3.

inflected NOM and ACC forms of this numeral. That is to say, even when the latter part of the compound numeral is NOM/ACC *tri* ‘three’, then the first part can, as it were, take no inflection.¹⁶¹

To conclude this subsection, I have shown that, under certain circumstances, a monosyllabic constituent can be exceptional. This, along with the monosyllabicity exception discussed in the preceding subsection, suggests that there may be certain phenomena whose distributions are limited to single-syllable constituents. I should emphasize that the syllable is a prosodic category, with no direct analogue in the syntax or morphology. I have also investigated other size limitations such as single-word limitations. There are both prosodic and syntactic instantiations of “word”. The size restriction on *s*+ACC, as argued in the preceding section, must be a **syntactic** word. The other single-word phenomena above in this section are consistent with a syntactic definition of “single word”.

This is not to say that no single-prosodic-word interactions with syntax exist in Russian. I show in the next chapter that the quantified noun in approximative inversion must be a single prosodic word. In chapter 6 I show that the landing site of the quantified noun in approximative-inversion constructions is sensitive to the prosodic word. I also show elsewhere (Billings 1995c) that the order of prepositions with the negation clitic *ni* in negated prepositional phrases is sensitive to the prosodic word. In yet other work (Billings 1994a; 1995a; 1995b) I show that the distribution of

¹⁶¹ The form *dvadcati* is also listed in (114a) because there is a tendency for some complex numerals to simplify to “a two-way contrast between a direct-case form [i.e., the NOM/ACC *dvadcat*’ in (114a)] and a **single** oblique-case form equivalent to the genitive” (Mayer 1978:217, an empirical study of how modern speakers of Russian pronounce various oblique-case forms of compound numerals). The phenomenon described here, however, appears to be based on a different diachronic trend: “toward a total analytical development of numerals, whereby the numeral becomes an indeclinable attribute” (*Ibid.*). It is therefore reasonable for Mel’čuk to attribute this phenomenon only to oblique cases (Mayer’s “two-way contrast” above). My extension of his explanation to all the cases follows Mayer’s “total analytical development”. (Following Mayer 1976:27, I synchronically treat *dvadcat*’ ‘twenty’ and *tridcat*’ ‘thirty’ as single simplex syntactic units; see also my fn. at the end of chapter 1 above.)

nasal-initial third-person pronouns after prepositions in Russian was originally sensitive to a single-prosodic-word environment and has since been reanalyzed to a single-syntactic-word environment.

To summarize this section, I have shown that the Russian language places restrictions on the size of various constituents. The *učit'sja na+ACC* construction (§4.6.1) requires a single-word complement as does *s+ACC*. Not just “constructions” have this type of constraint. When there are two different forms taken by a particular case-and-number combination, as in choosing between GEN-2 and GEN.SG (§4.6.3), or between ADPAUC and GEN.SG and between COUNT and GEN.PL (§4.6.4), then the marked member of that pair of forms can be restricted to appearing in environments where it is the only word. Finally, a diachronic phenomenon is taking place whereby certain monosyllabic elements are exceptional (§4.6.5). In short, *s+ACC* is not alone in placing a size restriction on its complement.

I conclude this chapter on multi-word complements of *s+ACC* by repeating the primary criterion for ruling out multi-word ACC-case complements of *s*: Limit the complement in the *s+ACC* construction to a single SnWd [= (81c)]. There remain, therefore, three types of data that are exceptions to this restriction: *s*-adjective-noun sequences in which the adjective delimits the measure (§4.2.3), the *s-pol*-noun data (in §4.3.5), and complements of *s* which consist of a measure noun with its own noun-phrase complement (§4.4). I return to these exceptions in the final chapter where I present Optimality-theoretic constraints to account for them.

Chapter 5 Other approximate-measure constructions:

Whereas *s+ACC* has not been investigated in depth, two other approximative constructions have been extensively researched: the preposition *okolo*, which can mean either ‘approximately’ or ‘near’, and approximative inversion, reversing the order of numeral and noun to express approximation. I present the relevant details of these two constructions here in order to contrast them with *s+ACC*. In the first section I also show that *s+ACC* and *okolo* are distinct kinds of prepositional quantifiers; of the two, only *s* heads a PP and has a full NP as its complement. In the second section I investigate how *s+ACC* and this approximative inversion interact in the same utterance. As I mention above in the introduction, much of the work on these two constructions is in Babby (1985; 1987) and Mel’čuk (1985), respectively. In a brief section I also briefly discuss *ètak* ‘approximately’ and the property it shares with *s+ACC*, requiring approximative inversion. I conclude the chapter with an analysis of *neskol’ko* ‘several’; I assess several indicators which suggest that this word is a numeral. This chapter does not exhaust all means of expressing approximation, but it does show how approximative mechanisms interact, especially in connection with the *s+ACC* construction.

5.1 On approximate-measure constructions with *okolo*

One way to express approximate measure in modern Russian is to use the GEN-assigning preposition *okolo*.¹⁶² Babby (1985) shows that *okolo* can have either a **proximate** meaning (as in ‘near’) or an **approximative** meaning (as in ‘about/approximately’). He explains these two meanings using the structures corresponding to the same three words in (115a-b), arguing that it is the two different bracketings which cause the difference in meaning.¹⁶³

(115a) Locative reading

okolo [*desjati sosen*]
 near ten_{GEN} pine-trees_{GEN.PL}

(115b) Quantificational reading

[*okolo desjati*] *sosen*
 about ten_{GEN} pine-trees_{GEN.PL}
 [≈ exx. 14a-b in Babby (1985:98)]

I should point out that the structures in (115a-b) are by no means agreed upon. Other studies, namely Neidle (1988:160-65) and Franks (1995:143-44) disagree with the structure in (115b), arguing that both usages of *okolo* have essentially the phrase structure in (115a). In the next section I show that these two interpretations of *okolo* have different orders when there is approximative inversion (cf. exx. (131) and (133) below), thus providing empirical evidence that Babby is correct. Below in this section I repeat some of Babby’s rationale for the structure in (115b).

One of Neidle’s and Franks’s arguments against Babby’s proposal is about the source of GEN case on the quantified noun. Babby writes the following:

¹⁶² It appears that in the formal/standard language *okolo* means ‘approximately’, while in the informal/conversational register *okolo* can mean ‘nearly/not quite’. Ušakov (1938:788) and Mel’čuk (1985:362) both list the meaning ‘almost, not quite’ for *okolo* (the former adds that this meaning is “conversational”). Dal’ (1989b:665) does not, however, report this ‘not-quite’ meaning. None of my own informants, however, are aware of this other meaning. Interestingly, the meaning of English *nearly* is synonymous with ‘not quite’ and is morphologically related to *near* just as the quantificational and non-quantificational meanings of *okolo* are lexically related.

¹⁶³ Incidentally, unlike the two bracketings of English examples like *beautiful girl’s dress* (which Babby (1985:98) uses as another example of structural ambiguity), the Russian bracketings in (115a-b) do **not** appear to be prosodically distinct from each other.

“[... In (115a)], the locative reading, the genitive case marking on *sošen* is due to the preposition *okolo*, which governs the genitive case, and not the quantifier *desjati*, since the latter is **optional**, and *sošen* is marked genitive even when it [the numeral] is absent (e.g., *Daču postroili okolo sošen* ‘The cottage was built near the pines’). ... recall that a quantifier can impose genitive marking on a N in its scope only if the NP dominating them is in a direct case [specifically, only if the numeral bears morphological nom case/LAB ...]

“[... In the quantificational reading in (115b)] the genitive case marking on *sošen* is not due to the preposition *okolo*, but to the entire constituent *okolo desjati* ‘about ten’, and this can be easily demonstrated: if *okolo* is removed (which is possible only under the [reading in (115b)], *sošen* remains in the genitive (e.g., *On posadil desjat’ sošen za domom* ‘He planted ten pines behind the house’). Thus *okolo* in [115b] governs only *desjati* [...]”

[Babby (1985:98-99); underlined notation updated to conform to Babby (1987)]

Both Neidle and Franks counter-argue that if the numeral after *okolo* is *dvux* ‘two_{GEN}’, *trex* ‘three_{GEN}’ or *četyrex* ‘four_{GEN}’ (instead of *desjati* ‘ten_{GEN}’ in (115) above), then the quantified noun must nonetheless be in the GEN.PL, not the GEN.SG as expected of the so-called paucal numerals. Their argument is even apparently bolstered by the fact that if the quantified noun in (115a-b) is one which shows a distinct COUNT form (e.g., *čelovek* ‘people’), whether or not the numeral is paucal, then the COUNT is required: *okolo* {*dvux/trex/četyrex/desjati*} {[√]*čelovek*/**ljudej*} ‘near/approximately {two/three/four/ten} people_{COUNT}’. Their argument is at best inconclusive, however. The only circumstances under which a quantified noun takes the GEN.SG (or special ADPAUC form) is if this noun is quantified by a paucal numeral *with morphological nom case* (cf. §4.3 above). The fact that the quantified noun in (115b) is in the GEN.PL only indicates conclusively that it is quantified. Various other kinds of quantifiers trigger COUNT forms: measure nouns (§4.3.3), non-numeral number nouns like *million* (§4.3.4), and even so-called adverbial quantifiers like *neskol’ko* ‘several’, which I discuss below (in §5.4). If such a varied array as these can all trigger the GEN.PL, then why is it not possible for the combined constituent

okolo desjati in (116b) to also trigger the COUNT form?¹⁶⁴ In both structures of (115a-b) above the **numeral**, be it *desjati* ‘ten’ or one of the paucal numerals, clearly receives GEN case from *okolo*. This is not true, however, with regard to the GEN marking on the noun (115b), be it the simple GEN.PL *sošen* ‘pines’ or the special COUNT *čelovek*. I return now to my comparison of *okolo* and *s+ACC*:

Koka (1955), a study that deals exclusively with time expressions in Russian (and hence does not consider all of the uses of *s+ACC* or *okolo*), compares the applicability of replacing *s* with *okolo* and vice versa:

“[...] When expressing approximation in a particular period of time the synonymous constructions *s+ACC* and *okolo+GEN* are used; for example: [...] *s čas* ‘about an hour’ — *okolo časa* ‘about an hour’. [...] Compare also the following:

[116a] *S četvert’ časa deržal on obeimi rukami ruku Čičikova.*
 about quarter_{ACC} hour_{GEN.SG}

‘He held Čičikov’s hand with his own two hands for about a quarter of an hour.’

[116b] *Okolo četverti časa ... provozilsja on s kuznecami.*
 about quarter_{GEN} hour_{GEN.SG}

‘He hung around with the blacksmiths for about a quarter of an hour.’

The *s+ACC* construction, however, in the contemporary standard language does not have a wide distribution, since constructions with this preposition are extremely limited as to their formation by its lexical properties. As for constructions with the preposition *okolo*, they do not have such limitations.”

[Koka (1955:111), quoting N. Gogol’ (no cit.) in both exx.; my translation/LAB¹⁶⁵]

¹⁶⁴ The other main prepositional quantifier in Babby (1985), distributive *po* ‘each/apiece’, likewise triggers COUNT forms: *po pjat’ čelovek* ‘[five_{ACC} people_{COUNT}]_{ACC} apiece’ (**po pjat’ ljudej*_{NON-COUNT}). My informants could not judge archaic *po* which assigns DAT to the numeral. Because *po* takes a direct case, the ADPAUC is also attested: *po dva čaSA* ‘[two_{ACC.nom} hours_{ADPAUC}]_{ACC} each’. This stress would suggest, following my arguments in §4.3 that the phrase structure is [*po* [*dva čaSA*]]. In order to preserve Babby’s (1985; 1987) bracketing — [*po dva*] *čaSA* — I must resort to a relativized-head model discussed below in this section. The PP headed by *po* inherits the paucal feature and in turn triggers the ADPAUC form in this noun. This assumes, reasonably, that *po* is not marked for the paucal feature.

¹⁶⁵ This extended quote in the original: „Pri vyraženii priblizitel’nosti togo ili drugogo otrezka vremeni upotrebljaetsja sinonimičeskie konstrukcii s predlogom «s» v sočetanii s imenem v forme vinitel’nogo padeža i konstrukcii s predlogom «okolo» v sočetanii s imenem v forme roditel’nogo padeža, naprimer: [...] *s čas* — *okolo časa*. [...] Sr. takže: [exx. (116a-b)] Odnako konstrukcii s predlogom «s» v sovremennom literaturnom jazyke ne imejut širokogo rasprostraneniya, t.k. konstrukcii s etim

Footnote continued on next page

Koka’s statement is a bit misleading in a few ways: First, it is not coincidental that *četvert´* is used for the comparison in (116a-b): *četvert´* and *tysjača* ‘thousand’ are the only two number words that can be used in both examples (and still leave *okolo* with only an approximate-measure interpretation).¹⁶⁶ Whereas *okolo dva časov* ‘around two_{GEN} hours_{GEN.PL}’ is licit, **s dva časa* ‘about two_{ACC} hour_{GEN.SG.ADP}’ is not, because *s* forbids multi-word complements (cf. §4.3). Next, Koka’s claim regarding the lack of limitations on *okolo* is likewise deceptive: As Babby (1985:98) points out, *okolo* cannot have an approximative-quantificational meaning without some sort of quantifier—either a numeral, as in (115b), or a measure noun, such as *čas-* in (116b)—as its complement. Lacking a quantifier *okolo*+GEN has only a locative-proximative meaning, while *s*+ACC has both approximative and quantificational semantics. Koka’s claim that *okolo*’s distribution is not limited has to do with the limited scope of that study, which deals only with time expressions, which invariably include a measure noun of some sort as the complement of *okolo*.¹⁶⁷ Finally, Koka fails to examine the stress of *časa* in (116a-b). While it is not certain where the stress was pronounced when Gogol´ wrote both examples in the 1800s, my elicitations from contemporary speakers indicate that the stress is non-ADNUM ČAsa in both.

predlogom krajne ograničeny v svoem obrazovanii leksičeskim sostavom. Čto že kasaetsja konstrukcii s predlogom «okolo», to oni ne imejut takogo ograničenija.”

¹⁶⁶ If *tysjača* ‘thousand’ were used, then the noun would have to be GEN.PL *časov* in both.

¹⁶⁷ There is one type of word, which might be referred to as a “time expression”, which is not a measure noun: There are two words in the modern language that begin with *pol* which do not synchronically involve the numeral *pol* ‘half’. These are *polden´* ‘noon’ and *polnoč´* ‘midnight’. Unlike the numerical uses of *pol*, here the main word stress is initial: *POLden´*, *POLnoč´* (in §4.3.5 I show that *pol* has secondary stress, with main stress on a syllable in the word it quantifies); here *pol* does not trigger the GEN.SG in the following stem; finally, unlike the numerical uses of *pol*, the oblique-cases of these words cause the first part to be *polu* and the second part to be in that oblique case (e.g., *do polunoči* ‘before midnight’; in these cases the stress is on the second syllable: *poLUnočĭ*). Note also that these words are illicit as ACC-case complements of *s*: **s polden´*, **s polnoč´*. And after *okolo* they only have the locative meaning: *okolo poludnja* ‘about noon’, *okolo polunoči* ‘about midnight’ (specifically, a proximate-temporal one—i.e., ‘near’ that point in the timeline—despite the ‘about’ in the English gloss). Thus, I should exclude these two words from my definition of “time expression”.

part after *i* ‘and’ in example (118) to be conjoinable with the non-approximative numerical expression before the conjunction, it is reasonable to assume parallel structures (namely, one in which *okolo* is sister of *desjatka*), thus also ruling out (117c, g), leaving only (117d, h):¹⁶⁹

(118) Conjoined quantified expressions

[[Vosem´(QP).NOM	[krepostnyx sten](N´)GEN](NP)NOM	i	[[okolo desjatka](QP)NOM
eight	fortified walls	and	about unit-of-ten
[nebol´šix fortov](N´)GEN](NP)NOM](NP)NOM	zaščiščajut(V)PRES.3PL	gorod(np)ACC](VP)
small forts		defend	city

‘Eight fortified walls and about a dozen small forts defend the city.’ [= (59a) above]

If this particular argument still appears insufficient, I present another empirical argument in support of this structure for quantificational *okolo* in the next section; cf. (131) and (133) below.

Example (118) is most opportune to the present study because *desjatka* is not the numeral form of ‘ten’ (cf. *desjati* in (115a-b) above), but rather the measure noun, which I gloss, following Babby, as ‘unit-of-ten’.¹⁷⁰ Above (in §4.3.2-§4.3.3) I show that the noun version of *čertvert´* ‘quarter’ corresponds to this measure noun *desjatka* (here in the GEN.SG). Thus, not only does (118) argue for the phrase structure in (117b, d, f, h)—that is, the right-hand column of (117)—it also proves that a measure noun can appear in this structure. This leaves only one structure with two part-of-speech labelings for *okolo čertverti časa*, namely just (117d, f). Based on the apparent fact that either a noun or a numeral can occupy the sister position of *okolo* (as shown

¹⁶⁹ My informants reject the corresponding structure with s+ACC instead of *okolo*: *Vosem´ krepostnyx sten i s desjatok nebol´šix fortov zaščiščajut goroda. Below in this section I show that only some prepositional quantifiers can be conjoined with numerical expressions; s+ACC is not one of them.

¹⁷⁰ I deviate, however, from Babby’s sentential gloss by rendering *okolo desjatka* as ‘about a dozen’, which preserves the measure-noun tenor, if not its numerical accuracy.

empirically in (118) and (115b), respectively), I conclude that *okolo četverti ČAsa* in (116b) can have either of the structures in (117d, h). In either event the stress is non-final on *ČAsa* (i.e., not **okolo četverti čaSA*).

The preceding arguments and the data in (57) lead me to propose the following quite divergent structures corresponding to the examples in (116a-b), respectively (both presented in terms of the NP model in Babby 1987):

- (119a) [[*okolo*_{P°} *četverti*_{N° or Num°}] PP [*ČAsa*_{N°}] N°] NP
 GEN<SG> (NOM/ACC) GEN.SG GEN.SG (NOM/ACC)
 + PROX – PROX + PROX – PROX – PROX + PROX
 ∅ + Q + Q + Q + Q
- (119b) [*s*_{P°} [[*četvert*_{N° only}] [*ČAsa*_{NP}] NP] PP
 ACC.SG GEN.SG ACC.SG (NOM/ACC)
 + PROX – PROX – PROX – PROX + PROX
 + Q + Q + Q + Q + Q

Several comments on the notation used in (119) are in order:

First, these two examples have slightly different semantics: Both *okolo* and *s* share one semantic feature: [+ PROX(imate)].¹⁷¹ They differ, however, with regard to the other semantic feature [Q(uantification)]: *okolo* is not marked for the Q feature, while *s* is [+ Q]. At this point the complement of *s* need not be [+ Q] (or [– Q], for that matter). The Q feature percolates all the way to the PP node in both structures.

¹⁷¹ Macdonald (1972), a study restricted to the semantics of the prepositions of time in Russian, uses the feature PROXIMATE (based on his earlier work on English prepositions) to describe *s*+ACC and *okolo*. In fact, these two prepositions are the only ones in his article that have this feature. They also share the feature INCIDENT (as opposed to PRIOR or SUBSEQUENT) in Macdonald's system. Moreover, both *okolo* and *s* have only these two features. I would object to the suggestion that *s*+ACC and *okolo* share exactly the same features. Macdonald assesses only prepositions of time and does not discuss any Q feature. Invariably, in my system these two prepositions share the features [+ Q, + PROX] if they are in time expressions (because *okolo* is sister to a [+ Q] element: either a numeral or a measure noun). I would also object to the proposal that either of these two prepositions inherently has any INCIDENT feature, since they are both attested with non-time uses. Maybe there is some feature, perhaps my Q, which allows for an "incident" interpretation when these prepositions are in some structural relationship with a measure noun of time. In any event, I agree with Macdonald's PROXIMATE feature, with the proviso that it extend to non-time/-quantificational uses, as in (115a).

Because *okolo* must have a [+ Q] sister to obtain the approximative meaning, the phrase structures of the two prepositions are decidedly different as well.

As for their syntactic specifications, *okolo* and *s* differ in the case they assign: *okolo* assigns GEN; *s* assigns ACC. This distinction is realized morphologically only on the word for ‘quarter’: *čtvrť*_{GEN} and *čtvrť*_{ACC}. The GEN .SG *ČAsa* in each receives its case from different sources: In (119a) the [+ Q] PP assigns the GEN of quantification to its N'' sister *ČAsa*; in (119b) the N°, which happens to be [+ Q], assigns the adnominal genitive to its NP sister *ČAsa*. It is inconsequential that *čtvrť* in (119b) is [+ Q], because the adnominal GEN does not depend on it.

Before continuing with the comparison of *s* and *okolo* it is actually worth showing that *s* is very much a preposition, not just a clitic which precedes an element and adds approximative meaning: It has been argued in the literature that *s* does not really assign any case. Gladney (1986:141) argues that “because *s* [...] does not determine the case form of the accompanying NP,” there is no reason to call this a prepositional phrase. His argument hinges on the apparent assumption that this construction’s distribution is that of ACC-case adverbials of time duration only, as the following type of example shows: *Saša rabotal (s) čas* ‘*Saša*_(MASC) worked_{(V)MASC.SG} (about) hour_{(MASC)ACC.SG}’ [Gladney (1986:141); my glosses/LAB].

Such an assumption is mistaken: *s*+ACC’s distribution is the same as that of other prepositional-quantifier constructions (as well as the GEN of negation,¹⁷²

¹⁷² In one example *s*+ACC there is also supposedly GEN or negation:

Molodec"	s" voz",	a	uma	s" nakopyl'nika	nětu. (??)
youth	cart	but	mind	running-board	not
(N.MASC)NOM.SG	(N.MASC)ACC.SG	(CONJ)	(N.MASC)GEN.SG	(N.MASC)GEN.SG	(NEG)

‘The young man is as big as a cart, but (his) brains don’t even reach the running board.’
[Dal’ (1991:373), no citation given]

Dal’ (1989b:427) has a slightly different version of what appears to be the same example. Dal’ (1991:373) adds that the GEN case is due to of negation [„rodit. pad. po otrican’ju”]. It is unclear which noun (*uma* or *nakopyl’nika*—or both) is supposedly affected by the GEN of negation. It would

Footnote continued on next page

comparatives without *čem* ‘than’,¹⁷³ and the GEN case assigned by quantification¹⁷⁴)—namely, (surface) subject of intransitive or copula verbs, direct object of transitive verbs, and (ACC-case) adverbials of time duration.¹⁷⁵ While I have not found a clear case of an intransitive verb with an subject containing an *s*+ACC phrase,¹⁷⁶ I did find the following types of examples:¹⁷⁷

appear that Dal’ has the complement of *s* in mind, since there is no other explanation for the GEN case in that word. The GEN case of *uma can* be explained perhaps as the partitive-existential. In any event, this example was strange at best to my (modern-Russian) informants.

¹⁷³ See the second half of example (109) above for an example of a *čem*-less comparative.

¹⁷⁴ All three constructions—prepositional quantifiers, GEN of negation, and non-*čem* comparatives—share this distribution for a reason: Babby’s (1987:116) Syntactic Case Hierarchy:

Lexical Case > GEN of quantification > (structural) NOM/ACC [his ex. 51]

Babby’s hierarchy can be stretched by inserting all **four** of these limited-distribution constructions in the middle portion of this hierarchy, because lexical case has precedence over any of these. The only problem is that prepositions and *čem*-less comparatives have no way of exhibiting lexical case, and GEN of negation or quantification merely gets overridden by lexical case. (It is interesting to note that Babby’s hierarchy looks very much like an Optimality hierarchy. I leave for future research whether it can be translated into Optimality terms.)

¹⁷⁵ Cf. Babby (1985:100, 114 n. 6); Baš *et al* (1959:165-66) essentially say the same thing. Mel’čuk (1985:367) mentions a fourth structural position in which prepositional approximative expressions can occur, apparently appositives of other NPs: *pri temperature okolo tysjači gradusov* (which, in its English gloss, requires *of*) ‘at (a) temperature_{PREP.SG} (of) about 1000_{GEN(SG?)} degrees_{GEN.PL}’. I use this test below in this section to distinguish between two kinds of prepositional quantifiers.

¹⁷⁶ The following example, although in *Sintaksis* (1980), is not part of Gladney’s corpus:

Prošlo	/ prošel	s	mesjac	vremeni
passed _{(V)PAST.NEUT.SG}	passed _{(V)PAST.MASC.SG}	about _(P)	month _{(N.MASC)NOM.SG}	time _{(N.NEUT)GEN.SG}
‘About a month’s time passed’			[<i>Sintaksis</i> (1980:245), emphasis added/LAB]	

I assume that this example is intended to mean that either the NEUT.SG or the MASC.SG form of the verb may be used prior to *s mesjac vremeni*. This is incorrect! Only *prošlo*, the NEUT.SG-agreeing verb, is allowed in this sentence. This is clearly an error in the Academy grammar. (As S. Franks pointed out to me, the remaining licit verbal form no longer supports the subjecthood of the *s*+ACC phrase in this particular example.) I am aware of no register or dialect of Russian that allows such predicate agreement with the object of *s*.

¹⁷⁷ In fact, both of (120) and (121) are in Gladney’s corpus, which he describes as “the 197 items listed in the index to volume 2 of *RG* [= *Sintaksis* (1980)] under the heading ‘prepositions and prepositional formations.’ [Gladney (1986:133)]”.

reasoning would suggest that in examples (121)-(122) the noun *desjatok*, which has homophonous NOM and ACC forms, is in the NOM case. However, replacing *desjatok* with the word ‘unit-of-hundred’ quickly dispels that reasoning: *s* { $\sqrt{\text{so}t\text{n}j\text{u}}/*\text{so}t\text{n}j\text{a}$ } *učenikov bol’ny* ‘about unit-of-hundred_{{\check{V}ACC/*NOM}} schoolchildren are sick’ or ... *s* { $\sqrt{\text{so}t\text{n}j\text{u}}/*\text{so}t\text{n}j\text{a}$ } *ljudej dvigalis’* ... ‘about unit-of-hundred_{{\check{V}ACC/*NOM}} people moved’.¹⁸¹ It is clear, therefore, from examples (121)-(122) that *s*+ACC phrases can appear in sentential positions usually assigned NOM case as well, thereby requiring *s* to be the only possible source of the ACC case on its complement in such examples.

I do not deny that *s* is a clitic. As a prosodically light preposition, as I show in the next section, *s* obligatorily procliticizes to the first word of its complement. Like most any other preposition, *s* triggers nasal-initial forms of a third-person, personal-pronoun complement: *budu l’ ja s nego?* (literally ‘will_{1.SG} Y/NCL I_{NOM} about_(P) him_{ACC}’) ‘Will I be about as big as him?’ [Ivšić (1950:364), quoting Krylov’s *Ljaguška i vol*; also in Ušakov (1940:15)].¹⁸² Thus, *s* plays very much an active syntactic role, obligatorily assigning ACC case and discharging other preposition-like roles, such as requiring nasal-initial pronouns.

Continuing the comparison of the properties of *s* and *okolo*, neither preposition in (119) selects a particular part of speech for its complement: If *okolo*’s complement happens to be a [+Q] nominal expression, then the PP headed by *okolo* is likewise marked as [+Q]. Likewise, *s* can take a [+Q] nominal complement, but does not have to do so. The similarities end here, however: Crucially, the configurations differ in

form adjectives as predicates must show PL agreement with a quantified subject if the subject precedes it, the reasons for which awaits further study.

¹⁸¹But cf. the W. Ukrainian example in (123), in which *s* appears to assign no case whatsoever.

¹⁸² Cf. Hill (1977) for the circumstances under which prepositions take nasal-initial pronouns.

whether or not the PP is subordinated to the NP. I resort to a set of syntactic features and to the notion of “relativized head” in order to explain this difference.

Fowler (1988:254-70) deals with hybrid syntactic categories of Russian quantificational words by enhancing a familiar set of syntactic features, [$\pm N, \pm V$]. That is, nouns are [$+ N, - V$]; verbs are [$- N, + V$]; adjectives are [$+ N, + V$]; and prepositions are [$- N, - V$]. This exhausts the permutations of these two binary features. He adds [$\pm Q$], a **syntactic** feature, allowing for quantificational and non-quantificational variants of the four parts of speech. In his treatment of prepositional quantifiers, Fowler (1988:321-25) largely re-analyzes the data in Babby (1985) in terms of this expanded syntactic-feature matrix.

In order to account for some unique quantifiers, Fowler proposes a model in which certain syntactic heads lack particular features. Instead of a relativized-head model, Fowler’s specialized categories remain lacking in one or the other feature. Recall also that my Q feature is strictly semantic, while Fowler’s (1988:262) is both semantic and syntactic.

Here I argue that *okolo* does not have a specification with regard to the semantic Q feature, hence the “Ø” on the third row under *okolo* in (119a); *okolo* is, therefore neither non-quantificational nor quantificational in this regard.¹⁸³ When its complement is [$+ Q$], then the PP headed by *okolo* assumes the [$+ Q$] specification, which further allows this [$+ Q$] PP to quantify N’’ within NP. When *okolo* happens to take a [$- Q$] complement, then the PP of which it is the head is likewise [$- Q$] and it therefore cannot enter into the quantifier position in the NP and is limited to a non-quantificational (but still [$+ PROX$]) semantic interpretation. By contrast, *s* is

¹⁸³ The other prepositional quantifier which Babby (1985, 1987) proposes to be sister of N’’ is *po* ‘each/apiece’, which doesn’t appear to have a non-quantificational counterpart. I suggest, without working out the details, that unlike *okolo*, which can discharge its [$+ PROX$] feature regardless of whether it is quantificational, *po* must be [$+ Q$] as a precondition for discharging its distributive feature.

inherently [+ Q] and will therefore project a [+ Q] PP regardless of its complement's specifications with regard to this feature.

In (119a-b) the PP node in each structure is categorially prepositional and has the semantic feature [+ Q] (and [+ PROX]). So far I have not proposed any distinction to prevent the PP headed by *s* from being the quantificational sister of N''. This does not mean that the Q feature is not needed; I abide by Q as a strictly **semantic** feature—a necessary, although not sufficient, condition being sister of N'' within NP. This means, however, that a **syntactic** distinction between *s* and *okolo* is still required (aside from the different cases they assign), which in turn predicts that only one of these—*okolo*—will be within the NP as sister of N''.

If the specification [+ Q] is a precondition for being the sister of N'', then the vast majority of Russian prepositions are excluded by merely being [– Q]. PPs can be sister of N'' only by having the specification [+ Q], either by having a [+ Q] preposition head, or by having a [+ Q] sister to a prepositional head, like *okolo*, which is not marked for the Q feature.

This leaves *s*, which cannot be in the sister of N'' despite being [+ Q]. I again resort to a relativized-head solution, extending Fowler's syntactic-feature approach: Like *okolo*, *s* can also be in the sentential subject, as evidenced most clearly by the PL predicative agreement in (61c)/(121): *S desjatok učenikov bol'ny* 'about_(P) unit-of-ten_{(NUM)ACC} schoolchildren_{(N.MASC)GEN.PL} (are) sick_{(V.SHORT-FORM)PL}'. I propose, however, that *s* is the **head** of the maximal projection in subject position. Rather than attribute an *ad hoc* feature on *s*, it seems more prudent to consider an **impoverished** feature setup for *s*. Following the feature matrix used by Fowler, it is possible that one of the two features of this preposition is left unspecified. Prepositions are [– N, – V]. I consider deleting one or the other of these two feature specifications.

If s is $[-N]$ with no statement of V , then the following occurs in conjunction with an NP $[+N, -V]$ complement: The result is either $[-N, -V]$, a PP, if s is the head, or $[+N, -V]$, an NP, if s is **not** the head. If, on the other hand, s is $[-V]$ with no statement of N , and has the same $[+N, -V]$ complement, then the result will be $[+N, -V]$, an NP, regardless of which sister is the head. Which is preferable?

Actually there are four permutations: I place a subscript “s” after any feature projected from s and a subscript “n” for any feature projected from its nominal sister: (i) $[-N_s, -V_n]$ is a PP with s as relativized head; (ii) $[+N_n, -V_n]$ is an NP with NP as head, gaining no syntactic features from s ; (iii) $[+N_n, -V_s]$ is an NP with NP as relativized head; and (iv) $[+N_n, -V_n]$, an NP with N as head, again drawing no features from s . Possibilities (ii) and (iv) result in the same feature combination, but I consider them separately for the sake of completeness.

I rule out possibility (i), with the features of a PP, because an $s+ACC$ phrase can be the subject of a clause;¹⁸⁴ the only constructions with prepositions in subject position, according to Babby (1987), are PPs embedded within NPs, as in (119a). The resulting feature combination should be able to bear syntactic, if not morphological, case. PPs bear neither.

I likewise rule out (ii) and (iv) because s assigns ACC case; if both of the resulting NP node’s syntactic features come from the sister of s , then it is unlikely that s would be a case-assigner. There is evidence that such structures exist, however. Shevelov (1963:56, fn. 2) reports that Western Ukrainian dialects **optionally** allow exactly such sentences, with the complement of s in the nominative case (z in Ukrainian is cognate to Russian s):

¹⁸⁴ I am assuming that the sentential subject must be an NP. Cf., however, Jaworska (1986).

(123) Rada povoli sxodylasja. Vže **bula** z polovyna radnyx.
 already was about half councilors
 (ADV) (V)PAST.FEM.SG(?) (N.FEM)NOM.SG (ADJ)GEN.PL

‘The council was gathering slowly: about one half of the councilors were already there.’
 [Shevelov (1963:56, fn. 2), quoting V. Stefanyk (no cit.); my word-glosses, bold-facing /LAB]

This example is most opportune in that the noun is of the *-a* declensional class and shows unmistakably NOM—i.e., morphologically nom—case. To those detractors who might suggest that the noun *polovyna* ‘half’ is somehow caseless, the **verb** also shows unmistakably FEM.SG agreement. I am unsure of the categorial status of *z* here, but I would suggest that it is either $[-N, ()V]$, $[-V, ()N]$, or even $[()N, ()V]$, but in which *z* is **not** the head of the NP due to the inertness of *z* to case-assignment. This further suggests that none of the features of *z* are projected upward to the next higher node; all of the features come from the NP sister of *z*. The indirect evidence provided by this Western Ukrainian example at the very least corroborates my suggestion that *s* in Russian has impoverished properties.¹⁸⁵

This leaves only possibility (iii), in which *s* has the lone syntactic feature specification $[-V]$ with the resulting relativized $[+N_n, -V_s]$ feature specifications: $[s_{[-V]} [NP]_{[+N, -V]}]_{[+N_n, -V_s]}$. These features are, of course, those of any NP. This allows an *s+ACC* phrase to occupy an NP position. Recall, however, from (59d), repeated here as (124a), that an *s+ACC* phrase cannot be conjoined with a numerically quantified NP. I have added a number of other syntactic configurations in which an *s+ACC* phrase is and is not acceptable:

¹⁸⁵ It would be far from prudent to attempt to draw too many conclusions from this one datum. Ukrainian ACC-assignment is far from sufficiently understood. For example, unlike (standard) Russian, Ukrainian passive verbs allow the underlying direct object to remain in the ACC case (cf. Billings and Maling 1995). In addition, in (mostly northern) dialects of Russian and one north-central area of Ukrainian there are attested NOM-case objects of infinitives (cf. Matvijak 1984 and Timberlake 1974a; 1974b). Nonetheless, (123) is a worthwhile rough indicator for this study.

- (124a) *Vosem´ krepostnyx sten i s *desjatok* *nebol´šix* *fortov* zaščičajut gorod.
 eight fortified walls & about unit-of-ten small forts defend city
 NOM GEN.PL (P) (N)ACC.SG (ADJ) GEN.PL(V)3.PL ACC.SG
- (124b) *s *desjatok* *studentov* {*dumajut/dumaet*} ...
 about unit-of-ten students think
 (P) (N.MASC)ACC.SG (N.MASC)GEN.PL (V)PRES.3PL{3.PL/3.SG} [elicited/LAB]
- (124c) kupit´ tri marki i s *desjatok* *otkrytok*
 buy three stamps and about unit-of-ten postcards
 (V)INFIN (NUM)NOM(N.FEM)GEN.SG (P) (MASC)ACC.SG (FEM)GEN.PL
 ‘to buy two stamps and about ten postcards’ [cf. (120) above]
- (124d) Dva učitelja i s *desjatok* *učenikov* bol´ny.
 two teachers and about unit-of-ten pupils sick
 (NUM)NOM(N.MASC)GEN.SG (P) (N.MASC)ACC.SG (N.MASC)GEN.PL (ADJ-SF)PL
 ‘Two teachers and about ten schoolchildren are sick.’ [cf. (121) above]
- (124e) Tri sobaki i s *desjatok* *košek* šli vmeste.
 three dogs & about unit-of-ten cats walked
 (NUM)NOM(N.MASC)GEN.SG (P) (N.MASC)ACC.SG (N.FEM)GEN.PL(V)PAST.PL
 ‘Three dogs and about a dozen cats were walking together.’ [cf. (122a) above]
- (124f) Stošnilo pjat´ členov komandy i s *desjatok* *passažirov*.
 made-ill five members and about unit-of-ten passengers
 (V)PAST.NEUT.SG (NUM)ACC (N.MASC)GEN.PL (P) (N)ACC.SG (N.MASC)GEN.PL
 ‘Two members of the crew and about a dozen passengers got sick.’ [elicited/LAB]

I qualify my statement following (59d) as follows: An s+ACC phrase cannot be in the external argument. That is, s+ACC phrases **cannot** be either the subject of a transitive verb (124a) or the subject of an unergative verb, a verb with only a lone external argument (124b). In (124c-e) I show that s+ACC phrases **can** be, respectively, the direct object of a transitive verb, the subject in a predicate-adjective clause, or the subject of an unaccusative verb (a verb with a single **internal** argument). I also elicited (124f), which has a verb that requires an internal argument which cannot be realized as subject (cf, Babby 1989 for details). In each of (124a-f) I have shown an s+ACC phrase conjoined with a numerically quantified NP. Removing the latter, and adjusting the verbal agreement in some cases, has no effect on the grammaticality of these sentences: (124a-b) remain unacceptable, while the rest are acceptable. These

sentences show that the only limitation on *s+ACC* in modern Russian is external-argument (underlying-subject) position.

This distribution is identical to that of the GEN of negation. When there is so-called clausal negation an underlyingly VP-internal argument gets GEN case. This phenomenon is extensively studied and just as complicated. I refrain from delving into the details here. I provide only a few key references: Fowler (1988:294-319) is a clear summary of the generative-syntactic problem; Timberlake (1975) describes the peculiarities in detail; and Corbett (1986) is a bibliography of works on this construction. Babby (1980) and Chvany (1975) also treat GEN of negation in detail. Pesetsky (1982) uses the GEN of negation as a test of unaccusativity: Only VP internal NPs exhibit this phenomenon.

As to the reasons why *s+ACC* and the GEN of negation have the same distribution, I can only suggest one idea here: As several of the works above show, the GEN of negation is becoming more and more restricted. Whereas the GEN of negation is not subject to any single-word restriction discussed in the previous chapter, it could be that VP-internal generation is another way for a construction to be incrementally restricted. Franks (1995:107) reports that the unaccusativity distinction is not observed in South or West Slavic. In Polish (W. Slavic) the GEN of negation is very extensive and, unlike Russian, obligatory (cf. Franks 1995:204-09). It is possible that Russian utilizes the VP as a domain for limiting both GEN of negation and *s+ACC*. I leave for future research the details of how this restriction is formalized.

One final point on the relativized-head notation: I have assumed so far in this study, following the general outlines of Babby (1985; 1987), that numeral-noun sequences are [[[numeral]_{NumP} [[noun]_{N°} N']_{N''}]_{NP}, while a measure noun followed by a GEN.PL noun has the following structure: [[[measure noun]_{N°} [noun]_{NP} N']_{NP}. That is, measure nouns have NP complements, while numerals have N'' sisters. What

is to keep the hybrid category headed relatively by *s* from being in the same position as a numeral or prepositional-quantifier phrase (as sister of N´)? I assume that a [+ N] specification is not allowed in the quantifier position as sister of N´´. I further assume that numerals have no [+ N] specification (probably no statement of the N feature). Prepositional-quantifier phrases are PPs—i.e., [– N, – V]—and do not have a [+ N] specification. An *s*+ACC phrase, as I propose above, has the features [+ N, – V]; the [+ N] specification is not allowed within the NP except as the complement of N°. This then accounts for why an *s*+ACC phrase is not within the projection of the noun it quantifies.

For ease of exposition I will continue to show the structure in (119b), with a PP dominating NP. Crucially to the discussion in the next chapter, there are two maximal projections with a specifier position under each. I return now to the remainder of my explanations of the structures in (119a-b).

Above (in §4.3.2) I point out that *čtvrt´* can be either a noun or a numeral in the modern language. Nouns in Russian have full six-case paradigms in both the SG **and** PL numbers. Numerals, on the other hand, have historically lost the morphological-PL portion of their paradigms, leaving only a single six-case paradigm. This is perhaps due to the common-sense notion that [+ Q] nominals are semantically PL. Thus, in (119a), the structure which optionally takes either a noun or a numeral, I have placed angled brackets around “SG”, to indicate that if a numeral occupies this position, then morphological number is no longer applicable. The underlined SG in (119b) indicates something slightly different: *s*, as a [+ Q] preposition, need not have a complement with one morphological number or the other; in fact, as I discuss above in chapter 3, the only complements of *s* that can appear with PL morphology are *pluralia tantum*. It follows, therefore, from the structure in (119b)—with *čtvrt´* as a noun—that a [+ Q] preposition, like *s*, might actually syntactically **prohibit**

morphological-number features from (being **added** to) its complement. I will assume non-crucially that this is the case. Except for *pluralia tantum* nouns, therefore, in any of the structures in (119), the preposition's complement will take SG number.

In light of the syntactic approach in (81c) above—"limit the complement in the s+ACC construction to a single syntactic word"—it would be worthwhile to consider whether *s* really subcategorizes for just an N°, not a full NP. Firstly, in many examples the only word uttered after *s* is a numeral. See, for example, (8)-(14) above. This suggests that *s* does not pre-select an N° constituent.

Consider also the following examples, with possessive pronouns:

(125a) Sobaka rostom **s** **moju.**
 dog height about my
 (N.FEM)NOM.SG (N.MASC)INST.SG (P) FEM.ACC.SG
 '(The/That) dog is **about the size of mine.**' [Ušakov (1940:15)]

(125b) Èta komnata širinoj **s** **moju.**
 this room width about my
 FEM.NOM.SG (N.FEM)NOM.SG (N.FEM)INST.SG (P) FEM.ACC.SG
 'This room is **about the width of mine.**' [Baš *et al.* (1959:166)]

(125c) — My barščinnye! **S** **naše-to** Poprobuj, poterpi!
 we vassals about our try endure
 NOM.PL (ADJ)NOM.PL (P) NEUT.ACC.SG(V.PERF)IMPERATIVE (V.PERF)IMPER
 'We're sharecroppers! Try enduring **something like what we have** (endured).' [Slovar' (1962:20), quoting Nekrasov (1959:586)]

(125d) Poživite-ka **s** **moe.** 'Try enduring something like what I have.'
 live about my
 (V.PERF)IMPERATIVE (P) NEUT.ACC.SG [Ušakov (1940:15)]

(125e) [...] žil-to ja ne **s** **tvoe** i gorja-to vidal pobol'še [...]
 lived I not about your
 (V)PAST.MASC.SGNOM (NEG) (P) (SG)ACC.SG
 '... what I've lived is nothing **like what you have**, and (I've) seen more misery ...' [Slovar' (1962:20), quoting Ostrovskij [1974:431]]

(125f) — To, čto ty perenes, ešče cvetočki.
 that which you endured still flowers

‘What you’ve endured is a bed of roses.’

A vot ty s moe poživi.
 and here you about my live
 (SG)NOM (P) NEUT.ACC.SG (V.PERF)IMPERATIVE

‘So how about trying to go though something like what I have (gone through).’
 [*Slovar*´ (1962:19), quoting Rešetnikov (1890:237)]

(125g) — Ty sperva s moe poživi [...], a togda uže i osuđdaj menja.
 you first about my live
 (SG)NOM (ADV) (P) NEUT.NOM.SG (V.PERF)IMPERATIVE

‘First try going through **something like what I have** ..., then go ahead and judge me’
 [Kapanadze (1991:117), quoting Turgenev’s *Kontora* (no cit.)]

(125h) Razve že ja s vaše vspašu?
 really I about your will-plow
 (ADV) (CL)EMPH NOM (P) NEUT.ACC.SG (V.PERF)FUT.1.SG

‘Will I really plow **about as much as you have**?’
 [Bukatevič (1958:132), quoting Šoloxov’s *Podnjataja celina*]

Each of the examples in (125a-h) has a single possessive pronoun as the complement of *s*. There are three distinct types: Examples (125a-b) are both clear cases of ellipsis, with the possessive pronoun agreeing in case number and gender with the elided noun, which both happen to be FEM. The examples in (125c-g) each include the same possessive pronoun, but in the NEUT.ACC.SG, suggesting that this may be default agreement with an empty category of some sort in the N° position. In each of (125c-g) the pronoun represents ‘that which {I/we/you} have endured’. Example (125h) is somewhat unique, with the same NEUT.SG possessive pronoun which means ‘that which you have plowed’. Whatever the exact nature of ellipsis, which is quite extensive in Russian, these examples show that a single-word pronominal form is used as an apparent means of adhering to the single-word restriction. Assuming that these words are the specifier of the NP with an inaudible head, they also prove that the grammar cannot specifically select a noun but rather a single word within NP. Also in light of some of the modified complements of *s* above (in §4.2.4)—especially example

(48), with a possessive specifier even—or (25), with the personal pronoun *Vas* ‘you’ as the complement of *s*—it remains necessary to allow for a full NP complement of *s* in (119b).

There is, however, one final point on the notation in (119): Both structures that are shown stand in for an NP that is assigned either NOM or ACC case (as either clausal subject or time adverbial/direct object, respectively). I assume that the matrix NP node in (119a), the PP node—actually a hybrid node just discussed above—in (119b) and the once-embedded PP node in (119a) are all assigned syntactic ACC or NOM case. Yet on none of these nodes is this NOM or ACC case morphologically realized, hence the parenthesized “(NOM/ACC)”. I assume that direct syntactic case percolates downward to any node not already assigned idiosyncratic/lexical case.¹⁸⁶ The PP nodes, including the hybrid node in (119b), cannot exhibit morphological case under any conditions; the highest NP node in (119a-b) likewise cannot exhibit morphological direct case because the head N° is exhibiting GEN case, assigned ACC case by *s* in (119a) and assigned adnominally by the noun *četvertʹ* in (119b).

At this point it is necessary to define some of the terms used in (119) with more precision: “Numeral” stands for the following cardinal numbers: *četvertʹ* ‘quarter’ (in most cases in modern Russian), *tretʹ* ‘(one-)third’ (rarely),¹⁸⁷ *pol* ‘half’ (always), *poltory/poltora* ‘one and a half_{FEM/NON-FEM}’, *dve/dva* ‘two_{FEM/NON-FEM}’, *tri*

¹⁸⁶ This is in keeping with Babby’s (1987:116) “Syntactic Case Hierarchy” for Russian [= his ex. 51], which he derives from the Projection Principle:

Lexical case > GEN of quantification > (structural) NOM/ACC.

¹⁸⁷ Worth (1959) argues for the numeral-hood of *tretʹ* and Mel’čuk (1985:322-25) argues against it. The fractions smaller than 0.25 are rendered using adjective stems. For example, *pjataja* ‘fifth_{(ADJ)FEM,NOM,SG}’ means both ‘fifth’ as the fraction and ‘fifth’ as the ordinal adjective, because the understood head noun is *častʹ* ‘part/portion_{(N,FEM)NOM,SG}’. Tolbert (1974:12) has an ingenious explanation for the split between noun/numeral and adjective fraction words: The non-adjectives *pol(ovina)* ‘half’, *tretʹ* ‘third’ and *četvertʹ* ‘quarter’ are fractions in which the denominator is a paucal number, while adjectival fractions are required when the denominator is a non-paucal numeral.

‘three’, *četyre* ‘four’, *pjat’* ‘five’, then each integer onwards up through *sto* ‘hundred’,¹⁸⁸ *poltorasta* ‘150’,¹⁸⁹ and *tysjača* ‘thousand’, which almost always functions as a numeral in the modern language. In each of these (except for *pol*) there is a full six-case paradigm, but no separate morphological-PL paradigm (Mel’čuk 1985:267).¹⁹⁰

As far as case-assignment is concerned, numerals are of hybrid design: If the overall nominal expression containing the numeral is assigned either ACC or NOM case, then the numeral itself exhibits that morphological case and assigns GEN to the noun it quantifies. If, on the other hand, the overall nominal expression is syntactically assigned an oblique case, then both the numeral and the quantified noun take that oblique case morphologically (*Ibid.*, p. 291).¹⁹¹ There are also properties that

¹⁸⁸ It is possible to get *neskol’ko sot* ‘several_{NOM/ACC} hundred_{GEN(PL?)}’ [Corbett (1978a:44), Skoblikova (1959:111), Sorokin (1977:58-59)], suggesting that *sto* retains some old nominal properties. I stay away from complex numerals in this study, assuming (perhaps incorrectly) that their combination is not productive and therefore each complex numeral is stored separately in the lexicon. I have, however, found the following examples in which an oblique-case form of *sto* quantifies a GEN.PL noun:

- (i) [...] s dvumjastami rabočix ‘... with 200 workers’
with_(P) two-hundred_{INST} workers_{(ADJ)GEN.PL}
[DePerno (1991:ch.5:8), citing Suprun (1964:71-71), quoting Trofimov’s *Studenty*.]
- (ii) [...] s četymjastami tonn ‘... with 400 tons’
with_(P) four-hundred_{INST} tons_{(N.FEM)GEN.PL} [from ex. 32’ in Mel’čuk (1985:154)]
- (iii) raven primerno devjatistam {kilogrammov/kilogramm}
equal approximately_(ADV) nine-hundred_{DAT} kilograms_{(N.MASC)GEN.PL(NON-ADNUM/ADNUM)}
‘equal to approximately 400 kilograms’ [Ibid.]

In each of (i)-(iii) the non-GEN form is also acceptable. This suggests that *sto* has still not fully transitioned to numeral-hood. This is not unexpected, since the apparent universal trend, according to Corbett (1978a:44-45), is for the smaller numbers to have transitioned soonest and for the largest numbers to do so incrementally later. Drovnikova (1985:66), citing Matveeva (1954:146-47), reports that in Russian dialects there is inversion of the type *sta tri* ‘hundred_{GEN.SG} three_{NOM/ACC}’ (= ‘about 300’).

¹⁸⁹ This word literally means ‘one and a half hundred’ and is somewhat archaic (cf. an ex. of it in (108f), along with *polsta* ‘50’, literally: ‘half a hundred’; cf. Mel’čuk (1985:37-39).

¹⁹⁰ Likewise, numerals have neither of the PREP-2 and GEN-2 “cases” discussed in §4.6.3 above.

¹⁹¹ There are some data that appear to obscure this definition. Recall that *tysjača* ‘thousand’ can either be a noun or a numeral, but not both at the same time:

Footnote continued on next page

numerals do not possess: They appear unable to assign so-called adnominal GEN: \sqrt{tret} *ix* ‘third_{(NOUN)NOM/ACC} them_{GEN.PL}’ (= ‘a third of them’), vs. **tri ix* ‘three_{(NUM)NOM/ACC} them_{GEN.PL}’ (*Ibid.*, p. 268; where *tret* is argued not to be a numeral).

In addition to numerals there is a class of words I have been calling “measure nouns”.¹⁹² These consist of two kinds of nouns: (*čajnaja*) *ložka* ‘(tea)spoon’, (*kilo*)*gramm* ‘(kilo)gram’, *list* ‘leaf/sheet (of paper)’, etc. are used to measure substances not usually countable (i.e., liquids, powder, weights, bulk, etc.); while *sotnja* ‘unit-of-hundred’, *djužina* ‘dozen’, *desjatok* ‘unit-of-ten’ and *pjatok* ‘unit-of-five’ express set quantities of items. The two subgroups differ as to whether the quantification is by counting or not, but this group (as I show below) nonetheless acts syntactically as a cohesive set.

Finally, in addition to numerals and measure nouns there are simple nouns which are nouns with absolutely no quantificational designation. Examples of simple

-
- (i) k kakim-nibud’ tysjače rubej
to something-like_{(ADJ)DAT.SG} thousand_{DAT(SG?)} roubles_{GEN.PL}
‘to something like a thousand roubles’
- (ii) s tysjač’ju rubej v karmane
with thousand_{INST(SG?)} roubles_{GEN.PL} in pocket_{PREP.SG}
‘with a thousand roubles in (his/her) pocket’
- (iii) Delo [...] ne v ètix žalkix tysjače rubej.
it’s-not-a-matter-of these_{PREP.PL} pitiful_{(ADJ)PREP.PL} thousand_{PREP(SG?)} roubles_{GEN.PL}
‘It’s not a matter of these pitiful thousand roubles.’
[≈ exx. 10a-b-v in Mel’čuk (1985:291); glosses added/LAB]

In (i) and (iii) the PL adjectives *kakim-nibud’* and *ètix žalkix* imply a numeral function of *tysjač-*, but the GEN.PL of the noun *rubej* in each implies that *tysjač-* is functioning as a noun. In (ii) the special numeral-only INST.SG form of this word is used; nonetheless it governs *rubej* in the GEN.PL, implying that the former is a noun. Mel’čuk (1985:292) points out that the reverse—SG adjectives and a morphologically oblique quantified noun is ruled out completely.

¹⁹² I. Mel’čuk has suggested the English term “numerical nouns” to me to correspond to his Russian terms *imja edinicy izmerenija* ‘unit-of-measure noun’ (e.g., *metr* ‘meter’; cf. Mel’čuk 1985:29) and *imja množestva opredelennoj moščnosti*, literally ‘noun of plurality of a particular capacity’ (e.g., *djužina* ‘dozen’; *Ibid.* p. 27).

nouns from this study so far include (*greckij orex* ‘(wal)nut’, *vorota* ‘gate’, *bulavoč-naja golovka* ‘pinhead’, and even names, including complex ones like *Petr Velikij* ‘Peter the Great’ or pronouns like *Vy* ‘you_{SG.POLITE/FORMAL}’. While everyone who can recognize a walnut will also know its general size, or while most everyone who knows how big a gate is, or—as most Russians today do—that Peter the Great stood a phenomenal two meters or so tall, these nouns still do not function as measure nouns (**Poltora Petra Velikogo ètogo materiala, požalujsta*. ‘*One and a half Peter the Greats of this fabric, please.’) Thus, simple nouns are not marked for the Q feature. It is important to add that it is necessary to leave simple nouns altogether unmarked, not [– Q], because such nouns head the NP and as such [– Q] would be incorrectly percolated regardless of whether there is a quantificational phrase sister of N′′, as in (119a).

Numerals have the following properties: Those which assign GEN.SG trigger the special ADPAUC end-stressed form *časA* ‘hour’; those which assign GEN.PL trigger the special COUNT GEN.PL form *čelovek* ‘people/persons’ (instead of the ordinary GEN.PL form *ljudej*) if these two words are sisters (cf. ex. (91) above). Semantically, numerals, for these purposes, are always [+ Q] (and, apparently, always [– PROX]).¹⁹³

The hybrid part of speech “measure noun” is categorially a noun but always [+ Q] (and probably always [– PROX]). Thus the difference between simple and measure nouns is semantic, not lexico-syntactic: they are both nouns, they both assign adnominal GEN case, and have full SG and PL morphological-number paradigms. Of the two types only measure nouns are [+ Q].

¹⁹³ I also hypothesize above that numerals are [– N] in order to be allowed to be sister of N′′. I do not work out the other syntactic features of numerals here. Fowler (1988:254-270) convincingly shows that numerals are syntactically adjectival in the oblique cases and nominal in the direct cases. He thus proposes that numerals share the [+ N] specification with both nouns and adjectives and lack any V specification. His syntactic [+ Q] feature is assigned to numerals as well. Clearly some **syntactic** feature is needed to set numerals apart from the other parts of speech; these details remain to be solved.

I conclude this section by reiterating that *okolo*, in order to have a quantificational interpretation, requires a [+Q] complement. This predicts that either a numeral or a measure noun, but not a simple noun, can be the complement of quantificational *okolo*: Only *s+ACC* is inherently quantificational. I also distinguished between numerals, which I consider a distinct part of speech, and nouns. Numerals are invariably quantificational while only some nouns are quantificational: measure nouns. Thus, *okolo* can have either a numeral or a measure noun as its sister in order to have quantificational force. Like *s+ACC*, *okolo* has the inherent semantic feature [PROX(imate)], which, coupled with a quantificational sister, results in the combined approximate-measure semantics inherent in *s+ACC*. This explains why only *s+ACC*, and not *okolo*, can have a non-quantificational (simple) noun as its complement and still have an approximative interpretation. I also showed that *s+ACC* phrases are not within NP (as the quantificational sister of N´). Nor are they a full PP above their NP complement. Instead, the node over *s* and its complement is relativized, drawing features from its NP daughter as well. This, in turn, results in distributional restrictions of *s+ACC*, which is not attested in an external-argument position.

5.2 Approximative inversion

Another very frequent mechanism for expressing approximation in Russian is to invert the order of a numeral and the noun which it quantifies, as shown in (125):¹⁹⁴

(125a)	tri three _{NOM/ACC}	časa hour _{(MASC)GEN.SG (ADPAUC)}	(125b)	časa hour _{(MASC)GEN.SG (ADPAUC)}	tri three _{NOM/ACC}
	‘three hours’			‘about three hours’	

¹⁹⁴ The term “approximative inversion” was first used, to my knowledge in Franks (1994), following roughly equivalent Russian terms in Mel’čuk (1985).

In this section I investigate several aspects of this phenomenon. I surely do not exhaust all of the peculiarities having to do with approximative inversion. To do so, a detailed comparison with the other two languages that have this phenomenon—Ukrainian and Belarusian—would be necessary.¹⁹⁵ Instead, I cover those aspects which are significant to this study of *s+ACC*.

Inversion is not a simple juxtaposition of the numeral and noun, as the following excerpt shows:

“Russian has an interesting syntactic process by which a [...] numeral and a noun, which normally occur in that order, as in [(125a) above], are inverted, as in [(125b) above], adding the extra shade of meaning ‘approximately ...’. When this process occurs in a prepositional phrase, as in [(126a) below], the preposition falls between the noun and the numeral [as in (126b) below], even though the numeral and noun form a constituent. A logical way to deal with this construction would be for *Move- α* to apply [...]” [Fowler (1988:39-40)]

¹⁹⁵ The behavior of the inversion is slightly different in the other two languages. I show only a few of these differences here, as listed in Franks (1995:173-74, 216):

- (i) Ukrainian: Pryjšlo čolovik visim ‘about eight people arrived’
arrived_{PAST,NEUT.SG} people_{GEN.PL} eight_{NOM} [his ex. 115b, also in Mel’čuk (1985:156)]
- (ii) Ukrainian: dva dolary ‘two dollars’
two_{NOM/ACC} dollars_{NOM.PL} [his ex. 116a, citing W. Browne (personal communication)]
- (iii) Ukrainian: dolariv dva ‘about two dollars’
dollars_{GEN.PL} two_{NOM/ACC} [his ex. 116b, citing W. Browne (p.c.)]
- (iv) Belarusian: (praz) dni dva ‘(in) about two days’
in (elapsed)_(P) days_{ACC.PL} two_{ACC} [Franks (1995:216, n. 66), citing W. Browne (p.c.)]

Paucal integers in both these languages assign **NOM.PL** (with certain ADPAUC forms distinguished by stress); non-paucals assign **GEN.PL**. Russian has **GEN.SG** and **GEN.PL** (resp.). Ex. (i) shows that non-paucals take the **COUNT** form regardless of approximative inversion (cf. *visim čolovik* ‘eight people’). Examples (ii)-(iii) show **NOM.PL** is assigned when there is no inversion; when there **is** inversion the noun takes the **GEN.PL**. The Belarusian example in (iv), originally in *Sučasnaja* (1980:109), shows that paucals assign the **NOM.PL** even with inversion. (This example also shows that the **P** does not undergo inversion, just the numeral and noun.) I hope to do a comparative study of inversion in these three languages. That study cannot be done now, however. The preposition *z(iz, zi, zo)+ACC* ‘about’ appears to have quite distinct properties as well; cf. (123). Cf. also the following literature on Ukrainian and Belarusian approximative inversion: Akiner (1983:62), Arašonkava & Lemcjuhova (1994: esp. 152), Carlton (1972:20-21), Hurski *et al.* (1955:199-200), Janowski (1961:9, 28-29), Lapaw (1962), Mayer (1971), Shevelov (1963:56-57, 209, 239-43), Staniševa (1966:152), Švačko (1981:99, 115-19), and Tymčenko (1928:35). I, too, have W. Browne to thank, for informing me of some of these titles.

- (126a) na tri čaSA (126b) čaSA na tri
 for_(P) three_{ACC} hour_{(MASC)GEN.SG (ADPAUC)} hour_{(MASC)GEN.SG (ADPAUC)} for_(P) three_{ACC}
 ‘for three hours’ ‘for about three hours’

Approximative inversion interacts with s+ACC in a peculiar way, at least according to some of the linguistic literature on s+ACC:

“[...] If along with the noun there is a numeral (for example, *dva goda* ‘two years’, *tri kilometra* ‘three kilometers’, *desjat’ rublej* ‘ten roubles’), then the approximation [or] impreciseness is expressed by positioning the numeral after the noun: *goda dva* [‘about two years’] (that is, ‘priblizitel’no dva goda’ [‘approximately two years’]), *kilometra tri* [‘about three kilometers’], *rublej desjat’* [‘about ten roubles’] (that is, ‘priblizitel’no tri kilometra’ [‘approximately three kilometers’], ‘priblizitel’no desjat’ rublej’ [‘approximately ten roubles’]).” [Baš *et al.* (1959:165); translation mine/LAB¹⁹⁶]

This passage (and the last clause in the Bukatevič quote in chapter 1 above) implies that approximative inversion and s+ACC are mutually exclusive. This is not so, as the following example shows:

- (127) Tom, projdja uže šagov s **pjat’desjat**“, obernulsja i posmotrel na nee. (✓)
 paces about fifty
 GEN.PL(P) ACC

‘Tom, having traversed **about fifty paces** already, turned around and looked at her.’
 [= ex, 27a in Mel’čuk (1985:374), quoting Dostoevskij’s *Prestuplenie i nakazanie* (no cit.)]

See also examples (8) though (14) above.

Mixing inversion with another approximative construction “is not a doubling [of approximation] and does not ‘raise the degree’ of approximation, but is an **overlay of approximation on the hypotheticality of the number**” [Suprun (1962:2); emphasis

¹⁹⁶ This extended quote in the original: “[...] Esli že rjadom c suščestvitel’nym stoit čislitel’noe (naprimer, *dva goda*, *tri kilometra*, *desjat’ rublej*, priblizitel’nost’, netočnost’, vyražetsja postanovkoj čislitel’nogo posle suščestvitel’nogo: *goda dva* (t.e. ‘priblizitel’no dva goda’), *kilometra tri*, *rublej desjat’* (t.e. ‘priblizitel’no tri kilometra’, ‘priblizitel’no desjat’ rublej’).” This passage is copied, nearly word-for-word in Nikolaev (1968:203), without citing Baš *et al.* (1959) at all.

added/LAB].¹⁹⁷ Mel'čuk (1985:158) also implies that prepositional approximation and approximative inversion have distinct semantics,¹⁹⁸ but quickly adds, "In any event, the semantics of all the means shown here of expressing imprecise quantities is still in need of specialized investigation."

Note that Fowler's characterization of inversion with prepositional phrases applies directly to *s*+ACC, with the noun moving to a position before the preposition and numeral. Fowler refers to this as an instance of Move-Alpha, the term common in generative syntax to describe movement transformations. Tolbert (1974:127) calls this "a late (meaning-changing!) rule that reverses the order of the quantifier and its immediate complement. The rule is presumably of late origin because it appears to apply after the normal rules of government and concord." I point out a few further specifications to Fowler's characterization:

My approach to approximative inversion will not assume syntactic movement but rather considers the possibility of prosodic reordering and other non-syntactic means of juxtaposing two constituents. I eventually conclude that there is syntactic movement but not before considering other mechanisms and showing that there is a single-word limitation on the size of the constituent which moves.

First, in support of Fowler's Move-Alpha suggestion, note that the numeral clearly still governs the noun in (125b), as evidenced by the special end-stressed, ADPAUC form *čaSA*. Stem stressed *ČAsa* is ungrammatical in both of the examples in (125a-b). This suggests that the noun has apparently moved after receiving ADPAUC

¹⁹⁷ In that passage Suprun is not specifically discussing *s*+ACC.

¹⁹⁸ Specifically, Mel'čuk is discussing the permissibility of combining three types of approximative means in the same utterance, as in (131c), where there is a preposition (*okolo*), approximative inversion, and a number range (*pjati ili šesti* 'five or six'). He says that the semantics (*smysl*) of the number range is separate from either approximative inversion or "lexemes like *priblizitel'no* ['approximately'] or *okolo*." This presumably means that the meanings of all three mechanisms are distinct from each other.

to—some other prosodic word (PrWd). And in this respect *s* is no exception, as example (127) above shows. Thus, in those examples so far where inversion takes place with prepositions—*exx.* (126) and (127) above—the inversion might very well be described as the inversion of two consecutive PrWds, as shown in (129a-b):

- (129a) [na tri]^{PrWd} [čaSA]^{PrWd} --> [čaSA]^{PrWd} [na tri]^{PrWd}
 for three_{ACC} hour_{(MASC)GEN.SG.ADPAUC} hour_{(MASC)GEN.SG.ADPAUC} for three_{ACC}
 ‘for three hours’ ‘for about three hours’
- (129b) (*) [s tri]^{PrWd} [čaSA]^{PrWd} --> [čaSA]^{PrWd} [s tri]^{PrWd}
 about three_{ACC} hour_{(MASC)GEN.SG.ADPAUC} hour_{(MASC)GEN.SG.ADPAUC} about three_{ACC}
 ‘about three hours’

Note the asterisk in parentheses in the left-hand part of (129b). In modern Russian, as I have shown repeatedly above, the overt order *s* + numeral_{ACC} + noun_{GEN} is not allowed. It appears, therefore, that with *s*+ACC approximative inversion is obligatory. Since *s*+ACC always expresses approximation, [*s* + numeral_{ACC}]^{PrWd} + [noun_{GEN}]^{PrWd} must always invert to the order [noun_{GEN}]^{PrWd} + [*s* + numeral_{ACC}]^{PrWd}. The often simplistic statements repeated in the literature (see Baš *et al* quote above in this chapter, as well as the ones by Koka in §5.1 and by Lomtev in chapter 1), that *s* is “limited” somehow, is explained by the requirement that inversion must be used to express approximate measure.

If a preposition is too heavy to procliticize, as is the case with *otnositel’no* ‘regarding’, then the inversion involves switching only the numeral and noun, as shown in (130a-c):²⁰⁰

²⁰⁰ Cf Hill (1977:216) for empirical confirmation that *otnositel’no* is synchronically a preposition.

- (130a) Otnositel'no tridcati očkov i govorit' ne stoit.
 regarding_(P) thirty_{(NUM)GEN} points_{(N,NEUT)GEN.PL} [sic.]
 'It is not even worth having a discussion regarding thirty points.'
- (130b) * Očkov otnositel'no tridcati i govorit' ne stoit.
 points_{(N,NEUT)GEN.PL} [sic.] regarding_(P) thirty_{(NUM)GEN}
 ('It is not even worth having a discussion regarding **approximately** thirty points.')
- (130c) Otnositel'no očkov tridcati i govorit' ne stoit.
 regarding_(P) points_{(N,NEUT)GEN.PL} [sic.] thirty_{(NUM)GEN}
 'It is not even worth having a discussion regarding **approximately** thirty points.'
 [≈ ex. 25 in Mel'čuk (1985:153)²⁰¹]

That is, the non-approximative example in (130a) can be inverted only as in (130c), without pied-piping, after a fashion, the prosodically heavy preposition *otnositel'no* along with the numeral *tridcati* in the inversion.²⁰² I should add that my informants do not accept either of (130b-c), but, when asked to choose between them, they consistently prefer (130c) to (130b).

There is, however, one notable exception to the generalization that heavy prepositions do not pied-pipe: *okolo* 'about/approximately' (discussed above in §5.1).

Recall from the paragraph before (129) that any preposition which procliticizes is no

²⁰¹ In (130b-c) I use the gloss 'approximately' rather than 'about' due to the confusion between the English words *regarding* and *about* in its other (non-approximative) meaning, hence the somewhat stilted sentential glosses. The noun /očko-/ 'point' is one of a handful of NEUT nouns that take the GEN.PL ending *-ov*, usually found only on MASC nouns of the *-Ø* declensional class. Cf. Zaliznjak (1987:54).

²⁰² It **is** possible, though rare, in quite colloquial standard Russian, according to Mel'čuk (1985:159), for prosodically light (i.e., proclitic) prepositions to take the same order as (130c):

- (i) **V** **šagax** **pjati** on zamer.
 away_(P) paces_{(N,MASC)PREP.PL} five_{(NUM)PREP} he_{NOM.SG} froze (momentarily)_{(V)PAST.MASC.SG}
 'About five paces away he froze for a moment.'
 [= ex. 48 in Mel'čuk (1985:156), citing K. Fedin's *Neobyknovennoe leto*; glosses added/LAB]
- (ii) Rastjani **na** **minut** **sem'**.
 stretch-out_{(V)IMPERATIVE.SG} for_(P) minutes_{(N,FEM)GEN.PL} seven_{(NUM)ACC}
 'Stretch (it) out for about five minutes.'
 [Overheard in March, 1995/LAB]

I return to this colloquial register in chapter 6 in order to clarify the relative rankings of constraints.

more than two syllables in size; *okolo* is trisyllabic and as such cannot procliticize.²⁰³ Thus, in any of the examples of *okolo* + numeral_{GEN} + noun_{GEN} above, for example (116b) or (119a), there are three PrWds, which is not an acceptable input for inversion (at least according to the model in (129a-b) above), because there are now three PrWds: [*okolo*^{PrWd} numeral^{PrWd} noun^{PrWd}]. Recall also that the *okolo* data do not have the same **syntactic** structure as in the other prepositional-inversion data above in this chapter, including the *s*+ACC data in (127) and (128) above. As the examples in (126) show, the inverted noun_{GEN} with the non-*okolo* prepositions is in the distinctively ADPAUC form (as evidenced by end-stressed GEN .SG form *čaSA*); if the numeral assigns GEN.PL then the COUNT form *čelovek*—as opposed to the ordinary *ljudej*—must be used). I have found the following approximative-inversion examples with *okolo*:²⁰⁴

²⁰³ Recall from §4.3.5 that another test for full-PrWd-hood of a constituent is whether there is word stress. If the constituent is monosyllabic, as is *pol* ‘half’, then a further test, if the vowel is /o/, is whether that vowel is pronounced with lip-rounding (unstressed /o/ is not pronounced with rounding in standard Russian). If the constituent is disyllabic or larger, the further test is whether there is a syllable with stress. In the case of *okolo*, there is not only a stressed syllable (the first one), but the /o/ in it is invariably pronounced with lip-rounding, proving that *okolo* is a stand-alone PrWd, and therefore cannot procliticize.

²⁰⁴ Mel’čuk (1985:152) reports that “non-primary” prepositions cannot appear between the inverted noun and the numeral. Strangely, he appears to avoid committing to whether *okolo* is primary or not. (On p. 353, n. 16, he lists several “primary” prepositions, all of which are monosyllables or lighter, and several “non-primary” prepositions, all of which happen to be disyllabic or heavier), without listing *okolo* in either set. On p. 158 he writes that if the quantified noun is governed by a primary preposition, then this preposition generally appears between the noun and numeral. He then refers to several examples in his preceding discussion, but again does not refer to any of his examples of this kind with *okolo* shown in (131a, c-d) below. Elsewhere, on p. 152, Mel’čuk writes that in approximative-inversion structures the noun and numeral must be adjacent with certain exceptions, including the following one:

“(i) A preposition [which governs the noun can appear between the noun and numeral], conditional on the preposition being either primary [...] **or an indicator of approximation** [referring here to his excursus on *okolo*, pp. 362ff]—namely, also a preposition or a special adverb in the comparative:

[three examples—one with primary *na*, (131a), and one with a comparative—here]

“**All types of non-primary prepositions** between the [noun] and [numeral] in this construction are impossible [...]” [Mel’čuk (1985:152); my translation/LAB]

Footnote continued on next page

PrWd; the quantifier portion—be it a single numeral or the combined constituent consisting of *okolo* and the numeral—need not be a single PrWd.

Before leaving the *okolo*-inversion data, I should add that I have not found a single example of approximative inversion in which quantificational *okolo* precedes both the noun and the numeral (i.e., the order in (130c) above). Nor was I able to elicit such examples. As with the examples in (130b-c), my informants do not fully accept any of (131a-e), but prefer these to the corresponding *okolo*-noun-numeral order. Assuming that a dialect or register of Russian exists in which the orders *otnositel'no* + noun + numeral in (131c) and noun + *okolo* + numeral in (131a-e) are tolerated, but not either of noun + *otnositel'no* + numeral, as in (131b), or quantificational-*okolo* + noun + numeral (not shown), then this might constitute an additional argument in favor of Babby's (1985) analysis of quantificational *okolo*: [[*okolo* numeral] noun] (i.e., distinct from other prepositions, which have the structure [P [numeral noun] ; cf. (115a-b), irrespectively. In order to achieve an approximative interpretation, the quantified noun must precede the element which quantifies it. If the noun appears between quantificational *okolo* and the numeral (**okolo*-noun-numeral), then, according to Babby's model in (115b), the noun is not actually preceding the entire constituent which quantifies it. It is as though the quantified noun changes places with the minimum number of PrWds to be in front of the (entire) constituent which quantifies it. Assessing the optimal position of the quantified noun must be performed, crucially, in terms of the number of **prosodic** words that lie between the two positions (where the noun would appear with and without inversion). I return to this task in the Optimality-theoretic models in chapter 6.

In the preceding discussion I refer strictly to **quantificational** *okolo*. Steven Franks, while discussing an earlier draft of this study with me, suggested that it would follow—assuming Babby's model—that quantificational and locative *okolo* should

invert differently: Quantificational *okolo* would be between the noun and numeral, as in (131), and locative *okolo* should appear before both the noun and numeral. I have confirmed this prediction with my informants. I should point out that concepts like ‘near approximately ten pines’ are difficult pragmatic concepts and thus cause many informants to balk at making judgments.²⁰⁶ Nonetheless, those who can conceive of such an utterance prefer the order in (133a).

(133a) Locative *okolo* [i.e., ex. (115a) above] with added approximative inversion:

okolo	sošen	desjati	‘near about ten pines’
near	pinés	ten	
(P)	(N.FEM)GEN.PL	(NUM)GEN	

(133b) Quantificational *okolo* [i.e., ex. (115b) above] with added inversion:

sošen	okolo	desjati	‘about ten pines’
pinés	about	ten	
(N.FEM)GEN.PL	(P)	(NUM)GEN	

They consider the other order, in (133b), to mean only ‘approximately/about ten pines’ (i.e., the already approximative reading from quantificational *okolo* plus approximative inversion). The fact that there is approximation from two sources— inversion and quantificational *okolo* —is not considered redundant; Mel’čuk (1985:158), citing Suprun (1962:2), reports that combining approximative inversion with other mechanisms for expressing approximation such as quantificational *okolo* merely creates an “overlay of approximation on the conjecture-hood of the number” (*nakladyvanie priblizitel’nosti na predpoložitel’nost’ čísla*). Numerous approximation devices co-occur in Russian quite frequently.²⁰⁷ The asymmetry in (133a-b) is empirical evidence to support Babby’s (1985) proposal that quantificational *okolo*

²⁰⁶ M. Yadroff informs me that approximative inversion may require that the noun phrase be non-referential and the locative reading of *okolo* requires reference, hence the problems my informants had.

²⁰⁷ Cf. (151a) below where there are three separate approximative mechanisms employed: a preposition (*s*), an adverb (*étak*) and approximative inversion.

‘about’ does not have the same phrase structure as locative *okolo* ‘near’ (and against the arguments in Franks 1995:143-44 and Neidle 1988:160-65).

Additional support for the prosodic model of approximative inversion in (129b)—as modified in (132) and the preceding paragraph—is the invertability of numerical expressions in which a single noun is quantified by an apparent multi-PrWd numeral. Mel’čuk (1985:36 n. 5, 150) reports that a noun and numeral followed by *s polovinoj* ‘and a half’ or *s četvert’ju* ‘and a quarter’ (literally: ‘with_(P) half_{(N.FEM)INST.SG}’ and ‘with_(P) quarter_{(N.FEM)INST.SG}’, respectively) **do** undergo approximative inversion:

- (134a) kilometrov šest’ s polovinoj ‘about six and a half kilometers’
 Kilometers_{GEN.PL} six_{NOM/ACC} with_(P) half_{(N.FEM)INST.SG}
- (134b) kilometrov šest’ s četvert’ju ‘about six and a quarter kilometers’
 Kilometers_{GEN.PL} six_{NOM/ACC} with_(P) quarter_{(N.FEM)INST.SG}
 [= ex. 12 in Mel’čuk (1985:150); cf. also exx. (94a-c) above]

(Note that the *s* in (134a-b) is the INST-assigning preposition meaning ‘with’, not *s+ACC!*²⁰⁸) The list of allowable additional elements to a numeral which do not affect invertability seems to be limited to just *s polovinoj* and *s četvert’ju* (i.e., it is impossible to invert the analogous phrase meaning, e.g., *kilometrov šest’ s tremja četvertjami* ‘about six and three quarters kilometers’). Nonetheless, the fact that any extra PrWd is allowable in the numeral portion of the approximative-inversion model, added to the fact that the noun does not appear **within** the string *šest’ s polovinoj* or *šest’ s četvert’ju*, is additional support for the revisions of (129b) in (132) and the preceding paragraph.

Additional evidence to show that the noun in approximative inversion must consist of just one word is the following: Mel’čuk (1985:15, 96) reports that sentences

²⁰⁸ Cf. similar examples in (94b-c) and the footnote following that example.

like (135a) **cannot** undergo approximative inversion by juxtaposing the numeral with the rest of the nominal expression—i.e., from non-approximative (135a) to approximative (135b)—if the noun is modified by an adjective:

(135a) On kupil desjat´ starinnyx knjig.
 he_{NOM.SG} bought_{(V)PAST.MASC.SG} ten_{ACC} antique_{(ADJ)GEN.PL} books_{(FEM)GEN.PL}
 ‘He bought ten antique books.’

(135b) * On kupil starinnyx desjat´ knjig.
 he_{NOM.SG} bought_{(V)PAST.MASC.SG} antique_{(ADJ)GEN.PL} ten_{ACC} books_{(FEM)GEN.PL}
 (‘He bought about ten antique books.’)

(135c) On kupil štuk desjat´ starinnyx knjig.
 he_{NOM.SG} bought_{(V)PAST.MASC.SG} items_{GEN.PL} ten_{ACC} antique_{(ADJ)GEN.PL} books_{(FEM)GEN.PL}
 ‘He bought about ten antique books.’

(135d) On kupil knjig desjat´, starinnyx i očen´ dorožix
 he bought books ten [antique and very expensive]
 NOM.SG (V)PAST.MASC.SG (FEM)GEN.PL ACC GEN.PL
 ‘He bought about ten antique and very expensive books.’

[all from ex. 58 in Mel’čuk (1985:96); punctuation modified/LAB; cf. Mel’čuk (1985: 151)]

The unacceptability of (135b) supports the model in (132): It is impossible to invert a numerical expression if the noun is modified by an adjective. More specifically, the inversion is limited to single-word (N´´) complements of the numeral.

To be most clear, the quantified N´´ constituent is not allowed to exceed a word in size under any circumstances. Specifically, this constituent must be a single **prosodic** word; two-PrWd syntactic compounds, discussed above (in §4.2.2) cannot invert to pre-numeric position: *čajnyx ložki tri ‘about three teaspoons’. I conclude, therefore, that approximative inversion, unlike s+ACC (and the other four single-SnWd constructions in §4.6.1-§4.6.4), is subject to a single-**prosodic**-word restriction.

Examples (135c-d) show other ways of performing inversion that circumvent this problem. Example (135c) utilizes a pleonastic count noun štuk ‘items_{GEN.PL}’. Yet another option, shown in (135d), is to shift a prosodically heavy or contrastive adjectival phrase to the right (cf. example (153) and preceding fn.). I do not discuss

the structure in (135d) further. As regards (135c), I discuss structures of this type above in (108a-i), which I repeat here as (136a-i), respectively:²⁰⁹

(136a) odnaždy **čelovek** desjat' našix oficerov obedali u Sil'vio
 people ten our officers
 GEN.PL (NUM)NOM GEN.PL (N.MASC)GEN.PL

‘at one time **about ten** (of) *our officers* dined at Silvio’s’

(136b) (Dymov) ... proiznes **štuk** pjat' nexorošix slov
 items five bad words
 GEN.PL (NUM)ACC (ADJ)GEN.PL (N.NEUT)GEN.PL

‘Dymov ... uttered **about five obscene words**.’

(136c) [...] vpolzali **štuk** desjat' malen'kix devoček s knižkami
 items ten little girls
 GEN.PL (NUM)NOM (ADJ)GEN.PL (N.FEM)GEN.PL

‘... **about ten little girls** with books would creep into (the gates of her house).’

(136d) On kupil **štuk** desjat' starinnyx knig.
 he bought items ten antique books
 NOM.SG (V)PAST.MASC.SG GEN.PL (NUM)ACC (ADJ)GEN.PL (N.FEM)GEN.PL

‘He bought **about ten antique books**.’

[also = (135c)]

(136e) [...] sidelo **čelovek** sem'desjat slučajnyx posetitelej [...]
 sat people seventy chance spectators
 (V)PAST.NEUT.SG GEN.PL (NUM)NOM (ADJ)GEN.PL (N.MASC)GEN.PL

‘... there sat **about seventy chance spectators** ...’

(136f) [...] **čelovek** poltorasta anglijskix soldat ostalis' [...]
 people hundred-fifty English soldiers remained
 GEN.PL (NUM)NOM (ADJ)GEN.PL (N.MASC)GEN.PL (V)PAST.NEUT.SG

‘... **about a hundred and fifty English soldiers** remained ...’

(136g) V nebol'soj komnate prisjažnyx bylo **čelovek** desjat' raznogo sorta ljudej.
 jurors was people ten of-various-kinds people
 (ADJ) (V) (N.MASC) (NUM) (NP)GEN.SG (N.MASC)
 GEN.PL PAST GEN.PL. NOM GEN.PL
 .NEU COUNT .NON-COUNT
 T.SG

‘There were **about ten jurors of various kinds** in the small room.’

²⁰⁹ See my discussion following example (108) above in §4.6.4 of yet another function the special use of *štuk* and *čelovek*. The two uses of the words *čelovek* and *štuk* should not be confused.

(136h) [...] stojalo **čelovek**pjat´ *skromno odetyx* ljudej.
 stood people five modestly dressed people
 (V)PAST.NEUT.SG GEN.PL (NUM)NOM (ADV) (ADJ)GEN.PL (N.MASC)GEN.PL

‘... there stood **about five modestly dressed people.**’

(136i) Okolo nego tolpilos´ **čelovek**pjat´ *dvorovyx* ljudej
 around him crowded people five court people
 (P) ACC.SG (V)PAST.PL GEN.PL (NUM)NOM(ADJ)GEN.PL (N.MASC)GEN.PL

‘**About twenty court servants** crowded around him.’

What each of these examples show is a numeral (underlined in each) quantifying a constituent consisting of a noun modified by an adjective (or, in the case of (136g), with an adnominal-NP complement) in italics. Such structures, as I have shown in the preceding discussion, cannot simply invert to show approximation. Instead, a GEN.PL pleonastic noun (bold-faced), either *čelovek* ‘people’ or *štuk* ‘items’ is uttered immediately prior to the numeral. Deleting either *čelovek* or *štuk* from any of (136a-i) merely deletes the meaning of approximation. I interpret this use of *čelovek* and *štuk* to be a means of achieving the structure of approximative inversion, with either of these words occupying the position that the noun would occupy if it weren’t modified. That is to say, only when the numerically quantified noun is modified can *čelovek* or *štuk* be used to fill the would-be position of the moved noun. In my discussion of ADPAUCs and COUNTs (in §4.6.4) above I specify that using *čelovek* or *štuk* followed by a numeral and single noun **not** only has approximative meaning, but also a special “postquantifier” interpretation, as defined in DePerno (1990; 1991: chapter 9); only when an adjective modifier prevents normal approximative inversion does the use of *čelovek* and *štuk* have a non-postquantifier interpretation.

The primary exception to my assertion that approximative inversion is obligatory with s+ACC is structures with *pol* ‘half’. As I show above in my discussion of this unique numeral (in §4.3.5), *pol* must immediately precede the noun which it quantifies at all times; *pol* forms a morphological stump compound with that noun. If this is so, then constructions with *pol* likewise do not fit the prosodic criteria for

approximative inversion, because the numeral and noun are not located in separate PrWds (whether or not *s+ACC* is involved): [(S)POL[časa]_{PrWd}]_{PrWd} ‘(about) half an hour’. In this structure there are no separate matrix PrWds, meaning that approximative inversion cannot take place.²¹⁰ Put otherwise, a quantified noun is not allowed to extract from within a PrWd. This explains the preponderance of *s+ACC* examples with *pol*.

Additionally, as Mel’čuk (1983; 1985:280-88) and others point out, it is problematic in modern Russian to assign an oblique case to *pol*, requiring a choice between either extremely bookish or extremely substandard structures, thus making it likewise difficult to quantify *pol* using *okolo+GEN*. There is no (better) way, as it were, to express approximate measure with *pol* than to merely prepose it with *s*.

I should also qualify my discussion of approximative inversion by distinguishing it from a deceptively similar phenomenon: It is possible for a noun, even with an adjective modifier, to precede a numeral by uttering the noun at the beginning of the clause, as shown in the following examples:

²¹⁰ Mel’čuk (1985:148) calls the restriction against approximative inversion and *pol* “strictly syntactic” because there is no semantic limitation on applying other approximative operations/constructions to numerical expressions with *pol* as the numeral. I agree that the limitation is not semantic, but disagree with his syntactic characterization based on the prosodic/morphological explanation I give here. He also provides counterexamples to this restriction, however:

- (i) Esli svobodny — časa na pol zaparxivajte na aviabol.
 if free_{(ADJ,SF)PL} hour_{GEN.SG} for half_{ACC} flit-on-over to the air-show
 ‘If (you’re) free, flit on over to the air show for about half an hour.’
 [= his ex. 4, Mel’čuk (1985:148), citing V. Majakovskij]
- (ii) Banka nebol’šaja, tak litra na pol budet
 jar_{(FEM)NOM.SG} not-big_{(ADJ)NOM.SG.FEM} oh liter_{(MASC)GEN.SG} for half_{ACC} will-be_{3.SG}
 ‘The jar’s small, so it would be, oh, about half a liter.’
 [= ex. 5, Mel’čuk (1985:148), no citation]

Mel’čuk considers these examples “beyond the limit” of standard Russian. Note that my informants likewise consider them to be somewhat strange. The stress in (i), in its most acceptable form, is on the second syllable of *časa*. In both examples the postposed *pol* appears to be pronounced—along with *na*—as a separate PrWd from the quantified noun (either *časa* or *litra*).

(137a) My različaem rovno tri takix slučaja
 we distinguish exactly three such occurrence
 NOM.PL (V)PRES.2.PL ADV ACC (ADJ)GEN.PL (N.MASC)GEN.SG

(137b) My različaem takix slučajev rovno tri
 we distinguish such occurrence exactly three
 NOM.PL (V)PRES.2.PL (ADJ)GEN.PL (N.MASC)GEN.PL ADV ACC

(137c) Takix slučajev my različaem rovno tri
 such occurrence we distinguish exactly three
 (ADJ)GEN.PL (N.MASC)GEN.PL NOM.PL (V)PRES.2.PL ADV ACC

‘We distinguish exactly three such occurrences.’
 [= exx. 3, 2, 1 (respectively) in Mel’čuk (1985:144); glosses added/LAB]

None of (137a-c) involves approximation, as evidenced overtly by the word *rovno* ‘exactly’. I will refer to this construction as “emphatic-thematic inversion” following Mel’čuk (1985:143-46).²¹¹ Sentence (137a) does not exhibit emphatic-thematic inversion, while the examples in (137b-c) both do.

Note that in (137a) the noun quantified by the numeral *tri* ‘three’ must be in the GEN.SG, as normally expected of the so-called paucal numerals (even with approximative inversion!). In (137b-c) the same noun must appear in the GEN.PL.²¹² House (1982) explains that this construction is not an instance of syntactic movement; instead, the noun is in an initial NP which must be in the GEN.PL. Moreover, this NP cannot show the COUNT form. That is, whereas the quantified N’’ constituent **must** bear the respective ADPAUC, COUNT, GEN.SG or GEN.PL when there is approximative

²¹¹ House (1982) shows that this construction cannot be a simple case of movement by the noun leftward from post-numeral position and prefers to call such constructs “genitive-initial sentences” instead. That is, there is no “inversion” as such in this construction. Franks & House (1982) call the initial elements in such sentences “genitive themes”. I refrain from outlining the phrase structure here.

²¹² Note, however, that the adjective in both (137a) and (137b-c) is in the GEN.PL; paucal numerals in modern Russian, except for emphatic-thematic inversion, assign GEN.SG to the noun, while non-paucal numerals assign the GEN.PL to nouns they quantify. Adjectives modifying such quantified nouns, however, are usually in the GEN.PL regardless of whether the numeral is paucal. I say “usually” because it is possible for the adjective modifying a noun quantified by a paucal numeral to be in the NOM/ACC.PL. In any event, the adjective modifying a noun quantified by a numeral will always show PL morphology, while the quantified noun, under certain circumstances, can show SG morphology.

inversion, in emphatic-thematic inversion with a numeral only the non-COUNT GEN.PL is attested.²¹³

In most cases, from my observation, this construction involves uttering the noun at the front of the clause, as in (137c), sufficiently far from the numeral to be distinguishable from approximative inversion. It is apparently also possible, however, for the noun to be uttered after the subject and verb, as in (137b), and still have the same emphatic-thematic interpretation. Thus, (137b) has word order which is deceptively similar to that of approximative inversion, where it is possible for certain adverbs to appear between the inverted noun and numeral.²¹⁴ Another striking difference between these constructions is that the emphatic-thematic construction allows a modifier adjective to accompany the (near-)initial noun.

Since approximative and emphatic-thematic inversion are distinct phenomena, it is therefore possible for the two to be present in the same clause:

- (138) Rybešek bylo [štuk desjat']
 smallfry was items ten
 (N.FEM)GEN.PL (V)PAST.NEUT.SG (N.FEM)GEN.PL (NUM)NOM
 'There were about ten smallfry.' [= ex. 106 in House (1982:59); glosses mine/LAB]

It should likewise be possible to construct examples with emphatic-thematic inversion and *s+ACC* both in the same clause; one such example is (12) above, repeated here as (139a). Example (139b) shows another non-initial instance of emphatic-thematic inversion with *s+ACC* and *pol*, which further quantifies a measure noun:²¹⁵

²¹³ If the NP is not countable—i.e., headed by a mass noun—then it bears GEN.SG (and the post-verbal quantifier is not a numeral). In no event will there be a GEN.SG (or ADPAUC) NP with a post-verbal paucal numeral in an emphatic-thematic-inversion sentence.

²¹⁴ Of course, approximative-inversion constructions would not tend to have adverbs that mean 'exactly'. Chey (1967:54) mistakenly requires the inverted noun to “immediately precede” the numeral.

²¹⁵ Another example involving *s+ACC* in a lexically frozen expression is shown in (49b) above.

- (139a) **Pušek** **polkovyx"** u vas" budet" **s"** **dvadcat'.** (✓)
 cannons regimental about twenty
 (N.MASC)GEN.PL (ADJ)MASC.GEN.PL (P) (NUM)ACC
 ‘(As for) **regimental cannons**, you will have **about twenty**.’ [= (12) above]

- (139b) So vse^x dvorov **sobak** sbežalosja **s** **pol** **sotni**
 dogs came-running about half unit-of-hundred
 GEN.PL (V)PAST.PL (P) ACC (FEM)GEN.SG
 ‘There were **about fifty** dogs that came running from all the yards.’
 [Skoblikova (1959:93) and *Sintaksis* (1960:503), quoting Krylov’s *Proxožie i sobaki* (no cit.)]

The following two examples are not as straightforward: Because of *s* there is approximation. Is there also approximative inversion or emphatic-thematic inversion?

- (140a) [...] Zajcev **s** **desjatok** spasalos' na nem.
 hares about unit-of-ten
 (MASC)GEN.PL (P) (MASC)ACC.SG
 ‘... **About a dozen hares** were on it (a floating log) to keep from drowning.’
 [Skoblikova (1959:93) and *Sintaksis* (1960:503), quoting Nekrasov (1971:203)]

- (140b) Rublej **s"** **pjatok"** izderžal" [...]
 roubles fiver spent
 GEN PL ACC (= NOM) MASC SG
 ‘He spent **about a fiver** ...’ [≈ ex. (20) above]

Mel'čuk (1985:143-44) mentions that there is distinct prosody in the emphatic-thematic construction, which presumably distinguishes it from approximative inversion. These two examples, encountered in print, were not provided with such prosodic indicators. There is one other relatively reliable test for emphatic-thematic inversion: The quantificational element—be it a numeral, measure word, or some other fixed expression like *s gul'kin nos* ‘very little’ in (49)—must be post-verbal. As (137b) shows, the GEN-case (adjective and) noun **can** be—but the quantifier **must** be—post-verbal. In **none** of the emphatic-thematic examples in Franks & House (1982), House (1982) and Mel'čuk (1985:143-47) is the quantifier pre-verbal. Thus, (140a-b) are examples of approximative inversion. This test works, obviously, only if the verb is overt; see (144a) and (145) below. This means that (140a-b) are examples of

approximative, and not emphatic-thematic, inversion. I show below, however, in (143) through (146), that approximative inversion is generally not allowed with measure nouns, such as *desjatok* and *pjatok* in (140a-b).

The following is an example with the two different stems meaning ‘people’: The initial non-COUNT GEN.PL *ljudej* is the initial element of the emphatic-thematic formula (which cannot be in the COUNT, as Mel’čuk 1985:146 shows); the COUNT *čelovek* is functioning as a postquantifier (as defined above in §4.6.4) with added emphasis on the sheer quantity, hence ‘all of’ in my gloss:²¹⁶

- (141) Ljudej bylo devjat’ čelovek.
 people was nine people
 GEN.PL.NON-COUNT (V)PAST.NEUT.SG (NUM)NOM (MASC)GEN.PL.COUNT
 ‘As for people there were all of nine.’ [Bukatevič (1958:145), quoting Peter (1893:50)]

It is possible to omit *čelovek* in (141), but not redundant to leave it in, because of the added semantics it conveys.

The preceding excursus has shown that approximative inversion is distinct from emphatic-thematic inversion. While approximative inversion can front only a lone quantified noun, in the emphatic-thematic construction there can be modifier with the noun. Moreover, the numeral or other quantifier must be post-verbal in emphatic-thematic inversion. Additionally, only in approximative inversion can a paucal numeral trigger GEN.SG in the noun; in emphatic-thematic inversion the noun must be in

²¹⁶ Cf. also the following sentence, in which the initial GEN-case element is more than just *ljudej*. The numerical expression at the end is deceptively complicated: Because the complex numeral ends in *odin* ‘one’, the noun must take the NOM.SG form, which is homophonous with the COUNT form:

- Vsex zdes’ ljudej dvadcat’ odin čelovek.
 all here people twenty one person
 GEN.PL (ADV) (N.MASC)GEN.PL.NON-COUNT (NUM)NOM (ADJ)MASC.NOM.SG (N.MASC)NOM.SG
 ‘As for all the people here there are twenty-one.’ [*Sintaksis* (1980:331), quoting L. Tolstoj (no cit.)]

the (non-COUNT) GEN.PL. There is also a prosodic difference between the two constructions. The two constructions are thus only similar looking on the printed page.

Returning, then, to **approximative** inversion, is it possible for non-numeral [+ Q] nouns (i.e., measure nouns) to undergo approximative inversion?²¹⁷ The picture is by no means clear: Mel'čuk (1985:159) reports that non-numerals cannot undergo approximative inversion as in (142a), recommending the *okolo* construction in (142b) instead (he does not consider *million* to be a numeral):

- (142a) * Ètot ostrov naseľjaet čelovek million.
 this island populates people million
 (PRON-ADJ) (MASC) (V) (N.MASC) (N.MASC)
 ACC.SG.MASC.INANIM ACC.SG PRES.3.SG GEN.PL.COUNT ACC.SG
- (142b) Ètot ostrov naseľjaet okolo milliona čelovek.
 this island populates about million people
 (PRON-ADJ) (MASC) (V) (P) (N.MASC) (N.MASC)
 ACC.SG.MASC.INANIM ACC.SG PRES.3.SG ACC.SG GEN.PL COUNT
- 'About a million people populate this island.' [≈ ex. 57a in Mel'čuk (1985:159)]

Recall from the discussion of large-number words (in §4.3.4) that the distinctive COUNT form is not a conclusive indicator of numeral-hood in the quantifier that triggers it; *million* is a noun, and as such appears unable to undergo approximative inversion.

I have, however, uncovered examples (143) and (144a) in which quantified expressions involving measure nouns have indeed undergone approximative inversion. The quantificational element in (143) is post-verbal, allowing this to be either type of inversion. The noun, however, is in the distinctive COUNT-GEN.PL form, excluding the possibility that this is emphatic-thematic inversion.

²¹⁷ Within the various kinds of numerals (many of which I do not discuss here) approximative inversion is not limited to cardinal numerals. Mel'čuk (1985:77-78, 99-100) discusses ordinal numerals in such inversions: *v godu tysjača sem'sot vos'midesjatom* 'in year thousand seven-hundred eightieth' (= 'in about 1780') [p. 100, quoting M. D'jakonov (no cit.)]; cf. also (151d) below. Inversion with the so-called collective numerals is limited to *pluralia tantum* nouns (Mel'čuk 1985:149 and Pete 1984:76).

(145e)	Tumboček nightstands _{(MASC)GEN.PL}	—	pjat´. seven _{(NUM)NOM}	‘There are five nightstands.’ [Suprun (1962:5, fn. 3)]
(145f)	Veder buckets _{(MASC)GEN.PL}	—	dva. two _{(NUM)NOM}	‘There are two buckets.’ [Suprun (1962:6, fn. 3)]
(145g)	Velosipedov bicycles _{(MASC)GEN.PL}	—	pjat´. five _{(NUM)NOM}	‘There are five bicycles.’ [Suprun (1962:6, fn. 3)]
(145h)	Učenic schoolgirls _{(MASC)GEN.PL}	—	tri. three _{(NUM)NOM}	‘There are three schoolgirls.’ [= ex. 33 in House (1982:16); my glosses/LAB]

A dash is often used, as in each of (145a-h), to render the null copula orthographically. Note that in (145b, d, f, h) the numeral happens to be paucal, but the noun is nonetheless GEN.PL, showing conclusively that these are emphatic-thematic. I have actually modified Suprun’s examples in (145b-e), which he shows as follows: *Stolov — dva, stol´ev — šest´, ètažerok — tri, tumboček — pjat´*, which I would instead gloss as ‘There are seven tables, six chairs, three bookcases, (and) five nightstands.’ The fact that (145c, e) can be conjoined with (145b, d) further shows that these are emphatic-thematic.²¹⁸ None of (145b-h), Suprun emphasizes, imparts approximation.

Returning, then, to the recalcitrant example in (144a), if this sentence did **not** have *s*, then the meaning would be ‘There are a dozen postcards.’ Thus, (144a) is not an actual instance of approximative inversion, but of emphatic-thematic inversion, discussed above in this section.

As I propose in the preceding section, measure nouns (like *desjatok*) are not numerals, but nouns with a [+ Q] semantic feature. Crucially, *s* + measure noun_{ACC.SG} + noun_{GEN.PL} is not required to undergo approximative inversion. Several other factors are involved, which Mel´čuk (1985:159-60) valiantly attempts to tease apart. The language appears to be moving away from allowing approximative-inversion with

²¹⁸ Often the overt conjunction *i* ‘and’ is also left out of lists in Russian. Suprun writes that such sentences are characteristic of “official, business lists”.

non-numerals. Example (143) is judged to be perfect if *desjatok* is replaced with its numeral counterpart (*√Sobralos' čelovek s desjat*). Without discussing example (143) further I must conclude that it is archaic, perhaps originally from a source that long predates Ušakov's 1940 publication date.²¹⁹

As for the examples in (144a-b), my informants have no objections. The *s+ACC* portion of (144b) is structurally identical to (119b). The initial noun (in 144a) is not inverted to express approximation, but rather is emphatic-thematic. Without the intonation shown, however, it is impossible to determine this conclusively (cf. Mel'čuk 1985:143-44).

The remaining examples of approximative inversion with measure nouns in (146a-e) all come from Mel'čuk (1985:159, exx. 58a-b-v and 59 a-b, respectively):

(146a) [...] prošli oni šagov sotnju.
traversed_{(V)PAST.PL} they_{NOM.PL} paces_{GEN.PL} unit-of-hundred_{ACC.SG}
'they (they) traversed about a hundred paces.' [citing F. Dostoevskij's *Prestuplenie i nakazanie*]

(146b) — I kupi ešče buloček desjatok!
and buy_{IMPERATIVE} another rolls_{GEN.PL} unit-of-ten_{ACC.SG}
'And buy about ten more rolls.'

(146c) On kak-to s"el v odin prisest jaic djužinu.
eggs_{GEN.PL} dozen_{ACC.SG}
'One day he consumed about a dozen of eggs in one sitting.'

²¹⁹ It is clear that (143) is not an instance of emphatic-thematic inversion (as in (137b-c) above) for the following reason: Mel'čuk (1985:146) argues convincingly that emphatic-thematic inversion does not allow special COUNT GEN.PL forms:

√Ljudej/*Čelovek	nužnogo	namtipa	my	najdëm	rovno desjat'
people	needed	us kind	we	will-find	exactly ten
(N.)GEN{NON-COUNT/COUNT}	(ADJ)MASC.GEN.SGDAT	(N.MASC)GEN.SG	NOM.PL(V)FUT.2.PL	(ADV)	(NUM)ACC
'Of the kind of people we need we'll find exactly ten.'			[= his ex. 14; glosses added/LAB]		

Example (143) uses the COUNT form, meaning that it cannot be an instance of emphatic-thematic inversion. Cf. other examples of PL agreement with *četvert'* in exx. 108a-b, Crockett (1976:398).

(146d) [...] Ja frontov desjatok peresěk [sic].
 I_{NOM.SG} fronts_{GEN.PL} unit-of-ten_{ACC.SG} infiltrated_{(V)PAST.MASC.SG}
 ‘I (have) infiltrated about ten fronts/battles.’

(146e) A ešče let čerez desjatok uznali oni, čto [...]
 years_{GEN.PL} through unit-of-ten_{ACC.SG}
 ‘And after another ten years or so they found out that ...’
 [citing A. Solženicyn’s *Bodalsja telenok s dubom* (no cit.)]

Mel’čuk is careful to point out that all of these are **approximative**-inversion examples, with the right prosodic and other requisites (1985:160-61, n. 3). He adds that all these sentences have an added conversational tenor, but are “not outside the bounds of standard Russian”—especially (146d-e). He also admits that the factors involved in determining whether this specific type of example is acceptable are not fully understood, citing illicit data with some of the same measure nouns as in (146). It may well be that they are becoming archaic (as I suggest for (143) above). My informants consider each of (146a-e) to be “somewhat bookish”. If sentences like (143) and (146) are less than fully acceptable, then this is predicted by my suggestion that approximative inversion takes place only when there is a quantificational element as sister of N’. Measure nouns are in N°. ²²⁰ As the emphatic-thematic-inversion examples in (137) show, there are other phenomena that alter sub-clausal constituents.

²²⁰ It may also be that words like *desjatok* ‘unit-of-ten’, which are morphologically derived from a numeral stem—in this case form *desjat* ‘ten’—were at some point numerals and have progressed diachronically toward noun-hood. This would explain the decreasing acceptability of such words in approximative inversion. As nouns, the element they quantify also has a very different structure, that of an adnominal NP, while numerals are sister of N’ (within the conception of the quantified NP in Babby 1987). One problem is that such measure nouns have made the opposite transition with regard to the distribution of COUNT (GEN.PL) forms. In §4.6.4 I cite Crockett (1976) who writes that words like *desjatok* have only recently become able to trigger COUNT forms. Thus, while words like *desjatok* have lost the ability to invert, they have gained the ability to trigger COUNT forms. I believe that the ability to trigger COUNT forms was once limited to numerals and now has extended to other quantifiers which express counted quantity. Approximative inversion, however, has been the sole purview of numerals. Recall from §4.3.1 that all numerals were originally nouns or adjectives in Old Russian. Franks (1994:661-62, fn. 73)—citing W. Browne, L. Langlois [DePerno], Bogusławski (1966:92, 109) and Drovnikova (1985:66)—reports that the earliest attested examples of approximative inversion correspond to the dates when noun and adjective number words became reanalyzed as numerals.

For example, inverting adjective-noun order (as in (51) above) has other semantics, none of them approximative.

Recall from the discussion of the large-number words above (in §4.3.4) that *tysjača* ‘thousand’ can be either a noun or a numeral. This predicts that this word optionally undergoes approximative inversion. Specifically, the measure noun generally prohibits inversion (147a), and the numeral *tysača* requires approximative inversion in (147b).²²¹ Note that the *s* means ‘about’, not ‘with’; and *tysjaču* is the ACC-case form of ‘thousand’ (not the INST form *tysjač’ju*):²²²

<p>(147a) <i>s tysjaču litrov</i> thousand liters (N.FEM)ACC.SG (N.MASC)GEN.PL ‘about a thousand liters’ [≈ ex. (65) above]</p>	<p>(147b) <i>litrov s tysjaču</i> liters thousand (N.MASC)GEN.PL (NUM)ACC ‘about a thousand liters’ [<i>Sintaksis</i> (1980:448, 71)]</p>
---	---

An even more striking contrast is shown in the following minimal pair:

²²¹ Recall that *tysjača* ‘1000’ is only beginning to function as a numeral in modern Russian; higher numbers don’t function as numerals. Mel’čuk (1985:103, 149) reports that *tysjača* undergoes approximative inversion only under certain circumstances; it does not if it is part of a complex integer: **rublej dve tysjači pjat’ sot* ‘roubles_{GEN.PL} two_{NOM/ACC} thousand_{GEN.SG} five-hundred_{NOM/ACC}’; nor does *tysjača* invert when it is in an oblique case (non-NOM or -ACC), regardless of whether there are prepositions:

(i) **rubljami (s) tysjač’ju* ‘roubles_{INST.PL} (with) thousand_{INST}’ [Mel’čuk (1985:149)]

NB: This is a special numeral form of ‘thousand’; *tysjačej* is the INST.SG of the corresponding noun (see fn. above in §5.1 on *tysjača*). This measure noun does, however, invert in direct (NOM or ACC) cases (again, regardless of whether there is a preposition):

(ii) *My vstretili čelovek tysjaču* ‘We met about a thousand people.’
we_{NOM.PL} met_{(V)PAST.PL} people_{GEN.PL(COUNT)} 1000_{ACC} [= ex. 6 in Mel’čuk (1985:149)]

(iii) [...] *vėrst za tysjaču* ‘about a thousand versts away.’
versts_{GEN.PL} away_P 1000_{ACC}
 [= ex. 7 in Mel’čuk (1985:149), citing V. Majakovskij’s *150.000.000* (no cit.); glosses added/LAB]

²²² This does not mean that measure words are excluded from approximate inversion altogether. They are perfectly acceptable in the **noun** position. Pete (1984:76) lists three such examples, including the following one: [...] *zarabatyvat’ sotni poltory* ‘earn_{(V)INFIN} unit-of-hundred_{(N.FEM)GEN.SG} one-and-a-half_{(NUM)ACC}’ (= ‘to earn about 150 roubles a month’) [quoting M. Gor’kij (no cit.)]

(148a) s četvert´ ČAsa (148b) čaSA s četvert´
 quarter hour hour quarter
 (N.FEM)ACC.SG (N.MASC)GEN.SG (NMASC)GEN.SG.ADP AUC (NUM)ACC
 ‘about a quarter hour’ ‘about a quarter hour’

(149a) *s četvert´ čaSA (149b) *ČAsa s četvert´
 quarter hour hour quarter
 (NUM)ACC.SG (N.MASC)GEN.SG.ADP AUC (N.MASC)GEN.SG (N.FEM)ACC
 [elicited/LAB; cf. (56b)]

Note the stress on *časa* in each. Only *ČAsa* is acceptable in the grammatical **uninverted** order in (148a); in the grammatical **inverted** order in (148b) only *čaSA* is acceptable. The other stress on either example is not acceptable (149a-b). This is due to the fact that the numeral *četvert´*, which triggers ADPAUC stress (*čaSA*) in the noun it quantifies, requires inversion, while the noun *četvert´*, which triggers the non-ADPAUC stress, *ČAsa*, generally prohibits approximative inversion.

One final question needs to be answered: Why is *s+ACC* required to undergo approximative inversion? The answer, in my view, lies in the synonymy of *s+ACC* and approximative inversion, coupled with *s+acc*’s one-word-complement requirement. That is, inversion is a means to achieve the one-word requirement. Measure nouns, which generally cannot invert, are left in their canonical position, after both *s* and the numeral. As for *okolo*, this preposition is not proclitic and has semantics distinct from approximative inversion. What of other proclitic approximative prepositions? The closest one semantically is *v+ACC* of identity, discussed above (in §3.2), which only optionally undergoes approximative inversion. My explanation for this, based on my discussion of the different semantics of *s* and *v* (in §3.2), is that *v+ACC* has a slightly different meaning from *s+ACC*.²²³ The additional question of why approximative

²²³ This is one of the reasons why I chose to discuss *v+ACC* outside of chapter 5; it does not exactly fit the description of “other approximative constructions” fully .

inversion is possible without *s*+ACC being required I put off until the next chapter, where I also formalize the structural position of the inverted noun.

There are many more restrictions on and characterizations of approximative inversion (see Mel'čuk 1985:147-61). For the purposes of this study it is sufficient to say the following: First, approximative inversion is the inversion of a numeral (or other [+ Q] sister of *N'*) and the noun which it quantifies, a reversal of the numeral-noun precedence. Next, if the complement of *s* is a numerically quantified nominal expression, then approximative inversion is obligatory in the modern language. Finally, approximative inversion does not occur if there is more than one word in *N''*. In such cases a pleonastic count noun is inserted before the numeral.

5.3 Regarding *ètak* 'about/approximately'

One other point is worth mentioning in connection with approximative inversion: Mel'čuk (1985:364) lists *ètak* and its variant *èdak* as—*inter alia*—yet another way of saying 'about/approximately' in Russian. Mel'čuk points out that this word is primarily used in approximative-inversion constructions:

(150a)	(150b)	(150c)
rublej ètak dvadcat´	ètak rublej dvadcat´	? ètak dvadcat´ rublej
roubles about twenty	about roubles twenty	about twenty roubles
GEN.PL (ADV) NOM/ACC	(ADV) GEN.PL NOM/ACC	(ADV) NOM/ACC GEN.PL
		[all from Mel'čuk (1985:363)]

Following Mel'čuk, I consider *ètak* to be an adverb of approximation; he mentions elsewhere (1985:152) that adverbs are optionally ordered between the noun and numeral in approximative inversion; cf., e.g., *tam* 'there' in ex. (2) above. Pete (1984:76) points out this word is used "to emphasize the idea of approximation", adding that *tak*, which usually means 'like so' or 'thusly', can also be used for this function. Apparently *ètak* can either participate in the inversion, as in (150a), or not, as in (150b). But it is odd for *ètak* to appear in a numerical phrase without any

inversion, as shown in (150c). It is even possible to combine *ètak* and *s+ACC* in the same example, as shown in (151a). Examples (151b-c) show other prepositions with *èdak/tak*; (151d) shows *ètak* with inversion and without prepositions (there happen to be two ordinal numerals conjoined).

(151a) Lěnja polučal togda rublej ètak s pjat´ sot.
 Lěnja received then roubles about about five-hundred
 (MASC)NOM.PL (V)PAST.MASC.SG ADV (MASC)GEN.PL (ADV) (P) (NUM)ACC
 ‘Lěnja was receiving (a monthly salary of) about five hundred roubles then.’
 [= ex. 5a in Mel’čuk (1985:363)]

(151b) Ja starše tebja, let èdak na dvadcat´ pjat´ ...
 I older you years about by twenty-five
 (SG)NOM (COMPAR) (SG)GEN (N)GEN.PL (ADV) (P) (NUM)ACC
 ‘I’m older than you, by about twenty-five years.’
 [Pete (1984:76), quoting M. Gor’kij (no cit.); my glosses/LAB]

(151c) èto obojdětsja rublej tak v tysjaču
 this will-cost roubles about into thousand
 (SG)NOM (V)FUT.3.SG (N.MASC)GEN.PL (ADV) (P) (NUM)ACC
 ‘This (expense) will run into about a thousand roubles’ [Pete (1984:76); my glosses/LAB]

(151d) Byl čas èdak vos´moj ili devjatyj [...]
 was hour about eighth or ninth
 (V)PAST.MASC.SG (N.MASC)NOM.SG (ADV) (ADJ)MASC.NOM.SG (ADJ)MASC.NOM.SG
 ‘It was between 7 and 9 o’clock.’ [Pete (1984:76), quoting Kuprin (no cit.); my glosses/LAB]

It appears that, like *s+ACC*, this word is **required** to undergo approximative inversion. Predictably, based on the behavior of *s+ACC*, this would predict that the one numeral that cannot invert, *pol* ‘half’ (based on my discussion in §4.3.5), **is** allowed not to invert after this adverb: *ètak (s) polčaSA* ‘about half an hour’.²²⁴

²²⁴ Mel’čuk (1985:363) lists three ways to express approximate measure in Russian (which he treats elsewhere in the book): (i) adverbs like *priblizitel’no* ‘approximately’ and *ètak*, (ii) prepositions like *s* and *okolo*, and (iii) comparatives (without *čem*) which assign GEN case like *bol’še* ‘more’. He adds that whereas approximative adverbs can coexist in the same construct as either approximative prepositions or *čem*-less comparatives, all three are not allowed to coexist. Crucially, apparently, approximative prepositions and *čem*-less comparatives are what cannot coexist. Perhaps due to the organization of his book, Mel’čuk unfortunately does not assess whether any of these three can coexist with approximative inversion. Apparently, one adverb, *ètak*, and one preposition, *s+ACC*, each can (indeed must if possible) coexist with approximative inversion (cf. (105) above).

I argue in this section that the semantics of approximative inversion constitute a subset of the semantics of *s+ACC* and as such, inversion is free to occur. I would argue that the semantics of *ètak*, which probably includes an emphatic component, as Pete (1984:76) suggests, has the semantics of approximative inversion as a proper subset as well. What requires movement is probably a requirement that *ètak* be adjoined to the inverted noun's landing site, which, as I argue in the next chapter, is a specifier position. Whereas adjoining to specifier position is rare, it is not unheard-of in the literature (cf. Rudin 1988).

I conclude this brief section by summarizing the facts of *ètak*: This approximative adverb is like *s+ACC* in requiring approximative inversion. Like many other adverbs, *ètak* can either precede or follow the inverted quantified noun. As in the *s+ACC* construction, the one numeral that is allowed to keep its noun complement without inversion is *pol* 'half'. Thus, two odd-looking approximative constructions share the property of requiring approximative inversion to take place if possible.

5.4 Regarding *neskol'ko* 'several'

One last means of expressing indefinite quantity which deserves mention is *neskol'ko* 'several/a-few/some'.²²⁵ So far in this dissertation I have referred to this word a few times: For example, I mention that *neskol'ko* and *skol'ko* 'how many?' require the COUNT form of the noun they quantify (§4.6.4). I also mention the calcified *s*-plus-adjective *wh* word *skol'ko* 'how many', from which this word is derived—by means of the indefinite proclitic *ně-* (§4.2.4). In (101a) the phrase *neskol'ko grekov* 'several Greeks' is conjoined with numerical expressions, suggesting the same structure as a numerically quantified noun phrase. Example (101b) shows that *neskol'ko* can also be

²²⁵ There is another, quite distinct meaning of *neskol'ko*, 'somewhat_(ADV)', which I do not discuss.

(152d) [...] *neskol'ko* *pričin* **oxvatili** *lager'* [...]

 several reasons gripped camp

 NOM (N.FEM)GEN.SG (V)PAST.PL (N.MASC)ACC.SG

‘... several reasons gripped the camp ...’

 [DePerno (1991:ch.6:12), quoting Pasternak (1959:368)]

In (152a-c) various adjectival elements show PL agreement with *neskol'ko*: the universal quantifier, a demonstrative determiner, and a (de-participial) modifier adjective, respectively. When a noun phrase with *neskol'ko* is in syntactic-subject position, then PL verbal agreement is possible as well, as shown in (152d).²²⁷ Several other examples and similar argumentation are also in Chey (1967:63, 65-67). A PL adjective is not generally triggered by non-numerals, which can take either NOM/ACC.PL or GEN.PL adjectives.

Another peculiarity of *neskol'ko* is that it requires the COUNT form in the noun it quantifies. Unlike other “adverbial” quantifiers, DePerno (1991:ch.6:5) reports that this word can only quantify countable items:²²⁸

(153) [...] *uvidel* [...] *ob'javlenie* *i* **neskol'ko** *čelovek*, *čitavšix* *ego* [...]

 saw announcement and several people reading it

 (V)MASC.SG (N.NEUT)ACC.SG ACC COUNT (PRT)ACC.PL (SG)ACC

‘... he saw ... an announcement and several people reading it.’

 [DePerno (1991:ch.6:12), quoting Solženicyn (1968:424)]

²²⁷ When there is a PL determiner, then the predicative agreement **must** be PL; cf. (53a-d) and fns.

²²⁸ The quantified noun in (153) is modified, by the participial phrase headed by *čitavšix*. I show above in §4.6.4 that such structures prohibit the COUNT form. Indeed, the same is true for *neskol'ko*:

[...] *neskol'ko* *prišlyx* *ljudej* [...]

 several_{NOM} newly-arrived_{(ADJ)GEN.PL} people_{(N.MASC)GEN.PL.NON-COUNT}

‘... several newly arrived people ...’ [DePerno (1991:ch.6:12), quoting (1965:236)]

Approximative inversion, which requires a single-**PrWd** N'', cannot take place with most modifiers. Franks (1995:167-688), following Mel'čuk (1985:96, 151), shows that contrastively emphatic and prosodically heavy modifier phrases apparently undergo some sort of heavy-**AP** shift out of N'', which in turn allows approximative inversion. The same sort of rightward movement seems to apply to (153).

Crockett (1976:319) and Rožkova (1966:37) also list which adverbial quantifiers take COUNT form. As I show, however, in chapter 4 (§4.3.4), several types of constituents trigger the COUNT form. Essentially any element which denotes a countable quantity will do so. The fact that *neskolʹko* triggers COUNT forms is a necessary, but far-from-sufficient condition for the numeral-hood of *neskolʹko*. What **is** significant, in my view, is the fact that this word cannot quantify non-countable nouns. Numerals also have this limitation. Then again, so do nouns like *million* ‘million’. The COUNT argument merely shows that *neskolʹko* is a semantically special type of quantifier.

In addition to triggering PL agreement and requiring COUNT forms, *neskolʹko* shows an asymmetry between the direct and oblique cases. As I show in my introduction to the properties of numerals (in §4.3.1) above, all numerals show a distinction between the NOM/ACC cases and all the others. When the NP is assigned a direct case, then the numeral bears that case and triggers GEN case (or ADPAUC/COUNT if applicable) in the noun it quantifies. If the NP is assigned an oblique case, then both numeral and noun bear that oblique case morphologically. The same is true of *neskolʹko*: In all of the examples in (152) and (153) the overall NP is assigned either NOM or ACC case and triggers GEN.PL—in (153) specifically the COUNT form—on the noun it quantifies. Compare this to *ot neskolʹkix družej* ‘from_(P) several_{GEN} friends_{(N.MASC)GEN.PL}’, in which the GEN-assigning preposition *ot* ‘from’ requires both words to be in this oblique case. Alas, this argument does not conclusively support the numeral-hood of *neskolʹko*; other “adverbial” quantifiers, like *mnogo* ‘plenty’, which do not even trigger the COUNT form in the noun they quantify, nonetheless do exhibit such an asymmetry.

Yet another property of *neskolʹko* is that it shares declensional properties with the so-called collective numerals. DePerno (1991:ch.6:5), citing Vinogradov (1947:314), shows that stems formed from the *wh*-interrogative root /kolʹ-/, including

skol'ko ‘how many?/what a lot!’ (which I discuss somewhat above in §4.2.4), decline like collective numerals; cf. (32e) and (102a) above. That is, collective numerals and (*ne*)*skol'ko* have the /-o/ ending in the NOM and ACC cases, and adjectival (long-form) endings in the oblique cases. The spelling of the collective numerals obscures this, however: *troe* ‘three’ is underlyingly /troj-o/, with initial stress keeping the final /o/ from being pronounced with lip rounding (cf. §4.3.5). The corresponding GEN forms are (*ne*)*skol'kix* (cf. preceding paragraph) and *troix*, both underlyingly /(*ne*)skol'k-ix/ and /troj-ix/, respectively.²²⁹

Another property that *neskol'ko* shares with collective numerals is the ability to modify *pluralia tantum* nouns (other examples of which are shown in §3.1 above). As I show in (101) above, the use of collective numerals has become quite limited in modern Russian. (See Mel'čuk 1985:376-405 for fuller details.) Virtually the only structure where collective numerals are **required**, however, is with *pluralia tantum* nouns. It is therefore significant that *neskol'ko* can quantify such a noun as well:

- (154) [...] menja i neskol'ko sutok sosedi ne obnaružat.
 me even several days neighbors not discover
 (SG)ACC ACC (N.PL)GEN (N.MASC)NOM.PL NEG (V)FUT.3.PL
 ‘... the neighbors won't even find me for several days.’
 [DePerno (1991:ch.6:12), quoting Solženicy'n (1968:367)]

The noun *sutk-* means ‘24-hour period’ and is only used in the PL. I have elicited such structures with other adverbial quantifiers, however: *mnogo sutok* ‘many days’. This property, then, is likewise inconclusive.

²²⁹ These two word-types also share accentuational declensional similarities. This particular feature is not observable on *neskol'ko*, however, because prefixal *ne-* is inherently accented, requiring initial stress throughout the paradigm of *neskol'k-*; for this reason I show only *skol'k-*. In the direct cases both stems show initial stress: *SKOL'ko*, *TROe*. In the oblique cases the stress is on the declensional ending: *skol'KIX*, *troIX* (both shown here in the GEN case). Alas, this accentuational pattern is only attested in Russian according to Vinogradov (1947:314). Aside from fixed expressions like *k skol'KIM* ‘toward_(P) how-many_{DAT}’ (= ‘by what time’), in standard Russian this word has fixed stem stress.

One last property of *neskol'ko* is that it undergoes approximative inversion:

- (155) Vyzvali oni [...] odnu devušku, potom druguju
 summoned they one girl then other
 (V)PAST.PL NOM.PL (ADJ)FEM.ACC.SG (N.FEM)ACC.SG (ADV) (ADJ)FEM.ACC.SG
- i raspekali časov po neskol'ko .
 and upbraided hours apiece several
 (V)PAST.PL (N.MASC)GEN.PL (P) ACC

'They summoned one girl and then the other and upbraided (them) for several hours each.'
 [DePerno (1991:ch.6:12), quoting Solženicy'n (1968:339)]

I have hardly discussed the distributive prepositional quantifier *po* in this study except for the occasional footnote and a few examples—cf. (32e), (34e), (59c), and (100c)—primarily because *po* does not express approximation. This preposition is discussed at length in Babby (1985; 1987) and Franks (1994; 1995). Babby argues that *po* has the same structure as quantificational *okolo* discussed at the beginning of this chapter—[NP [N' [N' [PP[+Q] *po* [NumP numeral]] [N' [N' N°]]]]] (with my slight modifications of the *po*-numeral structure). I have implied that only numerals allow approximative inversion (because measure nouns are gradually becoming unable to trigger inversion). I have also shown, however, that the quantificational PP headed by *okolo* triggers inversion. The same appears to have taken place in (155). Note however that if the distributive component of (155) is removed—i.e., one 'girl' instead of 'two girls'—then approximative inversion, according to my informants, is not acceptable: **Vyzvali oni devušku i raspekali časov neskol'ko* 'They summoned a girl and upbraided her for several hours.' This suggests that it is the properties of *po* and not of *neskol'ko* that allow approximative inversion: *po* heads a [+Q] PP which quantifies the noun. The sister of *po* within that PP can be a numeral or some other quantificational element. Thus, the fact that *neskol'ko* can undergo inversion in the

structure in (155) does not show conclusively that *neskol'ko* is a numeral, just that it can be the quantifier sister of *po*.²³⁰

The combination of indicators—PL agreement, requiring COUNT forms, only quantifying countable nouns, accentuational and distributional similarities with collective nouns, and the ability to undergo approximative inversion each suggest that *neskol'ko* is a numeral. I have shown, however, that each of these indicators is inconclusive. The only thing that can be said categorically about *neskol'ko* is that it quantifies only **countable** nouns, and as such, requires the COUNT form if possible. I've shown that other quantificational elements likewise trigger the COUNT form, not just numerals (cf. §4.3.4) and certain non-numerals require the COUNT form, so there is no solid case to be made for the numeral-hood of *neskol'ko*.

I conclude this chapter on non-*s* constructions that express approximative measure with a brief summary: Quantificational *okolo* (i.e., the use of *okolo* that means ‘approximately’) must be the sister of a [+Q] constituent in order to have a quantificational interpretation. Approximative inversion reverses the order of a numeral and the noun which that numeral quantifies; this type of inversion is required if it is possible, even requiring *s* constructions to invert if the complement of *s* consists of a numeral and a noun. The only exception is the (otherwise unique) numeral *pol* ‘half’. Unlike *s*+ACC and several other constructions investigated in the previous

²³⁰ DePerno shows other examples of *po* with the somewhat archaic DAT-case form *neskol'ku*. Babby (1985) and Franks (1995) list similar examples in which numerals take the DAT after *po* in what is by now decidedly archaic Russian. Note that the quantified noun after *neskol'ku* is **not** in the DAT:

[...] *po neskol'ku čelovek* [...] ‘... several people each ...’
apiece_(P) several_{NOM} people_{(N.MASC)GEN.PL(COUNT)}
 [DePerno (1991:ch.6:12), quoting Solženicy'n (1968:171)]

This example is consistent with the behavior of numerals in that archaic period: (non-paucal) numerals after distributive *po* bore DAT case but the quantified noun remained invariably in the GEN. Babby and Franks have differing analyses of these archaic-Russian facts, which I do not reproduce here. What is important is that *neskol'k-* here occupies the same structural position as the numeral does.

chapter, which are subject to single-syntactic-word constraints, approximative inversion is subject to a single-**prosodic**-word constraint. I also briefly investigated one other approximative word: *ètak* ‘about/approximately’, showing that it too, like *s+ACC*, requires approximative inversion. In the last section I showed that *neskol’ko* ‘several’, while appearing to be numeral-like, is not conclusively a numeral.

Chapter 6 Optimality-theoretic treatment of *s*+ACC:

I have deferred many of the formal mechanisms until this final chapter due to the complexity of the data. I begin with an introduction to the data that must be accounted for (§6.1), followed by a brief summary of Optimality Theory as applied to syntax (§6.2). I then propose an Optimality model of approximative inversion (§6.3), followed by a model of *s*+ACC itself (§6.4). I conclude the chapter with a brief commentary on the universal viability of the constraints I propose for Russian (§6.5).

6.1 A summary of the crucial data

The surface ACC complement of *s*—i.e., the overt material after *s*—must consist of a single syntactic word. That single syntactic word can either be an unmodified noun, as in (156a), or a numeral and noun, which has undergone approximative inversion, as in (156b), or a numeral with an elided noun, as in (156c):

- (156a) Prošlo **s** **nedelju.** ‘About a week passed.’
 passed about week
 (V)PAST.NEUT.SG (N.FEM)ACC.SG [= (1a) above]
- (156b) **časov s pjat’** ‘about five hours’
 hours about five
 GEN.PL ACC [≈ (10) above]
- (156c) Pušek u vas budet s **dvadcat’.**
 cannons at you will-be about twenty
 (N.FEM)GEN.PL (P) (PL)GEN (V)3.SG ACC (NUM)ACC
 ‘(As for) cannons, you will have **about twenty.**’ [simplification of (12) and (139a) above]

That is, either approximative inversion or ellipsis is required if it is at all possible.

In chapter 4, specifically (81c), I concluded that the complement of *s* must be a single syntactic word (SnWd). The following three distinct types of exceptions, however, remain unexplained:

(157a) **s polversty** ot nix
 about [half [verst]] from them
 (P) [ACC [(N)GEN.SG]] (P) GEN.PL
 ‘about half a verst from them’ [≈ (37d) above]

(157b) **so šljapku sopožnogo gvozdika**
 cap shoe nail
 (N.FEM)ACC.SG (ADJ)MASC.GEN.SG (N.MASC)GEN.SG
 ‘about-the-size-of the head of a cobbler’s nail.’ [≈ (21c) above]

(157c) Rodničok vsego-to — **s detskuju ladon´**
 about child’s palm
 (P) (ADJ)FEM.ACC.SG (FEM)ACC.SG
 ‘The spring is only about the size of a child’s palm.’ [= (43a) above]

In (157a) the numeral *pol* ‘half’ is a special stump compound (as defined in §4.3.5 above), one of the properties of which is the requirement that its complement not undergo approximative inversion or ellipsis. A constraint is needed to allow the noun which *pol* quantifies to remain immediately after it. Another exception to the single-word restriction is shown in (157b): certain adnominal complements are allowed after the ACC-case complement of *s* (cf. §4.4 above). As I showed in the preceding chapter (§5.2), nouns, as opposed to numerals, generally cannot undergo approximative inversion. Thus, another constraint is needed to account for complements with adnominal genitives. In (157c), quite similarly to the rationale for examples like (157b), the adjective is licensed, as it were, because it further delimits the semantics of **measure**. That is, *ladon´* ‘palm (of the hand)’ is not sufficiently accurate to depict the size of the ‘spring’ being described in this example (cf. §4.2.3 above). A mechanism is needed to allow such adjectives which further delimit size.

The three example types in (157a-c) are the **only** actual exception types to the requirement that the complement of *s* be a single syntactic word. Interestingly, these three example types constitute exceptions from quite distinct grammar components: The morphology justifies (157a). The rationale for (157b-c) must come from the

semantic sphere but their structure is defined in syntactic terms. Only within a theory like Optimality could such seemingly disparate data be dealt with uniformly.

One additional problem that must be explained is the complicated set of facts surrounding approximative inversion, as discussed (in §5.2) above. Because approximative inversion or ellipsis is required when the *s+ACC* complement includes a numeral it will be necessary to devise constraints that accurately account for the data. Because this is not a study specifically of inversion I will only outline a possible Optimality-theoretic approach to this phenomenon.

Constraints are needed to explain the following: First, when a nominal expression consists only of a noun and the numeral (\neq *pol* ‘half’) which quantifies it, then the noun-numeral order carries an approximative reading:

- (158) **časa tri** ‘about three hours’
 hours three
 GEN.PL (NUM)ACC [≈ (125b) above]

Second, when these two words are the complement of a prosodically light preposition, the approximative-reading order is noun-preposition-numeral:

- (159) **časa na tri** ‘(designated) for about three hours’
 hour for three
 (MASC)GEN.SG (ADPAUC) (P) (NUM)ACC [≈ (126b) above]

Third, if the preposition is prosodically heavy and does not mean ‘approximately’, then the approximative order is preposition-noun-numeral:

- (160) **Otnositel’no očkov tridcati** ‘regarding approximately thirty points.’
 regarding points thirty
 (P) (N.NEUT)GEN.PL (NUM)GEN [≈ (130c) above]

Fourth, if that heavy preposition is quantificational, then the approximative inversion is expressed by the order noun-preposition-numeral:

Optimality Theory is most useful as a theory of human language because, in addition to Gen, there are constraints which are violable. An individual language is defined as a unique ranking of these universal constraints. In each language the constraints are ranked in such a way that the most highly ranked one is not violated.²³² Other, less highly ranked constraints can be violated in order to satisfy a higher constraint. Thus, an individual language's grammar is defined as its constraint hierarchy. Languages differ because the same universal inventory of constraints is ranked in a different order in each.

Although Prince & Smolensky (1993) deal primarily with phonology, their theory lays claim to the entire grammar and has been applied to other grammar components. Grimshaw (1993; 1995),²³³ an Optimality-theoretic treatment of English clausal syntax, argues that the set of competing candidates is the one in which each candidate begins with the same "input", which is defined roughly as the same lexical and semantic material. I make use of some of Grimshaw's proposed constraints in my analyses of approximative inversion and *s+ACC*.

Grimshaw's work is quite applicable for this study because it deals with another kind of inversion: Among other things, Grimshaw accounts for why matrix clauses with a *wh* subject require the order *Who will read the books?*, with the inflected verb after the subject, while the corresponding clause with a *wh* object *What will she read?*, with the same inflected verb before the subject pronoun. She argues

²³² In many cases there can be more than one "superordinate" constraint. This does not necessarily mean that these most highly ranked constraints are not ranked with respect to each other, it only means that these constraints do not interact. For example, the requirement that the complement of *pol* 'half' immediately follow it, as in (157a), does not ever come into direct conflict with the requirement that a noun immediately precede its adnominal complement, as in (157b). Thus, ranking either of these above the other yields the same results. The theory assumes, however, that such constraints are ranked, but that the actual ranking cannot be determined conclusively.

²³³ I specify both these works because the yet unpublished 1995 version differs considerably from the 1993 version. Many of the constraints in Grimshaw (1995) do not appear in Grimshaw (1993).

that while questions like *Who will read the books?* have a matrix I[nflectional]P[hrase], ones like *What will she read?* have an added C[omplementizer]P. That is, if the *wh* word is in a specifier position—*who* is in Spec of IP—then there need not be a CP.

Grimshaw proposes that *wh* phrases, as operators, must be in a specifier position. She calls this constraint OP-SPEC.²³⁴ The subject, which occupies the Spec-of-IP position already satisfies this constraint and does not require a CP projection to satisfy OP-SPEC. The direct object, which is not already in a specifier position, must be given a new specifier position. A new projection, CP, is constructed over IP in order to satisfy OP-SPEC: the direct object moves to Spec of CP.

Furthermore, Grimshaw proposes that each projection requires a head: OB-HD. The new CP projection occupied by *what* thus requires a head. The inflected element *will* is positioned in C° in order to satisfy this constraint, leaving a trace in I°. Additionally, in order to limit the overapplication of movement, Grimshaw (1995:1) also proposes STAY, which rules out traces.

Of the microgrammar I've mentioned here, the relative ranking in English is OP-SPEC » OB-HD » STAY. This entails that OP-SPEC is not violated, but OB-HD is violated when its requirements conflict with those of OP-SPEC. Furthermore, STAY is violated when its requirements conflict with either of OP-SPEC or OB-HD.

To “prove” that any one constraint outranks another the convention is to use a “tableau”. I explain each of the notational conventions below this first tableau:

²³⁴ Unless I note otherwise, the constraints are defined on page 1 of both Grimshaw (1993; 1995).

(163) Matrix interrogatives in English [\approx Tableau 1 in Grimshaw (1995:5)]

Candidates	OP-SPEC	OB-HD	STAY
a. \$ [CP <i>what</i> _j will _i [IP she <i>t</i> _i [VP read <i>t</i> _j]]]			**
b. [CP <i>will</i> _i [IP she <i>t</i> _i [VP read what]]]	*!		*

In an Optimality tableau the candidates, which I have labeled (a) and (b), are listed along the left column. The remaining columns are used for assessing each constraint. The square where each column intersects with a candidate's row is called a cell. An asterisk is used to signify a violation by that candidate of the corresponding constraint listed above it. For example, candidate (163b) violates OP-SPEC once and STAY once, while candidate (163a) violates only STAY, but twice. That is, candidate (b) has one operator that is not in specifier position, the *wh* direct object within the verb phrase and one trace (the STAY violation); candidate (a) has two traces, indicating two STAY violations. This tableau proves that OP-SPEC dominates STAY. In shorthand: OP-SPEC » STAY. There do not happen to be any OB-HD violations in either of the candidates of this tableau.

Following Grimshaw (1993) I use a dollar sign (\$) to indicate the optimal candidate. Candidate (163a) is optimal for the following reason: The most highly ranked constraint, OP-SPEC, is assessed first: Candidate (163b) has more violations of this constraint than does candidate (a), making (a) the optimal form. The optimality determination is made at the highest constraint at which there is no tie. A tie is whenever the attested candidate and the candidate with which it is being compared fare equally with regard to a constraint—i.e., have the same number of asterisks, be it none, one, two, etc. The particular violation that makes the crucial determination is marked with an exclamation point after the asterisk. Once an optimality determination has taken place, all other cells to the right of that column are immaterial to the

determination. Following Prince & Smolensky (1993), such cells are shaded to indicate that the asterisks in them are immaterial.

Following conventional generative-syntactic notation, “t” indicates a trace. For clarity, I italicize the constituent extracted from a trace.

The fact that candidate (163a) actually has more violations of the more lowly ranked constraint STAY is also immaterial to the optimality determination. It does, however, prove the relative ranking of OP-SPEC » STAY. As Prince & Smolensky (1993) show, the crucial tableau configuration for proving that one constraint outranks another is shown in (164), in which candidate (a), the attested form, violates constraint A and the candidate (b), an unattested form, violates constraint B. *Ceteris paribus*, based on which candidate is attested, it is possible to more highly rank the constraint violated by the **unattested** candidate, CONSTRAINT-A:

(164) The classic arrangement of violations to prove relative rankings of constraints

Candidates	CONSTRAINT-A	CONSTRAINT-B
a. \$ Attested Candidate		*
b. Unattested Candidate	*!	

It is this crucial configuration of four cells in which asterisks appear on diagonally opposite cells only that constitutes proof that one constraint outranks another.²³⁵

I return to the tableau in (163): Ignoring the OB-HD column in (163) for the moment, it is possible to conclusively rank OP-SPEC » STAY in English. Note that it is

²³⁵ Note that both candidates violate STAY in (163). The crucial factor is that one candidate violates this constraint **more** times than the other candidate does. Crucially, Optimality Theory does not count; it merely compares which candidate has more violations in a given constraint’s column.

necessary that all other factors remain equal, that no third constraint interfere with the comparison.

The data in (163) do not make any determination about the ranking of OB-HD relative to the other two constraints. It is often the case that different data are needed to prove different rankings. The following tableau shows that OB-HD » STAY:

(165) Matrix negative-induced inversion [≈ Tableau 18 in Grimshaw (1995:30)]

Candidates	OP-SPEC	OB-HD	STAY
a. \$ [XP <i>Never</i> _j <i>will</i> _i [IP she <i>t</i> _i [VP read this <i>t</i> _j]]]			**
b. [XP <i>Never</i> _j <i>e</i> [IP she will [VP read this <i>t</i> _j]]]		*!	*

“XP” in (165) stands for a maximal projection of some sort (which Grimshaw does not actually name) required by preposing the element *never* to the front of the sentence.²³⁶ In this tableau both candidates tie with regard to OP-SPEC, thus requiring consideration of the next most highly ranked constraint, OB-HD. Only candidate (165a) satisfies this constraint, making it the more optimal one of the two. This tableau also proves that OB-HD » STAY, because of the classic configuration, as in (164), in which the attested candidate violates OB-HD **less** times than the unattested one, while the reverse is true for the other constraint, STAY, which the attested candidate violates **more** times than the unattested one does.

Having proven that OB-HD » STAY in tableau (165), and that OP-SPEC » STAY in tableau (163), it is further possible to prove that OP-SPEC » OB-HD:

²³⁶ Grimshaw actually uses *under no circumstances* in her tableau, which might make it clearer for some readers why the negative element is preposed. I use *never* because it fits better in the tableau.

(166) Subordinate interrogatives [≈ Tableau 15 in Grimshaw (1995:25)]

Candidates	OP-SPEC	OB-HD	STAY
a. \$ I know [_{CP} <i>what</i> _j <i>e</i> _i [_{IP} she will [_{VP} read <i>t</i> _j]]]		*	*
b. I know [_{IP} she will [_{VP} read what]]]	*!		

By itself, tableau (166) only proves that OP-SPEC dominates either one or the other of OB-HD or STAY. But, since the ranking OB-HD » STAY is determined using different data (i.e., the preceding tableau), it is further possible to rank OP-SPEC » OB-HD.

Furthermore, the combined rankings from tableaux (163), (165) and (166) show that OP-SPEC » OB-HD » STAY. These are just three of twelve constraints proposed in Grimshaw (1995). I chose these three as an example to illustrate how an Optimality approach to syntax works. I also use two of these constraints below in my analysis of the Russian data. I have also selected only two candidates from each of Grimshaw's tableaux for simplicity's sake. I show some other of her constraints below in my application of Optimality Theory to Russian.

In this section I have presented an introduction to Optimality Theory, specifically how it has been applied to Syntax. I have shown the following crucial features of the theory: First, constraints are violable; if a constraint comes into direct conflict with a more highly ranked constraint then there can be a violation of the lower constraint in the attested utterance. I have also shown the various notational conventions of an Optimality-theoretic tableau. Finally, I have also shown how constraint rankings are proven. In the following sections I will use the same types of constraints and notation to show how *s+ACC* works. First, however, I must formalize how approximative inversion works, because *s+ACC* is so closely related to this phenomenon.

6.3 Formalizing Russian approximative inversion

As I mention above, this is not a study primarily devoted to approximative inversion and therefore the Optimality-theoretic analysis is not a conclusive one. In order to fully understand approximative inversion, a comparative study of the slightly different corresponding phenomena in Ukrainian and Belarusian (the only other languages with this phenomenon) is needed. The following is one **possible** approach that yields the correct results.

I repeat the crucial data in (158) through (162) shown again here as (167) through (171), respectively:

(167) časa tri 'about three hours'
 hours three
 GEN.PL (NUM)ACC [≈ (125b) and (158) above]

(168) čaSA na tri 'for about three hours'
 hour for three
 (MASC)GEN.SG (ADPAUC) (P) (NUM)ACC [≈ (126b) and (159) above]

(169) Otnositel'no očkov tridcati 'regarding approximately thirty points.'
 regarding points thirty
 (P) (N.NEUT)GEN.PL (NUM)GEN [≈ (130c) and (160) above]

(170) časov okolo dvux 'approximately two hours'
 hours about two
 (N.MASC)GEN.PL (P) (NUM)GEN [≈ (131e) and (161) above]

(171) štuk desjat' starinnyx knig 'about ten antique books'
 items ten antique books
 GEN.PL (NUM)ACC (ADJ)GEN.PL (N.FEM)GEN.PL [≈ (135c), (136d) and (162) above]

Fowler (1988:39-40) suggests that this is a Move-Alpha phenomenon; Franks (1995:165-74) argues convincingly that approximative inversion is specifically head movement, namely, movement of the noun to the left. Franks assumes that the noun adjoins to some constituent, weighing the pros and cons of whether the adjunction site

is a head or a maximal projection. Since I use a different phase-structure model of the quantified noun phrase, following Babby (1987), I will assess one possibility here.²³⁷

The properties of approximative inversion can be accounted for by an “approximative operator”: As shown in the preceding section, Grimshaw proposes a constraint requiring operators to be in a specifier position (OP-SPEC). Like many operators, this approximative operator originates in a canonical position within the lexical projection and moves to the closest available specifier position. This operator, unlike the *wh* phrases in tableaux (163) and (166) above, which move within verbal projections, undergoes movement within a **nominal** projection, to Spec of NP. That is, the noun moves to its own specifier position to render the approximative meaning.

There is evidence for and against this suggestion: First, there is very little empirical support for a determiner position in Russian, be it the Spec of NP in the framework I use here or a separate D[eterminer] phrase, as Franks (1994; 1995) and others use, following recent theory. There are no articles in Russian, and many elements analogous to an NP ending in *'s* in English actually appear **after** the noun. Rappaport (1992) proposes that possessive NPs are generally ordered pre-nominally if they agree with the noun and post-nominally if an adnominal-GEN structure is required. Moreover, multiple determiner-like elements can co-occur in the same nominal expression: *Ja čital ètu ego poèmu*, literally ‘I read **this his** poem’ [Franks (1994:645, fn. 49), citing Avrutin (1992)], which suggests that only the demonstrative is the determiner and possessives are not in Spec of NP. The only overt material that can translate to the concept of “determiner” in pre-nominal position are demonstrative pronouns like *èt-* ‘this’ —see (21f), (48), (53), (64), (125a) and (142) above—or an NP

²³⁷ For the purposes of this study I, too, will assume that the noun’s position in approximative inversion does not entail a new maximal projection.

ending in a very minimally productive suffix, *-in-*, which translates roughly as 's in English. Borsley & Jaworska (1988) discuss the determiner position in Polish.

Despite the preceding evidence against a determiner position in Russian, there is some reason to believe that one exists in the language nonetheless: Franks (1994:645, fn. 49), again citing Avrutin (1992), shows that while extracting possessive *wh* phrases is usually possible in Russian, extracting it out of the nominal expression is impossible when there is a demonstrative pronoun: **Č'ju ja čital ètu poèmu?* (literally: 'Whose I read this book?'), suggesting that the demonstrative is in the determiner position, which then blocks extraction.²³⁸

As M. Yadroff has pointed out to me personally, approximative inversion apparently requires the nominal expression to be non-referential. At the very least, approximative inversion is impossible with an overt determiner.²³⁹ It therefore follows that the landing site of the quantified noun in approximative inversion is the same position usually occupied by determiners.²⁴⁰

The proposals so far in this section account for examples (167), (169) and (170), repeated as (172) through (174) with the relevant phrase structures added.

²³⁸ There is indirect evidence for a determiner in Russian: In (53b-c) I show that if there is a demonstrative—which invariably shows agreement with its head noun—in an NP that's the clausal subject, then the predicate must show agreement. In (53b), the determiner is FEM.SG and the verb can only be {FEM/3}.SG; in (53c) the determiner is PL and the verb must show PL agreement. In neither case is default (i.e., {NEUT/3}.SG) verbal agreement allowed. (Cf. also (63a-b) incl. fn.) This suggests that if a quantified nominal expression is subject and the verb shows default agreement, then the determiner position is **not** filled (or there is no DP projection à la Franks). Similarly, the *wh* element *skol'ko* 'how many?' can extract to clause-initial position only if the verb has default agreement, which suggests an empty determiner. Cf. ex. 97 in Franks (1994:660) and Pesetsky (1982:229 n. 50, 399-420). This argumentation in Franks (1994) not only supports a determiner position, it makes a reasonably good case for the determiner being in a separate projection, in order to block movement because the D[eterminer]P keeps the moved item from antecedent-governing the trace left behind.

²³⁹ Franks (1994:646, fn. 50) assesses the issue of referentiality in the complement of *po* 'apiece', as in (32e), (34e) and (100c) above. That construction, too, excludes demonstratives in *èt-*.

²⁴⁰ I do not further justify the structure with the determiner in Spec of NP or even the proposal that the moved noun goes to Spec of NP. It is possible to translate the following proposals into a theory in which the noun adjoins to NP (or PP), but using a constraint other than OP-SPEC.

In all of the structures below I use the model of the NP in Babby (1987) amended as follows: I do not show the N^{''''} or N^{'''} nodes. It is understood that the NumP inside NP is always the daughter of N^{'''} which is not shown. Any constituent to the left of NumP within NP is in the Spec-of-NP position. I show N^{''} if it has an adjective daughter.

(172) [NP *časa_i* [NumP *tri*]NumP [N^{''}[N[°]*t_i*]N[°]]NP 'about three hours'
 hours three
 GEN.SG (NUM)ACC [≈ (125b), (158) and (167) above]

(173) [PP[P[°] *Otnositel'no*]P[°] [NP *očkov_i* [NumP *tridcati*]NumP [N^{''}[N[°]*t_i*]N[°]]NP]PP
 regarding points thirty
 (P) (N.NEUT)GEN.PL(NUM)GEN
 'regarding approximately thirty points' [≈ (130c), (160) and (169) above]

(174) [NP *časov_i* [PP [P[°] *okolo*]P[°] [NumP *dvux*]NumP]PP [N^{''}[N[°]*t_i*]N[°]]NP
 hours about two
 (N.MASC)GEN.PL (P) (NUM)GEN
 'approximately two hours' [≈ (131e), (161) and (170) above]

That is, the quantified noun moves to the Spec of NP in each of these structures, as indicated by the subscript "i" on the moved noun and the trace left behind in N[°] position. The structure of (174) is discussed extensively (in §5.1) above.

An explanation is still needed for the example with a prosodically light preposition, (168), and the one in which the quantified noun is modified, which requires a pleonastic pre-numeric noun, example (171). I deal first with the latter: As I argue above (in §5.2), the relative order of the numeral and quantified noun changes only if the noun is the only syntactic word in the constituent quantified by the numeral. That is, a single-word restriction similar to the one investigated in depth above for s+ACC in chapter 4 is apparently at work in approximative inversion. The following is a preliminary formulation of such a constraint:

- (175) LONE-WD [N'', approximative inversion, PrWd]:
There is no [approximative inversion] if the [N''] (i.e., constituent quantified by the numeral) consists of more than one [prosodic word].

The formula in (175) is written in such a way for other phenomena in Russian, in other languages even, to make use of this constraint by filling in the three variables: the constituent assessed, the operation, and the kind of constituent the first variable must be. I make use of this same constraint to formalize the single-word restriction in *s+ACC* as well below. Based on the empirical fact that LONE-WD is unviolated in approximative inversion, and following Grimshaw's OP-SPEC constraint—"syntactic operators must be in specifier position"—the following conflicting requirements result: OP-SPEC requires an operator in a specifier position while LONE-WD prohibits the quantified noun in (171) from moving. The solution arrived at in Russian is for a pleonastic noun to fill the specifier position. Grimshaw discusses a similar phenomenon, English *do*-support, proposing that *do* is "a semantically and functional empty verbal head" (Grimshaw 1993:28; 1995:8) inserted into structures to fill functional-head positions because of OB-HD (discussed in §6.2 above). Grimshaw (1995) proposes the FULL-INT[erpretation] constraint ("lexical conceptual structure is parsed") to limit the use of *do* support. Essentially, using **any** word in a syntactic position without also parsing its lexical conceptual structure²⁴¹ violates FULL-INT, but selecting *do*, the verb with the least LCS to begin with, is favorable to using any other verb, which inevitably involves more LCS and thus incrementally more violation of FULL-INT.

In the model of Russian approximative inversion I'm pursuing, a specifier position must be filled, which is motivated by OP-SPEC, but cannot be filled by the

²⁴¹ Lexical conceptual structure can be defined simply as the "meaning" that is lexically associated with a word. That is, *do* means something and using *do* in a *do*-support role is using the word without parsing its meaning.

adjectivally modified noun in (171). Like the English verb *do*, the Russian nouns *človek* ‘people GEN.PL.COUNT’ and *štuk* ‘items_{GEN.PL}’ seem to be the most semantically generic.²⁴² I repeat (171) as (176) with phrase structure added:

(176) [NP *štuk* [NumP *desjat*']_{NumP} [N'' *starinnyx* [N' [N°*knig*]N°]N']N'']NP
 items ten antique books
 GEN.PL (NUM)ACC (ADJ)GEN.PL (N.FEM)GEN.PL
 ‘about ten antique books’ [≈ (135c), (136d), (162) and (171) above]

The structure in (176) essentially shows that the Spec-of-NP position is filled by a pleonastic noun when the numeral quantifies more than one prosodic word, which, because of LONE-WD, cannot move to Spec of NP itself.

This leaves only one more type of example within approximative inversion to be described structurally: example (168), in which a light preposition is between the inverted noun and numeral. My primary idea here is that the Spec-of-NP position is not the final landing site because this would break up the prosodic word which consists of the proclitic preposition and the numeral. I therefore propose the following constraint:

(177) PR-CNTG: Maintain prosodic-word contiguity in approximative inversion.

Because there is a PP projection in (178) it is possible for the noun to move to the Spec of PP:

(178) [pp *čaSA*_i [P° *na*]_{P°} [NP*t*_i [NumP° *tri*]_{NumP} [N' [N°*t*_i]N°]N']NP]PP
 hour for three
 (MASC)GEN.SG.ADP AUC (P) (NUM)ACC
 ‘for about three hours’ [≈ (126b), (159) and (168) above]

²⁴² Cf. Sussex (1976), DePerno (1990; 1991) for further discussion on the semantically depleted status of these words. DePerno also discusses why the features [+ human]/[– human] are kept intact.

In (178) propose that the noun moves first to Spec of NP and then onward to Spec of PP, leaving two traces behind. This structure is preferable to the one in which the noun moves only as far as Spec of NP, because that landing site entails breaking up a prosodic word consisting of the proclitic preposition *na* and the word that follows it in the structure, *tri* ‘three’.

With the structures in (172)-(174) and (177)-(178) it is now possible to construct Optimality tableaux. I make use of LONE-WD in (175), PR-CNTG in (177), as well as Grimshaw’s OP-SPEC, STAY (both discussed in the preceding section²⁴³), and FULL-INT. First I assess the simplest structure:

(179) Simple numeral-noun inversion [cf. (125b), (158), (167) and (172) above]

‘about five hours’	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [časov _i [pjat’] _{NumP} [[t _i] _{N°}] _{N’}] _{NP}					*
b. [[pjat’] _{NumP} [[časov] _{N°}] _{N’}] _{NP}			*!		
c. [štuk [pjat’] _{NumP} [[N° časov] _{N°}] _{N’}] _{NP}				*!	

Whereas each of the tableaux above includes only two candidates, this one has three. Crucially, any time there are more than two candidates in a tableau, the comparisons must be between the attested candidate and one other. Tableau (179), therefore, is equivalent to two tableaux, one with candidates (179a-b), the other with candidates (179a, c).

Comparing (179a-b) proves that OP-SPEC » STAY in Russian. Comparing candidates (179a, c) proves that FULL-INT » STAY.

²⁴³ Grimshaw also lists a constraint NO-LEX-MVT, “a lexical head cannot move”, which would appear to restrict the quantified noun from moving. I proceed with the understanding that STAY and NO-LEX-MVT are ranked consecutively, and therefore functionally as one constraint in Russian.

(180) Approximative inversion with heavy P [cf. (130c), (160), (169) and (173) above]

'regarding approximately thirty points'	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a.\$ [[Otnositel'no] _P [očkov _i tridcati [[_{t_i}] _{N°}] _{N'}] _{NP}] _{PP}					*
b. [[Otnositel'no] _P [tridcati [[očkov] _{N°}] _{N'}] _{NP}] _{PP}			*!		
c. [[Otnositel'no] _P [štuk tridcati [[očkov] _{N°}] _{N'}] _{NP}] _{PP}				*!	
d. [očkov _i [Otnositel'no] _P [_{t_i} tridcati [[_{t_i}] _{N°}] _{N'}] _{NP}] _{PP}					*!*

In candidate (183d) I assume cyclic movement first to Spec of NP, then onward to Spec of PP, thus two violations of STAY.²⁴⁴ Comparing (180a-b) proves once more that OP-SPEC » STAY in Russian; comparing (180a, c) proves again that FULL-INT » STAY. Comparing candidates (180a, d) shows that movement all the way to Spec of PP constitutes an unnecessary violation of STAY..

(181) Approximative inversion with *okolo* [cf. (131e), (161), (170) and (174) above]

'approximately two hours'	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [časov _i [okolo [dvux] _{NumP}] _{PP} [[_{t_i}] _{N°}] _{N'}] _{NP}					*
b. [[okolo [dvux] _{NumP}] _{PP} [[časov] _{N°}] _{N'}] _{NP}			*!		
c. [štuk [okolo [dvux] _{NumP}] _{PP} [[časov] _{N°}] _{N'}] _{NP}				*!	

In (181) I assume that other constraints restrict movement to the Spec of PP or Spec of NumP, perhaps because either of these movements would constitute movement downward into a maximal projection. I argue above (in §5.2) that the noun must move

²⁴⁴ Other superordinate constraints (or Gen) presumably keep the noun from moving directly to Spec of PP without an intermediate trace in Spec of NP.

to the left of both *około* and the numeral in such structures. Assuming as I am that the landing site of the noun is an operator position, it makes sense for that landing site to obligatorily also c-command the trace. In either Spec of PP or Spec of NumP this is not so. Otherwise, tableau (181) proves no new rankings.

(182) Approximative inversion with adjective [cf. (135c), (136d), (162), (171), (176) above]

'about ten antique books'	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [štuk[desjat´] _{NumP} [starinnyx[[knig] _{N°}] _{N´}] _{N´´}] _{NP}				*	
b. [[desjat´] _{NumP} [starinnyx [[knig] _{N°}] _{N´}] _{N´´}] _{NP}			*!		
c. [knig _i [desjat´] _{NumP} [starinnyx [[t _i] _{N°}] _{N´}] _{N´´}] _{NP}	*!				*
d. [[starinnyx [[knig] _{N°}] _{N´}] _{N´´} [desjat´] _{NumP} t _i] _{NP}	*!				*

In this tableau the LONE-WD constraint comes into force. It does not prove anything about the relative ranking between it and OP-SPEC, because comparisons must always be between the **attested** constraint and one other, but does show that OP-SPEC » FULL-INT and that LONE-WD » FULL-INT.

(183) Approximative inversion with light P [cf. (126b), (159), (168) and (178) above]

'for about five hours'	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [časov _i [na] _{P°} [t _i [pjat´] _{NumP} [[t _i] _{N°}] _{N´}] _{NP}] _{PP}					**
b. [[na] _P [[pjat´] _{NumP} [[časov] _{N°}] _{N´}] _{NP}] _{PP}			*!		
c. [[na] _{P°} [časov _i [pjat´] _{NumP} [[t _i] _{N°}] _{N´}] _{NP}] _{PP}		*!			*
d. [štuk [na] _P [[pjat´] _{NumP} [[časov] _{N°}] _{N´}] _{NP}] _{PP}				*!	*
e. [[na] _P [štuk [pjat´] _{NumP} [[časov] _{N°}] _{N´}] _{NP}] _{PP}		*!		*!	

In both of candidates (196c, e) an overt element breaks up the contiguity of the prosodic word *na tri*, thus violating PR-CNTG. I assume that the trace in (196a), though situated between the two parts of this prosodic word, nonetheless does not cause it to be disrupted in the surface utterance, meaning no PR-CNTG violation. Tableau (183) adds a few more new rankings, proving—using candidates (196a, c)—that PR-CNTG » STAY. This tableau does not, however, show the relative ranking between PR-CNTG and FULL-INT, which means that either one of these two constraints rules out candidate (196e); it cannot be determined, however, from this tableau which constraint of the two actually does this, hence the asterisks in both columns.

The preceding five tableaux have shown the relative rankings in (184a):

(184a) OP-SPEC	»	STAY	cf. (179), (180), (181), or (183)
FULL-INT	»	STAY	cf. (179), (180), (181), or (183)
OP-SPEC	»	FULL-INT	cf. (182)
LONE-WD	»	FULL-INT	cf. (182)
PR-CNTG	»	STAY	cf. (183)

(184b) LONE-WD » STAY

(184c) {PR-CNTG › { { OP-SPEC › LONE-WD } » FULL-INT } } » STAY

Transitively, combining the results of (182) with any of (179), (180), (181) or (183), the additional ranking in (184b) can further be determined. Finally, the precise ranking in (184c) can be derived, where “›” means that the two constraints (or groups of constraints) on either side of this symbol cannot be ranked conclusively. That is, STAY is dominated by each of the other four constraints; additionally, OP-SPEC and LONE-WD each dominate FULL-INT. It is often the case that the entire consecutive ranking cannot be determined fully (cf. Grimshaw 1995:2). In my treatment of *s+ACC*

itself (§6.4.3) I return to these constraints, adding to them and further defining some of their relative rankings.

In this section I have devised one possible set of Optimality constraints that yields the attested approximative-inversion data. In addition to Grimshaw's constraints OP-SPEC, STAY and FULL-INT, I have proposed the following two constraints: LONE-WD, which disallows inversion when the quantified noun is modified, and PR-CNTG, which disallows breaking up a prosodic word in the process of inversion. In the next section the *s*+ACC construction itself is formalized. At the very end of this chapter (§6.5) I reconsider the universal viability of the constraints I've proposed.

6.4 The treatment of *s*+ACC proper

In this section I propose an Optimality-theoretic model to account for *s*+ACC, the primary aim of this dissertation. To account for the three example types that override the single-word restriction listed above at the beginning of this chapter I propose three new constraints. I divide the section into three parts: One of these deals with multi-word complements of *s* without numerals, examples (157b-c). Another deals with the one numeral, *pol* 'half', which overrides the single-word restriction. First, however, I formalize *s*+ACC's single-word restriction in Optimality-constraint form.

6.4.1 Formalizing the single-word constraint: I propose that the single-word constraint is a universal one, which a language may make use of for certain lexically marked constructions. In the case of Russian, this constraint is used in connection with approximative inversion **and** *s*+ACC (as well as several other constructions discussed in §4.6 above). The following is its *s*+ACC version:

- (185) LONE-WD [NP, s+ACC, SnWd]:
 There is no [s+ACC] if the [NP] (i.e., complement of *s*) consists of more
 than one [syntactic word]. [≈ (81c) ; see also tableau (175) above]

This constraint is identical to the single-word restriction on approximative inversion in (175), except that the three variables have been changed. I underline this one to distinguish it from (175).

Note that I specify a single **syntactic** word, which keeps all the other types of attested multi-word complements—prequantifier adjective (§4.2.1), syntactic compounds (§4.2.2), calcified expressions (§4.2.4)—from being restricted against. Other constructions in Russian, discussed above (in §4.6), are likewise subject to this constraint. For example, the ADPAUC is limited to a single-word environment: LONE-WD [N'', ADPAUC, SnWd], which limits the N'' to a single SnWd if it is to receive the distinctive ADPAUC morphology.

Thus, one constraint is used for various constructions which, for one reason or another, have single-word restrictions. In the following subsections I apply this constraint, along with others, to generate the s+ACC data.

As it so happens, it is possible to merge the LONE-WD constraints for approximative inversion and s+ACC into one and still get the same results. I do this in the tableaux below primarily due to width restrictions. It would not be at all problematic, however if, confronted with additional data, these two constraints are require to be ranked separately.

6.4.2 Formalizing non-numerical exceptions to the single-word constraint: In the following discussion I account for s+ACC complements consisting of a noun and either an adnominal-NP complement or an adjective modifier.

The following two sentences, repeated from (157b-c), are examples of each :

(186) **so šljapku** **sopožnogo** **gvozdika**
 cap shoe nail
 (N.FEM)ACC.SG (ADJ)MASC.GEN.SG (N.MASC)GEN.SG
 ‘about-the-size-of the head of a cobbler’s nail.’ [≈ (21c) and (157b) above]

(187) Rodničok vsego-to — **s detskuju** **ladon’**.
 about child’s palm
 (P) (ADJ)FEM.ACC.SG (FEM)ACC.SG
 ‘The spring is only **about the size of a child’s palm**.’ [≈ (43a) and (157c) above]

I am not claiming that just any adjective or adnominal-GEN complement is allowed. Rather, additional words are licensed because they are essential to the meaning of **measure** in the *s+ACC* complement. For example, omitting *sopožnogo gvozdika* in (186) would leave only the meaning of ‘about the size of a cap/small hat’. Likewise, removing the adjective *detskuju* in (187) changes the meaning to ‘about the size of a palm’ (i.e., ‘a handful of water’), the original size, that of a **small** handful, is lost. The common-sense notion here is that more than a word has to be uttered in order to convey certain approximate measures. In other words, the semantics must be parsed. I propose the following constraint:

(188) P-MEAS: Fully parse the measure in the *s+ACC* complement.

This constraint is specialized in that it allows extra material only if that material further delimits the measure represented by the noun in the complement of *s*.

I begin with a single-noun example to show that it doesn’t violate any of the constraints proposed so far:

(189) Single-noun complement [cf. (1a) and (183) above]

‘about a week’	P- MEAS	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [[[s]P° [[[nedelju]N°]N°]NP]PP						

I don't show any competing candidates because no other form could fare better than this one, which has absolutely no violations of any of the constraints so far. In order to save space in the remaining tableau in this subsection, I do not list PR-CNTG, OP-SPEC, FULL-INT or STAY. This is because these four constraints do not incur any violations in any non-numerical candidates I list here.

(190) Complement with adnominal NP [cf. (21c), (157b) and (186) above]

'about the size of the head of a cobbler's nail'	P- MEAS	LONE -WD
a. \$ [[so] _{P°} [[[šljapku] _{N°} [sopožnogo gvozdika] _{NP}] _{N'}] _{NP}] _{PP}		*
b. [[so] _{P°} [[[šljapku] _{N°}] _{N'}] _{NP}] _{PP}	*!	

The same arrangement of asterisks is incurred by the measure-adjective example:

(191) Complement with adnominal NP [cf. (157c) and (187) above]

'about the size of a child's palm'	P- MEAS	LONE -WD
a. \$ [[s] _{P°} [[[detskuju [[ladon´] _{N°}] _{N'}] _{N''}] _{NP}] _{PP}		*
b. [[s] _{P°} [[[[ladon´] _{N°}] _{N'}] _{N''}] _{NP}] _{PP}	*!	

The P-Meas constraint is not entirely interesting theoretically because it essentially acts as an override mechanism for LONE-WD, a stipulative constraint. I nevertheless need P-MEAS to account for the data in (190) and (191).

In this subsection I have proposed a constraint to insure that certain additional words be allowed so long as they contribute to the semantics of measure. This constraint is more highly ranked than the single-word constraint, causing ACC-case complements of *s* to be generated with either adjective modifiers or adnominal-NP complements under very specific circumstances. In the next subsection I combine the

constraints used to generate approximative inversion in the preceding section with the single-word constraint for *s+ACC* defined in this subsection to then limit the size of an *s+ACC* complement which contains a numeral.

6.4.3 Modeling numerical s+ACC complements: Here I devise a set of constraints which require approximative inversion when the *s+ACC* complement consists of a numeral and a noun, but which does not require inversion if the numeral is *pol* ‘half’.

I rely on the constraints developed in the preceding section, on approximative inversion. For convenience I repeat these constraints’ definitions and rankings::

(192) PR-CNTG: Maintain prosodic-word contiguity in approximative inversion. [= (177)]

(193) OP-SPEC: Syntactic operators must be in specifier position. [Grimshaw (1993:1; 1995:1)]

(194) LONE-WD [N´´, approximative inversion, PrWd]:
 There is no [approximative inversion] if the [N´´] (i.e., constituent quantified by the numeral) consists of more than one [prosodic word]. [= (175)]

(195) FULL-INT(erpretation): Lexical conceptual structure is parsed. [Grimshaw (1995:1)]

(196) STAY: Trace is not allowed (also known as the economy of movement). [Grimshaw (1995:1)]

(197) {PR-CNTG › {{ OP-SPEC › LONE-WD } } » FULL-INT } } » STAY [= (197C)]

In addition, in this section I have also introduced the following two constraints and their ranking relative to each other:

(198) P-MEAS: Fully parse the measure in the *s+ACC* complement. [= (188)]

(199) LONE-WD [NP, s+ACC, SnWd]:
 There is no [s+ACC] if the [NP] (i.e., complement of *s*) consists of more than one [syntactic word]. [= (81c) and (185); see also tableau (175) above]

(200) P-MEAS » LONE-WD [cf. tableaux (190)-(191)]

The rankings in (197) are, for the moment, still separate from the ranking in (200), because the two sets have not been used in the same tableau yet.

(201) **Rankings so far for (standard) Russian:** [Combining (197) and (200)]
 {P-MEAS » LONE-WD} » {{{PR-CNTG » OP-SPEC} » LONE-WD} » FULL-INT » STAY }

In the following tableaux I determine whether these sets of constraints interact. I also propose another application of one of these constraints.

As I point out in my main discussion of approximative inversion (in §5.2) above, the *s+ACC* construction is unique in requiring approximative inversion. Other prepositions with similar semantics, such as *v+ACC* of identity (discussed in §3.2) and quantificational *okolo* (in §5.1) both undergo approximative inversion only **optionally**. I propose that the semantics of *s+ACC* **includes** the semantics of approximative inversion, but not vice versa. Thus, if there is a numeral involved in the complement of *s*, then there can be no *s+ACC* without approximative inversion. There can, however, be approximative inversion without *s+ACC*. Approximative inversion also conveniently eliminates one word from the overt complement of *s*.

I begin with a numeral that allows inversion. Because of space limitations I do not write brackets or labels around the numeral (i.e., “**pjat**” = “[**pjat**’]_{NumP}”).²⁴⁵

²⁴⁵ In candidate (202e) I show the option of having the syllabic variant *so* instead of *s* because the preposition would procliticize to a word with the initial cluster *št*; under such conditions the syllabic *so* is usually attested: *so štopku* ‘about as (thick) as darning thread’ [elicited/LAB].

(202) Complement consisting of non-paucal numeral and noun [cf. (10), (156b) above]

'about five hours'	P- MEAS	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [[časov _i [s] _{P°} [t _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}						**
b. [[s] _{P°} [časov _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*	*!			*
c. [[s] _{P°} [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		*		*!		
d. [štuk [s] _{P°} [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		*			*!	
e. [[s(o)] _{P°} [štuk pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		**	*!		*!	

In this tableau the attested form, candidate (202a), competes with each of the other candidates (202b-d). I have placed exclamation points after each asterisk in the PR-CNTG, OP-SPEC, LONE-WD and FULL-INT columns because each of these four constraints has been proven to dominate STAY, as summarized most recently in (197). Because at least one of PR-CNTG, OP-SPEC, LONE-WD and FULL-INT dominates STAY, there can be no ranking of P-MEAS or LONE-WD relative to any of these constraints. Nor is either of P-MEAS or LONE-WD even needed in this tableau; deleting the LONE-WD column from tableau (202) would achieve the same results. I merely list these two constraints to the left of all the others for now.

Note that P-MEAS, as I've defined it, does not exclude numeral-noun sequences, as in (202b, d-e), because both the numeral and the noun contribute to the semantics of measure. Note also that moving the quantified noun to the Spec of NP, as in candidate (202b), still incurs a LONE-WD violation, because there are still too many words in the complement of *s*.

In order to explain the relative ranking of LONE-WD with the candidates to its right in tableau (202) I must make a brief excursus into the colloquial register: An alternative to PR-CNTG is to require the noun to move to the specifier of the highest

projection, to the Spec of PP. Franks (1995:170) lists colloquial variants of the type *na časov pjat'* ‘for about five hours’ (literally: ‘for_(P) hours_{(N.MASC)GEN.PL five_{(NUM)ACC}’), with the light preposition preceding both the noun and numeral (cf. the footnote in ex. (130c) above). He suggests that only the standard register requires movement to the highest projection. Unfortunately such a requirement would generate the wrong results for prosodically heavy prepositions, as in tableau (180). I would propose that in the colloquial register the constraint PR-CNTG is ranked lower than STAY, essentially rendering it irrelevant.}

I was unable, however, to elicit colloquial examples like **s časov pjat'*—i.e., the would-be colloquial counterpart of (202a). I have a brief explanation for this: Assume that colloquial Russian has a grammar with the same rankings as I’ve shown so far, **except** with PR-CNTG ranked **below** STAY:

(203) **Rankings in colloquial Russian** [\neq rankings in (197) and (200)]²⁴⁶

{P-MEAS » LONE-WD} , {{{OP-SPEC » LONE-WD} » FULL-INT} » STAY » PR-CNTG}

Unlike the standard register in tableau (183) above, the optimal candidate in this grammar is the one in which the quantified noun moves once, to Spec of NP:²⁴⁷

²⁴⁶ In the remaining **standard**-Russian tableaux I list the constraints with PR-CNTG first, which is consistent with the ordering in (203). This is in anticipation of the data in tableau (206) and ff.

²⁴⁷ There is one other difference between tableau (204) and its standard-Russian counterpart in (183). I use a non-paucal numeral to remove a complication which I deal with below: Pleonastic count nouns like *štuk* are not attested if the numeral is paucal.

(204) **Colloquial**Russian with prosodically light P [compare with tableau (183) above]

'for about five hours'	P- MEAS	LONE -WD	OP- SPEC	FULL -INT	STAY	PR- CNTG
a. [časov _i [na] _P [t _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}					*!*	
b. [[na] _P [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}			*!		*	*
c. \$ [[na] _P [časov _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}					*	
d. [štuk [na] _P [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}				*!		
e. [[na] _P [štuk pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}				*!		*

In tableau (204) the constraints P-MEAS and LONE-WD have no violations because there is no *s+ACC* construction. There is still approximation, as the gloss shows informally; thus OP-SPEC applies, ruling out candidate (204b). Since there is only a lone noun quantified by the numeral, there is no chance of violating LONE-WD in these candidates. FULL-INT rules out the two candidates with pleonastic nouns in (204d-e). Of the two remaining candidates (204a) violates STAY more times than (204c) does, leaving candidate (204c) as the optimal one, and therefore the attested form.

Tableau (205) now uses *s+ACC* instead of *na*, but still in the colloquial register:

(205) **Colloquial**Russian with *s+ACC* [compare with (202)]

'about five hours'	P- MEAS	LONE -WD	OP- SPEC	FULL -INT	STAY	PR- CNTG
a. \$ [časov _i [s] _P [t _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}					**	
b. [[s] _P [časov _i pjat' [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*!			*	*!
c. [[s] _P [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		*!	*!			
d. [štuk [s] _P [pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		*!		*!		
e. [[s(o)] _P [štuk pjat' [[časov] _{N°}] _{N'}] _{NP}] _{PP}		*!*		*!		*!

Concentrating first on candidates (205a-b), it is evident that if it were not for the LONE-WD constraint, the grammar would incorrectly yield (205b) as the optimal candidate. If there were no constraint dominating STAY ruling out candidate (205b), then the optimality decision would be left to the STAY constraint, which is violated more times by (205a) than by (205b). Comparing any of (205c-e) to the attested candidate in (205a) is inconclusive as to the relative ordering of LONE-WD with any of the other candidates except STAY.

This excursus into colloquial Russian provides additional evidence that PR-CNTG is the type of constraint needed, rather than a constraint requiring movement to the highest projection. It also shows that a constraint like LONE-WD is needed in order to rule out candidates like (205b). In the standard register the LONE-WD constraint actually does no real work when there are numerals, since at least one of PR-CNTG, OP-SPEC or FULL-INT each take care of ruling out candidates (202b-e). This does not mean that LONE-WD is completely redundant in standard Russian; it rules out all sorts of extra words (except for measure-semantics adjectives and adnominals; cf. tableaux (190)-(191) in §6.4.2 above).

Returning now to the standard register for the remainder of the study, there is one more major problem to solve: In *s*+ACC examples with a complement consisting of a **paucal** numeral and quantified noun there never appears to be the option of inserting a pleonastic noun. This does not cause problems with most of the paucal numerals, as the following tableau shows:

(206) Complement consisting of paucal integer and noun

'about half an hour'	P- MEAS	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. \$ [čaSA _i [s] _P [t _i tri [[t _i] _{N°}] _{N'}] _{NP}] _{PP}						**
b. [[s] _P [čaSA _i tri [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*	*!			*
c. [[s] _P [tri [[čaSA _i] _{N°}] _{N'}] _{NP}] _{PP}		*		*!		
d. [štuk(i) [s] _P [tri [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}		*			*!	
e. [[s(o)] _P [štuk(i) tri [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}		**	*!		*!	

As tableau (206) shows, the constraints PR-CNTG, OP-SPEC and FULL-INT, each of which are proven to dominate STAY using previous tableaux, are sufficient to rule out any of the unattested candidates, in (206b-d). That is, regardless of which pleonastic noun is used in candidates (206d-e), the GEN.PL *štuk* or the GEN.SG *štuki*, these two candidates are ruled out independently.²⁴⁸

The problem is a bit more complicated when the morphologically unique numeral *pol* 'half' is used instead:

²⁴⁸ I know for certain that it is the GEN.SG that is used when a pleonastic noun is inserted before the numeral to achieve the effect of approximative inversion. It does not happen, however, in structures like (205)-(207). Moreover, GEN.SG is used only when there is a paucal **integer**—e.g., *tri* 'three'—and the constituent it quantifies consists of more than one prosodic word—e.g., and adjective and noun:

Iz ètogo materiala vyjdet vsego **štuki** tri novyx plat'ev.
 item three new dresses
 GEN.SG (NUM)NOM (ADJ)GEN.PL (N.NEUT)GEN.PL

'From this material only **about three new dresses** can be made.' [Elicited/LAB]

Note that the noun is not in the GEN.SG, the form expected when a morphological-nom numeral governs it. Note also, however, that the constituent quantifying *novyx plat'ev* 'new dresses' is not the numeral alone, but rather the combined element *štuki tri*, meaning that there is not, technically speaking, numerical quantification here. When I say "integer" I really mean "numeral greater than 1". It appears to be possible for *poltora/poltory* 'one and a half' to likewise take such pleonastic nouns in these structures, which otherwise behaves as a numeral—i.e., *poltor-* triggers the ADPAUC, undergoes approximative inversion and passes other tests of numeral-hood. Replacing the whole number here appears to be acceptable, thus leading me to believe that the crucial factor is that fractions are excluded.

(207) Complement consisting of *pol* ‘half’ and noun (preliminary)

‘about half an hour’	P- MEAS	LONE -WD	PR- CNTG	OP- SPEC	FULL -INT	STAY
a. [čaSA _i [s] _{P°} [t _i pol [[t _i] _{N°}] _{N'}] _{NP}] _{PP}			*!			**
b. [[s] _{P°} [čaSA _i pol [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*	*!*			*
c. \$ [[s] _{P°} [pol [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}		*		*!		
d. [štuk(i) [s] _{P°} [pol [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}		*			*!	
e. [[s(o)] _{P°} [štuk(i) pol [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}		**	*!		*!	

As I show in my primary discussion of this unique numeral (in 4.3.5) above, *pol* ‘half’ is distinct from all other numerals in being lexically required to be part of a morphological “stump” compound. Specifically, *pol* must have the following morphological structure:

(208) [*pol*_{STUMP} [noun]_{M_rW_d}]_{M_rW_d} [= (82b)]

That is, *pol* is lexically specified as morphologically deficient and must adjoin overtly to its complement. I also show that *pol*’s prosodic structure is the following:

(209) [*pol*_{P_rW_d} [noun]_{P_rW_d}]_{P_rW_d} [= (82a)]

Assuming that some sort of superordinate LX≈PR constraint²⁴⁹ requires that the morphological structure in (208) correspond to the prosodic structure in (209), it is possible to use PR-CNTG again. I repeat its definition once more:

²⁴⁹ Prince & Smolensky (1993) propose the following constraint:

LX≈Pr (MCat): A member of the morphological category *MCat* correspond[s] to a PrW_d.
[= ex. 52 in Prince & Smolensky (1993:43)]

Footnote continued on next page

(210) PR-CNTG: Maintain prosodic-word contiguity in approximative inversion.
[= (177) and (192)]

Note that there are two PrWds in the structure in (209), one embedded within the other. Moving the noun to a specifier position to the left of *pol* constitutes a violation of PR-CNTG because the matrix PrWd—the one indicated with bold-faced brackets and label—would no longer be contiguous. Hence any movement by *čaSA*, as in (207a-b), is a violation of PR-CNTG. Other, more familiar uses of PR-CNTG are shown in candidates (207b, d), in which the landing site of the moved noun (or the insertion point of the pleonastic noun) breaks up the prosodic word formed by a proclitic preposition and the next word. (Candidate (207b) thus violates this constraint twice: once by moving out of a PrWd and again by moving into another.)

That said, tableau (207) is nonetheless **insufficient** to generate candidate (207c), the attested form. The reason for this is the following: First, the ranking OP-SPEC » FULL-INT » STAY was established independently of this tableau. Second, P-MEAS and LONE-WD have no bearing on this tableau, since all candidates fare equally well with regard to each of these constraints.²⁵⁰ Third, the LONE-WD having to do with *s+ACC* has not been proven to be ranked above or below any of the other constraints (in the standard register) except P-MEAS. Fourth, the only constraint which PR-CNTG has been proven to dominate definitively is STAY. Thus, in none of the possible orderings of LONE-WD and PR-CNTG with regard to the already established sequence of OP-SPEC » FULL-INT » STAY can the attested form in (207c)

For the purposes of this study *MCat* = *MrWd*. This constraint thus requires all morphological words to be prosodic words. This constraint is further refined in their manuscript, but this simple version is sufficient for these purposes.

²⁵⁰ Actually, only candidates (207b-d) fare equally with regard to LONE-WD. If ranked higher than PR-CNTG it predicts—incorrectly—candidate (207a). If below PR-CNTG, then this constraint is moot.

be generated correctly. In any of the ranking possibilities the optimal candidate is incorrectly predicted to be either (207a) or (207d), not the attested form (207c).

In order to fix tableau (207) I rely on the following idea, suggested in DePerno (1990; 1991:ch. 9): Pleonastic nouns like *štuk(i)* are used only with countable items. I extend her suggestion as follows: Pleonastic count nouns cannot be used when the quantity is not a countable one. Such nouns do not appear, for example, when the numeral is a fraction. I assume that this restriction is in Gen, but if not, then a constraint could be fashioned easily enough to handle this. I revise tableau (207) as follows. Note that candidates with pleonastic count nouns are not an option, hence candidates (207d-e) are not repeated in tableau (211):

(211) Complement consisting of *pol* ‘half’ and noun (**final**) [revision of (207)]

‘about half an hour’	P- MEAS	PR- CNTG	LONE -WD	OP- SPEC	FULL -INT	STAY
a. [čaSA _i [s] _P [t _i pol [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*!				**
b. [[s] _P [čaSA _i pol [[t _i] _{N°}] _{N'}] _{NP}] _{PP}		*!*	*			*
c. \$ [[s] _P [pol [[čaSA] _{N°}] _{N'}] _{NP}] _{PP}			*	*		

Comparing candidates (211a, c)—proves that PR-CNTG dominates both OP-SPEC and LONE-WD; if not, then (211a) would be generated incorrectly. Because no choices remain after the PR-CNTG column, all columns to the left of it have been shaded. This tableau does not allow a relative ranking between LONE-WD and OP-SPEC; tableau (182), however, has shown that LONE-WD » FULL-INT.

This concludes the tableaux. I have determined the following ranking:

(212) **Final ranking for (standard) Russian:** [Revision of (197c) and (197)]
 {P-MEAS » LONE-WD} › {PR-CNTG » {OP-SPEC › LONE-WD} » FULL-INT » STAY}

The underlined LONE-WD is the one having to do with approximative inversion. This ranking generates all of the data presented in the standard-Russian tableaux above. The only noticeable difference between the standard and colloquial registers is that PR-CNTG is ranked below STAY (cf. (203) above).

In this section I have shown that the mechanics of *s*+ACC can be accounted for using only two more constraints. One formalizes the need to limit the size of *s*+ACC's complement to a single syntactic word; the other allows for additional words in the complement if they contribute to the semantics of measure. Curiously, these two constraints are crucial in standard Russian only to *s*+ACC complements in which there are no numerals. I also investigated a slightly different set of data in the colloquial register and determined that the two registers differ in the ranking of one constraint relative to the others. The colloquial data does show that a separate single-word constraint is needed in numerical complements of *s*. In the next and final section I briefly speculate about the universal viability of the constraints I've proposed.

6.5 The universal viability of the constraints proposed above

In addition to Grimshaw's constraints OP-SPEC, FULL-INT and STAY, I have proposed four of new ones: LONE-WD, P-MEAS, LONE-WD, and PR-CNTG. In this section I assess briefly the universal viability of each:

Single-word phenomena are relatively common across human language. They are often accounted for by linguists using different mechanisms, such as non-branching. Surely there is a constraint which allows a language to impose a stricter size restriction on particular lexical items or constructions. I believe that LONE-WD from approximative inversion and LONE-WD from *s*+ACC can be derived from such a universal constraint, although I do not do so here. I merely point out that if the two

are the same constraint then this is fully consistent with the consolidated rankings as shown in (213):

(213) If LONE-WD and LONE-WD are fused into a single constraint:

P-MEAS › { PR-CNTG › { OP-SPEC › **LONE-WD** } › FULL-INT › STAY }
(with the proviso that P-MEAS › **LONE-WD**)

[Poss. simplification of (212)]

I do not try to re-write the two constraints, just mention that both approximative inversion and *s*+ACC are subject to the same size limitation.

P-MEAS resembles other constraints in the Optimality literature which often seem to allow for a specialized type of exception. For example, Prince & Smolensky (1993), in their treatment of the phonology of the Australian language Lardil, propose the following constraint:

(214) FREE-V: Word-final vowels must not be parsed (in the nominative).

[= ex. 152 in Prince & Smolensky (1993:101)]

They justify FREE-V as follows:

“Although FREE-V takes the bull by the horns, it would not perhaps be put forth as the canonical example of a universal markedness principle. [...] Any theory must allow latitude for incursions of the idiosyncratic into the grammar. What is important for our program is that such incursions are best expressible as *constraints*; that they are (slightly) modified versions of the universal conditions on phonological form out of which core grammar is constructed; and that they interact with other constraints in a manner prescribed by the general theory.” [*Ibid.*; underlining added]

It would seem that my P-MEAS fails by the last (underlined) criterion: I have not shown that P-MEAS, too, is dominated by other constraints.

This leaves PR-CNTG: Much of the Optimality theory dealing with prosodic morphology also deals with contiguity. Grimshaw’s Optimality-syntax work also restricts movement into a particular type of constituent:

(215) PROJ-PRIN: No adjunction to subordinate clauses; and no movement to the head of a subordinate clause [Grimshaw (1995:1)]

Nevertheless, writing a constraint in terms of restricting against what might or might not result is a problematic exercise. Perhaps PR-CNTG should be worded in terms of the number of embedded PrWds which result from each of the candidates.

In this brief section I have speculated about the viability of the constraints I have proposed, both within the theory and empirically. While some appear ad hoc or even tenuous, I nonetheless find them necessary in the hierarchy I propose.

In this chapter I have applied Optimality Theory to *s+ACC* and other related constructions in Russian. I began by summarizing the data that needed to be explained (§6.1), followed by a brief introduction to Optimality Theory, specifically as applied to syntax (§6.2). I then proposed a set of constraints to account for approximative inversion (§6.3), followed by a model of *s+ACC* itself (§6.4). In the final section (§6.5), I considered the viability of the constraints I had proposed in the preceding two sections. In all, I have devised a model that accounts for and even explains the seemingly fickle behavior of approximative inversion and *s+ACC*. I do not consider this to be the definitive solution, just a viable approach to the problem. At the very least, this framework makes a strong argument for the **type** of approach that is needed, one that looks at factors in various grammar components and deals with them as re-rankable, violable constraints.

Conclusion:

I conclude this dissertation by summarizing the points made above:

First, a comprehensive description is provided of the diachronic change that has caused *s*+ACC to become significantly restricted in distribution. Whereas *s* could at one point take an overt complement consisting of a numeral and noun, as a result of this change such a construction is no longer possible, requiring numerical complements to undergo approximative inversion.

Next, in chapter 2, I correct several mistaken characterizations in the literature about a construction deceptively similar to *s*+ACC, which in fact assigns GEN case. This similarity has eluded more than one author in the past, including Isačenko, Stang, and Ušakov.

In chapter 3 I then analyze several properties which *s*+ACC has in common with other ACC-assigning quantificational prepositions. I show that two of *s*+ACC's properties—no pluralized (non-numerical) complements and no animate-ACC paucal complements—are really properties shared with prepositional quantifiers in general.

Then, in chapter 4, the primary property that distinguishes *s*+ACC from other prepositions—a single-word restriction on the complement of *s*—is investigated. I assess several types of apparent multi-word complements, as well as look at a few other single-word constructions in the language. I specify that the restriction is against complements consisting of more than one **syntactic** word.

In chapter 5 other constructions which also express indefinite quantity are investigated: the GEN-assigning preposition *okolo*, which, in one of its uses, means 'approximately', and approximative inversion, a phenomenon unique to East Slavic, in which a noun and quantifier are juxtaposed to achieve an added meaning of approximation. By comparing *s*+ACC and *okolo* I have elucidated the exact phrase structure of *s*+ACC. My investigation of approximative inversion makes a number of

discoveries, including a better understanding of Russian quantifier constructions overall. I show that there is also a single-word constraint pertaining to approximative inversion, which specifically requires a single **prosodic** word. I argue that approximative inversion is required with *s+ACC* because the semantic component of approximative inversion is a proper subset of the semantics of *s+ACC*. Two other words with peculiar syntactic properties, *ètak* ‘about’ and *neskol’ko* ‘several’ are also investigated in this chapter.

In the final chapter I construct a model in the framework of Optimality Theory to account for the data in the preceding chapters. This theory proposes a hierarchy of violable output constraints which generates the grammatical data. I argue that approximative inversion is the movement of the quantified noun to the specifier of the noun phrase or the specifier of the prepositional phrase. Movement to specifier position is required because the quantified noun is an operator. When movement is not allowed, then a pleonastic count noun is inserted in that specifier position. I also propose constraints which allow specific violations of the single-word restriction and which select the proper specifier-position landing site. I show that *s+ACC* and approximative inversion are closely intertwined phenomena.

In all, this dissertation improves overall understanding of Russian quantificational constructions, especially those facets having to do with approximate measure. This study also elucidates the interaction of various grammar components—syntax, semantics, morphology, and prosody—in a closely related set of constructions.

References:

- Akiner, Shirin (1983) "The syntax of the numeral in Byelorussian, compared with Ukrainian, Russian and Polish." *Slavonic and East European review* 61:1. 55-68.
- Aksenov, Vasilij P. (1965) *Pora, moj drug, pora*. [Moskva]: Molodaja gvardija.
- Aleksandrov", A. [A. Aleksandrow] (1923) *Polnyj russko-anglijskij slovar'/Complete Russian-English dictionary*. Berlin: Glücksmann.
- Arašonkava, H.U. & V.P. Lemcjuhova (1994) *Karotki slovník belaruskaj movy*. Minsk: Mastackaja literatura.
- Arpolenko, H.P; K.H. Horodens'ka & H.X. Ščerbatjuk (1980) *Čislivnyk ukajins'koji movy*. Kyjiv: Naukova dumka.
- Avrutin, Sergey (1992) "Movement of bound variables." In *Third Leiden conference for junior linguists*. 1-16.
- Avvakum, Protopop [1934] *Žitie protopopa Avvakuma, im samym napisannoe i drugie ego sočinenija*. N.K. Gudzija (ed.). [Moskva]: Academia [Publisher's name in Latin letters].
- Avvakum, Protopop (1960) *Žitie protopopa Avvakuma*. Moskva: Gosudarstvennoe izdatel'stvo xudožestvennoj literatury.
- Babby, Leonard H. (1980) *Existential sentences and negation in Russian*. Ann Arbor: Karoma. (= *Linguistica extranea, Studia* 8.)
- Babby, Leonard H. (1985) "Prepositional quantifiers and the Direct Case Condition in Russian." In *Issues in Russian morphosyntax*. M.S. Flier & R.D. Brecht (eds.). Columbus, Ohio: Slavica. 91-117.
- Babby, Leonard H. (1987) "Case, prequantifiers, and discontinuous agreement in Russian." *Natural language and linguistic theory* 5:1, 91-138.
- Babby, Leonard H. (1989) „Subjectlessness, external subcategorization, and the Projection Principle." *Zbornik Matice srpske za filologiju i lingvistiku* 9, 1-55.
- Babby, Leonard H. (1991) "Noncanonical configurational case assignment strategies." In *Non-canonical case-assignment*. A. Toribio & W. Harbert (eds.) (= *Cornell working papers in linguistics* 9.) 1-55.
- Babov, Kiril G. (1968) *Ruskite predlozi. Nauka i izkustvo*. Sofija: G. Dimitrov.
- [Barannyk, D.X.] (1969) „Čislivnyk." In *Sučasna ukrajins'ka literaturna mova* 2. Kyjiv: Naukova dumka.
- Bardin, A.V. (1940) *Fol'klor Čkalovskoj oblasti*. Čkalov: OGIZ, Čkalovskoe izdatel'stvo.

- Barsov, A.A. (1980) *The comprehensive grammar of A.A. Barsov / Obstojaťel'naja grammatika A.A. Barsova*. Lawrence W. Newman (critical ed.) Columbus: Slavica.
- Baš, Eva G.; N.K. Venediktova; M.V. Vsevolodova; I.V. Tolstoj & L.N. (1959) *Bespredložnoe i predložnoe upravlenie. Sbornik upražniniť dlja inostratev*. [Moskva]: Izdatel'stvo moskovskogo universiteta.
- Belinskij, V.G. (1948a) *Sobranie sočinenij v trex tomax 1*. Moskva: GIXL.
- Belinskij, V.G. (1948b) *Sobranie sočinenij v trex tomax 2*. Moskva: GIXL.
- Bernard, Roger. (1954) "Alcune osservazioni sul plurale secondo in bulgaro moderno." *Richerche slavistiche* 3, 30-42.
- Billings, Loren A. (1993a) "Prosodic interactions with syntax and other grammar components: The s + accusative construction in Russian." Talk presented at Harvard University.
- Billings, Loren A. (1993b) "On interactions between phonology, morphology and syntax: The S + accusative construction in modern Russian." Abstract printed and talk presented at the Slavic-phonology panel, AATSEEL [American Association of Teachers of Slavic and East European Languages] conference, Toronto.
- Billings, Loren A. (1994a) "Prosodic limitations in marked syntactic constructions." Talk presented at the pre-session on Slavic linguistics, Georgetown University Round Table on Linguistics (GURT), Washington, D.C.
- Billings, Loren A. (1994b) "An Optimality Approach to the Syntax and Prosody of Discourse Particles." Abstract printed and talk presented at the Slavic-phonology panel, AATSEEL [American Association of Teachers of Slavic and East European Languages] conference, San Diego.
- Billings, Loren A. (1995a) "Syntactic and prosodic constraints on pronoun suppletion in Russian." Abstract printed and talk presented at the historical-linguistics panel, Linguistic Society of America, New Orleans.
- Billings, Loren A. (1995b) "Sandhi phenomena and language change." Abstract printed and talk presented at the Conference on Interfaces in Phonology, Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Berlin. [To appear in a proceedings volume in the *Studia Grammatica* series from Akademie Verlag, Berlin.]
- Billings, Loren A. (1995c) "The interaction of scope, prosody, and sisterhood in Slavic negated prepositional phrases." Abstract printed and talk presented at FASL [Formal Approaches to Slavic Linguistics] 4 conference, Ithaca, NY. [To appear in a proceedings volume.]
- Billings, Loren A. & Joan Maling (1995) "Accusative-assigning participial (-no/-to) constructions in Ukrainian, Polish and neighboring languages: An annotated bibliography." To appear in *Journal of Slavic linguistics* 3.

- Billings, Loren A. & Catherine Rudin (1994) "Optimality and Superiority: A new approach to overt multiple-*wh* ordering." Abstract printed and talk presented at FASL [Formal Approaches to Slavic Linguistics] 3. [To appear in proceedings volume, Jindřich Toman (ed.). Ann Arbor, MI: Michigan Slavic Publications.]
- Bitextina, G.A. & N.N. Luckaja (1960) *Predložnoe i bezpredložnoe upravlenie v russkom jazyke. Praktičeskoe posobie dlja studentov-bolgar.* Moskva: Vysšaja škola.
- Blažev, B.I. (1962) „Za sintaktičnata struktura na ščetanijata ot tipa *poslednie polveka* v ruski ezik.” *Ezik i literatura* 17:3, 64-68.
- Blažev, B.I. (1966) „Ob odnoj osobennosti kategorii oduševlennosti-neoduševlennosti u čislitel'nyx *odin, dva, tri i četyre*.” *Russkij jazyk v škole* 27:2, 103-105.
- Borras, F.M. & R.F. Christian (1971) *Russian syntax: Aspects of modern Russian syntax and vocabulary.* 2nd ed. Oxford: Clarendon Press.
- Borsley, R.D. & Ewa Jaworska (1988) "A note on prepositions and Case marking in Polish." *Linguistic inquiry.* 19:4, 685-691.
- Bukatevič, N.I. (1957) *Opyt istoričeskogo izučenija predlogov i predložnyx sočetanij v russkom literaturnom jazyke* 1. Odessa: Odesskij gosudarstvennyj universitet.
- Bukatevič, N.I. (1958) *Opyt istoričeskogo izučenija predlogov i predložnyx sočetanij v russkom literaturnom jazyke* 2. Odessa: Odesskij gosudarstvennyj universitet.
- Bulaxovskij, Leonid A. (1958) *Istoričeskij kommentarij k russkomu literaturnomu jazyku.* Kiev: Radjans'ka škola.
- Butorin, D.I. (1968) „Sočetanie suščestvitel'nyx s čislitel'nym *poltora*.” *Russkaja reč* (no.) 6.
- Carlton, T[erence] R. (1972) *The numeral in Ukrainian: Its forms and uses.* Edmonton: Department of Slavic Languages, University of Alberta.
- Česnokova, L.D. (1981) „Kategorija količestva i sintaksičeskie struktury.” *Voprosy jazykoznanija* (no.) 2, 45-52.
- Čexov, Anton P. (1935) *Višnevyy sad.* (= Izbrannye proizvedenija 3. 403-431.) Moskva: Izdatel'stvo detskoj literatury.
- Chey, Youn-Cha Shin (1967) *Russian quantifiers.* Unpublished Yale University Ph.D. dissertation.
- Chvany, Catherine (1975) *BE-sentences in Russian.* Cambridge, MA: Slavica.
- Comrie, Bernard (1978) "Genitive-accusatives in Slavic: The rules and their motivation." *International review of Slavic linguistics* 3, 27-42.
- Comrie, Bernard & Gerald Stone (1978) *The Russian language since the Revolution.* Oxford: Clarendon Press.

- Corbett, Greville G. (1978a) "Numerous squishes and squishy numerals in Slavonic." *International review of Slavic linguistics* 3, 43-73.
- Corbett, Greville G. (1978b) *Predicate agreement in Russian*. (= Birmingham Slavonic studies monographs 7.) Birmingham: University of Birmingham.
- Corbett, Greville G. (1980) "Neutral agreement." *Quinquereme. New studies in modern languages* 3:2, 164-170.
- Corbett, Greville G. (1986) "The use of the genitive or accusative for the direct object of negated verbs in Russian: A bibliography." In *Case in Slavic*. R.D. Brecht & J.S. Levine (eds.). Columbus, OH: Slavica. 361-372.
- Corbett, Greville G. (1988) "Agreement: A partial specification, based on Slavonic data." In *Agreement in natural language: Approaches, theories, descriptions*. M. Barlow & C. Ferguson (eds.). Stanford: Center for the Study of Language and Information. 23-53.
- Corbett, Greville G. (1991) *Gender*. Cambridge: Cambridge University Press.
- Crockett, Dina B. (1976) *Agreement in contemporary standard Russian*. Cambridge, MA: Slavica.
- Čukovskij, K[ornej] (1914) „Ègo-futuristy i kubo-futuristy.” *Literaturno - xudožestvennye al'manaxi izdatel'stva „Šipovnik”* (kniga) 22, 95-135.
- Dal', Vladimir" (1989a) *Tolkovyj slovar' živago velikoruskago jazyka 1 (A-Z)*. Moskva: Russkij jazyk. (Identical to 1955 ed., which is based on 1880 (2nd.) ed., S-P & M: Vol'f".)
- Dal', Vladimir" (1989b) *Tolkovyj slovar' živago velikoruskago jazyka 2 (I-O)*. Moskva: Russkij jazyk. (Identical to 1955 ed., which is based on 1881 (2nd.) ed., S-P & M: Vol'f".)
- Dal', Vladimir" (1990) *Tolkovyj slovar' živago velikoruskago jazyka 3 (P)*. Moskva: Russkij jazyk. (Identical to 1955 ed., which is based on 1882 (2nd.) ed., S-P & M: Vol'f".)
- Dal', Vladimir" (1991) *Tolkovyj slovar' živago velikoruskago jazyka 4 (R-Ÿ)*. Moskva: Russkij jazyk. (Identical to 1955 ed., which is based on 1882 (2nd.) ed., S-P & M: Vol'f".)
- Demidova, G.I. (1976) „Predložnye konstrukcii, vyražajuščie sravnitel'no-sopostavitel'nye otnošenija v dialekte.” In *Voprosy grammatičeskogo stroja i slovoobrazovanija v russkix narodnyx govorax*. A.S. Gerd et al. (eds.). Petrozavodsk: Petrozavodskij gosudarstvennyj universitet. 92-104.
- DePerno, Laurie L. [Laurie S. Langlois] (1990) "Numeral classifiers in Russian." Handout and text of talk presented at the American Association of Teachers of Slavic and East European Languages conference, Chicago.

- DePerno, Laurie L. [Laurie S. Langlois] (1991) *Quantifiers and quantifier-modifiers in Russian: A morphosyntactic analysis*. Draft of Cornell University Ph.D. dissertation in progress.
- Di Sciullo, A.M. and Edwin Williams (1987) *On the definition of word*. Cambridge, MA: MIT Press.
- Drovnikova, Larisa N. (1985) *Istorija čislitel'nyx v russkom jazyke*. Vladivostok: Izdatel'stvo Dal'nevostočnogo universiteta.
- Elenskij, Jordan (1977) „O kategorii oduševlennosti v russkom jazyke.” *Bolgarskaja rusistika*. (no.) 6, 41-45.
- Elenskij, Jordan (1978) „Oduševlennost' pri čislitel'nyx v petrovskuju èpoxu.” *Bolgarskaja rusistika*. no. 4, 57-67.
- Ermolinskaja lëtopis'*. (1910) (= Polnoe sobranie russkix" lëtopisej (tom") 23.) S.-Peterburg": Tipografija M.A. Aleksandrova.
- Ferrell, James (1958) “The systems of the third person pronoun with particular reference to the category of the preposition in Slavic languages.” In *American contribution to the Fourth International congress of Slavists. Moscow, September 1958*. (= Slavistic printings and reprintings 21.) 's-Gravenhage: Mouton. 75-101.
- Fowler, George H. (1988) *The syntax of the genitive case in Russian*. Revised version of 1987 unpublished University of Chicago Ph.D. dissertation.
- Franks, Steven L. (1994) “Parametric properties of numeral phrases in Slavic.” *Natural language and linguistic theory* 12:4, 597-674.
- Franks, Steven L. (1995) *Parameters of Slavic morphosyntax*. (= Oxford studies in comparative syntax 8.) New York: Oxford University Press.
- Franks, Steven L. & Richard House (1982) “Genitive themes in Russian.” In *Papers from the Eighteenth regional meeting*. Chicago: Chicago Linguistic Society.
- Fryšćák, Milan (1969) *The morphology of Slavic numerals*. Unpublished Ohio State University Ph.D. dissertation.
- Gercen, A.I. (1919) *Doktor" Krupov". Povëst'*. (= Polnoe sobranie sočinenij i pisem" (tom) 5.) Petrograd": Literaturno-izdatel'skij otdel" Narodnago komissariata po prosvëšćeniju.
- Gercen, A.I. (1955) *Doktor Krupov. Povest'*. (= Sočinenija (tom) 1.) Moskva: Gosudarstvennoe izdatel'stvo xudožestvennoj literatury.
- Gladney, Frank Y. (1986) “Prepositional case government in Russian.” In *Case in Slavic*. R.D. Brecht & J.S. Levine (eds.). Columbus, OH: Slavica. 130-151.
- Gončarov, I.A. (1965) *Obyknovennaja istorija. Roman v dvux častjax*. Frunze: Mektep.

- Grannes, Alf (1984) "Impersonal animacy in 18th century Russian." *Russian Linguistics* 8. 295-311.
- Grannes, Alf (1986) "*Rodi mne tri syna*: Animacy in Russian numerals—norm and usage." In *Festschrift für Wolfgang Geseman*. (= Beiträge zur slawischen Sprachwissenschaft und Kulturgeschichte 3; = Slavische [sic.] Sprachen und Literaturen 8.) München: Hieronymus. 107-117.
- Grimshaw, Jane (1993) "Minimal projection, heads and optimality." Handout to talk given at MIT. [A similar work is available as a technical report from the Rutgers University Center for Cognitive Science, Piscataway, NJ]
- Grimshaw, Jane (1995) "Projection, heads and optimality." Manuscript, Rutgers University. To appear in *Linguistic inquiry*.
- Grosberg, John (1957) "The numeral in idiomatic Russian." *Études Slaves et Est-Européennes/Slavic and East-European studies* 2:3, 177-79.
- Henriksen, John (1993) "The combinability of 'indeterminate numerals' with the nouns *chelovek* [sic.] and *ljudi*." Handout to presentation given in Slavic 269 *Grammar for instructors*, Harvard University.
- Hockett, Charles F. (1958) *A course in modern linguistics*. New York: Macmillan.
- House, Richard C. (1982) *The use of genitive initial sentences for the specification of quantity in Russian*. Unpublished Cornell University Ph.D. dissertation
- Hurski, N.I.; M.H. Bulaxaw & M.C. Marčanka (1955) *Leksikalohija, fanetyka i marfalohija*. (= Belaruskaja mova (časťka) 1.) Minsk: Džjaržawnae vučebna-pedahahičnae vydavectva Ministertva asvety BSSR.
- Ickovič, V.A. (1971) „Sovremennye abbreviatury.” *Russkaja reč'* (no.) 2. 74-79.
- Ickovič, V.A. (1980) „Suščestvitel'nye oduševlennye i neoduševlennye v sovremennom russkom jazyke (norma i tendencija).” *Voprocjazykoznanija* (no.) 4. 84-96.
- Il'f, Il'ja & Evgenij Petrov (1961) *Dvenadcat' stul'ev*. (= Sobrańie sočinenij v pjati tomach (tom) 1. 25-382.) Moskva: Gosudarstvennoe izdatel'stvo xudožnoj literatury.
- Isačenko, A.V. (1962) *Formenlehre*. (= Die russische Sprache der Gegenwart (Teil) 1.) Halle (Saale): VEB Max Niemeyer Verlag.
- Isengalieva, V.A. (1959) *Russkie predlogi i ix èkvivalenty v kazaxskom jazyke*. Alma-Ata: Izdatel'stvo Akademii nauk Kazaxskoj SSR.
- Ivanov, Valerij V. (1990) *Istoričeskaja grammatika russkogo jazyka*. 3rd ed. Moskva: Prosveščenie.
- Ivanova, V.F. (1969) „Čelovek s tysjač'ju lic ili čelovek s tysjač'ju licam?” *Russkij jazyk v škole* (no.) 2, 66-73.

- Ivšić, Stjepan (1950) „O značenju prijedloga s' s akuzativom u slavenskim jezicima.” In *Franu Ramošu za šestdesetletnico*. (= *Slavistična revija*. 3:3.) Ljubljana. 360-368 (incl. French résumé).
- Jakobsmeier, Brian (1994) “Infant formula: One-fisted prose for small-minded writers. *Spy*. January, p. 21.
- Jakobson, Roman O. (1921) „Vliv revoluce na ruský jazyk. Poznámky ke [Mazon (1920)].” *Nové Atheneum* (roč.) **2** [Praha: Tiskem Edvarda Leschingra v Praze]. 3-32. [MIT Archives (Jakobson Papers): MC 72 Box 7 Folder 21 (offprint)]
- Jakobson, Roman O. (1958) „Morfologičeskie nabljudenija nad slavjanskim sklonením (sostav russkix padežnyx form).” In *American contributions to the Fourth International Congress of Slavists. Moscow, September 1958*. (= Slavistic printings and reprintings **21**.) 's-Gravenhage: Mouton. 127-156.
- Janowski, F.[M.] (1961) *Pytanni kul'tury movy*. Minsk: Vydavectva Ministërstva vyšějšaj, sjarědnjaj specyjal'naj i prafesijanal'naj adukacyi BSSR.
- Jaworska, Ewa (1986) “Prepositional phrases as subjects and objects.” *Journal of linguistics*. 22, 355-374.
- Kačevskaja, G.A (1969) „K istorii složnyx suščestvitel'nyx s *pol-* i *-polu-* v russkom jazyke.” In *Voprosy teorii i istorii jazyka. Sbornik statej, posvjaščennyj pamjati B.A. Larina*. 322-29.
- Kapanadze, I.L. (1991) *Učebno-metodičeskij slovar'-spravočnik russkix predlogov i sojuzov i ix gruzinskix èkvivalentov*. Tbilisi: Izdatel'stvo Tbilisskogo universiteta.
- Karcevskîj, S.Î. (1923) *Jazyk", vojna i rvoljucîja*. Berlin": Russkoe universal'noe izdatel'stvo.
- Koka, A.A. (1955) “Konstrukcii s vremennym značením v sovremennom russkom jazyke.” In *Voprosy izučenija russkogo jazyka. Sbornik lingvističeskix statej*. X.X. Maxmudov (ed.). Alma-Ata: Izdatel'stvo Akademii nauk Kazaxskoj SSR. 67-111.
- Koneski, Blaže. (1954) *Gramatika na makedonskiot literaturnen jazyk (del) 2*. Skopje: Prosvetno delo.
- Korolenko, V.G. (1914) *Bez" jazyka*. (= Polnoe sobranîe sočinenîj V.G. Korolenko (tom") **4**.) S.-Peterburg": Peterburg": Izdanîe T-va A.F. Marks" [Priloženîe k" žurnalu „Niva” na 1914 g.].
- Kotošixin, Gr. (1906) *O Rossii v carstvovanii Alekseja Mixajloviča*. 4th ed. S.-Peterburg": Tipografîja Glavnago upravlenija udělov". [1969 photoreprint (= Slavistic printings and reprintings **126**.) The Hague & Paris: Mouton, in cooperation with Vaduz: Europe Printing.]
- Krylov, I.A. (1946) *Sočinenija 3*. Moskva: OGIZ.

- [Lapaw, B.S.] (1962) „Ličebnik.” In *Marfalahija*. K.K. Atraxovič & M.H. Bulaxaw (eds.). (= Hramatyka belaruskaj movy (tom) 1.) Minsk: Vydavectva Akadèmii navuk BSSR. 254-70.
- Levin, Maurice I. (1978) *Russian declension and conjugation: A structural description with exercises*. Columbus, OH: Slavica.
- Lomonosov, M.V. (1755) *Rosijskaja grammatika*. Sanktpeterburg: Imperatorskaja akademija nauk. (1975 photoreprint, Leipzig: Zentralantiquariat der Deutschen demokratischen Republik.)
- Lomonosov, M.V. (1952) *Rosijskaja grammatika*. (= Polnoe sobranie sočinenij 7.) Izdatel'stvo AN SSSR.
- Lomtev, T.P. (1956) *Očerki po istoričeskomu sintaksisu russkogo jazyka*. Moskva: Izdatel'stvo Moskovskogo universiteta.
- Lubensky, Sophia (1995) *Random House Russian-English dictionary of idioms*. New York: Random House.
- Macdonald, R. Ross (1972) “Prepositions of time in Russian.” *Languages and linguistics working papers* (no.) 6 [Washington, DC: Georgetown University Press], 65-74.
- Maksimov, V.I (1973) „Polovina pervogo, polpervogo?” *Russkaja reč'* (no.) 6, 51-53.
- Matveeva, G.I. (1954) *Čislitel'nye v russkix govorax* Dissertacija na soiskanie učenoj stepeni kandidata filologičeskix nauk. Moskva.
- Matvijas, I.H. (1984) „Slovospolučennja typu vydno xatu, čuty pisnju.” Map 267 of *Polissja, serednja naddnprjanščina i sumižni zemli*. (= Atlas ukrajins'koji movy v tr'ox tomax 1.) Kyjiv: Naukova dumka. (commentary) 55, (map) 380.
- Mayer, Gerald L. (1967) *A comparative study of the syntax of the cardinal numeral in the Slavic languages*. Unpublished University of Pennsylvania Ph.D. dissertation
- Mayer, Gerald L. (1971) “Ukrainian dva (dvi), try, čotyry + noun: A reinterpretation.” *Canadian Slavonic papers* 13:1, 73-79.
- Mayer, Gerald L. (1973) “Common tendencies in the development of ‘two,’ ‘three,’ and ‘four,’ in Slavic.” *Slavic and East European journal* 17:3, 308-14.
- Mayer, Gerald L. (1974) “The morphological development of compound numerals in East Slavic.” *Canadian Slavonic papers* 16:1, 62-72.
- Mayer, Gerald L. (1976) “Towards a new categorization of Russian numerals.” *Russian language journal* 30:105, 25-30.
- Mayer, Gerald L. (1978) “The use of Russian numerals in oblique cases.” *Canadian Slavonic papers* 20:2, 208-17.

- Mazon, André (1920) *Lexique de la guerre et de la révolution en Russie (1914-1918)*. (= Bibliothèque de l'Institut français de Petrograd (tome) 6.) Paris: Librairie ancienne honoré Champion.
- McCarthy, John & Alan Prince (1993a) *Prosodic morphology I: Constraint interaction and satisfaction*. Manuscript. [Available as a technical report from the Rutgers University Center for Cognitive Science, Piscataway, NJ]
- McCarthy, John & Alan Prince (1993b) "Generalized alignment." Manuscript. [Available as a technical report from the Rutgers University Center for Cognitive Science, Piscataway, NJ]
- Mel'čuk, Igor [sic.] A. (1980) "Animacy in Russian cardinal numerals and adjectives as an inflectional category." *Language* 56:4, 797-811.
- Mel'čuk, Igor [sic.] A. (1981) "The grammatical case of Russian numerical expressions of the type (*bol'she*) *na dva mal'čika* or *po troe bol'nyx*." Steven Franks (transl.) *Forum linguisticum* 6:2, 116-140.
- Mel'čuk, Igor [sic.] A. (1983) "The numeral *pol'* 'half' in modern Russian: An autonomous word or a bound morpheme." Paul Gorgen (transl.) *Journal of linguistic research* 2:1, 49-62.
- Mel'čuk, Igor' A. (1985) *Poverxnostnyj sintaksis russkix čislovyx vyraženiij* [The surface syntax of Russian numeral expressions]. (= *Wiener slawistischer Almanach* Sonderband 11.) Wien: Institut für Slawistik der Universität Wien.
- Mel'nikov", P.I. [Andrej Pečerskij] (1897a) *V" lěsax". Roman" v" četyrex" častjax"* (čast') 1. (= *Polnoe sobranie sočinenij P.I. Mel'nikova (Andreja Pečerskago)* 3.) S.-Peterburg" & Moskva: Vol'f".
- Mel'nikov", P.I. [Andrej Pečerskij] (1897b) *V" lěsax". Roman" v" četyrex" častjax"* (čast') 2. (= *Polnoe sobranie sočinenij P.I. Mel'nikova (Andreja Pečerskago)* 4.) S.-Peterburg" & Moskva: Vol'f".
- Nakhimovsky, A.D. & R.L. Leed (1980) *Advanced Russian*. Columbus: Slavica.
- Naylor, Keneth E. (1977) "A note on the 'quantification form' in Russian." *Folia slavica* 1:1. 89-95.
- Neidle, Carol (1988) *The role of case in Russian syntax*. (= *Studies in natural language and linguistic theory* 10.) Dordrecht, Boston, London: Kluwer Academic Press. [Revision of author's Ph.D. dissertation, Massachusetts Institute of Technology, 1982.]
- Nekrasov, Nikolaj A. (1959) *Komu na rusi žit' xorošo*. (= *Izbrannye proizvedeniija*. 391-627) Moskva: Moskovskij rabočij.
- Nekrasov, Nikolaj A. (1971) *Deduška Mazaj i zajcy*. (= *Stixotvorenija i poëmy*. 200-04.) Moskva: Xudožestvennaja literatura.

- Nikolaev, I.L. (1968) „O nekotoryx leksiko-grammatičeskix osobennostjax russkix predlogov.” *Sbornik rabot aspirantov Tadžikskogo gosudarstvennogo universiteta* 6 [Dušanbe]. 197-204.
- Nosov, Nikolaj N. (1987) *Priključenija Neznajki i ego družej. Roman-skazka dlja mladšego škol'nogo vozrasta*. Kiev: Veselka.
- Orfoèpičeskij slovar' russkogo jazyka. Proiznošenie, udarenie, grammatičeskie formy. Okolo 65 000 slov.* (1989) 5th ed. Moskva: Russkij jazyk.
- Orfografičeskij slovar' russkogo jazyka. 106 000 slov.* (1989) Moskva: Russkij jazyk.
- Ostrovskij, A.N. (1974) *Grex da beda na kogo ne živet.* (= Polnoe sobranie sočinenij v dvenadcati tomax 2. 391-627.) Moskva: Iskusstvo.
- Ožegov, S.I. (1983) *Slovar' russkogo jazyka*. 14th ed. Moskva: Russkij jazyk.
- Palašti, L. (1957) „Upotreblenie variantov predlogov s beglym o v russkom jazyke.” *Russkij jazyk v nacional'noj škole* (no.) 4.
- Parrott, Lillian A. (1992) “‘Agumentative že’ in discourse.” *Harvard studies in Slavic linguistics* 1, 83-104.
- Pasternak, Boris (1959) *Doktor Živago*. Ann Arbor: University of Michigan Press.
- Patton, Helen B. (1969) *A study of the agreement of the predicate with a quantitative subject in contemporary Russian*. Unpublished University of Pennsylvania Ph.D. dissertation.
- Perlmutter, David M. (1978) “Impersonal passives and the Unaccusative Hypothesis.” In *Proceedings of the annual meeting of the Berkeley Linguistics Society* 4. 157-189.
- Pesetsky, David M. (1982) *Paths and categories*. Unpublished Massachusetts Institute of Technology Ph.D. dissertation.
- Peškovskij, A.M. (1956) *Russkij sintaksis v naučnom osveščanii*. 7th ed. Moskva: Gosudarstvennoe učebnoe izdatel'stvo Ministerstva prosveščeniya RSFSR.
- Pete István (1984) “Semantic types of approximate quantitative relations in Russian and Hungarian.” In *Contrastive-studies: Russian-Hungarian*. Papp Ferenc (ed.). Budapest: Akad. Kiado. 69-85.
- Peter I, Emperor of Russia (1887) *Pis'ma i bumagi imperatora Petra Velikogo* 1. Sankt-Peterburg: Gos. tip.
- Peter I, Emperor of Russia (1893) *Pis'ma i bumagi imperatora Petra Velikogo* 3. Sankt-Peterburg: Gos. tip.
- Peter I, Emperor of Russia (1948) *Pis'ma i bumagi imperatora Petra Velikogo* 8:1. Moskva: Izdatel'stvo AN SSSR.

- Pircxalava, G.A. (1971) „K voprosu o sinonimii imen suščestvitel'nyx so značeniem množestva v russkom literaturnom jazyke XVIII v. (Na materiale M.V. Lomonosova).” *Naučnye doklady Vyššej školy. Filologičeskie nauki.* 63:3, 106-111.
- Popova, Zinaida D. (1969) *Sistema padežnyx i predložno-padežnyx form v russkom literaturnom jazyke XVII veka (strukturno-sematičeskoe opisanie).* Voronež: Izdatel'stvo Voronežskogo universiteta.
- Potapova, N.P. (1987) „Derivacionnye svjazi v prostranstvennoj i vremennoj semantike padežnyx i predložno-padežnyx form v dialektnoj reči v sravnenii s faktami drevnerusskogo jazyka.” *Derivacija i istorija jazyka: Mežvuzovskij sbornik naučnyx trudov.* Perm': Permskij gosudarstvennyj universitet. 79-83.
- Potebnja, A.A. (1941) *Glagol. Mestoimenie. Čislitel'noe. Predlog.* 1st ed. (= Iz zapisok po russkoj grammatike 4.) Moskva & Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Prince, Alan & Paul Smolensky (1993) *Optimality Theory: Constraint interaction in generative grammar.* Ms. (= Tech. rpt. #2, Rutgers University Center for Cognitive Science, Piscataway, NJ.) To appear from MIT Press.
- Puškin, Aleksandr S. (1942) *Vîstrěl [Vystrel].* Maro Beath Jones (transliterator and ed.). Claremont, CA: [Claremont College?].
- Radiščev, Aleksandr N. (1961) *Putešestvie iz Peterburga v Moskvu.* Leningrad: Gosudarstvennoe Izdatel'stvo xudožnoj literatury.
- Rappaport, Gilbert C. (1988) “On the relationship between prosodic and syntactic properties of pronouns in the Slavic languages.” In *American contributions to the tenth international congress of Slavists. Sofia, September 1988.* Alexander M. Schenker (ed.). Columbus: Slavica. 310-327.
- Rappaport, Gilbert C. (1992) “On the adnominal genitive and the structure of noun phrases in Russian and Polish.” In *Linguistique et slavistique. Melanges offerts a Paul Garde.* M. Guirard-Weber & C. Zaremba (eds.). (= Travaux publiés par l'Institut d'études slaves. 36.) Aix-en-Provence: Publications de l'Université de Provence & Paris: Institut d'études slaves. 241-262.
- Rěšetnikov", F.M. (1890) *Glumovy.* (= Sočinenîja F.M. Rěšetnikova v" dvux" tomax". 93-264) S.-Peterburg": Tipografîja gazety «Novosti».
- Rozental', Ditmar E. (1974) *Praktičeskaja stilistika russkogo jazyka.* 3rd ed. Moskva: Vyssšaja škola. [Previous editions, 1965 and 1968.]
- Rozental', Ditmar E. (1977) *Praktičeskaja stilistika russkogo jazyka.* 4th ed. Moskva: Vyssšaja škola.
- Rozental', Ditmar E. (1987) *Praktičeskaja stilistika russkogo jazyka.* 5th ed. Moskva: Vyssšaja škola.

- Rožkova, G.I. (1966) „Glavnye členy predloženiya.” In *Praktičeskoe posobie po sintaksisu russkogo jazyka*. 2nd ed. N.A. Lobanova (ed.) Moskva: Vysšaja škola.
- Rudin, Catherine (1988) “On multiple questions and multiple wh fronting.” *Natural language and linguistic theory* **6**:4, 445-501.
- Rybnikov, P.V. (1959) *Pesni, sobrannye P.N. Rybnikovym. Narodnye byliny, stariny i pobyval ščiny*. Moskva: Izdatel'stvo Moskovskogo universiteta.
- Sadovnikov, Dmitrij N. (1959) *Zagadki russkogo naroda. Sbornik, sostavlennyj D.N. Sadovnikovym*. [Moskva]: Izdatel'stvo Moskovskogo universiteta.
- Sajkiev, X.M. (1955) „Konstrukcii s vinitel'nyx padežom v sovremennom russkom jazyke.” In *Voprosy izučeniya russkogo jazyka. Sbornik lingvističeskix statej*. X.X. Maxmudov (ed.). Alma-Ata: Izdatel'stvo Akademii nauk Kazaxskoj SSR. 11-66.
- Šarolapova, Ju.A. (1956) *Punktuačija serediny XVIII veka (Po proizvedenijam M.V. Lomonosova)*. Avtoreferat dissertacii na soiskanie učenoi stepeni kandidata filologičeskix nauk. Kujbyšev: Kafedra russkogo jazyka, Kujbyševskij gosudarstvennyj pedagogičeskij universitet.
- Ščerbakov, Ju.I. (1969) *Upotreblenie sobiratel'nyx čislitel'nyx v sovremennom russkom jazyke*. Avtoreferat dissertacii na soiskanie učenoi stepeni kandidata filologičeskix nauk. Kujbyšev: Kafedra russkogo jazyka, Kujbyševskij gosudarstvennyj pedagogičeskij universitet.
- Schütz, Joseph (1986) “Der Archaismus *sorok sorokov cerkvej moskovskix*.” In *Festschrift für Herbert Bräuer zum 65. Geburtstag am 14. April 1986*. R. Olesch & H. Rothe (eds.). (= Slavistische [sic.] Forschungen **53**.) Köln & Wien: Böhlau Verlag. 505-514.
- Seliščev, A.M. (1928) *Jazyk revoljucionnoj èpoxi. Iz nabljudenij nad russkim jazykom poslednix let (1917-1926)*. 2nd ed. Moskva: Rabotnik prosvěščenija.
- Shevelov, George Y. (1963) *The syntax of modern literary Ukrainian*. (= Slavistic printings and reprintings **38**.) The Hague: Mouton.
- Sintaksis*. (1960) V.V. Vinogradov & E.S. Istrina (eds.). (= *Grammatika russkogo jazyka*. 2:1.) Moskva: Izdatel'stvo Akademii nauk SSSR.
- Sintaksis*. (1980) N.Ju. Švedova et al. (eds.). (= *Russkaja grammatika* **2**.) Moskva: Nauka.
- Širokova, Natal'ja A. (1963) *Iz istorii predložnyx konstrukcij, vyražajuščix sravnitel'no-sopostavitel'nye otnošenija*. Kazan': Izdatel'stvo Kazanskogo universiteta.
- Skoblikova, E.S. (1959) „Forma skazuemogo pri podležaščem, vyražennom količestvenno-imennym sočetaniem.” *Voprosy kul'tury reči*. (vyp.) **63**, 91-116.

- Slovar' sovremennogo russkogo literaturnogo jazyka. 13 (S-snjat'sja).* (1962) Moskva & Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Slovar' sovremennogo russkogo literaturnogo jazyka. 15 (T).* (1963) Moskva & Leningrad: Izdatel'stvo Akademii nauk SSSR.
- “Snošenija Carja Ioanna. IV Vasil'eviča s" Imperatorom" germanskim" Rudol'fom" II, s" napojem" Grigorijem" VIII i s" Veneciju, 1580-1582.” (1851) In *Pamjatniki diplomatičeskix" snošenij drevnej Rosii s" deržavami inostrannymi. Čast' pervaja: Snošenija s" gosudarstvami evropejskimi 1. Pamjatniki diplomatičeskix" snošenij s" Imperijeju rimskoju (s" 1488 po 1594 god)*. S. Peterburg": II otdělenie sobstvennoj E.I.V. Kancelarii. 765-906.
- Solženicyn, Aleksandr I. (1968) *Rakovyj korpus*. Pariž: YMCA Press.
- Solženicyn, Aleksandr I. (1969) *V krugě pervom*. New York: Harper and Row.
- Sorokin, Boris (1977) „K voprosu o trudnosti russkix čislitel'nyx.” *Russian language journal* 31:109. 51-60.
- Spojky, částice, zájmena a zájmenná adverbia.* (1980) Bohuslav Havránek et al. (eds.) (= Etymologický slovník slovanských jazyků. Slova gramatická a zájmena (svazek) 2.) Praha: Academia.
- Sreznevskij, Izmail" I. (1906) *Materialy dlja slovarja drevne-russkago jazyka po pis'mennym" pamjatnikam" 3.* Sanktpeterburg": Otd-ě. russkago jazyka i slovesnosti. Imperatorkaja akademija nauk".
- Stang, Chr[istian] S. (1956) “Slavonic s' with the accusative in expressions of measure.” In *For Roman Jakobson. Essays on the occasion of his sixtieth birthday. 11 October 1956.* Morris Halle et al. (eds.). The Hague: Mouton. 514-517.
- Staniševa, Dina S. (1966) *Vinitel'nyj padež v vostočno slavjanskix jazykax.* Sofija: Izdatel'stvo Bolgarskoj akademii nauk.
- Strelkov, P.G. (1950) „Čislitel'noe v russkom jazyke.” *Russkij jazyk v škole* (no.) 3. 27-35.
- Strugackij, Arkadij N. & Boris N. Strugackij (1975) *Polden', XXII vek.* (= Polden', XXII vek. Malyš. Dve povesti.) 2nd ed. Leningrad: Detskaja literatura.
- Strugackij, Arkadij N. & Boris N. Strugackij [Arkady Strugatsky & Boris Strugatsky] (1978) *Noon: 22nd century.* (= Macmillan's best of Soviet science fiction.) Patrick L. McGuire (transl.). New York: Macmillan.
- Suprun, A[dam E.] (1962) „Oboznačenie netočnyx (priblizitel'nyx) količestv pri pomošči opredelenno-količestvennyx čislitel'nyx.” In *Voprosy leksiki i grammatiki russkogo jazyka.* (= Učenyje zapiski Filologičeskogo fakul'teta (vyp.) 8.) Frunze: Kirgizskij gosudarstvennyj universitet. 5-14.

- Sussex, Roland (1976) "The numeral classifiers of Russian." *Russian linguistics* 3:2. 145-55.
- Suzdal'skaja letopis' po Akademičeskomu spisku.* (1927) (= Polnoe sobranie russkix letopisej (tom) I (vyp.) 2. 2nd ed. Akademija nauk SSSR.
- Švačko, S.A. (1981) *Jazykovye sredstva vyraženiya količestva v sovremennom anglijskom, ruskom i ukrainskom jazykax.* Kiev: Vyšča škola.
- Timberlake, Alan (1974a) "The nominative object in North Russian." In *Slavic transformational syntax.* R. Brecht, C. Čhvanj (eds.). (= Michigan Slavic materials 10.) Ann Arbor: Michigan Slavic Publications. 219-243. [Rutgers U.: ALEX PG13 .M46 no. 10]
- Timberlake, Alan (1974b) *The nominative object in Slavic, Baltic, and West Finnic.* (= Slavistische Beiträge 82.) München: Otto Sagner.
- Timberlake, Alan (1975) "Hierarchies in the genitive of negation." *Slavic and East European journal.* 19, 123-38.
- Tixonravov", Nikolaj S. (1863) *Pamjatniki otrečenoj ruskoj literatury* 2. Moskva: Universitetskaja tipografija (Katkov" i K^o). [(1973) photoreprint (with vol. 1). (= Variorum reprint S12.) London: Variorum Reprints..]
- Townsend, Charles E. (1990) *A description of spoken Prague Czech.* Columbus, OH: Slavica.
- Učebnyj slovar' sočetaemosti slov russkogo jazyka. Okolo 2500 slovarnyx statej.* (1978) P.N. Denisova & V.V. Morkovkina (eds.) Moskva: Russkij jazyk.
- Unbegaun, Boris O. (1935) *La flexion des noms.* (= La langue russe au XVI^e siècle (1500-1550) 1.) (= Bibliothèque de l'Institut Français de Léningrad (tome) 16.) Paris: H. Champion.
- Unbegaun, Boris O. (1960) *Russian grammar.* Oxford: Clarendon Press.
- Ušakov, D.N. (1935) *Tolkovyj slovar' russkogo jazyka* 1 (A—Kjuriny). Moskva: Sovetskaja ènciklopedija.
- Ušakov, D.N. (1938) *Tolkovyj slovar' russkogo jazyka* 2 (L—Ojalovet'). Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarěj.
- Ušakov, D.N. (1940) *Tolkovyj slovar' russkogo jazyka* 4 (S—Jaščurnyj). Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarěj.
- Vinogradov, S.I. (1979) „Para nožnic, para slov.” *Russkaja reč'* (no.) 2. 81-83.
- Vinogradov, Viktor V. (1947) *Russkij jazyk. Grammatičeskoe učenie o slove.* Moskva: Gosudarstvennoe učebno-pedagogičeskoe izdatl'stvo Ministerstva prosveščeniya RSFSR.

- Vinokurov, Evgenij M. (1964) „Korotko o sebe.” In his *Stixotvorenija*. Moskva: Xudožestvennaja literatura. 5-10.
- Vlaxov, Sergej I. & L.P. Muckov (1974) *Neglagol'naja leksiko-sintaksičeskaja sočetaemost'.* *Slovar'-spravočnik*. Sofija: Narodna prosveta.
- Vostokov", Aleksandr" (1831) *Russkaja grammatika Aleksandra Vostokova, po načertaniju ego že sokraščennoj Grammatiki polnče izložennaja*. Sanktpeterburg": I. Glazunov.
- Vostokov", Aleksandr" (1839) *Russkaja grammatika Aleksandra Vostokova, po načertaniju ego že sokraščennoj Grammatiki polnče izložennaja*. 4th ed. S. Peterburg": Tipograija Imperatorskoj rossijskoj akademii.
- Ward, Dennis (1965) *The Russian language today. System and anomaly*. London: Hutchinson University Library.
- Williams, Edwin (1994) *Thematic structure in syntax*. Cambridge, MA: MIT Press.
- Worth, Dean S. (1959) "Grammatical and lexical quantification in the syntax of the Russian numeral." *International journal of Slavic linguistics and poetics* 1-2. 117-132. (Reprinted in his *On the structure and history of Russian*. (1977) München: O. Sagner. 43-58.)
- Xajzer, L. (1976) „O sočetaemosti neopredelennyx čislitel'nyx s suščestvitel'nyx čelovek i ljudi.” *Russkij jazyk v škole* 4.
- Zaliznjak, A.A. (1967) *Ruskoe imennoe slovoizmenenie*. Moskva: Nauka.
- Zaliznjak, A.A. (1987) *Grammatičeskij slovar' russkogo jazyka. Slovoizmenenie. Okolo 100 000 slov*. 3rd ed. Moskva: Russkij jazyk.
- Zalucky, Henry K. (1991) *Compressed Russian. Russian-English dictionary of acronyms, semiacronyms and other abbreviations used in contemporary standard Russian*. Amsterdam etc.: Elsevier.
- Zolotova, G.A. (1988) *Sintaksičeskij slovar'.* *Repertuar èlementarnyx edinic russkogo sintaksisa*. Moskva: Nauka.