How derivational links affect lexical access: evidence from Russian verbs and nouns

Natalia Slioussar[1,2], Anastasia Chuprina[2]

[1] *School of Linguistics, Higher School of Economics, Staraya Basmannaya street 21/4, Moscow, 105066, Russia* <slioussar@gmail.com>
[2] *Department of Liberal Arts and Sciences, St. Petersburg State University, Galernaya street 58/60, St. Petersburg, 190000, Russia* <a.o.chuprina@gmail.com>

We report two lexical decision experiments analyzing Russian prefixed and unprefixed verbs and deverbal nouns. These experiments show that the summed frequency of direct derivatives influences access time to the base word. We demonstrate that this effect cannot be explained by semantic or phonological similarity and is not due to the fact that derived words undergo morphological decomposition during lexical access.

KEYWORDS**:** morphological processing, derivational links, morphological decomposition, Russian

1. Introduction

In this paper, we report two lexical decision experiments on Russian. We study how words derived from a given word influence lexical access to this word, showing that this influence is not due to decomposition. In the following section, we provide some background information on lexical access to morphologically complex words, and then outline the purpose of the present study.

1.1. How morphologically complex words are accessed and stored in the mental lexicon

One of the most discussed questions in the domain of lexical access concerns morphologically complex words: are they decomposed into morphemes or accessed as a whole? Most studies are dedicated to inflectional morphology, but derivational morphology is widely analyzed as well.

Existing models can be divided into three groups. Firstly, some authors assume full-form storage for all words; these forms are used for initial access, and then constituent morphemes may be activated and may influence postlexical processing (e.g., Butterworth 1983; Giraudo & Grainger 2000; Mannelis & Tharp 1977; Rueckl et al. 1997). Secondly, some models presuppose that all morphologically complex words are decomposed (e.g., Taft & Forster 1975; Taft 1979, 2004). They do not necessarily exclude full-form representations. For example, in Taft & Forster's (1975) model these representations exist, but cannot be found directly: they are organized into 'file drawers', which can be accessed via the root morpheme.

Thirdly, there are dual-route models positing two routes of access to the mental lexicon for morphologically complex words (e.g., Caramazza, Laudanna, & Romani 1988; Deutsch, Frost & Forster 1998; Dominguez, Cuetos & Segui 2000; Pollatsek, Hyönä & Bertram 2000; Baayen, Dijkstra & Schreuder 1997; Baayen & Schreuder 1999; Schreuder & Baayen 1995). These models differ between each other in many ways: in particular, the two routes can be activated serially or simultaneously, the order of activation or the outcome of competition depending on different factors. Finally, connectionist models (e.g., Seidenberg & McClelland 1989) do not assume morphological representations at all, but rely on multi-level (formal and semantic) distributed representations for all words.

The majority of experimental studies aiming to discriminate between these models examined whether the time to identify a morphologically complex word depends on the whole word frequency or on the frequency of the constituent morphemes. The former can be taken as a piece of evidence for whole-word storage, while the latter would support decomposition. Among the factors that influence access to morphologically complex derived words, the role of orthographic and semantic transparency was analyzed (e.g. Juhasz 2007; Pollatsek & Hyönä 2005; Zwitserlood 1994).

A number of studies are dedicated to the interplay between different frequency counts and word length. For example, Burani & Thornton (2003), in their study of Italian derivational morphology, demonstrated that for words with highly frequent roots, morpheme frequency plays a role, but when root frequency is lower, whole word frequency matters. Several authors showed that for shorter complex words, whole word frequency plays a more important role, while for longer ones, root frequency does (e.g. Bertram & Hyönä 2003; Hyönä 2012; Niswander-Klement & Pollatsek 2006). They argue that long words, composed of eight letters or more, typically require more than one eye-fixation, and the duration of the first fixation is influenced by the frequency of the first morphemic constituent.

However, studies of derivational morphology are not limited to complex words. Schreuder & Baayen (1997) found that simplex words that have many derivatives (a big morphological family) are accessed faster, and termed this facilitation "family size effect". This effect was analyzed in a number of subsequent experiments (e.g. Bertram, Baayen & Schreuder 2000; De Jong et al. 2002; Lüdeling & De Jong 2002).

In particular, Moscoso del Prado Martín et al. (2004), analyzing Finnish, Dutch and Hebrew, showed that the family size effect is present in all three languages. However, in Finnish, unlike in Dutch and Hebrew, it is not the complete morphological family of a word that codetermines lexical access time, but only the subset of words directly derived from this word, or its "dominated family". This is explained by the fact that morphological families tend to be very large in Finnish, while they are relatively small in Dutch and very small in Hebrew. For example, in Finnish there are about 7000 words derived from *työ* 'work' (e.g. *työkalu* 'tool', *työskennellä* 'to work', *urotyö* 'heroic deed'). But what matters for a word like *työläinen* 'worker' is the number of words like *käsityöläinen* 'craftsman' or *työläisluokka* 'working class', which are directly derived from it (possibly via composition as, e.g., in the last quoted example). Notably, all authors mentioned above believe that the family size effect is semantic in nature, i.e. it is explained by the fact that morphologically related words are usually semantically related.

1.2. The present study

In Russian, new words can be derived by prefixation (e.g. *prygat'* 'to jump' — *zaprygat'* 'to start jumping') or by suffixation (e.g. *prygat'* 'to jump' — *prygnut'* 'to jump once', *pryžok* 'jump'). Prefixation never changes the part of speech or the inflectional class a word belongs to, while suffixation may change the former and often changes the latter, especially for verbs. There is also a morpheme *-sja* called postfix, where the term stresses the fact that it is attached after inflectional morphemes, the postfix being a former enclitic pronoun (e.g. *brosat'* 'to throw' — *brosat'sja* 'to throw oneself'). It is used only with verbs and does not change their inflectional class.

The meaning of a derivative usually cannot be fully predicted. Consider the following examples. When the prefix *po-* is added to the verb *bež*at› 'to run', we get *pobež*at› 'to start running'. When it is added to *pet'* 'to sing', we get *popet'* 'to sing for a while'. Adding it to *rodit'* 'to give birth' yields *porodit'* 'to generate, to produce', adding it to *kinut'* 'to throw' yields *pokinut'* 'to abandon'. Some verbs, like *isčeznut›* 'to disappear' do not allow *po*-prefixation at all. Orthographically, the boundary between the prefix and the stem is salient, with prefixes being mostly monosyllabic. The boundary between stems and suffixes may be less salient.

In our study, we focused on the following cases. Russian has many deverbal nouns derived from unprefixed and prefixed verbs, for example *roždenie* 'birth' from *rodit'* 'to give birth' or *poroždenie* 'generation, production' from *porodit'* 'to generate, to produce'. Now let us look at the pair of verbs *rodit' — porodit'* and the pair of nouns *roždenie — poroždenie*. The relation between the verbs and the nouns is the same if semantics or constituent morphemes are considered, and, as we mentioned above, this is what is used to explain the role of morphological family in lexical access. But the verbs are connected by a direct derivational link (prefixed verbs are derived from unprefixed ones), while the nouns are not (prefixed deverbal nouns are derived from prefixed verbs, not from unprefixed deverbal nouns).[1]

In Experiment 1, our goal was to find out whether this difference plays a role. If it does, the family size effect cannot only be due to semantic similarity or morphemic overlap, but it rather requires that derivational links are represented in the mental lexicon and may be activated during lexical access. Having proved that this is the case, we demonstrated in Experiment 2 that the difference we observed was not due to the decomposition of prefixed words.

Thus, in Experiment 1 we aimed to determine whether having many prefixed counterparts (we cannot call them 'derivatives' because, in the case of nouns, there is no derivational link) would influence lexical access time to unprefixed verbs and deverbal nouns. It must be noted that instead of relying on family size, i.e. counting the number of prefixed words, we used a different measure in this study: the summed frequency of such words. According to the *Grammatical dictionary of the Russian language* (Zaliznyak 1977), many Russian verbs have numerous prefixed derivatives. However, only a fraction of them is frequent enough to be included in the *Frequency dictionary of modern Russian language* (Lyashevskaya & Sharoff 2009). Moreover, the *Grammatical dictionary* is not exhaustive: it does not contain some prefixed verbs that are used only occasionally. To avoid introducing some arbitrary threshold (for example, based on the word list of the *Frequency dictionary* or the *Grammatical dictionary*) in order to decide which prefixed words should or should not be counted in, we decided to rely on the summed frequency instead.

2. Experiment 1

2.1. Method

2.1.1 Participants

27 native speakers of Russian (19-52 years old, 20 females) volunteered to take part in Experiment 1. They were given no information about the specific purpose of the study.

2.1.2. Materials

We wanted to compare lexical decision times for unprefixed verbs and nouns that differ in terms of summed frequency of the prefixed counterparts. We selected nouns ending in *-anie*/*-enie* (denoting processes and events) for our study to have a homogeneous group of deverbal nouns. For verbs, we also counted prefixed derivatives with the reflexive postfix *-sja*.

Firstly, we took all verbs and all nouns ending in *-anie*/*-enie* from the *Frequency dictionary of modern Russian language* (Lyashevskaya & Sharoff 2009). Prefixes cannot be stripped off automatically: for example, in the verb *potet'* 'to sweat' *po-* is not a prefix, but part of the root. So together with Varvara Magomedova (SUNY, Stony Brook) we created program such that it stripped off initial segments that could be prefixes from all verbs and nouns in our list. Every time a potential prefix was stripped off, the program searched for the resulting sequence in the list to check whether it was a real word or a nonsense string (like **tet'* un case of *potet'* 'to sweat'). In the former case, the words were added to the list of prefixed-unprefixed pairs.[2] Many Russian words have two prefixes and some even have three or four, but we ran the program only twice because we needed approximate results. The summed frequencies of prefixed counterparts were then counted for every verb or noun.

Based on these data, we preliminarily selected verbs and nouns that were closely matched in frequency, length and CV structure, but sharply differed in terms of summed frequency of the prefixed counterparts. For all preselected stimuli, the last parameter was manually rechecked to avoid potential mistakes. For example, when the program strips off *so-* from the verb *solit'* 'to salt', the result is a real verb *lit'* 'to pour', but these verbs are not connected derivationally — *solit'* is derived from the noun *sol'* 'salt'.

The final list of stimuli included 18 triplets of unprefixed imperfective verbs and 12 pairs of unprefixed deverbal nouns. Detailed examples are provided in Tables 1 and 2. The list of all stimuli can be found in Appendix A.

Table 1. An example of verb triplet from Experiment 1.

GROUP	1	2	3
TARGET VERBS	*torčat'* (86.3) 'to stick out'	*dyšat'* (90.8) 'to breathe'	*platit'* (89.0) 'to pay'
PREFIXED and POSTFIXED VERBS	*protorčat'* (2.0) 'to stick out for a while'	*dyšat'sya* (3.2) 'to breathe in a certain way' *zadyšat'* (4.5) 'to start breathing' *nadyšat'sya* (1.8) 'to fill oneself with breathing' *otdyšat'sya* (12.2) 'to recover normal breathing' *podyšat'* (7.7) 'to breathe a little'	*vyplatit'* (6.5) 'to pay off' *zaplatit'* (43.6) 'to pay for smth' *oplatit'* (12.2) 'to make a payment' *otplatit'* (1.7) 'to pay back' *poplatit'sya* (3.4) 'to come into a debt' *rasplatit'sya* (12.4) 'to pay in full' *uplatit'* (6.5) 'to pay out'
summed pref. freq.	2.0	29.4	86.3

Table 2. An example of noun pair from Experiment 1.

GROUP	1	2
TARGET NOUNS	*roždenie* (98.5) 'birth'	*javlenie* (94.3) 'apparition'
PREFIXED and POSTFIXED NOUNS	*vozroždenie* (26.4) 'revival' *zaroždenie* (4.3) 'origin' *poroždenie* (5.1) 'production'	*vyjavlenie* (22.4) 'revelation' *zajavlenie* (109.0) 'declaration' *ob'javlenie* (29.8) 'announcement' *pojavlenie* (82.5) 'appearance' *pred'javlenie* (8.2) 'presentation' *projavlenie* (45.3) 'display'
summed pref. freq.	35.8	297.2

Verbs from every triplet were assigned to three groups (with low, medium and high summed frequency of prefixed counterparts). Nouns were assigned to two groups. Mean word frequency, length and summed frequency of prefixed counterparts for every group can be found in Tables 3 and 4 in the next section.

Finally, we created nonce words that closely matched the characteristics of the real stimuli: e.g. *gratit', dolot', bamanie, prunenie*. All nonce words were phonotactically licit from the point of view of Russian phonology. As a result, we had 54 real verbs, 54 nonce verbs, 24 real nouns and 24 nonce nouns in our experiment: 156 stimuli and fillers in total.

2.1.3. Procedure

We conducted a lexical decision experiment on a PC using *E-Prime* software (www.pstnet.com). Participants were asked to press one button if a letter string they saw on the computer screen was a Russian word and another button if it was not. After five practice trials, the main session began. In the beginning of every trial, an asterisk was shown in the middle of the screen (the duration was randomly chosen out of 1.0, 1.1, 1.2, 1.3, 1.4 or 1.5 seconds, so that participants would not know when exactly a stimulus would appear). Then a letter string appeared for 500 ms or until a response button was pressed. Verbs were presented in infinitive, nouns in the nominative singular form. The order of trials was randomized for every participant. If no button was pressed, participants saw a blank screen for up to 2 s. After a response was given or after these 2.5 s were over, the next trial began.

2.2. Results and discussion

We analyzed the participants' question-answering accuracy and reaction times. All participants gave at least 85% of correct answers (92.4% on average). We excluded any trial with incorrect answers from further analysis. We also discarded all RTs that exceeded 1.5 s or were less than 100 ms, as is customary in many such studies (e.g. Alegre & Gordon 1999). In total, 0.3% of reactions to real stimuli were discarded.

Average RTs for different groups of verbs and nouns are given in Tables 3 and 4. Using *IBM SPSS* software, repeated-measures ANOVAs were computed on participant mean RTs across items (F1) and on item means across participants (F2). The analysis revealed that RTs for verbs differ significantly depending on the summed frequency of corresponding prefixed verbs ($F1(2,52) = 8.66$, $p < 0.01$; $F2(2,34) = 4.99$, $p = 0.01$), but RTs for nouns do not ($p > 0.6$).

Table 3. Average RTs for verb stimuli in Experiment 1.

GROUP	MEAN FREQ. (IMP)	MEAN LENGTH (LETTERS)	MEAN SUMMED PREF. FREQ. (IPM)	MEAN RT (MS)
1	40.1	6.5	11.0	644.7
2	41.1	6.5	43.5	632.3
3	41.1	6.3	139.0	607.6

Table 4. Average RTs for noun stimuli in Experiment 1.

GROUP	MEAN FREQ. (IMP)	MEAN LENGTH (LETTERS)	MEAN SUMMED PREF. FREQ. (IPM)	MEAN RT (MS)
1	33.1	7.7	60.3	637.8
2	31.9	7.4	220.6	635.4

We believe that these results have the following explanation. The stems of the majority of Russian prefixed verbs and nouns are likely to be stored as a whole because even relatively transparent ones tend to have some aspects of meaning that cannot be predicted compositionally. Still, prefixed verbs have close connections with their unprefixed counterpart in the mental lexicon due to direct derivational links and therefore influence lexical access to it. Prefixed deverbal nouns (e.g. *poroždenie* 'generation, production') are not connected to their unprefixed counterpart (e.g. *roždenie* 'birth') in a similar way due to lack of derivational links, so the summed frequency of such prefixed nouns does not influence lexical access to the unprefixed noun. The fact that in terms of semantics or constituent morphemes the relation between prefixed and unprefixed verbs and nouns is the same shows that the role of derivational links cannot be reduced to these factors.

However, an alternative explanation can also be suggested: prefixed verbs are decomposed (and thus boost the frequency of their unprefixed counterpart), while the results for nouns are inconclusive. If they are decomposed, lexical access time should depend on the frequency of the prefixed verbs they are derived from, but we did not control for it. Experiment 2 was conducted to rule out this explanation.

3. Experiment 2

3.1. Method

3.1.1. Participants

24 native speakers of Russian (18-50 years old, 18 females) who did not participate in Experiment 1 volunteered to take part in Experiment 2. They were given no information about the specific purpose of the study.

3.1.2 Materials

We selected 18 pairs of prefixed verbs and 12 pairs of prefixed nouns whose unprefixed counterparts were analyzed in Experiment 1. In every pair, length, CV structure and frequency of the unprefixed counterpart were closely matched, while whole word frequency was different. Stimuli from every pair were assigned to two groups (with low and high word frequency). Examples of verb and noun stimuli are provided in Tables 5 and 6. The list of all stimuli can be found in Appendix B.

Table 5. An example of verb pair from Experiment 2.

GROUP	1	2
TARGET VERBS	*otplatit'* (1.7) 'to pay back'	*podyšat'* (7.7) 'to breathe a little'
UNPREFIXED VERBS	*platit'* (89.0) 'to pay'	*dyšat'* (90.8) 'to breathe'
SUMMED PREF. FREQ.	86.3	29.4

Table 6. An example of noun pair from Experiment 2.

GROUP	1	2
TARGET NOUNS	*poroždenie* (5.1) 'production'	*projavlenie* (45.3) 'display'
UNPREFIXED NOUNS	*roždenie* (98.5) 'birth'	*javlenie* (94.3) 'apparition'
SUMMED PREF. FREQ.	35.8	297.5

Moreover, we took care of the following. If verbs like *podyšat'* 'to breathe a little' and *otplatit'* 'to pay back' shown in Table 5 are accessed as a whole, their word frequency should matter, and consequently *podyšat'* (group 2) will be accessed faster. Now let us assume that they are decomposed, and so are many other prefixed verbs. Then lexical access time should not depend on the frequency of unprefixed *dyšat'* 'to breathe' and *platit'* 'to pay', but on the frequency of these verbs plus the summed frequency of their decomposed derivatives. As Table 5 shows, this value is greater for *platit'* than for *dyšat'*, so *otplatit'* (group 1) will be accessed faster. This was true for all other prefixed verb pairs in Experiment 2, so that the whole word access vs. decomposition scenarios would always give different predictions.

As the example in Table 6 shows, we could not find prefixed noun pairs with a similar distribution of frequencies in our materials. As it happens, however, no approach predicts that they could be decomposed by first stripping off their prefix. As a consequence, noun stimuli were included in order to make experimental materials more diverse, rather than to tease apart different lexical access scenarios. Mean word frequency, length and frequency of the unprefixed counterpart for every stimulus group can be found in Tables 7 and 8 in the next section.

We created nonce words that closely matched the characteristics of the real stimuli: e.g. *otgratit', zadolot', prebamanie, poprunenie*. All stimuli were prefixed, so nonce words contained pseudoprefixes. All nonce words were phonotactically licit from the point of view of Russian phonology. As a result, we had 36 real verbs, 36 nonce verbs, 24 real nouns and 24 nonce nouns in our experiment, 120 stimuli and fillers in total.

3.1.3. Procedure

The procedure was the same as in Experiment 1.

3.2. Results and discussion

We analyzed participants' question-answering accuracy and reaction times. All participants gave at least 85% of correct answers (92.0% on average). As in Experiment 1, we excluded trials with incorrect answers and all RTs that exceeded 1.5 s. In total, 0.4% of reactions to real stimuli were discarded.

Average RTs for the different groups of stimuli are given in Tables 7 and 8. Using repeated-measures ANOVAs (*IBM SPSS* software) we demonstrated that RTs for both verbs and nouns differed significantly depending on their word frequency ($F1(1,23) = 17.87$, $p < 0.01$, $F2(1,17) = 5.98$, $p = 0.03$ for verbs; $F1(1,23) = 21.27$, $p < 0.01$, $F2(1,11) = 7.88$, $p = 0.02$ for nouns).

Table 7. Average RTs for verb stimuli in Experiment 2.

GROUP	MEAN FREQ. (IMP)	MEAN LENGTH (LETTERS)	CORRESPONDING UNPREFIXED VERB	MEAN RT (MS)
1	2.0	8.3	high summed pref. frequency	704.0
2	16.3	8.3	low summed pref. frequency	670.0

Table 8. Average RTs for noun stimuli in Experiment 2.

GROUP	MEAN FREQ. (IMP)	MEAN LENGTH (LETTERS)	CORRESPONDING UNPREFIXED NOUN	MEAN RT (MS)
1	12.4	9.9	low summed pref. frequency	688.4
2	76.4	9.6	high summed pref. frequency	657.5

The results are indicative of the whole word lexical access. We can conclude that prefixed verbs influence lexical access to their unprefixed counterparts not through decomposition, but because they are closely connected in the mental lexicon via derivational links.

4. Conclusions

This paper focuses on the question of how morphologically related words are connected in the mental lexicon and how these connections may influence lexical access. A number of previous studies discussed so-called family size effect: words with large morphological families are accessed faster than words with smaller families (e.g. Bertram, Baayen, & Schreuder 2000; De Jong et al. 2002; Lüdeling & De Jong 2002; Moscoso del Prado Martín et al. 2004; Schreuder & Baayen 1997). In particular, Moscoso del Prado Martín et al. (2004) demonstrated that in Finnish, a morphologically rich language, unlike in English and Hebrew, it is not the complete morphological family of a word that codetermines lexical access time, but only the subset of words directly derived from this word. The family size effect is usually explained by semantic and phonological similarity between morphologically related words.

We conducted two lexical decision experiments with Russian prefixed and unprefixed verbs and deverbal nouns. They were chosen as stimuli because the relation between unprefixed and prefixed verbs and nouns is very similar in terms of semantic and phonological similarity: for example, *rodit'* 'to give birth' — *porodit'* 'to generate, to produce', *roždenie* 'birth' — *poroždenie* 'generation, production'. However, prefixed verbs are directly derived from their unprefixed cognates, while prefixed deverbal nouns are derived from prefixed verbs, rather than from unprefixed deverbal nouns.

In Experiment 1, we demonstrated that the summed frequency of prefixed counterparts influences lexical access time for unprefixed verbs, but not for unprefixed nouns. Therefore, the role of derivational links cannot be fully reduced to semantic and phonological similarity. In Experiment 2, we demonstrated that the effect found was not due to the decomposition of prefixed words. Since our results demonstrate that the actual words derived from a given word, rather than the morphological family as a whole, influence lexical access in Russian, they can be taken to support Moscoso del Prado Martín et al.'s (2004) previous findings for Finnish and suggest that similar patterns may be found in other morphologically rich languages with highly productive derivational processes.

In the present paper, we did not specify the mechanisms by which derivationally related forms are connected in the mental lexicon and how these connections are formed. In dual route models, it can be suggested that although decomposition normally does not win in cases such as the prefixed verbs and nouns analyzed in this paper, it may nevertheless take place. This mechanism could be used to create and support derivational links.

To solve these and other problems, many crucial questions that are outside the scope of the present study need to be answered. Which derivatives influence the lexical access time of a base word and to what extent? What is the role of semantic transparency and phonological similarity between the derivative and the base word? How important is it for their connection whether they belong to one part of speech or to one inflectional class? Would stress shifts and alternations influence our results? We hope to address some of these questions in our further research.

Appendix A. *Materials for Experiment 1*

Target words (unprefixed verbs and nouns) are given in triplets and pairs in which word frequency and length are balanced, while the summed frequency of corresponding prefixed words differs. Frequency counts rely on Lyashevskaya & Sharov's (2009) dictionary and are given in ipm (items per million).

TARGET	TRANSLATION	WORD FREQ.	SUMMED PREF. FREQ.	LENGTH	GROUP
torčat'	to stick out	86.3	2.0	7	1
dyšat'	to breathe	90.8	29.4	6	2
platit'	to pay	89.0	86.3	7	3
drožat'	to tremble	81.6	13.8	7	1
brosat'	to throw	83.7	70.5	7	2
terjat'	to lose	79.2	247.2	6	3
žalit'	to sting	71.4	2.6	6	1
guljat'	to walk	70.9	35.1	6	2
katit'	to roll	72.4	148.1	6	3
varit'	to boil	64.1	34.1	6	1
rešat'	to decide	64.9	66.4	6	2
lovit'	to catch	63.6	112.0	6	3
prygat'	to jump	57.4	14.0	7	1
kormit'	to feed	55.0	47.6	7	2
menjat'	to change	58.7	249.9	6	3
brodit'	to wander	52.0	7.7	7	1
šutit'	to joke	55.3	27.4	6	2
taščit'	to drag	56.7	206.3	6	3
šumet'	to make noise	44.6	18.2	6	1
pilit'	to saw	47.2	45.2	6	2
prjatat'	to hide	43.8	131.4	7	3
gudet'	to buzz	35.3	16.4	6	1
boltat'	to chat	36.9	39.2	7	2
xranit'	to keep	35.1	114.2	7	3
taskat'	to carry	33.7	12.3	7	1
tratit'	to waste	32.8	63.7	7	2
rezat'	to cut	34.5	194.8	6	3
gremet'	to rattle	30.5	19.3	7	1
šeptat'	to murmur	31.4	58.2	7	2
pugat'	to scare	31.9	130.6	6	3

TARGET	TRANSLATION	WORD FREQ.	SUMMED PREF. FREQ.	LENGTH	GROUP
zvenet’	to ring	27.1	13.8	7	1
temnet’	to darken	28.0	27.2	7	2
kačat’	to pump	28.1	96.6	6	3
kipet’	to seethe	22.0	10.4	6	1
*č*istit’	to clean	24.1	50.4	7	2
davit’	to crush	24.7	116.0	6	3
pljasat’	to dance	22.5	8.3	7	1
gladit’	to caress	25.0	42.5	7	2
krutit’	to twist	27.2	88.5	7	3
treščat’	to crack	21.7	9.7	7	1
kidat’	to throw	22.9	52.0	6	2
delit’	to give	22.5	254.7	6	3
rydat’	to cry	18.3	7.8	6	1
xvalit’	to praise	19.0	21.1	7	2
rubit’	to chop	19.6	60.2	6	3
glotat’	to swallow	18.3	2.2	7	1
belet’	to whiten	18.5	10.5	6	2
kopat’	to dig	19.6	57.9	6	3
migat’	to wink	17.4	3.6	6	1
putat’	to mix	16.7	80.5	6	2
gadat’	to guess	17.1	128.4	6	3
pylat’	to flare	17.0	2.1	6	1
šipet’	to hiss	16.3	16.7	6	2
kružit’	to turn	16.4	78.2	7	3
roždenie	birth	98.5	35.8	8	1
javlenie	apparition	94.3	297.5	7	2
davlenie	pressure	67.2	8.8	8	1
obščenie	communication	64.4	88.9	7	2
zvanie	title	47.2	163.0	6	1
učenie	study	46.7	224.2	6	2
myšlenie	thinking	40.8	29.7	8	1
padenie	falling	41.6	50.4	7	2
stradanie	sorrow	39.8	10.7	9	1
hranenie	storage	39.2	51.3	8	2
stroenie	building	35.5	127.5	8	1
pravlenie	administ	35.4	481.6	9	2
terpenie	patience	22.0	16.5	8	1
suždenie	thinking	20.5	88.2	8	2

TARGET	TRANSLATION	WORD FREQ.	SUMMED PREF. FREQ.	LENGTH	GROUP
pisanie	writing	15.6	110.9	7	1
delenie	division	14.1	384.2	7	2
vraščenie	twisting	10.7	106.6	8	1
selenie	settlement	8.4	200.6	7	2
sečenie	section	8.3	16.0	7	1
vlečenie	attraction	6.7	77.9	8	2
kreplenie	fastening	6.6	44.5	9	1
hoždenie	walking	7.3	100.1	8	2
veščanie	broadcast	5.2	54.2	7	1
nošenie	wearing	5.1	601.8	7	2

Appendix B. *Materials for Experiment 2*

Target words (prefixed verbs and nouns) are given in pairs in which the frequency of the corresponding unprefixed words and length are balanced. In group 2, the summed frequency of prefixed words is greater than in group 1. The frequency of the target word itself is greater in group 1 for verbs and in group 2 for nouns. Frequency counts rely on Lyashevskaya & Sharov's (2009) dictionary and are given in ipm (items per million).

TARGET	MEANING	WORD FREQ.	LENGTH	UNPREF. FREQ.	SUMMED PREF. FREQ.	GROUP
podyšat'	to breathe a little	7.7	8	90.8	29.4	1
otplatit'	to pay back	1.7	9	89.0	86.3	2
zadrožat'	to start trembling	13.8	9	81.6	13.8	1
uterjat'	to come to a loss	6.7	7	79.2	247.23	2
poguljat'	to have a walk	16.0	8	70.9	35.1	1
otkatit'	to roll away	1.0	8	72.4	148.1	2
svarit'	to boil down	14.2	7	64.1	34.1	1
slovit'	to catch down	2.8	7	63.6	112.0	2
nakormit'	to feed up	19.3	9	55.0	47.6	1
razmenjat'	to break down	2.5	9	58.7	249.9	2
pošutit'	to make a joke	24.1	8	55.3	27.4	1
dotaščit'	to drag up	2.8	8	56.7	206.3	2
vypilit'	to saw out	32.5	8	47.2	45.2	1
zaprjatat'	to hide down	4.2	9	43.8	131.4	2
zagudet'	to start buzzing	11.9	8	35.3	16.4	1
naboltat'	to chat a lot	1.0	9	36.9	39.2	2
utratit'	to waste up	25.4	8	32.8	63.7	1
izrezat'	to cut all down	1.4	8	34.5	194.8	2
zagremet'	to start to rattle	13.2	9	30.5	19.3	1
raspugat'	to scare away	1.2	9	31.9	130.6	2
potemnet'	to darken a little	15.6	9	28.0	27.2	1
otkačat'	to pump out	1.2	8	28.1	96.6	2
očistit'	to clean out	16.0	8	24.1	50.4	1
otdavit'	to crush up	1.4	8	24.7	116.0	2
pogladit'	to caress a little	27.3	9	25.0	42.5	1
vykrutit'	to twist out	1.1	9	27.2	88.5	2
zatreščat'	to start to crack	9.7	9	21.7	9.7	1
obdelit'	to give less	2.9	8	22.5	254.7	2

TARGET	MEANING	WORD FREQ.	LENGTH	UNPREF. FREQ.	SUMMED PREF. FREQ.	GROUP
poxvalit’	to praise a little	16.2	9	19.0	21.1	1
porubit’	to chop a little	1.6	8	19.6	60.2	2
pobelet’	to turn white	10.5	8	18.5	10.5	1
nakopat’	to dig out	1.2	8	19.6	57.9	2
sputat’	to mix up	10.5	7	16.7	80.5	1
vygadat’	to guess	1.1	8	17.1	128.4	2
zašipet’	to start to hiss	9.0	8	16.3	16.7	1
vskružit’	to turn sb’s head	1.1	9	16.4	78.2	2
poroždenie	production	5.1	10	98.5	35.8	1
projavlenie	display	45.3	10	94.3	297.5	2
podavlenie	suppression	8.8	10	67.2	8.8	1
soobščenie	message	74.5	9	64.4	88.9	2
prizvanie	calling	11.2	9	47.2	163.0	1
izučenie	studying	75.0	8	46.7	224.2	2
izmyšlenie	figment	1.04	10	40.8	29.7	1
napadenie	assault	27.5	9	41.6	50.4	2
sostradanie	sympathy	10.7	11	39.8	10.7	1
sohranenie	conservation	46.7	10	39.2	51.3	2
postroenie	lining	42.5	10	35.5	127.5	1
upravlenie	administration	256.5	10	35.4	481.6	2
neterpenie	impatience	16.5	10	22.0	16.5	1
obsuždenie	discussion	51.3	10	20.5	88.2	2
napisanie	scripting	9.0	9	15.6	110.9	1
otdelenie	separation	90.2	9	14.1	384.2	2
otvraščenie	revolt	17.6	10	10.7	106.6	1
naselenie	population	176.5	9	8.4	200.6	2
presečenie	suppression	7.3	10	8.3	16.0	1
uvlečenie	attraction	21.1	9	6.7	77.9	2
zakreplenie	fastening	7.7	11	6.6	44.5	1
prohoždenie	passing	15.6	11	7.3	100.1	2
zaveščanie	will	10.9	9	5.2	54.2	1
sootnošenie	correlation	36.8	11	5.1	601.8	2

Notes

[1] Similar examples can be found in English: e.g. *reformation* is derived from *reform* by adding a suffix, rather than from *formation* by adding a prefix. Notably, English prefix *re-* never attaches to nouns, while Russian prefix *po-* can do so (e.g. *bereg* 'coast' – *poberež'e* 'coastal area', *les* 'forest' – *poles'e* 'forest area').
[2] There were some additional complications. For example, Russian has prefixes like *po-* and *pod-,* or *na-* and *nad-*, so it was important to decide in which order the program strips them off etc. We will not go into these details here.

Bibliographical references

Alegre Maria & Gordon Peter 1999. Frequency effects and the representational status of regular inflections. *Journal of memory and Language* 40. 41-61.

Baayen R. Harald & Schreuder Robert 1999. War and peace: morphemes and full forms in a Noninteractive Activation Parallel Dual-Route Model. *Brain and Language* 68. 27-32.

Baayen R. Harald, Dijkstra Ton & Schreuder Robert 1997. Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language* 37. 94-117.

Bertram Raymond & Hyönä Jukka 2003. The length of a complex word modifies the role of morphological structure: evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language* 48. 615-634.

Bertram Raymond, Baayen R. Harald & Schreuder Robert 2000. Effects of family size for complex words. *Journal of Memory and Language* 42. 390-405.

Burani Cristina & Thornton Anna M. 2003. The interplay of root, suffix and whole-word frequency in processing derived words. In: Baayen R. Harald & Schreuder Robert (eds.), *Morphological structure in language processing*. Berlin: Mouton de Gruyter. 157-208.

Butterworth Brian 1983. Lexical representation. In Butterworth Brian (ed.), *Language production. Vol. II*. London: Academic Press. 257-294.

Caramazza Alfonso, Laudanna Alessandro & Romani Cristina 1988. Lexical access and inflectional morphology. *Cognition* 28. 297-332.

De Jong Nivja, Feldman Laurie B., Schreuder Robert, Pastizzo Matthew & Baayen R. Harald 2002. The processing and representation of Dutch and English compounds: Peripheral morphological, and central orthographic effects. *Brain and Language* 81. 555-567.

Deutsch Avital, Frost Ram & Forster Kenneth 1998. Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24. 1238-1255.

Dominguez Alberto; Cuetos Fernando & Segui Juan 2000. Morphological processing in word recognition: A review with particular reference to Spanish data. *Psicológica* 21. 375-401.

Giraudo Hélène & Grainger Jonathan 2000. Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes* 15. 421-444.

Hyönä Jukka 2012. The role of visual acuity and segmentation cues in compound word identification. *Frontiers in Psychology* 3. 1-8.

Juhasz Barbara 2007. The influence of semantic transparency on eye movements during English compound word recognition. In Van Gompel Roger, Fischer Martin, Murray Wayne & Hill Robin (eds.), *Eye Movements: A Window on Mind and Brain*. Amsterdam: Elsevier. 373-389.

Lüdeling Anke & De Jong Nivja 2002. German particle verbs and word-formation. In Dehé Nicole, Jackendoff Ray, McIntyre Andrew & Urban Silke (eds.), *Verb-Particle Explorations*. Berlin: Mouton de Gruyter. 315-333.

Lyashevskaya Olga N. & Sharoff Serge A. 2009. *Častotnyj Slovar' Sovremennogo Russkogo Jazyka* (The frequency dictionary of modern Russian language). Moscow: Azbukovnik.

MacWhinney Brian & Leinbach Jared 1991. Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 40, 121-157.

Manelis Leon & Tharp David A. 1977. The processing of affixed words. *Memory and Cognition* 5, 690-695.

Moscoso del Prado Martín Fermin, Bertram Raymond, Häikiö Tuomo, Schreuder Robert & Baayen R. Harald 2004. Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition* 30. 1271-1278.

Niswander-Klement Elizabeth & Pollatsek Alexander 2006. The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory and Cognition* 34. 685-702.

Pollatsek Alexander & Hyönä Jukka 2005. The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes* 20. 261-290.

Pollatsek Alexander, Hyönä Jukka & Bertram Raymond 2000. The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception & Performance* 26. 820-833.

Rueckl Jay G., Mikolinski Michelle, Raveh Michal, Miner Caroline & Mars Frank 1997. Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language* 36. 382-405.

Schreuder Robert & Baayen R. Harald 1995. Modeling morphological processing. In Feldman, Laurie B. (ed.), *Morphological aspects of language processing*. Hillsdale, NJ: Erlbaum. 131-154.

Schreuder Robert & Baayen R. Harald 1997. How complex simplex words can be. *Journal of Memory and Language* 37. 118-139.

Seidenberg Mark S. & McClelland James L. 1989. A distributed developmental model of word recognition and naming. *Psychological Review* 96. 523-568.

Stemberger Joseph P. & MacWhinney Brian 1988. Are inflected forms stored in the lexicon? In Hammond Michael & Noonan Michael (eds.), *Theoretical morphology: approaches in modern linguistics*. San Diego, CA: Academic Press. 101-116.

Taft Marcus & Forster Kenneth I. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14. 638-647.

Taft Marcus 1979. Recognition of affixed words and the word frequency effect. *Memory and Cognition* 7. 263-272.

Taft Marcus 2004. Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology* 57A. 745-765.

Zaliznyak Andrey 1977. *Grammatičeskij slovar' russkogo jazyka. Slovoizmenenie* (The grammatical dictionary of the Russian language. Inflection). Moscow: Russkij Jazyk.

Zwitserlood Pienie 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9. 341-368.