"*Nu*-drop" in Russian verbs: a corpus-based investigation of morphological variation and change

Tore Nesset and Anastasia Makarova, University of Tromsø, Norway

Abstract: In the present article we offer a corpus-based analysis of *nu*-drop in Russian verbs, the process whereby certain verbs with the suffix /nu/ omit this morpheme in past tense forms. We explore phonological, morphological and syntactic/semantic factors and show that inflectional and derivational morphology are most important for *nu*-drop. Our study of the inflectional and derivational morphological categories yields a polarized general picture; the categories display either close to 100% Ø-forms (i.e. forms without /nu/) or close to 0% such forms, while no categories are in the middle of the scale. Moreover, a diachronic survey of the development between the 19th and 21st centuries indicates increasing polarization, insofar as increasing percentages of Ø-forms are attested among forms with high percentages of Ø-forms, whereas decrease is characteristic of forms with low percentages of Ø-forms.

1. Overview

Many Russian verbs with the suffix /nu/ optionally leave out the suffix in past tense forms. For instance, *гаснуть* 'go out (about light)' displays vacillation between past tense forms like *гаснул* with the /nu/ suffix intact and forms like *гас*, which lack the suffix:[1]

(1) Верхний свет в ресторане не **гаснул**, и динамики воспроизводили сумасшедшее стаккато банджо Билла Хейли. [Валериан Скворцов. Каникулы вне закона (2001)]
'The ceiling light in the restaurant did not **go out**, and the loudspeakers played the crazy staccato of Bill Haley's banjo.'

(2) Он оставался один, доигрывал последнюю ноту в гордом одиночестве, и свет **гас**. [Сати Спивакова. Не всё (2002)]
'He stayed alone, played the last note in splendid isolation, and the lights **went out**.'

In this article, we present a corpus-based study of this phenomenon, which we refer to as "*nu*-drop". We address the following questions: what are the factors facilitating or inhibiting *nu*-drop, and what is their relative importance? These questions are discussed both from a synchronic and a diachronic perspective. We explore the following factors: phonology (the shape of the root), semantics (the meaning of the verb), inflectional morphology (paradigm cells) and derivational morphology (aspectual prefixation).[2] It is shown that all these factors have some impact on *nu*-drop, but they are not equally important. We propose that the best basis for predicting the distribution of Ø-forms and *nu*-forms is a morphological hierarchy that distinguishes between different inflected forms and prefixed/unprefixed verbs. This hierarchy enables us to distinguish between three groups: (a) categories where Ø-forms are virtually obligatory, (b) forms where Ø-forms dominate, but are not quite obligatory, and (c) forms where *nu*-forms dominate. Since there are no morphological categories with a roughly 50/50 distribution of Ø-forms and *nu*-forms, we demonstrate that the situation is polarized, and our diachronic investigation documents increasing polarization. In the beginning of the 21st century, the development has reached the point where Ø-forms are nearly obligatory for all finite verbs, except unprefixed masculines (and even for this category Ø-forms dominate strongly). The only morphological categories where *nu*-forms dominate in present-day Russian are gerunds and unprefixed active participles.[3]

Although *nu*-drop is well attested in major dictionaries and grammars of Contemporary Standard Russian (cf. e.g. Isačenko 1982; Ožegov and Švedova 2005; Švedova (ed.) 1980; Timberlake 2004 and Zaliznjak 1980) and discussed in a number of other scholarly works (Bulaxovskij 1950 and 1954; Černyšev 1915; Dickey 2001; Gorbačevič 1971 and 1978; Graudina et al. 1976, 2001 and 2007; Nesset 1998; Plungian 2000; Rozental' 1977; Vinogradov and Švedova (eds.) 1964), the only corpus-based investigation we are aware of is Graudina et al. (1976, 2001 and 2007). Their study is based on examples from a corpus of a total of 100,000 words culled from Soviet prose, newspapers, audio recordings of spontaneous speech and materials of a questionnaire, all from the 1960-70s. However, with the advent of large electronic corpora, it is possible to get a much more detailed picture of the situation. In order to shed new light on *nu*-drop we

excerpted all relevant examples from the Russian National Corpus, which contained more than 140 million words in November and December 2010, when the searches were carried out. The database was constructed as follows: corpus searches were performed for all verbs where *nu*-drop is possible according to Švedova (ed.) (1980) and Zaliznjak (1980) (see Nesset 1998, 129 for discussion; we included both forms with and without the *–sja* postfix). All examples with finite past tense forms as well as gerunds and past active participles were recorded and checked manually. In order to avoid skewed data due to multiple occurrences of a given variant in one author's work, we included only one example from each "document" in the Russian National Corpus. Although the corpus includes examples from the 18th century, these examples were removed from the database, since data from this period is sparse and therefore not suitable for statistical analysis. As a result, we ended up with a database of 34,026 examples representing the time span from 1800 to the beginning of the 21st century.

Table 1 provides an overview of the situation. The leftmost column lists all the verbs under scrutiny in the present study. We cite only unprefixed verbs, but the numbers include prefixations of these verbs as well. Verbs cited as starting with _ are only attested with prefixes in our database. We return to the effect of prefixation in section 5 below. Notice that we list verbs with the *–sja* postfix as separate entries in Table 1. The second column from the left gives the number of examples with the /nu/ suffix ("*nu*-forms"). In the column labeled "# Ø" we list the numbers of examples which lack the /nu/ suffix ("Ø-forms"), while the two rightmost columns provide the total number of examples for each verb and the percentage of examples without the suffix. The verbs are listed according to decreasing percentage of Ø-forms.

Verb	# nu	# Ø	# total	% Ø
_*меркнуться* 'get dark'	0	2	2	100
_*мозгнуть* 'freeze'	0	2	2	100
_*сохнуться* 'get dry'	0	239	239	100
_*хряснуть* 'get stuck'	0	11	11	100
_*жолкнуть* 'wither, yellow'	0	1	1	100
дохнуть 'die'	0	242	242	100
дряхнуть 'grow decrepit'	0	3	3	100
горкнуть 'become bitter'	0	14	14	100
обрыднуть 'make sick'	0	64	64	100
терпнуть 'become astrigent'	0	9	9	100
зябнуться 'feel chilly'	0	16	16	100
брякнуть 'swell'	1	155	156	99
брюзгнуть 'become a grumbler'	1	151	152	99
_*липнуться* 'stick oneself'	2	299	301	99
гибнуть 'perish'	13	1912	1925	99
пухнуть 'swell'	7	913	920	99
бухнуть 'swell'	6	573	579	99
мерзнуть 'be cold'	18	1315	1333	99
киснуть 'turn sour'	8	578	586	99
тухнуть 'fade away'	6	409	415	99
глохнуть 'go deaf, fade out'	12	777	789	98

_мерзнуться ‘be cold’	2	126	128	98
молкнуть ‘subside (about sound)’	14	880	894	98
гаснуть ‘fade out (about light)’	24	1432	1456	98
блекнуть ‘fade away’	6	288	294	98
_сякнуть ‘run dry, run out’	7	295	302	98
сипнуть ‘become hoarse’	3	117	120	98
липнуть ‘stick’	29	972	1001	97
_креснуть ‘resurrect’	10	323	333	97
мякнуть ‘become soft’	19	568	587	97
крепнуть ‘become hard’	33	926	959	97
жухнуть ‘shrivel’	3	79	82	96
вергнуться ‘plunge’	42	1101	1143	96
_верзнуться ‘fling’	7	176	183	96
_выкнуться ‘get used to’	13	318	331	96
грязнуть ‘get stuck’	12	286	298	96
слепнуть ‘become blind’	14	309	323	96
_слабнуть ‘become weak’	14	301	315	96
вязнуть ‘get stuck’	27	571	598	95
чахнуть ‘waste away, pine away’	8	168	176	95
хрипнуть ‘become hoarse’	9	173	182	95
вянуть ‘languish, wither’	17	322	339	95
зябнуть ‘feel chilly’	33	587	620	95
_скорузнуть ‘get rough, stale, harden’	1	16	17	94
меркнуть ‘become dark’	34	517	551	94
склизнуть ‘make a gliding sound’	1	15	16	94
стынуть ‘cool down’	11	158	169	93
_никнуться ‘droop’	35	482	517	93
_верзнуть ‘fling’	2	27	29	93
_торгнуться ‘intrude, extrude, tear away’	29	382	411	93
тускнуть ‘fade out’	1	13	14	93
дрябнуть ‘become shabby’	2	26	28	93
грузнуть ‘sink’	2	25	27	93
сохнуть ‘become dry’	31	374	405	92
пахнуть ‘smell’	78	897	975	92
_выкнуть ‘get into/out of the habit’	149	1677	1826	92
слизнуть ‘become slippery’	1	11	12	92
никнуть ‘droop’	215	2126	2341	91
_молкнуть ‘become quiet’	42	357	399	89
тихнуть ‘fade away (about sound)’	43	364	407	89
дрогнуть ‘feel cold’	17	131	148	89
_стигнуть ‘reach’	273	2001	2274	88
_киснуться ‘become sour’	1	7	8	88
_волгнуть ‘become wet’	1	7	8	88
_вергнуть ‘plunge’	205	1310	1515	86
виснуть ‘hang’	62	388	450	86
дрыхнуть ‘sleep’	14	72	86	84
двигнуться ‘move’	11	56	67	84

двигнуть 'move'	55	272	327	83
_чезнуть 'disappear'	312	1566	1878	83
_торгнуть 'intrude, extrude, tear away'	45	122	167	73
_бегнуть 'resort'	190	335	525	64
жухнуться 'shrivel'	4	1	5	20
_стигнуться 'reach'	1	0	1	0
Total	2288	31738	34026	93

Table 1: Overview of *nu*-drop in Russian verbs (finite past tense forms, active participles and gerunds, both prefixed and unprefixed verbs)

As shown in the bottom row of Table 1, Ø-forms represent 93% of the examples in our database. This is not unexpected: for example, Gorbačevič (1978, 164) states that retaining /nu/ in the past tense of the verbs in question belongs to the category of "residual phenomena" ("остаточные явления") in modern Russian. Although Ø-forms are dominant, Table 1 also shows that *nu*-forms have not been marginalized completely; the 2,288 attested *nu*-forms constitute 7% of our database. One must therefore ask under which conditions *nu*-forms occur. This will be the focus of our discussion in sections 2 through 7.

Ø-forms dominate not only when we count examples, but also if we count verbs. Of the 74 verbs in Table 1, 58 verbs have more than 90% Ø-forms. These verbs represent 78% of the listed verbs and 76% of all the examples in our database. Among the remaining 16 verbs, 12 display more than 80% Ø-forms. Of the four last verbs, *жухнуться* and *_стигнуться* are attested with very few examples in our database, so there are only two reasonably frequent verbs, *_торгнуть* and *бегнуть,* that have less than 80% Ø-forms.

2. Phonology: the root-final consonant

Is the phonological shape of the root of the verb relevant for *nu*-drop? This question has not received attention in the scholarly literature, although other cases of morphological variation in Russian verbs are sensitive to the shape of the stem (cf. e.g. Nesset's 2010 analysis of variation of the type *каплет* ~ *капает* 'drips'). In this section we show that the root-final consonant has a statistically significant, but relatively small impact on *nu*-drop, insofar as root-final labials favor Ø-forms more strongly than velars. However, diachronic analysis indicates that velar-final roots have shown increasing use of Ø-forms over the last 150 years, and have now virtually caught up with labial-final roots. Although relevant, the root-final consonant therefore does not appear to be a factor of major importance for *nu*-drop.

Table 2 shows that data are unequally distributed across natural classes of segments. For labials, only plosives are attested in root-final position, while for dentals, fricatives are dominant. Only for velars are both plosives and fricatives well attested. In view of this, only two comparisons are possible regarding place of articulation. First, for plosives we can compare labials and velars, i.e. verbs like *зябнуть* 'suffer from cold' and *меркнуть* 'grow dark':

(3) Штирлиц всю ночь **зяб** и топил камин. [Коллекция анекдотов: Штирлиц (1973-2000)]
'Štirlic was freezing cold all night and kept the fire going.'

(4) Звезды были четки и белы. **Меркнул** месяц в очень синем небе. [Б. А. Пильняк. Третья столица (1922)]

'The stars were bright and white. The moon **faded** in the very blue sky.'

In these examples, *зябнуть* has a Ø-form while *меркнуть* is represented by a *nu*-form, but as shown in Table 2, Ø-forms dominate both for roots ending in labials (97% Ø-forms) and velars (91% Ø-forms). This suggests that labials have a stronger preference for Ø-forms than velars. Statistical analysis demonstrates that the difference is highly significant, but that the effect size is small.[4]

Root-final C	# nu	# Ø	# total	% Ø
Labial plosive	152	5647	5799	97
Dental plosive	0	64	64	100
Dental fricative	490	6873	7363	93
Velar plosive	1406	13563	14969	91
Velar fricative	212	5111	5323	96
Total	2260	31258	33518	93

Table 2: Root-final place and manner (unprefixed+prefixed verbs)

The second comparison that can be made for place of articulation on the basis of Table 2 is between velar fricatives (96% Ø-forms) and dental fricatives (93% Ø-forms), i.e. between verbs like *чахнуть* 'waste away, pine' and *киснуть* 'turn sour'

(5) Юноша из царской семьи день ото дня **чах** от неизвестной болезни … [Владимир Леви. Искусство быть собой (1973)]
'The young man from the royal family **wasted away** day after day from an unknown disease.'

(6) Шурка мерз, **кис**, а во мне поднялся жар, я страстно ждал, чтобы еще покупали, еще. [Анатолий Кузнецов. Бабий яр (1965-1970)]
'Šurka was cold, he **languished**, and I got feverish and waited passionately for them to buy more.'

Statistical analysis shows that the difference is significant. However, the effect size is below the threshold for what can be considered a "small effect size".[5] In other words, our data does not allow us to conclude that the difference between dental and velar fricatives is of much importance for *nu*-drop. Since we cannot draw any conclusions about the role of dentals, therefore, the only claim about place of articulation that is backed up by our data is that verb roots in labials are more likely to undergo *nu*-drop than verbs with velars in root-final position. This finding is summarized in the following hierarchy, where the sign > indicates that the category to the left has a stronger tendency to undergo *nu*-drop than the category to the right:

(7) The phonological hierarchy:
labial > velar

For manner of articulation, the only possible comparison that can be made is between velar fricatives and plosives; for labials only plosives are attested, and for dentals plosives are too infrequent to permit statistical analysis. As shown in Table 2, velar fricatives display a higher proportion of Ø-forms (97%) than velar plosives (90%). Statistical analysis confirms that this difference is significant, but the effect size does not cross the threshold of what is regarded as a "small effect size".[6]

Let us now consider the situation from a diachronic perspective. Different hypotheses have been stated in the scholarly literature. Vinogradov and Švedova (1964, 173 et passim; see also Bulaxovskij 1954, 118 and Gorbačevič 1971, 207ff. and 1978, 164ff.), argue that, in general, the use of Ø-forms has increased, whereas Timberlake (2004, 105) claims that the "development is towards increasing use" of /nu/. For simplicity, we will refer to these hypotheses as the "Ø-increase hypothesis" and the "*nu*-increase hypothesis", respectively. The data in Table 3, which shows the total numbers of examples and the percentages for Ø-forms from 1800 to today, makes it possible to test these conflicting hypotheses. We have divided this time span into fifty-year periods, which provide sufficiently large numbers to facilitate statistical analysis. In the rightmost two columns, we have included data for the first decade of the 21st century, since contemporary data are well represented in our database. Dental plosives were not included in Table 3, since numbers are too small to make comparisons of different periods possible. Figure 1 visualizes the development over time.

	1800-49		1850-99		1900-49		1950-99		2000-	
	#total	%Ø	#total	%Ø	#total	%Ø	#total	%Ø	#total	%Ø
Labial plosive	163	93	591	96	1607	98	1921	97	1517	98
Dental fricative	273	92	760	94	1949	94	2640	94	1741	91
Velar plosive	691	72	1677	85	3334	91	4673	93	4594	93
Velar fricative	156	94	535	95	1527	95	1938	96	1167	97

Table 3: Root-final place and manner (unprefixed+prefixed verbs) over time

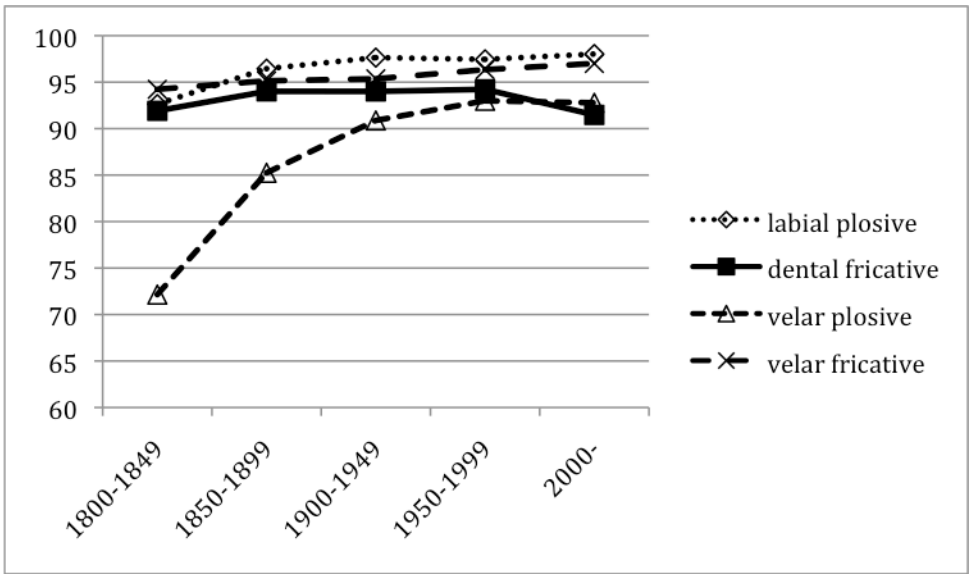

Figure 1: Root-final place and manner (unprefixed+prefixed verbs) over time

Table 3 and Figure 1 show that three out of four categories have flat contours between 90% and 100% for the whole time span. However, velar plosives differ from the other categories in showing a growth from 72% Ø-forms in the first half of the 19th century to 93% in the period after 1950. This difference is statistically highly significant, but the effect size is small to moderate.[7]

The development attested in Table 3 and Figure 1 suggests that the difference between root-final labials and velars referred to in (7) has decreased over time and is quite small in present-day Russian. In other words, over a period of 200 years velars have almost caught up with labials with regard to the use of Ø-

forms. While the phonological hierarchy in (7) represents a valid generalization over our database as a whole, this hierarchy seems to have lost its importance in present-day Russian.

As for the two conflicting hypotheses about the development of *nu*-drop, the data in Table 3 and Figure 1 do not provide support for the *nu*-increase hypothesis (Timberlake 2004), since none of the four categories display an increasing proportion of *nu*-forms. On the other hand, the Ø-increase hypothesis of Vinogradov and Švedova (1964) and others is also not supported fully. While one of the four categories under scrutiny shows increasing use of Ø-forms, stability rather than increase is characteristic of the remaining categories.

To summarize, statistical analysis shows that the root-final consonant is of limited importance for *nu*-drop. Although in our database labial-final roots are more prone to undergo *nu*-drop than velar-final roots, diachronic evidence indicates that this difference has been reduced over a period of 200 years, and is very small in present-day Russian. Our diachronic study furthermore shows that with the exception of velar-final roots, there have been remarkably small changes since 1800. This stability is at variance with both the *nu*-increase and the Ø-increase hypotheses.

3. Inflectional morphology: paradigm cells

It is often asserted in the literature on *nu*-drop that different inflected forms (i.e. the cells in a paradigm) respond differently to *nu*-drop (cf. e.g. Gorbačevič 1971, 208f.; Isačenko 1982, 251; Nesset 1998, 140f.; Rozental' 1977, 168ff.; Švedova (ed.) 1980, 652f. and Timberlake 2004,105). For instance, although the authors of the Russian Academy Grammar (Švedova (ed.) 1980, 652f.) are careful to point out that there are confounding factors such as prefixation (to which we turn in the following section), they argue that masculine forms are less prone to undergo *nu*-drop than other finite forms, and that participles and gerunds are even less likely victims of *nu*-drop than finite forms. Our database makes it possible to test this hypothesis empirically, i.e. to find out whether actual usage conforms to the hierarchy non-masculine finite > masculine finite > non-finite.

	# nu	# Ø	# total	% Ø
Masculine sg	315	8001	8316	96
Feminine sg	35	6686	6721	99
Neuter sg	19	4096	4115	100
Plural	56	6490	6546	99
Active participle	659	6312	6971	91
Gerund	1204	153	1357	11
Total	2288	31738	34026	93

Table 4: *Nu*-drop in various inflected forms

The data in Table 4, which conflates the numbers for all periods covered by our database, provides partial support for this hypothesis. As can be seen from the table, for the non-masculine finite forms *nu*-drop is virtually obligatory. The masculine sg forms have a somewhat lower percentage of Ø-forms (96%), followed by the participles (91%). The gerunds are in a different league with only 11% Ø-forms. Statistical analysis shows that the differences between the feminine singular, neuter singular and the plural are just barely significant. However, the effect size is far from crossing the threshold of a small effect size,

so for practical purposes these differences can be ignored. The other differences between the forms in Table 4 are statistically highly significant. Comparing all the non-masculine finite forms with the masculine yields a small effect size, and the same is true for comparisons of the masculine and the active participle. However, comparison of participles and gerunds gives an extremely large effect size.[8] In other words, Ø-forms dominate in the non-masculine finite forms, the masculine finite forms and active participles, while *nu*-forms occur in the majority of examples with gerunds, as illustrated by the following examples with *привыкнуть* 'get used to':

(8) Ирина постепенно **привыкла** к тому, что он уходит. [Токарева Виктория. Своя правда // «Новый Мир», № 9, 2002]
'Irina gradually **got used to** him leaving.'

(9) Я, как вы заметили, человек практический, к тому же бывший военный, **привык** к точности. [Светлана Бударцева. У хорошего хозяина метр зарабатывает (2002) // «Вечерняя Москва», 2002.03.14]
'As you have noticed, I am a practical man, even a former soldier, and I am **used to** punctuality.'

(10) Как человек, **привыкший** к гастролям, я собираюсь в дорогу легко. [Федор Чеханков: Ненависть меня разрушает (2002) // «Витрина читающей России», 2002.09.13]
'As a person **used to** touring, I easily pack for a new trip.'

(11) **Привыкнув** к темноте, я разглядел, что двое других — водитель и тот, что сидел рядом с ним, — ни в каких масках не нуждаются. [Евгений Прошкин. Механика вечности (2001)]
'**Having got used to** the darkness, I discerned that the two others, the driver and the person next to him, did not need any masks.'

The following hierarchy summarizes the situation:

(12) The inflectional hierarchy:
Non-masculine > masculine > active participle > gerund

The actual situation differs from the hypothesis mentioned in the beginning of this section in one important respect. Analysis of the data in Table 4 has shown that *nu*-drop is sensitive to the difference between participles and gerunds, and that this difference is much more important than the differences between the remaining forms in Table 4.

Let us now turn to the diachronic aspect of the situation. Vinogradov and Švedova (1964, 167ff.) claim that the use of Ø-forms has increased for finite forms and participles, while gerunds have displayed the opposite development. While for finite forms according to Vinogradov and Švedova (1964, 167ff.) *nu*-forms were used relatively widely in the 18th and early 19th centuries, they soon became stylistically marked and a gradual increase ("постепенный рост") in the use of Ø-forms started already in the first half of the 19th century. Vinogradov and Švedova (1964, 171) observe a parallel development for participles, although according to them in the 1700s and early 1800s *nu*-forms were more widely used in participles than in finite forms, and the increased use of Ø-forms started somewhat later among participles.

Vinogradov and Švedova (1964, 167ff.) do not comment on the difference between masculine and other finite forms with regard to *nu*-drop, although most of the examples they cite are masculine forms. However, Table 5, which is

organized in the same way as Table 3 above, shows that masculine forms confirm the hypothesis of increasing use of Ø-forms; the growth from under 80% before 1850 to percentages close to 100% in the 20th and 21st centuries represents a statistically significant change with a moderate effect size.[9] For other finite forms, Ø has been virtually obligatory at least since 1850, so for these forms the hypothesis of increased use of Ø-forms is not borne out by our data. Our data furthermore does not indicate increasing use of Ø-forms of participles; as shown in Table 5, the percentage of Ø-forms has remained relatively stable between 89% and 93% since 1900.[10] The data in Table 5 confirms the hypothesis of decreasing use of Ø-forms for gerunds. While the percentage of Ø-forms was around 15-20% up to 1950, it has sunk to under 10% in the two most recent periods documented in Table 5. The observed differences are statistically significant, and the effect size is small.[11] Although as shown in section 3 Timberlake's (2004, 105) *nu*-increase hypothesis does not receive support from *nu*-verbs in general, this hypothesis gives correct predictions for gerunds.

	1800-49		1850-99		1900-49		1950-99		2000-	
	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø
Masculine sg	353	79	943	90	2042	97	2964	99	2014	98
Other finite	671	94	1902	99	4392	100	5669	100	4748	100
Active participle	174	83	613	85	1808	89	2345	93	2031	91
Gerund	113	16	171	20	303	15	423	6	347	8

Table 5: Development of *nu*-drop in various inflected forms over time

To conclude, our discussion has shown that inflectional morphology is relevant for *nu*-drop insofar as different forms of the paradigm behave differently. Ø-forms are virtually obligatory for finite forms other than masculine sg. For masculine sg and participles, Ø-forms also dominate, while for gerunds *nu*-forms are by far most used. Our diachronic investigation suggests that the use of Ø-forms has increased in the masculine sg, but decreased in the gerund. The remaining forms have displayed a remarkable stability over time.

4. Derivational morphology: aspectual prefixation

A factor that is frequently commented on in the scholarly literature is aspectual prefixation. It is generally believed that prefixed verbs are more likely to undergo *nu*-drop than unprefixed verbs (cf. e.g. Švedova (ed.) 1980, 652; Isačenko 1982, 250 and Rozental' 1977, 168ff.). In the previous section, we demonstrated that Ø-forms are virtually obligatory in non-masculine finite forms. For these forms an investigation of prefixation would be futile. We will also not discuss gerunds, since all the 1357 gerunds in our database are prefixed. However, for the masculine sg forms and the active participle a discussion of prefixation is possible. Examples (13) and (14) indicate that in the masculine sg both *nu*-forms and Ø-forms are attested in our database:

(13) Невежество дико-восточного мира оскорбляло его, он в нем **чахнул** и рвался вон. [А. И. Герцен. Былое и думы. Часть седьмая. Вольная русская типография и «Колокол» (1866)]
'The ignorance of the wild eastern world offended him, he **pined away** in this world and longed [to go] away.'

(14) После ухода Эфроса театр на Малой Бронной быстро стал блекнуть, вянуть и **зачах** совсем. [Виктор Розов. Режиссер, которого я люблю (1990-2000)]

'After Efros left, the theatre on Malaya Bronnaya started fading, withering, and **wilted** completely.'

Examples (15) and (16) illustrate the variation between *nu*-forms and Ø-forms in participles:

(15) Когда он принял решение забрать **чахнувшего** в национальном парке кондора и появился с ним в аэропорту, служащие спрашивали: зачем ему эта некрасивая и даже неприятная с виду птица? [Марина Беляева. Кондор -- всевидящий бог инков (2000) // «Семья», 2000.01.19]
'When he decided to take the **languishing** condor from the national park, and appeared with it at the airport, the staff asked him, why would he need such an unattractive and even unpleasant-looking bird?'

(16) Не сумели сообразить, что изучение сохранившихся до настоящего времени диких племен, **зачахших** в голоде, болезнях и суеверии, практически ничего не дает для представления о наших подлинных предках. [И. А. Ефремов. Лезвие бритвы (1959-1963)]
'It was not understood that research on the wild tribes that had survived until now, but that had **languished** in hunger, disease and superstition, does not give us any information about our real ancestors.'

	# nu	# Ø	# total	% Ø
Unprefixed masculine sg	57	368	425	87
Prefixed masculine sg	258	7633	7891	97
Unprefixed active participle	136	9	145	6
Prefixed active participle	523	6303	6826	92

Table 6: *Nu*-drop and aspectual prefixation

Table 6 indicates that unprefixed verbs have much lower frequencies than prefixed verbs. This is true not only of masculine forms and participles, but holds of our database in general. Of the 34,026 examples in our database, only 2555 (about 8%) are unprefixed. Despite this skewed distribution, however, meaningful comparisons of *nu*-drop in prefixed and unprefixed verbs are possible. Table 6 demonstrates that prefixed verbs have higher percentages of Ø-forms than unprefixed verbs. For participles, the difference is dramatic (92% Ø-forms for prefixed verbs vs. 6% for unprefixed verbs). Not surprisingly, this dramatic difference is statistically significant and the effect size is moderate to large.[12] For masculine forms, the difference is less dramatic (97% Ø-forms for prefixed verbs vs. 87% for unprefixed forms), but the difference is statistically significant with a small effect size.[13] In other words, our data corroborates the following hierarchy:

(17) The derivation hierarchy:
prefixed > unprefixed

With the derivation hierarchy in mind, let us now consider the diachronic situation. As can be seen from Table 7, we have very small numbers for unprefixed verbs in the earlier periods. Since percentages based on small numbers are of little value, we decided to disregard periods with a total number of examples (i.e. the sum of examples with /nu/ and Ø) smaller than 50. This means that we have reliable data for masculine forms from 1850 and for participles from 1900. The historical developments are visualized in Figure 2,

which for the convenience of the reader also includes gerunds and non-masculine finite forms discussed in the previous section. Table 7 and Figure 2 show that both prefixed and unprefixed masculine forms display an increasing use of Ø-forms over time. The percentage of Ø-forms is always lower for unprefixed verbs, but the difference becomes smaller over time. While until 1950 the difference was about 15 percentage points, after 1950 the difference was reduced to about 5 percentage points. However, although unprefixed masculine forms appear to be in the process of catching up with prefixed forms, the difference is still statistically significant. Therefore, the derivation hierarchy in (17) is still valid in present-day Russian.[14]

	1800-49		1850-99		1900-49		1950-99		2000-	
	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø
Unprefixed masc	20	85	54	74	103	79	161	92	87	94
Prefixed masc	333	78	889	91	1939	98	2803	99	1927	99
Unprefixed part	5	20	16	0	62	6	40	5	22	9
Prefixed part	169	85	597	87	1746	92	2305	95	2009	92

Table 7: Development of *nu*-drop for unprefixed and prefixed verbs

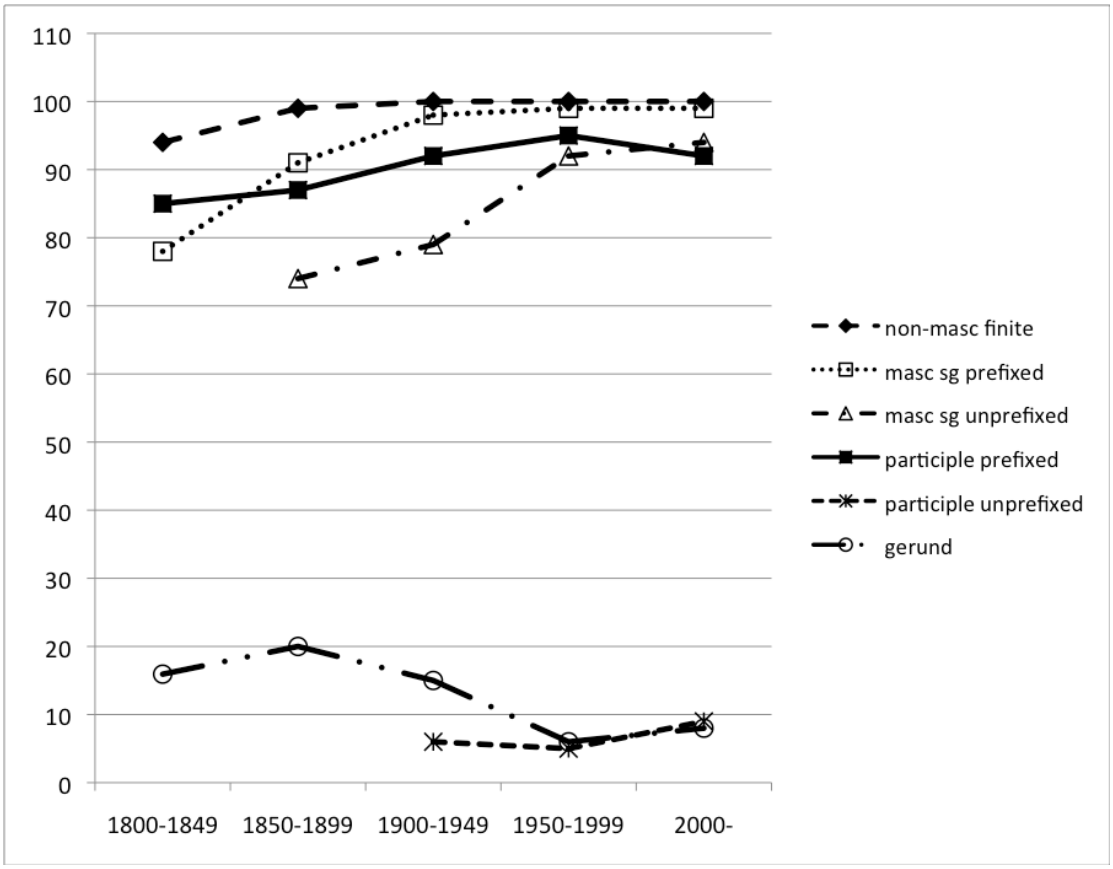

Figure 2: Development of *nu*-drop for unprefixed and prefixed verbs

As shown in Table 7 and Figure 2, unprefixed and prefixed participles display parallel contours indicating virtually no change since 1900. While for prefixed participles the proportion of Ø-forms has been stable above 90% since 1900, unprefixed participles remain stable below 10%. In other words, the dramatic difference between unprefixed and prefixed participles documented in Table 6 does not appear to be changing over time.

Summarizing, we have shown that aspectual prefixation facilitates *nu*-drop, while unprefixed verbs show lower percentages of Ø-forms. This generalization, however, is only valid for masculine sg forms and participles, since for non-masculine finite forms Ø-forms are virtually obligatory regardless of prefixation, while gerunds are only formed from prefixed verbs. The difference between unprefixed and prefixed verbs is most dramatic for participles, but even for masculine sg forms it is statistically significant. Diachronic analysis has demonstrated that the gap between unprefixed and prefixed participles remains stable over time, while unprefixed masculine sg forms are in the process of catching up with prefixed masculine sg forms.

5. Back to phonology: number of syllables

The finding that unprefixed verbs show a stronger tendency to retain /nu/ indicates that /nu/ is more frequent in combination with shorter forms, since unprefixed stems are shorter than prefixed stems. However, maybe the phonological parameter of number of syllables is a better measure of "shortness" than the morphological parameter of prefixation? In this section we investigate this question and demonstrate that this phonological hypothesis does not receive support from the corpus data under scrutiny in the present study.

In order to investigate the relevance of the number of syllables for *nu*-drop we divided our data material into four groups: unprefixed verbs, verbs with non-syllabic prefixes such as *v-,* verbs with monosyllabic prefixes such as *pri-,* and verbs with disyllabic prefixes such as *pere-.* In order to avoid confounding factors, we omitted verbs with *–sja* from the dataset, and we also disregarded the only verb root with more than one syllable, *_скорузнуть* 'get rough, stale, harden'. If verbs with fewer syllables are more likely to retain /nu/ than longer verbs, we predict that verbs with non-syllabic prefixes behave like unprefixed verbs. If, on the other hand, the decisive factor is not the number of syllables, but rather the presence or absence of a prefix, we predict that verbs with non-syllabic prefixes behave like other prefixed verbs.

Since, as shown in the previous section, prefixation is only relevant for masculine forms and participles, we restricted our investigation to these forms. Here are examples with *липнуть* 'stick' involving no prefix (18), non-syllabic prefix (19) and monosyllabic prefix (20):

(18) Спокойно натягивала на себя легкую рубашку, и шелк **лип**, впечатывался в тело и намокал. [Борис Васильев. А зори здесь тихие (1969)]
'She calmly pulled on a thin shirt, and the silk **stuck**, left an imprint on her body and soaked.'

(19) Женя-морячок все-таки **влип** в историю. [Виктор Астафьев. Веселый солдат (1987-1997) // «Новый Мир», 1998]
'Ženja the sailor nevertheless **got stuck** in a pretty mess.'

(20) Курчавый ореол волос развился и тонкими струйками **прилип** к голове, ко лбу. [Ю. П. Анненков. Дневник моих встреч (1966)]
'The curly halo of hair unfurled and stuck to the head and forehead in little streams.'

Table 8 summarizes the situation for the masculine forms and shows that non-syllabic prefixes have virtually the same percentage of Ø-forms as other prefixed

verbs, while unprefixed verbs are about 10 percentage points lower. Statistical analysis demonstrates that the difference between unprefixed verbs and verbs with non-syllabic prefixes are statistically significant, whereas the differences among prefixed verbs are not.[15] In other words, the data in Table 8 does not lend support to the phonological hypothesis that the number of syllables is relevant for *nu*-drop.

	# nu	# Ø	# total	% Ø
Unprefixed	57	368	425	87
Non-syllabic prefix	22	637	659	97
Monosyllabic prefix	234	6862	7096	97
Disyllabic prefix	2	133	135	99

Table 8: *Nu*-drop and number of syllables in masculine sg forms

Table 9 shows that the situation for participles is similar. Again, the percentage of Ø-forms for verbs with non-syllabic prefixes is much closer to that of other prefixed verbs than to unprefixed verbs. The difference is so dramatic that statistical analysis is superfluous. To sum up this section, both our analysis of masculine forms and participles show that the number of syllables is not a factor that influences nu-drop.

	# nu	# Ø	# total	% Ø
Unprefixed	136	9	145	6
Non-syll prefix	50	550	600	92
Monosyll prefix	464	5672	6136	92
Disyllabic prefix	9	54	63	86

Table 9: *Nu*-drop and number of syllables in active participles

6. Semantics and syntax: transitivity and change of state

Are the semantic and syntactic properties of a verb relevant for *nu*-drop? We will show that the question can be answered in the affirmative, insofar as transitivity has a small, but statistically significant effect on *nu*-drop. However, the difference between inchoative and stative intransitive verbs turns out not to be significant. Diachronic analysis shows that the difference between transitive and intransitive verbs was smaller in the 19th century, but has not decreased since 1900 and is still significant.

The verbs under scrutiny in the present study fall into three classes with regard to their semantic and syntactic properties (cf. Nesset 1998, 132 for discussion). First, there is a group of transitive verbs with agentive subjects, such as *двигнуть* 'move':[16]

(21) Шредингер ссылался в ней на Тимофеева-Ресовского, который **подвигнул** его на эту работу. [Даниил Гранин. Зубр (1987)]
'Schrödinger referred to Timofeyev-Resovsky, who **roused** him to this work.'

However, the majority of *nu*-verbs are intransitive verbs where the subject carries the role "patient", e.g. *гаснуть* 'go out (about light)' and *мерзнуть* 'be cold':

(22) Только свет **гас**, на скамейки укладывались и тут же засыпали до конца сеанса. [Вадим Сидур. Памятник современному состоянию (1973-1974)]

'As soon as the lights **went out**, one would lie down on the benches and fall asleep at once until the session was over.'

(23) Он не **мерз** и в тридцатиградусный мороз, только облачко пара висело у лица. [Юрий Дружников. Виза в позавчера (1968-1997)]
'He did not **feel cold** in minus thirty, only a cloud of steam would hang next to his face.'

Among the intransitive *nu*-verbs, most verbs denote a change of state. A case in point is *гаснуть* in (22) which describes the transition from light to darkness. For convenience, we will refer to verbs of this type as "inchoative". A smaller subgroup of intransitive verbs (e.g. *мерзнуть* 'be cold' in (23)) involves stable states, and these verbs are therefore called "stative". Here are full lists of the relevant types of verbs:[17]

(24) Transitive verbs: *_стигнуть, _торгнуть, _вергнуть, _верзнуть, бегнуть, двигнуть*

(25) Stative intransitive verbs: *дрогнуть, дрыхнуть, липнуть, мерзнуть, обрыднуть, пахнуть, виснуть, зябнуть*

(26) Inchoative intransitive verbs: all other verbs listed in Table 1

The question is now whether the syntactic/semantic classes behave differently with regard to *nu*-drop. Consider the data in Table 10, which shows that intransitive verbs display a stronger preference for Ø-forms than transitive verbs do. This difference is statistically significant and shows a small to moderate effect size, so the following hierarchy is supported by the data:[18]

(27) The syntactic/semantic hierarchy:
Intransitive > transitive

	# nu	# Ø	# total	% Ø
Transitive verbs	770	4067	4837	84
Intransitive verbs	1371	24466	25837	95

Table 10: *Nu*-drop and semantic classes

Among intransitive verbs, the opposition between inchoatives and statives is neutralized in the perfective aspect. The verbs we have classified as "stative" describe stable states only in the imperfective aspect, i.e. when they are unprefixed. When a perfectivizing prefix is added to a stative verb like *мерзнуть* 'be cold', the result is a verb that denotes a change of state, e.g. *замерзнуть* 'become cold' (cf. Zaliznjak and Šmelev 2000, 57). Table 11 therefore concerns unprefixed verbs only. Although the table indicates a small difference between stative and inchoative verbs and this difference is statistically significant, the effect size does not cross the threshold of a small effect. In other words, the distinction between stative and inchoative verbs does not have an impact on *nu*-drop.[19]

	# nu	# Ø	# total	% Ø
Stative intransitive verbs	120	1124	1244	90
Inchoative intransitive verbs	85	1203	1288	93

Table 11: *Nu*-drop and semantic classes (unprefixed verbs only)

Since the distinction between transitive and intransitive verbs appears to be relevant for *nu*-drop, the question arises as to whether the behavior of the two

semantic classes of verbs has changed over time. As shown in Table 12 and Figure 4, intransitives have been stable on 94%-96% Ø-forms since 1850, whereas transitive verbs display an increase from 62% Ø-forms in the first half of the 19th century to 89% after year 2000. Does this mean that we are witnessing a converging development, whereby transitives are in the process of catching up with intransitives? Statistical analysis demonstrates that such an interpretation of the data is not quite right – for two reasons. First, although the slight increase of Ø-forms among transitives since 1900 is just barely statistically significant, the effect size does not cross the threshold of a small effect.[20] Since both transitives and intransitives have been stable for more than a century now, our data does not indicate that the difference between the two verb types with regard to *nu*-drop is diminishing. Secondly, statistical analysis of the numbers for the 21st century indicates that the difference between the two verb types is still statistically significant, although the effect size is small.[21] In other words, the difference has not decreased for more than a century, and it is still statistically significant.

	1800-49		1850-99		1900-49		1950-99		2000-	
	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø	#tot	%Ø
Transitive	308	61	568	75	962	85	1346	87	1653	88
Intransitive	882	89	2641	94	6804	95	9012	95	6498	95

Table 12: *Nu*-drop and semantic classes – historical development

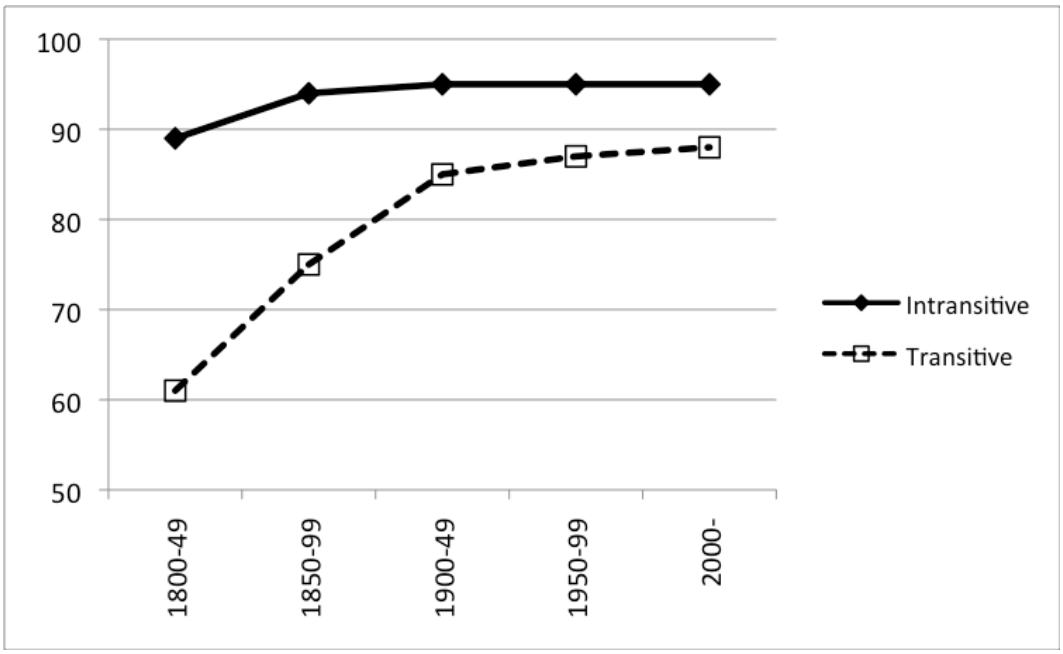

Figure 3: *Nu*-drop and semantic classes – historical development

Summing up the discussion of syntactic/semantic factors, we have shown that *nu*-drop is sensitive to the distinction between transitive and intransitive verbs, but not between stative and inchoative verbs. Diachronic analysis shows that the difference between the two groups was larger in the 19th century, but that it has not changed significantly since 1900, and is still statistically significant.

7. Interaction: what is the relative importance of the examined factors?

So far we have considered phonological, morphological and semantic/syntactic factors in isolation. How do these factors interact? What is their relative importance? In this section, we will see that morphology is more important for *nu*-drop than phonological and semantic/syntactic factors. We propose that the best basis for predicting the distribution of Ø- and *nu*-forms is a conflated hierarchy of derivational and inflectional morphological categories. In this hierarchy, the morphological categories fall into three groups: (a) categories where Ø-forms are virtually obligatory, (b) categories where Ø-forms are dominant (but not obligatory), and (c) situations where *nu*-forms dominate.

In the previous sections we have identified four hierarchies where the distribution of Ø-forms and *nu*-forms display statistically significant differences, i.e. where the differences are unlikely to be due to chance. However, even if an observed difference cannot be due to chance, it does not necessarily have a strong impact on *nu*-drop. For this reason, we have calculated Cramer's V-values, which measure the effect size of the relevant factors. In order to facilitate comparison, we repeat the four hierarchies in (28)-(31). As before, > indicates that the categories to the left are more likely to use Ø-forms than the categories to the right. In statistical terms, > represents a statistically significant difference with an effect size that crosses the threshold of what is considered a "small" effect size.

(28) Phonology: Labial > velar (0.1)
(29) Inflectional morphology: Non-masculine > masculine (0.1) > active participle (0.1) > gerund (0.7)
(30) Derivational morphology: Prefixed > unprefixed (participles: 0.4; masculines: 0.1)
(31) Syntax/semantics: Intransitive > transitive (0.1)

The numbers in parentheses are Cramer's V values, which measure the effect size of the difference between the two categories to the left of the number. Bear in mind that for Cramer's V values, 0.5 represents a large effect size, 0.3 a moderate effect size and 0.1 a small effect size (King and Minium 2008, 327-329). Since as shown in (28)-(31) only inflectional and derivational morphology involve large and moderate effect sizes, we can conclude that morphological factors are more important for *nu*-drop than phonology and semantics/syntax. Therefore, in the following we will limit ourselves to discussing the morphological factors, which evidently provide the best basis for predicting the distribution of Ø-forms and *nu*-forms.

The two morphological hierarchies in (29) and (30) interact in non-trivial ways, insofar as the derivational difference between prefixed and unprefixed verbs is only relevant for masculine finite forms and participles and display different effect sizes for these forms. In (32) we have conflated the two morphological hierarchies. Commas separate categories, for which differences are not statistically significant and/or the effect size is below what is considered a "small effect".[22] The percentages of Ø-forms are given in parentheses after each category. These numbers are taken from Tables 4 and 6 in sections 3 and 4.

(32) Conflated morphology (based on percentages of Ø-forms for entire database):
NON-MASC. (99%-100%) > MASC. PREFIXED (96%) > PARTICIPLE PREFIXED (92%), MASC. UNPREFIXED (87%) > gerund (14%), participle unprefixed (8%)

As we proceed from left to right in the hierarchy the likelihood of Ø-forms decreases. However, if we consider the percentages of Ø-forms, it becomes clear that the hierarchy does not report a gradual decrease from 100% to 0% Ø-forms. On the contrary, the categories fall into three broad groups occupying different areas of a scale from 100% to 0%. The use of small capitals and boldface captures this in (32). The first group (given in boldfaced small capitals) comprises non-masculine finite forms where Ø-forms are virtually obligatory (99%-100%). The second group, which is rendered in small capitals without boldface, includes categories where Ø-forms are dominant, but not quite obligatory (between 87% and 96% Ø-forms). In this group we find masculine finite forms (prefixed and unprefixed), as well as prefixed participles. In the third group, it is the *nu*-form that dominates (less than 15% Ø-forms). This pertains to gerunds and unprefixed participles, for which neither boldface nor small capitals are used in (32). The situation described in (32) is quite polarized in the sense that the relevant categories either have very high or very low percentages of Ø-forms, while no categories are in the middle part of the scale.

The hierarchy in (32) is based on our entire database, so it does not reflect changes in the distribution of Ø-forms and *nu*-forms between the 19th and 21st centuries. However, if we consider the numbers for the 21st century as an indication of the situation in present-day Russian we get a very similar conflated hierarchy for the morphological factors (percentages from Tables 5 and 7 in sections 3 and 4 in parentheses):[23]

(33) Conflated morphology (based on percentages of Ø-forms for 21st century):
NON-MASC. (100%), **MASC. PREFIXED (99%)** > MASC. UNPREFIXED (94%), PARTICIPLE PREFIXED (92%) > gerund (10%), participle unprefixed (9%)

Comparison of the hierarchies in (32) and (33) show that the only important difference is that the prefixed masculine forms have moved up into the group where Ø-forms are virtually obligatory, since the percentage of Ø-forms is now 99%. In other words, if our synchronic analysis of the entire database yields a polarized picture, our diachronic analysis of the development since 1800 shows that this polarization has increased over time.

9. Conclusion

Our investigation of *nu*-drop in Russian verbs based on 34,026 examples from the Russian National Corpus affords a number of conclusions. As shown in section 1, in general Ø-forms dominate over *nu*-forms, insofar as only 7% of the examples in our database involve *nu*-forms. In sections 2 through 7 we explored the impact of phonological, morphological and semantic/syntactic factors. With regard to phonology, it was demonstrated that the root-final consonant has a statistically significant, but small effect on *nu*-drop (section 2), whereas the number of syllables is not a relevant factor (section 5). As for syntax and semantics, *nu*-drop has been shown to be sensitive to the transitive-intransitive distinction, but not to the difference between inchoative and stative verbs (section 7). The morphological factors were shown to have the strongest impact

on *nu*-drop (sections 3 and 4). In section 8, we argued that the best basis for predicting the distribution of Ø-forms and *nu*-forms is a morphological hierarchy that distinguishes between different inflected forms and prefixed/unprefixed verbs. This hierarchy enabled us to distinguish between three groups:

(34) Situation in the database as a whole:
(a) Ø-forms are virtually obligatory:
Non-masculine finite forms
(b) Ø-forms are dominant, but not obligatory:
Masculine finite forms (prefixed and unprefixed) and prefixed active participles
(c) *Nu*-forms dominate:
Gerunds and unprefixed active participles

The situation in (34) is polarized; the categories show either a percentage of Ø-forms close to 100% or to 0%, while no categories are around 50%.

The conclusions above refer to our database as a whole, but our study has also enabled us to draw conclusions about the diachronic development from the 19th to the 21st century. First, with regard to the phonological shape of the root, we have shown that roots ending in velar plosives display an increasing percentage of Ø-forms over time, and that the difference between velar-final and other roots is in the process of disappearing. Second, our diachronic analysis of semantic/syntactic factors indicates that transitive verbs have had an increase in percentage of Ø-forms. However, after 1900 the growth has stopped, and the percentage of Ø-forms for transitive verbs is still significantly lower than for intransitive verbs. Finally, with regard to morphology, it has been demonstrated that stability over time is characteristic for most categories. The exceptions are the gerund, for which the percentage of Ø-forms decreases over time, and the masculine finite forms, which display increasing percentages of Ø-forms. The most important change is perhaps observed in prefixed masculines, since for this category Ø-forms became nearly obligatory in the 20th century. In other words, the diachronic development has created a situation in the beginning of the 21st century where Ø-forms are virtually obligatory for all finite verb forms. As can be seen from (35), the only exception is unprefixed masculines, but even for this category Ø-forms are strongly dominant:

(35) Situation in the beginning of the 21st century:
(a) Ø-forms are virtually obligatory:
All finite forms except unprefixed masculines
(b) Ø-forms are dominant, but not obligatory:
Unprefixed masculine finite and prefixed active participles
(c) *Nu*-forms dominate:
Gerunds and unprefixed active participles

Summarizing the diachronic development, we witness increasing polarization; increase is attested among categories with high percentages of Ø-forms, while decrease has been documented for categories with low percentages.

Even though this article has explored the phenomenon of *nu*-drop from a synchronic and diachronic perspective in great detail, many puzzle pieces have not yet fallen into place. First of all, we have not investigated all potentially relevant factors. For instance, a systematic study of homonymy avoidance and

nu-drop is yet to be carried out. Another potentially fruitful alley for further research is to compare *nu*-drop with other examples of morphological variation and change in Russian verbs. However, although these and other issues are beyond the scope of the present study, we hope to be able to address them in future research projects.

---

[1] All examples cited in this article are taken from The Russian National Corpus, available at http://www.ruscorpora.ru/. In each example, the relevant form is boldfaced. Some of the examples have been abbreviated. Notice that we do not consider variation in the infinitive of the type *достигнуть* ~ *достичь* 'reach'.

[2] Needless to say, these are not the only factors that are potentially relevant for *nu*-drop. For instance, Gorbačevič (1978:165) mentions homonymy avoidance; in order to avoid homonymy with, say, *слеп* 'blind' (the short form of the adjective) speakers may prefer the past tense form *слепнул* 'became blind' to *слеп* with the same meaning. Other potentially relevant factors are style and register (cf. e.g. Gorbačevič 1978:165ff.). However, since homonymy avoidance, style and register are not easily testable in a quantitative study, these factors are beyond the scope of the present article.

[3] We use the term "gerund" about forms like *достигнув*, *достигши* and *достигнувши* of *достигнуть*/*достичь* 'reach'. Alternative terms in English are "adverbial participle" and "converb".

[4] Since this is an article for linguists, and not for professional statisticians, we place information about statistical analysis in footnotes. In this article we use Pearson's Chi-squared test to check for statistical significance, and based on the results from this test we calculate Cramer's V-values as measures of effect size. All calculations are carried out in the software package R. Statistical significance measures the likelihood that the distribution of the data could be due to chance. According to standard practice, a result is considered statistically significant if the p-value < 0.05, which indicates that there is less than 5% likelihood that the observed distribution is due to chance. Notice that statistical significance is not the same as effect size, which measures the strength of the relationship between two factors. Even if a result is clearly not due to chance, this does not necessarily mean that the relevant factors have a strong impact. This is particularly true for large databases such as the one under scrutiny in the present study, where Pearson's Chi-squared test is able to identify very small differences as statistically significant. The data for labial and velar plosives in Table 2 illustrate the importance of supplementing chi-squares with Cramer's V-values. Pearson's Chi-squared test with Yates' continuity correction (X-squared = 275.2283, df = 1) yields the p-value < 2.2e-16, showing that the difference between labial and velar plosives is highly significant. (In fact, 2.2e-16, i.e. the number 0. … 22 with fifteen zeros after the decimal mark, is the smallest number the R software package operates with, so for all practical purposes the likelihood that the observed distribution can be due to chance is zero.) However, Cramer's V-value equals 0.1. Even though Cramer's V value can theoretically vary from 0 to 1, 0.5 is considered high, while 0.3 represents a moderate value and 0.1 a low value (cf. King and Minium 2008, 327-329). In other words, our statistical analysis enables us to conclude that the difference between velar and labial consonants in root-final position is relevant for *nu*-drop, but that this factor has a small effect.

[5] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 41.6919, df = 1) yields the p-value = 1.069e-10. Cramer's V = 0.06.

[6] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 166.0626, df = 1) gives the p-value < 2.2e-16. Cramer's V = 0.089.

[7] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 286.2948, df = 1) gives p-value < 2.2e-16. Cramer's V = 0.2.

[8] For non-masculine finite forms, Pearson's Chi-squared test (X-squared = 8.4189, df = 2) gave p-value = 0.01485. Cramer's V = 0.02. For the comparison of non-masculine and masculine sg forms, Pearson's Chi-squared test with Yates' continuity correction (X-squared = 342.3158, df = 1) yielded p-value < 2.2e-16. Cramer's V = 0.1. Comparing masculine sg and active participles, Pearson's Chi-squared test with Yates' continuity correction (X-squared = 203.0981, df = 1) provided p-value < 2.2e-16. Cramer's V = 0.1. Finally, for the comparison of participles and gerunds, Pearson's Chi-squared test with Yates' continuity correction (X-squared = 4105.707, df = 1) gave p-value < 2.2e-16. Cramer's V = 0.7.

[9] In order to investigate the historical development of *nu*-drop in the masculine sg we compared the numbers from 1800-1849 with the numbers from after year 2000. Pearson's Chi-squared test (X-squared = 260.7055, df = 1) gave p-value < 2.2e-16. Cramer's V = 0.3.
[10] Admittedly, Pearson's Chi-squared test (X-squared = 63.1782, df = 5) indicates significance (p-value = 2.674e-12), but the effect size does not cross the threshold of a small effect (Cramer's V = 0.09).
[11] For gerunds, we compared the numbers from 1800-1849 with the numbers from after year 2000. Pearson's Chi-squared test (X-squared = 4.534, df = 1) gave p-value = 0.03323. Cramer's V = 0.1.
[12] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 1220.505, df = 1) gave p-value < 2.2e-16. Cramer's V = 0.4.
[13] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 111.0614, df = 1) gave p-value < 2.2e-16. Cramer's V = 0.1.
[14] The statistical software package R provided warning messages for the Chi-squared test, so instead we employed Fisher's Exact Test, which works better for datasets involving small numbers. This test provided p-value = 1.146e-08. In order to obtain more reliable results we conflated the numbers for the second half of the 20th century and the 21st century.
[15] For unprefixed vs. non-syllabic prefixed verbs, Pearson's Chi-squared test with Yates' continuity correction (X-squared = 37.326, df = 1) gave p-value = 9.994e-10. For non-syllabic vs. monosyllabic prefixed, the same test (X-squared = 0.0034, df = 1) yielded p-value = 0.95.
[16] Notice that we use "transitive" in a wide sense so as to cover not only verbs with a direct object in the accusative, but also verbs like *dostignut'* 'reach' that govern the genitive case.
[17] Notice that for the purposes of the discussion semantic and syntactic factors we omitted all verbs with the postfix *–sja*, since this morpheme affects transitivity.
[18] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 705.122, df = 1) gave p-value < 2.2e-16. Cramer's V = 0.15.
[19] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 7.4912, df = 1) yielded p-value = 0.0062. Cramer's V = 0.05.
[20] We compared the numbers of *nu*-forms and Ø-forms for transitive verbs in the periods 1900-1949 and after year 2000. Pearson's Chi-squared test with Yates' continuity correction (X-squared = 7.5694, df = 1) gave p-value = 0.005937. Cramer's V = 0.05.
[21] Pearson's Chi-squared test with Yates' continuity correction (X-squared = 97.4567, df = 1) yielded p-value < 2.2e-16. Cramer's V = 0.1.
[22] In order to corroborate the conflated morphological hierarchy in (32), three additional statistical analyses were carried out. Comparison of the data for unprefixed masculines and prefixed participles (cf. Table 6 in section 4) shows that the difference is statistically significant, insofar as Pearson's Chi-squared test with Yates' continuity correction (X-squared = 18.7666, df = 1) yields p-value = 1.477e-05. However, Cramer's V = 0.05, which indicates that the effect size is far below the threshold of what is considered a small effect size. Comparison of gerunds (cf. Table 4) and unprefixed participles (cf. Table 6) indicates that the observed differences are not statistically significant. Pearson's Chi-squared test with Yates' continuity correction (X-squared = 2.9899, df = 1) gave p-value = 0.08. While Pearson's Chi-squared test checks the significance of individual factors, logistic regression incorporates all factors into one model, and therefore gives a more accurate picture of the interaction of the factors. Logistic regression indicates that there is a highly significant relationship between inflected form and prefixation on the one hand and the choice of Ø vs. /nu/ on the other. Consider the table below, which shows the odds ratio, 95%-Confidence Interval and the p-value for the significant predictors. The first four rows indicate that inflected forms are significant predictors of Ø vs. /nu/. Although the fifth row shows that prefixation per se is not significant, rows six and seven indicate that within the masculine sg and the participle, the difference between prefixed and unprefixed forms is highly significant. The bottom row shows that for gerunds the prefixed/unprefixed distinction does not apply, since all gerunds in our database are prefixed. (The statistical model was run both with and without the gerunds, and both versions gave the same results.)

Variable:	Odds ratio	95%-Confidence Interval		Pr(>|z|)	
FORMfinite non-masc (intercept)	6.20E-03	5.03E-03	7.54E-03	<2e-16	***
FORMgerund	1.27E+03	9.81E+02	1.66E+03	<2e-16	***
FORMmasc	5.45E+00	4.32E+00	6.94E+00	<2e-16	***
FORMpart	1.34E+01	1.08E+01	1.68E+01	<2e-16	***
PREFunprefixed	1.24E+00	6.89E-01	2.07E+00	0.442	

FORMmasc:PREFunprefixed	3.70E+00	2.02E+00	7.12E+00	4.24E-05	***
FORMpart:PREFunprefixed	1.47E+02	6.45E+01	3.73E+02	<2e-16	***
FORMgerund:PREFunprefixed	NA	NA	NA	NA	

Table 13: Statistical significance of the variables inflected form and prefixation and their interaction (data from the entire database, i.e. from 1800 to 2010)

[23] Three additional statistical analyses were carried out in order to establish the hierarchy in (33). First we compared the numbers for unprefixed masculines and prefixed participles after year 2000 (cf. Table 7 in section 4). Pearson's Chi-squared test with Yates' continuity correction (X-squared = 0.353, df = 1) showed that this difference is not statistically significant (p-value = 0. 5524). Second, we compared gerunds (cf. Table 5) and unprefixed participles after year 2000 (cf. Table 7). The statistical software package R gave a warning message for Pearson's Chi-squared test with Yates' continuity correction, and we therefore instead used Fisher's Exact Test, which is known to work better for datasets involving small numbers. This test gave p-value = 0.7, so the observed differences are clearly not statistically significant. In the same way as for hierarchy (32) we supplemented Pearson's chi-squared test and Fisher's Exact Test with logistic regression in order to get a better picture of the interaction of all the relevant factors. Logistic regression indicates that in the data from the 21st century there is a highly significant relationship between inflected form and prefixation on the one hand and the choice of Ø vs. /nu/ on the other. The table below shows the odds ratio, 95%-Confidence Interval and the p-value for the significant predictors. The results reported in this table are very similar to the those shown in the previous footnote, the only important difference being that the logistic regression analysis does not indicate a significant correlation between prefixation and masculine sg after year 2000. This is presumably due to the small number of unprefixed masculine forms in this period. However, since for masculine sg the percentage of Ø-forms has undergone little change since 1950 and Fisher's exact test shows that the difference between prefixed and unprefixed masculine sg forms is statistically highly significant for the period 1950-2010, we maintain our conclusion from section 4 (see footnote 14) that unprefixed masculine sg forms have not yet caught up with prefixed masculine sg forms with regard to *nu*-drop.

Variable:	Odds ratio	95%-Confidence Interval		Pr(>|z|)	
FORMfinite non-masc (intercept)	-6.429	0.3783	-16.996	<2.00E-16	***
FORMgerund	8.8238	0.4251	20.757	<2.00E-16	***
FORMmasc	2.2121	0.4235	5.224	1.75E-07	***
FORMpart	4.002	0.387	10.341	<2.00E-16	***
PREFunprefixed	1.1257	0.8035	1.401	0.16119	
FORMmasc:PREFunprefixed	0.2939	0.9455	0.311	0.75595	
FORMpart:PREFunprefixed	3.6039	1.0965	3.287	0.00101	**
FORMgerund:PREFunprefixed	NA	NA	NA	NA	

Table 14: Statistical significance of the variables inflected form and prefixation and their interaction (data from 21st century)

References

Bulaxovskij, L. A.: 1950, Istoričeskij kommentarij k russkomu literaturnomu jazyku, Kiev.

Bulaxovskij, L. A.: 1954, Russkij literaturnyj jazyk pervoj poloviny XIX veka, Moscow.

Černyšev, V. I.: 1915, Pravil'nosť i čistota russkoj reči. Izdanie 2-oe. Tom 2: časti reči, Petrograd.

Dickey, S. M.: 2001, '"Semelfactive" –nǫ and the Western Aspect Gestalt', Journal of Slavic Linguistics 9(1), 25-48.

Gorbačevič, K. S.: 1971, Izmenenie norm russkogo literaturnogo jazyka, Leningrad.

Gorbačevič, K. S.: 1978, Variantnosť slova i jazykovaja norma, Leningrad.

Graudina, L. K., Ickovič, V. A., and L. P. Katlinskaja: 1976, Grammatičeskaja pravil'nosť russkoj reči. Opyt častotno-stilističeskogo slovarja variantov, Moscow.

Graudina, L. K., Ickovič, V. A., and L.P. Katlinskaja: 2001, Grammatičeskaja pravil'nost' russkoj reči. Stilističeskij slovar' variantov, Moscow.
Graudina, L. K., Ickovič, V. A., and L. P. Katlinskaja: 2007, Slovar' grammatičeskix variantov russkogo jazyka. 3-e izdanie, Moscow.
Isačenko, A. V.: 1982, Die russische Sprache der Gegenwart. Formenlehre, Munich.
King, B. M. and E.W. Minium: 2008, Statistical Reasoning in the behavioral Sciences. Fifth ed., Hoboken, NJ.
Nesset, T.: 1998, Russian Conjugation Revisited, Oslo.
Nesset, T.: 2010, 'Suffix Shift in Russian Verbs: A Case for Markedness?', Russian Linguistics 34, 123-138.
Ožegov, S. I. and N. Ju. Švedova: 2005, Tolkovyj slovar' russkogo jazyka, Moscow.
Plungian, V. A.: 2000, ''Bystro' v grammatike russkogo i drugix jazykov' in Iomdin L. L. and L. P. Krysin (eds.), Slovo v tekste i v slovare: sbornik statej k semidesjatiletiju akademika Ju.D. Apresjana, Moscow, pp. 212-223.
Rozental', D. È.: 1977, Praktičeskaja stilistika russkogo jazyka, Moscow
Švedova, N. Ju. (Ed.): 1980, Russkaja grammatika. Tom 1, Moscow.
Timberlake, A.: 2004, A reference grammar of Russian, Cambridge.
Vinogradov, V. V. and N. Ju. Švedova. (eds.): 1964, Glagol, narečie, predlogi i sojuzy v russkom literaturnom jazyke XIX veka, Moscow
Zaliznjak, A. A.: 1980, Grammatičeskij slovar' russkogo jazyka, Moscow.
Zaliznjak, A. A. and A. D. Šmelev: 2000, Vvedenie v russkuju aspektologiju, Moscow.