A corpus-based study of English loanwords on –*ing* in Russian (2000-2011)

Benedikte Vardøy and Margje Post*

University of Bergen

Abstract

Until recently, large corpora of Russian had been lacking. This changed with the development of the Russian National Corpus (RNC) and the appearance of Integrum, a large commercial database that has shown to be a valuable tool for linguistic studies.

We have used both corpora to study the actual use of recent loanwords in Russian that have their origin in English -*ing* forms, such as *rejting* (*rating*) and *kasting* (*casting*). The RNC was used to compile a list of -*ing* words and their frequency. Integrum provided the data for a qualitative analysis of the semantics and possible developments between 2000 and 2011 of five -*ing* words. We also tried to use the corpora to study the changes in popularity of -*ing* words over time, which was possible to some extent. This corpus-based study of -*ing* words illustrates both possibilities and limitations of the two largest corpora for Russian, RNC and Integrum.

Keywords: corpus linguistics; loanwords; Russian

* Principal contact:
Margje Post, Associate Professor
Institute of Foreign Languages, University of Bergen, HF-bygget, 5020 Bergen, Norway
Tel.: +47 55 58 82 90
E-mail: margje.post@uib.no

1. Introduction

After perestroika, the number of English loanwords in Russian has increased enormously. A conspicuous group among them are loans from English ending in *-ing*, such as *miting* and *konsalting* (*meeting*, *consulting*). In his book entitled "Russian on the verge of a nervous breakdown",[1] Professor Maksim Krongauz complains about the excessive use of *-ing* words in Russian. They caught our attention after reading Viktor Grigor'ev's talk "The bright future of *-ing* forms in the language of Russian poetry" (2005).[2] Here, Grigor'ev experiments in combining Russian roots with *-ing*, as in *vining* "shopping tour for wine" and *xuliganing*[3] "hooliganism".

We wanted to know more about these *-ing* words in Russian, as we have chosen to call them. How is the *-ing* ending understood by Russians? Do the *-ing* words have common semantic characteristics? How many *-ing* words are actually in use, and are they frequent?

Both linguists and journalists have written about how *-ing* and words ending in *-ing* are used in Russian (Bobrova 1980; Sèšan 1996; Soelsuren 2003; Grigor'ev 2005; Krongauz 2007; Marinova 2008; Nikolaeva 2008), but they have not given these questions a satisfactory answer. There have been some earlier attempts to find common characteristics of *-ing* words. The Academy grammar *Russkaja grammatika* (Švedova et al. 1980) describes *-ing* as a suffix used to form verbal nouns with processual meaning. This might account for their English origin, but it is not very informative about their actual meaning in contemporary Russian. Marinova (2008) has studied the semantics of *-ing* words in the plural. Bobrova (1980) and Sèšan (1996) appear to have been more interested in semantic differences than similarities, and have studied their diverse thematic memberships. Bobrova (1980) and Sèšan (1996) do mention how many *-ing* words in Russian they have found, but they do not give actual lists of *-ing* words, and their numbers have become outdated (cf. section 5 below).

We also wanted to see whether individual *-ing* words have shown signs of any development since 2000, and to which degree the descriptions of the words in dictionaries of loanwords reflect their actual usage. Many loanwords are borrowed in only a specific meaning and undergo changes after entering a language. Apart from becoming more (or possibly less) frequent, their semantics and pragmatics can also continue to develop (cf. Ermakova 2008; Yoshizumi 2008; Marinova 2008).

In our attempt to answer these questions we have used two text corpora. Until recently, large corpora for the Russian language had been lacking. This changed with the development of the Russian National Corpus (RNC; available since 2003) and the database Integrum (available since 1996), a large, commercial database that was developed for other purposes, but that has been shown to be a valuable tool for linguistic studies.[4]

These two corpora will be described in section 3 and 4 below, after an introduction to Russian *-ing* forms (section 2). Section 5 is dedicated to the list of *-ing* words that was compiled on the basis of RNC, and section 6 to the semantics of Russian *-ing* words, based on data from Integrum and from dictionaries. Section 7 evaluates the ability of the two corpora to measure the frequency of *-ing* words. Finally, sections 8, 9 and 10 evaluate the possibilities and limitations of the corpora for studying Russian *-ing* words and make some suggestions for future research.

---

[1]*Russkij jazyk na grani nervnogo sryva*; Krongauz (2007).
[2]"Svetloe buduščee ingovyx form v russkom poètičeskom jazyke"; Grigor'ev (2005).
[3] We follow Comrie and Corbett's (1993) translitteration rules for Cyrillic.
[4]Other, smaller, corpora of Russian texts have been available for a long time; cf. the references on Nacional'nyj korpus russkogo jazyka (2003-2012), *Drugie korpusa.*

2. Russian *-ing* words

In our definition, Russian *-ing* words are loanwords from English that end in *-ing* and are ultimately derived from a verb form. Words like *konsalting*, *xolding* and *lizing* qualify as *-ing* words, as they have the English counterparts *consulting* (*to consult*), *holding* (*to hold*) and *leasing* (*to lease*). Words like *elling* and *string*, however, do not: *elling* is a loanword from Dutch,[5] while *string*, although being of English origin, does not contain any suffix *-ing* and cannot be associated with an English verb form.

In English, words ending in *-ing* have diverse functions. The Oxford English Dictionary (OED) contains four different *-ing* suffixes, which do not even have a common etymology. They can be used, among others, to create verbal nouns (gerunds), present participles and adjectives. In Russian, *-ing* words are exclusively used as nouns of the first declension, but, unlike their English counterparts, Russian *-ing* words can create derivatives, for instance, adjectives ending in *-ovyj* (*rejtingovyj*), verbs that end in *-ovat'* (*dempingovat'*) or nouns with a more abstract meaning than the *-ing* word, ending in *-ovanie* (*mitingovanie* "the activity of holding a (mass-) meeting". The *-ing* words can also be part of compounds, e.g. *trening-seminar*.

3. The Russian National Corpus

This and the following section describe our main sources, RNC and Integrum, and give illustrations of what they can be used for. Integrum and the RNC can complement each other in linguistic studies (Plungjan 2006, 83; Kopotev and Mustajoki 2008, 13-15).

The *Russian National Corpus* (*Nacional'nyj korpus russkogo jazyka*, RNC) was a long awaited project when it became available in 2003, although smaller corpora had been available for a long time (Nacional'nyj korpus russkogo jazyka (2003-2012), *Drugie korpusa*). The RNC is a tagged corpus, consisting of a main corpus and a number of subcorpora, including a corpus of newspaper texts.[6] The main corpus aims to be representative of the Russian standard language. Unlike many corpora, it contains transcriptions of oral speech, in addition to modern written texts (from the 1950s to the present day) and early texts (from the middle of the 18th to the middle of the 20th centuries). The texts have been morphologically tagged, and some of the homonyms have been disambiguated by hand.

The RNC is freely available online and easily accessible, without usernames and passwords. The RNC gives numerous search possibilities, online references, supporting sites, a discussion forum and much more; cf. ruscorpora.ru. One can search for words starting or ending in a certain letter combination. This made it possible to search for all words ending in *-ing*. The RNC is under continuous development. The corpora have been enlarged several times and new subcorpora and new search possibilities are added every now and then. The changes of content of the main corpus and the subcorpora has the drawback that searches can give different results when carried out at different times, but since January 2012, one can search in older versions of the corpora.[7] The number of linguistic studies based on the RNC is growing fast.[8]

---

[5]*Elling*, from Dutch *helling*, is a shipping term, meaning "slipway; boat ramp".

[6]RNC consists of the following subcorpora: a general, main corpus, a newspaper corpus, a corpus of spoken Russian, a corpus with stress marks, a poetry corpus, an educational corpus (a corpus of texts with disambiguated grammatical homonyms, adapted for the Russian school teaching program), several parallel corpora, a dialect corpus, a multimedia corpus, a deeply annotated corpus (containing sentences with full morphological and syntax structure markup) and a corpus with texts in Church Slavonic; cf. Nacional'nyj korpus russkogo jazyka (2003-2012), *Novosti proekta.*

[7]Cf. a news message, published on Nacional'nyj korpus russkogo jazyka (2003-2012), *Novosti proekta*.

[8]The website itself gives a bibliography of publications based on the RNC (Nacional'nyj korpus russkogo jazyka (2003-2012). *Publikacii po Nacional'nomu korpusu russkogo jazyka*), but the list has not been updated recently. Some examples of recent studies based on the RNC are B'ërklund (2010), Nesset and

3.1 The RNC newspaper corpus

For our study of recent *-ing* words we have chosen to use the newspaper subcorpus, which became available in 2010. There are several reasons for this choice. First, the language of the mass media is often the place where new loanwords are introduced to the standard language (Baš et al. 2012, 6; cf. Ferm 1994, 15; Kostomarov 1994, 88-89; Krysin 2004, 14). Second, the newspaper corpus contains far more texts from recent years (2000-2008) than the main corpus. As a consequence, it also contains far more occurrences of *-ing* words.

Whereas the main corpus contained 209 million words, from a period covering more than 250 years, the newspaper corpus of RNC contained 113 million words at the moment we used it,[9] from a period of only nine years. The texts of the newspaper corpus were published between 2000 and 2008 in four different newspapers and by three news agencies. The newspaper corpus is not part of the main corpus, as its inclusion – with its comparatively large number of newspaper texts from only a short period – would influence the representativeness of the main corpus for standard Russian in general. Queries for the word *bodibilding*, a recent loanword, illustrate this: whereas the main corpus gives 23 hits for *bodibilding*, the newspaper corpus contains more than 10 times as many occurrences of this word (245 hits; cf. section 7.3 below).[10]

We have used the RNC main corpus mainly to get an overview over the *-ing* words in Russian that are actually used in the 21st century; see section 5 below.

4. Integrum

Since the RNC is too small for some purposes, the database Integrum was used as well. Integrum is not a corpus that is made for linguistic research, but a very large, commercial, Russian electronic database with advanced search options, developed first of all as a source of information for e.g. tradesmen and journalists.

Its great advantage is its size. At present it contains, among others, the complete editions of more than 1200 central and local newspapers and periodicals appearing in Russia, some of them with all issues dating back from the early nineties (cf. Integrum 2012, *Database info*). It is a paid-for service, but researchers in Russia can get access at a reduced price. Integrum has a number of advantages over internet search engines. For instance, Integrum consists of complete texts and the number of misspellings and misprints is minimal (Mustajoki 2006). Furthermore, it provides much better search tools.

Integrum became available in 1996 and in the following years the searching possibilities were extended, e.g. searches for frequency and other statistics. Integrum is now used by researchers from various research fields (Nikiporec-Takigava 2006, 6-9). It can be used for a variety of purposes. In linguistics, it is an excellent tool for historical lexicography, as one can find, for instance, the first time a recent new word or phenomenon was mentioned in Russian press (Plungjan 2006, 84; Marinova 2008, 284). Integrum does not allow queries for specific morphological features, but it is possible to search for all forms of a certain word and for word combinations, for mentions in certain media and certain periods, and, since 2005, even for their frequency. More search options are listed in e.g. Nikiporec-Takigava (2006) and explained in Romanenko and Geršenzon (2006).

Kuznetsova (2011), Janda and Lyashevskaya (2011), Dobrovol'skij and Pëppel (2012) and a number of articles in Mustajoki et al. (2008) and Plungjan et al. (2009).

[9]It was extended to 173 million words in August 2012; cf. the news message of 08.08.2012 on Slavonic; cf. Nacional'nyj korpus russkogo jazyka (2003-2012), *Novosti proekta*.

[10]The searches of *bodibilding* were performed in October 2012, when the newspaper corpus had been enlarged from 113 to 173 million words. A search performed before August 2012 would have resulted in a lower number, but still many times larger than the number of hits in the main corpus.

Integrum gives several search options, among them the search engine *Artefakt*. *Artefakt* gives lists of search results organised by source, which can be useful if one's aim is to search in certain sources only, or to compare usage in several sources (Mustajoki 2006, 50-51). *Central press* covers over 200 newspapers published in Moscow, as opposed to *regional press* (over 1000 newspapers).[11] This option was used by Belikov (2007) in his study of the distribution of two widely used words for "cellular phone" in Russian, *sotovyj* and *mobil'nyj telefon*. Based on data from Integrum, he observed differences in popularity of the words over time and regional differences, *sotovyj* being more popular in Siberia than in newspapers published in Moscow (Belikov 2007). Apart from *Artefakt* we used the tool *Comparative and relative statistics*; cf. section 7.2 below. Examples of linguistic studies based on Integrum are found in Nikiporec-Takigava et al. (2006) and Mustajoki et al. (2008).

5. The list of *-ing* nouns

In order to get an overview of *-ing* words in Russian in use in the 21st century, we have composed a list of the *-ing* words that were actually used in the newspaper corpus of the RNC, covering the period 2000-2008.[12]

The list of *-ing* words is based on results from the query ***инг S** (further: *ing S). This query should provide a list of all nouns that end in *-ing* in the entire corpus. The query gave 69,640 words in 33,544 different documents. This search result had to be edited further. The RNC newspaper gives the option to select a certain amount of hits per document. The relative frequency of *-ing* words could be greatly influenced by a large amount of *-ing* words in a single document. Therefore, the search was restricted to only one hit per document.

Manual editing was needed in order to exclude the many hits of words that do not qualify as *-ing* words according to our definition. This was done in an Excel spreadsheet. Names and titles, e.g. *Notting (Hill)* and *boing* were removed manually, and also parts of names, e.g. *Xanting* from the film title *Umnica Uill Xanting* (*Good Will Hunting*). Other removed words are words of English origin that cannot be associated with a verb form, e.g. *string*, *ènising* (*anything*), and words of non-English origin, such as *viking* and *storting*. Some adjectives that appeared as results of the query *ing S, such as *rejtingovyj* and *mitingovyj* – which must be due to erroneous tagging – were also removed in the editing process. As a result, only 269 words in *-ing* remained. The query *ing S had resulted in 760 different words, so only 1/3 of them were actually *-ing* words of English origin. Since some words were names or parts of names in some, but not all cases, each single occurrence had to be checked, by performing additional searches on each word in the list. For example, the word *lizing* (*leasing*) occurred 570 times. In 56 of these occurrences, *lizing* is part of a name, often a company name. In the list of *-ing* words, *lizing* is therefore registered with 514 hits.

Some choices we made have had an impact on the results of our study. An example is the choice to limit the search results to one hit per document, for if two or more nouns in *-ing* occur in the same document, only one of them will be registered. Only after our analysis we found out that it is possible to exclude words starting with a capital letter from the search results.[13] This

---

[11] Cf. the list of sources published on Integrum (2012), *Database info*.

[12] As mentioned in section 1 above, Bobrova (1980) and Sèšan (1996) mention a certain number of *-ing* words, but their numbers are outdated, and they do not give complete lists. Bobrova reports that the Russian language has over 180 *-ing* words (Bobrova 1980), and Sèšan claims that there are at least 50 *-ing* words in addition to the 99 words she found in dictionaries (1996). Both Bobrova and Sèšan appear to include all words ending in *-ing*, not only those originating in English *-ing* forms.

[13] This option was found by accident, by selecting "capitalized word" under "additional features", and adding a hyphen in front of this option: -capital. It works for some options, e.g. -capital and -S, but not for the option "personal names" (famn | persn | patrn); on 31.10.2012.

would have taken out most names from the result list. At the same time, however, it would have excluded quite a lot of relevant *-ing* words at the beginning of sentences as well.

The edited list of *-ing* words contains 269 different *-ing* nouns from a total of 28,352 documents, and gives specifications on the number of documents in which each *-ing* word is mentioned.[14] The content of the list of *-ing* words is manifold. This concurs with the observation that nouns in *-ing* are found in different thematic fields (Bobrova 1980; Sèšan 1996; Soelsuren 2003; Marinova 2008; Nikolaeva 2008). The list contains many more *-ing* nouns than Bobrova (1980) and Sèšan (1996) have found earlier. This suggests that the number of *-ing* words in Russian has increased a lot over the last years, a tendency that is corroborated by the overall search for words ending in *-ing* in the RNC newspaper corpus (see section 7.1 below).

The total frequency of the three most frequent words, *rejting*, *xolding* and *miting*, is 15,162 documents. This means that 1.1 % of the words in the list make up 53.5 % of the entire number of documents where *-ing* words are mentioned. At the other end of the list, 89 out of 269 *-ing* words in the list are mentioned only once. These words make up 33 % of all words in the list, but only 0.3 % of the total number of documents containing *-ing* words. 119 words (44 %) are mentioned once or twice. The complete list can be accessed on Vardøy (2012b).

This shows us that the number of frequent *-ing* words is restricted, but that the group of infrequent *-ing* words must be large. Not all *-ing* words that are in use occur in the newspaper corpus. Since almost half of the *-ing* words occur only once or twice, they must be part of a much larger group of infrequent *-ing* nouns, and chance has played a role as to whether they occur in the corpus or not. It is questionable whether all of them can be said to be part of the Russian standard written language.

The list shows that the corpus is not large enough to contain all Russian *-ing* words. However, it also shows us that not all *-ing* words in use are frequent, and not all of them are taken up in the standard written language. A conspicuous group among these are the neologisms created with Russian roots: the list contains only two, incidental, *-ing* nouns with Russian roots (*vpendjuring* and *puting*). We had hoped to find more of them, as they are the most interesting *-ing* nouns from the point of view of semantics and morphological status of *-ing* as a derivational suffix in Russian: in these cases the ending is used actively by speakers of Russian.

6. Exploring the semantic properties of *-ing* words

The semantics of *-ing* words have been given little attention in the literature. We have used the database Integrum to study the use of five *-ing* words in detail, and compared our findings to the entries in English and Russian dictionaries. The following five words were chosen: *rejting* (*rating*), *kasting* (*casting*), *trening* (*training*), *bodibilding* (*body-building*) and *lifting* (*lifting*). These words are not representative for the entire group of *-ing* words, but by choosing thematically heterogeneous words, the study is not limited to a certain part of the vocabulary. All five words are among the 50 most frequent *-ing* words in the list.

As *-ing* words are often mentioned as a group, they might have common characteristics. Historically, they are all tied to verbal nouns (gerunds) or verb forms, so what they could have in common, is a tie to verbs. Apart from verbal characteristics, we have studied their connotations, possible semantic changes and secondary loans that were not mentioned in the dictionaries. In order to be able to study possible developments we have chosen to study use in two periods: in 2000 and in 2011. In both 2000 and 2011 a period was chosen so that the number of documents

---

[14]The query *ing S resulted in 760 different words, so only 1/3 of them are actually *-ing* words of English origin.

would be from 30 to 42. As a result of this, the periods were short for the most frequent words and longer for others.[15]

6.1 Five **-ing** words in English and Russian dictionaries

Loanwords are usually borrowed in a specific meaning (Ryazanova-Clarke and Wade 1999; Marinova 2008). A study of the entries of the five words in English[16] and Russian dictionaries[17] confirms that the words have far more meanings in English than in Russian. For instance, the meanings of *trening* registered in Russian dictionaries are related to improvement or development, while other meanings of the English words *training* or *to train* are absent in Russian. However, the words can develop more meanings in the target language than the one they started off in. They can develop new, but closely related meanings in the target language, either at the moment they are borrowed or at a later stage, or the word can be borrowed a second time in a new meaning (cf. Marinova 2008). Our studies of the five *-ing* nouns in Integrum and the dictionaries show signs of both types of new meanings. Some of these new meanings are mentioned in the dictionaries, others are found in the Integrum material only.

6.2 Closely related meanings

The development of closely related meanings can be noticed both in newspaper texts from Integrum and in Russian dictionaries, which mention some meanings that are absent for their English counterparts. An example mentioned by Baš et. al (2012) for *rejting* is "degree of popularity". Another meaning that is absent from English dictionaries is *kasting* – the act of selecting girls, e.g. for fashion shows, at beauty competitions (Bunimovič 2002; Krysin 2001; 2009; 2010).

6.3 Exploring the semantics of **-ing** words in Integrum

Integrum shows that most of the words have developed several closely related meanings, with different verbal characteristics. Some of them are countable and can be used in the plural, while others cannot; cf. section 6.5 below. These subtle differences are not always noticed in the Russian dictionaries.

6.4 New meanings: reborrowing

In the material in Integrum from 2000, *lifting* usually has the meaning "face-lift" or "face-lifting operation". In the contexts from 2011, we found that *lifting* also was used in two very different meanings: lifting of vehicles and lifting as part of a restoration method, whereby parts of a wooden building are lifted up in order to replace rotten logs of wood. This suggests that *lifting* has been borrowed several times, in distinct meanings.

6.5 Verbal characteristics

The study of the semantic characteristics of the five *-ing* words shows that some of the *-ing* words have retained some properties of verbal nouns in their original meaning of denoting the activity expressed by the verb, but in many cases, the meaning of verbal activity is distant. Four out of the five words occur as plural nouns, which means that they are used in a meaning not denoting an activity. The *-ing* words have an English verbal origin, but they have varying ties to verbal actions. Many of them are already used as a noun in English. We have looked up the 50 most frequent *-ing* words in our list in two English dictionaries (OED and Chambers). This shows that the majority of the English nouns corresponding to the Russian *-ing* words already have an

---

[15]*Rejting* was examined over a period of two days both in 2000 and 2011, whereas *lifting* was only mentioned in 16 documents in the entire year of 2000, which resulted in a smaller foundation for the study of its semantics, while the period that was studied in 2011 was May 1st to December 31st (36 documents).

[16] Chambers; OED.
[17] Krysin (2001); (2009); (2010); BISIS (2002); Bunimovič (2002); Zenovič (2002); Šagalova (2009); Baš et al. (2012).

entry as a noun in English, with a meaning different from verbal nouns. Therefore, *-ing* words usually do not originate directly from verb forms or gerunds.

Our study shows that the *-ing* words vary greatly as to which verbal characteristics they retain. One of the oldest *-ing* words in Russian, *miting*, usually refers to an event; *dansing* and *parking* are actually not activities at all, but objects, as they have their origin in *dancing hall* and *parking lot*. The five *-ing* words we have studied can show verbal properties in several ways: they can denote an activity or process, a single event or the result of an activity or process. Most of the words can denote several of these, in closely related meanings. An event and a result are countable, but most activities are not. Four out of the five words occur as plural nouns, which means that they are used in a meaning that is distant from verbal nouns, not denoting an activity.

Bodibilding is the only noun that has only the meaning of an activity or process, and is not used in plural in in our data from Integrum.[18] *Kasting* most often denotes an event, but it can also have a meaning of a process or the result of a process. It is often used in the plural. *Lifting*, "face-lift" or "face-lifting operation", also denotes a result or an event in most cases. *Trening* can in some cases express a number of events as a whole: a training course. *Rejting* mainly refers to the result of a verbal action, and, in consequence, has fewer qualities as a verbal noun than the other four words. The derivative *rejtingovanie* is used both in the RNC newspaper corpus and in Integrum. Russian nouns in *-ie* are deverbatives: abstract nouns of action or the result or product of an action (Townsend 1975, 152). Therefore, *rejtingovanie* can express the meaning of process and activity that *rejting* lacks.

6.6 Development of meaning in the period 2000-2011: more plurals

Kostomarov (1994, 192) and Sèšan (1996, 47) wrote that *-ing* nouns are usually only used in the singular. Marinova (2008, 195-196) shows that this is not the case in Russian today. She analysed 165 *-ing* nouns, and 75.8 % of these words appeared in the plural. Marinova claims that all abstract nouns in Russian today have a tendency to develop from abstract noun to more concrete usage, and *-ing* words seem to follow this general process. In our material, we found plural forms of all words except for *bodibilding*. We also found that more meanings of the *-ing* words *rejting* and *kasting* occur in the plural in 2011 than in 2000,[19] and the use of *trening* in the plural increased a lot in this period. This might be a sign of the process of becoming more concrete, mentioned by Marinova (2008).

6.7 An understanding of modernity

The dictionaries we studied do not provide information on usual connotations. We have searched for connotations in different contexts in Integrum. There seems to be a common understanding of the *-ing* words, and the objects and concepts they refer to, as modern, but all five *-ing* words can also be used without connotations. In our material, the five *-ing* words were related to phenomena that were only recently introduced into Russian society. According to Ryazanova-Clarke and Wade, the most important reason for borrowing in the period after perestroika was the introduction of a new concept in need of a word (1999, 138). Associations with modern phenomena can be evaluated both positively and negatively. In the majority of contexts, the five *-ing* words were not felt to have specifically positive or negative connotations, but their evaluation was more often positive than negative (cf. Vardøy 2012a).

6.8 Summary

Do the *-ing* words have common characteristics, or are their differences predominant? The study of the actual use of five *-ing* words in the data from Integrum shows that all five *-ing* words appear to have retained verbal characteristics, but in different ways. They tended to have

---

[18]Marinova (2008) mentions some other meanings, e.g. "bodybuilder".

[19]For a detailed account of the meanings of *rejting* and *kasting* in plural, cf. Vardøy (2012a, 81-82, 89).

developed closely related meanings, with a more distal relation to verbal actions: e.g. from activity to event, or from activity to the result of an activity. Not all meanings are mentioned in the Russian dictionaries. They tend to get more concrete meanings, and the use of plural forms has increased for most of the words from 2000 to 2011. The Russian words have developed meanings slightly different from their English counterparts. Some loanwords appear to have been borrowed more than once, in different meanings (e.g. *lifting*). All of the words refer to recently imported phenomena in Russian society.

7. Historical development of frequency

An interesting question as regards loanwords is their popularity over time: when were they used for the first time, when did they become more popular, did they become less frequent again, and how did their popularity develop relative to their near synonyms?

We have tried to use the main corpus and the newspaper corpus of the RNC and Integrum to study some aspects of the frequency development of -*ing* words. First, we studied the overall historical development of words on -*ing* in the RNC main corpus (7.1). We have also tried to use the RNC newspaper corpus and Integrum to study the changes in popularity of the five words we studied in more detail: *rejting*, *kasting*, *trening*, *bodibilding* and *lifting* (cf. section 6 above), and compared the popularity of *bodibilding* with its near synonym *kul'turizm*. This was possible to some degree (7.2 and 7.3).

7.1 Overall frequency of words ending in -ing after 1800 (RNC)

Within a few seconds, the main corpus of the RNC can show the popularity of all words ending in -*ing* that do not start with a capital letter for the period 1800-2010:

1) search for *инг S -capital. This results in 20,922 examples (31.10.2012);[20]

2) choose the function "*raspredelenie po godam*" ("annual distribution"). This results in Figure (1), which gives the number of hits per million words.

---
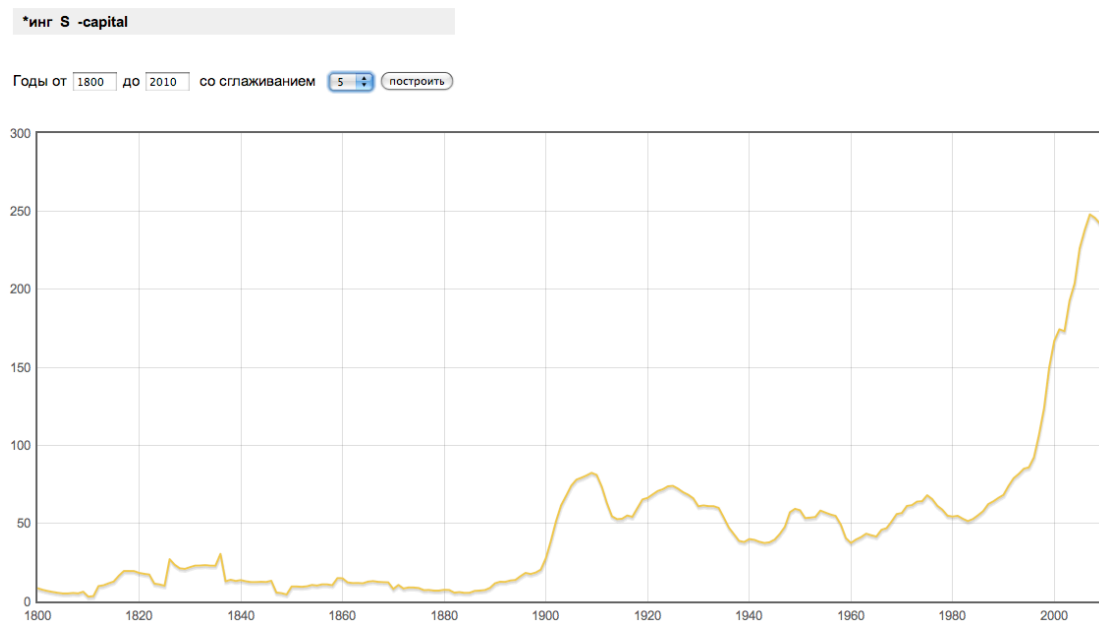
[20]We took out words with an initial capital, because the overwhelming majority of them were names. They make up approx. 1/3 of the total number of words ending in -*ing*: The query *инг S gives 32,727, so,11 805 of these start with a capital.

Figure 1. Number of nouns ending in *-ing* that do not start with a capital letter in the main corpus between 1800-2010, per million words.

The figure shows that the *-ing* words have gone through several stages of popularity. They have been part of the language at least since the 1830s – the first occurrence of *miting* in the corpus dates from 1839. They became more popular in the first half of the 20th century, with a peak around 1925, and experienced a real boost after perestroika. However, one should keep in mind that this figure includes a high percentage of nouns ending in *-ing* with an different origin than an English verb form (e.g. *šilling*, *sterling*, *landsgevding*), even though most names are taken out. This accounts for almost all occurrences in the 19th century.[21] The numbers vary much for individual words, as the main corpus can show us when using the same tools for the analysis of individual *-ing* nouns. Obviously, most *-ing* words turned up only after perestroika, but some *-ing* words have been part of the Russian language for a long time, such as *miting*. *Dansing* "dancing hall" is an example of a word that turned up in the 1930s, was at its most popular in the period 1940-1970, but is now almost out of use.

The newspaper corpus is better suited to studying recent loanwords, as it contains higher numbers of *-ing* words per annum, but it only covers the period 2000-2008, and lacks the search option of the main corpus for the annual distribution. The annual distribution can still be calculated, but not as easily (cf. Vardøy 2012a, 62-63).

7.2 Using RNC and Integrum for individual words

As part of our analysis of the usage of the five *-ing* words, we also tried to study their popularity development by using the RNC newspaper corpus and Integrum. RNC was found to be a good tool for the most frequent word, *rejting*. A pilot study of the relative frequency of *rejting* and *lifting* in the newspaper corpus (Fig. 2) – the most frequent and the least frequent among these five words – shows that most of the *-ing* words are not frequent enough in the newspaper corpus to give a meaningful overview of changes in frequency on a yearly basis. The numbers for *rejting* are based on a sufficient number of hits (N = 13,292), but *lifting* has far too few occurrences in the newspaper corpus (N = 48). For instance, in 2001, *lifting* was used only once in the corpus texts.

	2000	2001	2002	2003	2004	2005	2006	2007	2008
rejting	56.01	62.91	60.69	73.49	120.85	135.64	129.29	151.90	166.97
lifting	0.27	0.11	0	0.94	0.57	0.68	0.43	0.28	0

Figure 2. Relative frequency of *rejting* (13,292 hits) and *lifting* (48 hits) in the period 2000-2008 in the RNC newspaper corpus, in hits per million words.[22]

Integrum is a more reliable tool for studies in historical lexicography (cf. section 4 above), especially for low-frequency words, as it provides higher numbers, but they are still not very high for all words (cf. Vardøy 2012a, 64, Table 4.5 for the actual numbers). Fig. 3 provides these numbers in columns.

---

[21]As remarked upon in section 5 above, a manual check of the list showed that almost 500 out of the 760 words on *-ing* in the newspaper corpus had to be excluded, but this list still included words starting with a capital, and those 500 *-ing* words take up a much smaller part of the total number of hits, because they are less frequent than the *-ing* words of English verbal origin.

[22]These numbers are based on the absolute number of occurrences, and not on the number of documents the words are used in, so a single newspaper article with a large number of occurrences of an infrequent word like *lifting* has a huge effect on the relative frequency of the word in a certain year; cf. Vardøy (2012a, 62-63).
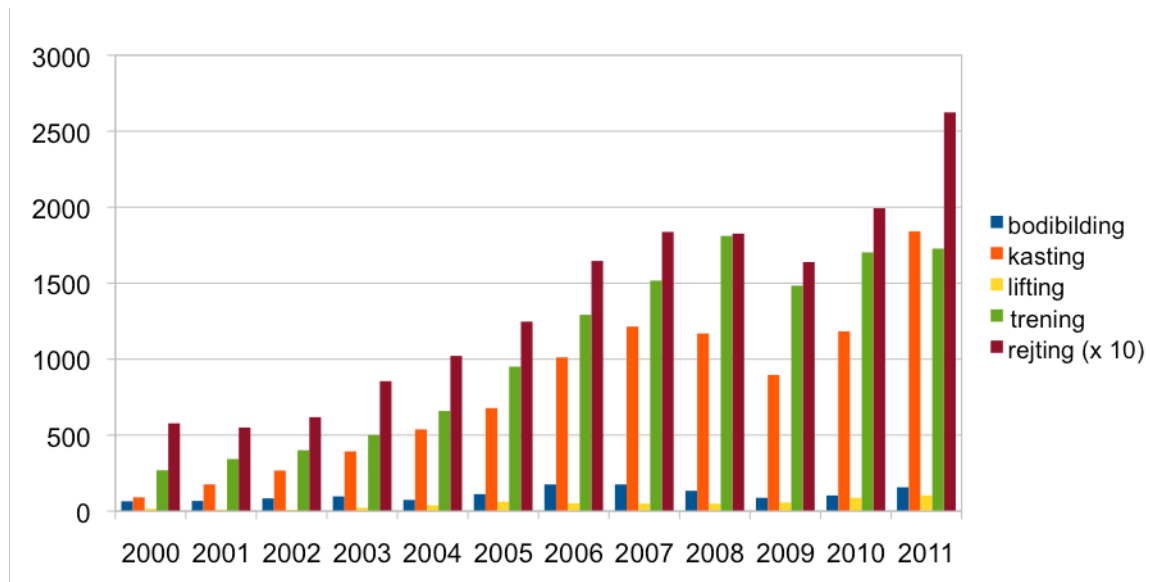
Figure 3. Absolute frequency in Integrum (*central press*) for five -*ing* words. The numbers for *rejting* must be multiplied by 10.

Fig. 3 gives an impression of the development of the five words. Note, however, that the figure does not give the number of occurrences per million words, like the RNC, but the absolute number of occurrences in Integrum. As the number of words in Integrum might vary from year to year, and probably increases, the figure does not necessarily show that *kasting*, *rejting* and *trening* have become much more frequent over the last years, but it does show their popularity, and the changes in popularity, compared to each other. For instance, *kasting* has become much more frequent over this short period of 11 years, and its popularity has increased much more than the popularity of the other words: it is used 12 times more in 2008 than in 2000, and 20 times as often in 2011 (cf. Vardøy 2012a, 64, Table 4.5 and 4.6).

7.3 Comparing the popularity of two near synonyms
RNC and Integrum also give possibilities to compare the popularity of two near synonyms. The study of the usage of *bodibilding* in Integrum showed that it is sometimes used together with its near synonym *kul'turizm*. Its non-English origin suggests that *kul'turizm* is an older loanword in Russian than *bodibilding*. What is their relative popularity, and has it changed in favour of the -*ing* word?

First, we searched for these two words in the RNC, as mentioned in section 3.1 above. The main corpus contains 23 occurrences of the word *bodibilding*; the newspaper corpus 245 (searches performed in October, 2012).[23] The corresponding numbers for *kul'turizm* are 10 in the main corpus and 16 in the newspaper corpus. These numbers suggest the declining popularity of *kul'turizm*. This is confirmed by a simple search in Integrum. Figure 4 gives the absolute frequency of *kul'turizm* and *bodibilding* in Integrum for the last 20 years, showing that the number of hits for *kul'turizm* is fairly stable, but that the absolute frequency of *bodibilding* has gone up enormously. It shows clearly that *bodibilding* has become more popular relative to *kul'turizm*.

---

s[23]In fact, a search performed before August 2012 would have resulted in a lower number, as the search of *bodibilding* was performed in October 2012, after the enlargement of the newspaper corpus to 173 millions words.
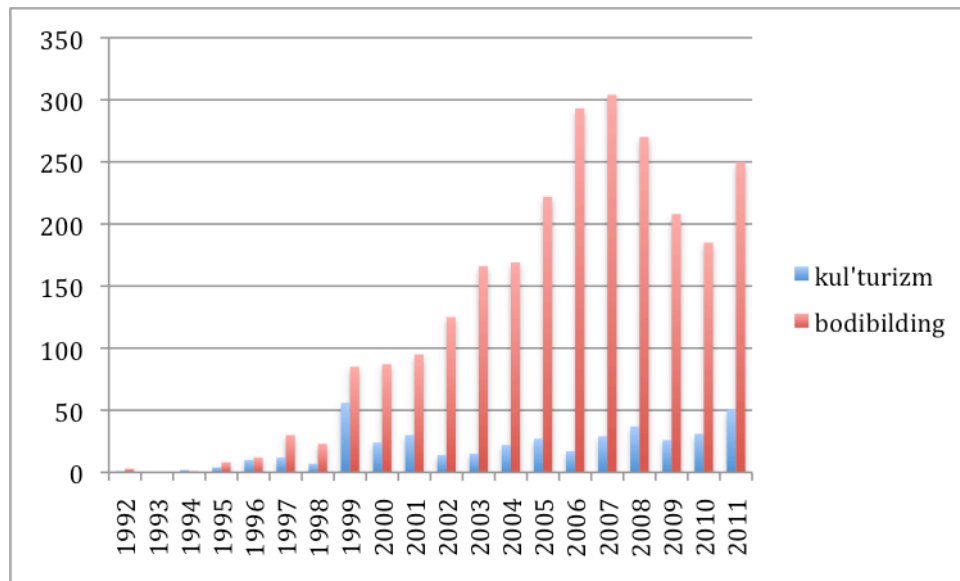
Figure 4. The absolute frequency of the synonyms *kul'turizm* and *bodibilding*, based on Integrum, *central press.*

8. Results

The development of large text corpora has enabled the study of the actual use of recent loanwords, and of their possible semantic development over the last years. After perestroika, many English words ending in *-ing* have entered the Russian language, such as *rejting* and *konsalting* (*rating*, *consulting*). Based on the RNC newspaper corpus we were able to provide a list of *-ing* words that are used in newspapers. Manual editing was needed, however, in order to exclude the numerous words ending in *-ing* that did not comply with our definition. Integrum provided the possibility to study a wider range of contexts where *-ing* words are used. Both RNC and Integrum provide tools to study changes in frequency over time, which we could use to some extent. The study exemplifies how RNC and Integrum can complement each other in linguistic studies.

9. Questions RNC and Integrum could not answer

The RNC is not large enough for all purposes. We wanted to use it for a study of changes in the frequency of individual words, but most of the words were not frequent enough in the corpus for this purpose. It also gave too few contexts to enable a comparison of actual contexts of the words from two different years. The main corpus can show developments over larger periods of relatively frequent words, such as *miting*. We also used it to analyse the development of all words ending in *-ing*, but these contain too many words that are names or do not originate from English *-ing* forms.

Only recently we found out that the RNC gives the possibility to *exclude* certain characteristics, like many other corpora, e.g. the RuN corpus of Russian and Norwegian parallel texts.[24] However, this option appears not to work for all features. It does work for the features *noun* and *capitalized*, but not for queries after personal names. It is not unlikely that this problem will be solved in the future, given the high speed in which the RNC is developing.

Integrum has advanced search tools, but it does allow queries for *-ing* words in general, as a search like **ing* is not possible, and one cannot search for morphological features. Because of its

---

[24] The RuN Corpus (and its extension: The RuN-Euro Corpus) (2011).

size, it might contain more *-ing* words with Russian roots, but one has to search for each individual word. Integrum can also calculate frequency. Its drawback is that is gives absolute numbers, not hits per million, like the RNC. Integrum has a relative frequency tool, but it was not applicable to our purposes.

An interesting result of our study is that *-ing* words with Russian roots appear to be almost absent from Russian newspapers and journals. The RNC newspaper corpus contains only a few, incidental occurrences, but not, for instance, *šubing*, a word with more than 10,000 hits on the internet.[25] In this case, the size of the corpus is not the main reason for their low frequency: A search for *šubing* in Integrum for the last 12 months gave 0 hits in the over 200 newspapers covered by *central press*, and only two occurrences in a single newspaper article in the over 1000 newspapers of *regional press*.[26]

The lack of *-ing* words with Russian roots shows that these words, although they are paid much attention to, are much less frequent than most of their brothers and sisters of English origin. This confirms previous claims in the literature that these words are not part of the standard language used in newspapers, or, at most, only a marginal part of it. In fact, neither Bobrova (1980) nor Sèšan (1996) report any *-ing* words made with Russian roots, and Sèšan (1996) even claims that such derivation does not take place. Marinova (2008), however, mentions some sporadic examples, but remarks that they have not been taken up by the standard written language (2008, 454).

It also shows that Integrum, although it is a large database, is not representative for all usage of the Russian language. Not all words used in informal language enter the standard written language. For these words one could use search engines for the internet, but they are far less user-friendly for linguistic research (e.g. Mustajoki 2006; Romanenko and Geršenzon 2006, 21).

10. Suggestions for future research

RNC and Integrum have been shown to provide interesting data about the use of the recent loanwords ending in *-ing* in Russian. In future studies, the same methodology could be used for the study of other neologisms. As to the *-ing* nouns themselves, the results from our study could be compared to studies from other text corpora. Comparison of our list of *-ing* words with lists based on other source material would give a broader insight into which *-ing* words are used and which ones are most frequent in the Russian newspaper language or in other genres. In order to get a broader image of the semantics of *-ing* words, more *-ing* words and more aspects of their semantics should be studied, especially *-ing* words with Russian roots, to find out if these words have signs of verbal characteristics, connotations to modernity, or any other common characteristics. Their use should also be studied in other genres, e.g. in other subcorpora of the RNC, such as the corpus of oral Russian. The analysis of the semantics of *-ing* words was based on a rather low number of contexts for each word, so we need to look for a higher number of contexts in other, and longer, periods than the ones we looked into in Integrum.

---

[25]A simple Google search for шубинг resulted in 13,900 hits; the Russian search engine Rambler provided 10.000 hits for "шубинг"; searches performed on 27.10.2012. Although the number of individual occurrences of the word is, of course, much lower, it must still be substantial. *Šubing* is associated with, and a hyponym of, *šopping*: it means a shopping tour (*šopping*) for fur coats (*šuby*). *Puting*, a demonstration in support of Putin, is also made by analogy: it is a combination of *Putin* and *miting*.
[26]In *Večernjaja Murmansk* of 26.10.2012; search performed on 29.10.2012 for the period 30.10.2011–12.11.2012.

References

B'ërklund [Björklund], Martina (2010). Konstrukcii s po + datel'nyj padež v značenii «sfera celenapravlennoj dejatel'nosti». *Scando-Slavica, 56 (1)*, 27-47.

Baš, L. M., A. V. Bobrova, G. L. Večeslova, R. S. Kimjagarova, E. M. Sendrovic (2012). *Sovremennyj slovar' nostrannyx slov*. Moscow: Veče.

Belikov, V. I. (2007). S sotovogo na mobil'nyj. In *Jazyk v dviženii. K 70-letiju L. P. Krysina*, edited by E. A. Zemskaja et al., 61-67. Moscow: Jazyki slavjanskoj kul'tury.

BISIS = Bol'šoj illjustrirovannyj slovar' inostrannyx slov (2002). Moscow: AST Russkie slovari.

Bobrova, Alina V. (1980). Suščestvitel'nye na «*-ing*» v russkom jazyke. *Russkij jazyk v škole*, *1980 (3)*, 87-91.

Bunimovič, N. T. (2002). *Slovar' sovremennyx ponjatij i terminov*. Moskva: Respublika.

Chambers = Chambers 21st Century Dictionary, online version, accessed in July 2012. www.chambers.co.uk.

Comrie, Bernard and Greville G. Corbett, eds. (1993). *The Slavonic Languages.* London – New York: Routledge.

Dobrovol'skij, Dmitrij and Ljudmila Pëppel [Pöppel] (2012). Revoljucija, vosstanie, perevorot: semantika i pragmatika. *Scando-Slavica*, *58 (1)*, 77-100.

Ermakova, Ol'ga P. (2004). Aktivnye processy v leksike i semantike. In *Sovremennyj russkij jazyk: Aktivnye processy na rubeže XX – XXI vekov*, edited by Leonid P. Krysin, 33-99. Moscow: Jazyki slavjanskix kul'tur.

Ferm, Ljudmila (1994). *Osobennosti razvitija russkoj leksiki v novejšij period (na materiale gazet)*. Uppsala: Almqvist & Wiksell.

Grigor'ev, Viktor (2005). Svetloe buduščee ingovyx form v russkom poètičeskom jazyke, http://www.vavilon.ru/textonly/issue13/grigoriev.html.

Integrum = Integrum World Wide (1998-2008). http://www.integrumworld.com.

Integrum (2012). *Database info*. http://aclient.integrum.ru/topics.aspx?si=cOkVDr2E&qu=0. [Accessed 31.10.2012]

Janda, Laura and Olga Lyashevskaya (2011). Aspectual Pairs in the Russian National Corpus. *Scando-Slavica 57 (2)*, 201-215.

Kostomarov, Vitalij G. (1994). *Jazykovoj vkus èpoxi*. Moscow: Pedagogika-Press.

Krysin, Leonid P. (2001). *Tolkovyj slovar' inojazyčnyx slov*. Moscow: Russkij jazyk.

Krysin, Leonid P. (2004). *Russkoe slovo, svoe i čužoe*. Moscow: Jazyki slavjanskoj kul'tury.

Krysin, Leonid P. (2009). *1000 novyx inostrannyx slov*. Moscow: AST-PRESS KNIGA.

Krysin, Leonid P. (2010). *Tolkovyj slovar' inojazyčnyx slov*. Moscow: Eskimo.

Kopotev, M. V. and A. Mustajoki (2008). Sovremennaja korpusnaja lingvistika. In *Instrumentarij rusistiki: korpusnye podxody. Slavica Helsingiensia 34*, edited by Arto Mustajoki et al., 7-24. Helsinki: Department of Slavonic and Baltic Languages and Literatures. Also available from http://www.helsinki.fi/slavicahelsingiensia/preview/sh34/pdf/index.htm.

Marinova, Elena V. (2008). *Inojazyčnye slova v russkoj reči konca XX – načala XXI v.: problemy osvoenija i funcionirovanija*. Moscow: ELIPS.

Mustajoki, Arto (2006). The Integrum Database as a Powerful Tool in Research on Contemporary Russian. In *Integrum: Točnye metody i gumanitarnye nauki,* edited by G. Nikiporec-Takigava, 50-75. Moscow: Letnij sad.

Mustajoki, Arto et al., ed. (2008). *Instrumentarij rusistiki: korpusnye podxody. Slavica Helsingiensia* 34. Helsinki: Department of Slavonic and Baltic Languages and Literatures. Also available from http://www.helsinki.fi/slavicahelsingiensia/preview/sh34/pdf/index.htm.

Nacional'nyj korpus russkogo jazyka (2003-2012). http://www.ruscorpora.ru.

Nacional'nyj korpus russkogo jazyka (2003-2012). *Drugie korpusa,* http://ruscorpora.ru/corpora-other.html.

Nacional'nyj korpus russkogo jazyka (2003-2012). *Novosti proekta*, http://www.ruscorpora.ru [Accessed 23.10.2012]

Nacional'nyj korpus russkogo jazyka (2003-2012). *Publikacii po Nacional'nomu korpusu russkogo jazyka*, http://www.ruscorpora.ru/corpora-biblio.html [Accessed 23.10.2012]

Nikiporec-Takigava, G., ed. (2006). *Integrum: Točnye metody i gumanitarnye nauki.* Moscow: Letnij sad.

Nikiporec-Takigava, G. (2006). *Integrum: Točnye metody i gumanitarnye nauki.* In *Integrum: Točnye metody i gumanitarnye nauki,* edited by G. Nikiporec-Takigava, 6-13. Moscow: Letnij sad.

Nesset, Tore and Julia Kuznetsova (2011). Stability and Complexity: Russian Suffix Shift over Time. *Scando-Slavica , 57 (2)*, 268-289.

Nikolaeva, Irina (2005). Polnyj šoking. *Moskovskij komsomolec*, May 5, 2008, p. 33. Available from Integrum, aclient.integrum.ru.

OED = *Oxford English Dictionary*, online version, accessed in July 2012. www.oed.com.

Plungjan, Vladimir A. (2006). «Integrum» i Nacional'nyj korpus russkogo jazyka v lingvističeskix issledovanijax. In *Integrum: točnye metody i gumanitarnye nauki*, edited by G. Nikiporec-Takigava, 76-84. Moscow: Letnij sad.

Plungjan, V. A., E. V. Raxilina and T. I. Reznikova, eds. (2009). *Nacional'nyj korpus russkogo jazyka: 2006 — 2008. Novye rezul'taty i perspektivy.* St. Petersburg, Nestor-Istorija. Available from http://ruscorpora.ru/corpora-biblio-2008.html.

Rambler, http://www.rambler.ru. [Accessed 27.10.2012]

RNC = The Russian National Corpus, *Nacional'nyj korpus russkogo jazyka* (2003-2012). Last modified 08.08.2012. http://ruscorpora.ru.

Romanenko, Fedor and Lev Geršenzon (2006). IPS «Integrum». Istorija sozdanija, opisanie, ispol'zovanie. In *Integrum: Točnye metody i gumanitarnye nauki,* edited by G. Nikiporec-Takigava, 19-49. Moscow: Letnij sad.

The RuN Corpus (and its extension: The RuN-Euro Corpus) (2011). Last modified 30.03.2011. http://www.hf.uio.no/ilos/english/research/projects/run/corpus/

Ryazanova-Clarke and Terence Wade (1999). *The Russian Language Today*. London and New York: Routledge.

Sèšan, Šarmila (1996). Suščestvitel'nye na *-ing* – simvol amerikanskoj jazykovoj ekspansii? *Russkaja reč', 1996 (3)*, 46-49.

Soelsuren, Baasanžav (2003). *Inojazyčnye slova s zaimstvovannymi suffiksami v sovremennom russkom jazyke: Na materiale slov s suffiksami -až, -er, -ing*. Avtoreferat kand. diss. Gosudarstvennyj institut russkogo jazyka im. Puškina, Moscow, http://www.dissercat.com/content/inoyazychnye-slova-s-zaimstvovannymi-suffiksami-v-sovremennom-russkom-yazyke-na-materiale-sl.

Šagalova, E. N. (2009). *Slovar' novejšyx inostrannyx slov*. Moscow: AST Astrel'.

Vardøy, Benedikte Fjellanger (2012a). *Ing-ord sin semantikk: Ein kvalitativ og kvantitativ studie av tydingar og tydingsutvikling hjå ing-ord i russiske aviser og i ordbøker (2000-2011).* Unpublished Master's thesis, University of Bergen.

Vardøy, Benedikte Fjellanger (2012b). *Russian words in -ing*, http://russianingwords.blogspot.com

Yoshizumi, Olga (2006). Anglicisms in Russian after Perestroika. PhD diss., Brown University, Providence, Rhode Island.

Zenovič, E. S. (2002). *Slovar' inostrannyx slov i vyraženij*. Moscow: Agenstvo KRPA Olimp.

Švedova, N. Ju. et al., eds. (1980). *Russkaja grammatika* I – II, Moscow: Nauka.