Word Sense Frequency Estimation for Russian: Verbs, Adjectives and Different Dictionaries

Anastasiya Lopukhina[1], **Konstantin Lopukhin**[2]

[1] National Research University, Higher School of Economics;
Russian Language Institute of the Russian Academy of Sciences, Moscow, Russia
[2] Scrapinghub, Moscow, Russia
E-mail: alopukhina@hse.ru, kostia.lopuhin@gmail.com

Abstract

In this paper we investigate several extensions to our prior work on sense frequency estimation for Russian. Our method is based on semantic vectors and is able to achieve good accuracy for sense frequency estimation trained on dictionary entries from the Active Dictionary of Russian and unannotated corpora. We apply our method to verbs and adjectives to obtain sense frequencies for 329 verbs and 256 adjectives in an academic corpus and a web-based corpus. We compare frequency distributions against dictionary sense ordering and between two corpora and find that the first dictionary sense is not the most frequent for almost half of the words we studied. Evaluation of verbs and adjectives shows that frequency estimation error is lower than 15%. We investigate the effect of sense granularity, evaluating how the accuracy of our method changes when applied to more coarse-grained senses. We also investigate if our method can be applied to other dictionaries with less elaborate sense descriptions, by evaluating its accuracy when training on dictionary entries from two other dictionaries.

Keywords: frequency; sense frequency; word sense disambiguation; semantic vectors; sense granularity

1. Introduction

When words have several senses, it is important that dictionaries describe them properly and exhaustively (see e.g. Pustejovsky, 1996; Apresjan, 2000; Iomdin, 2014). One of the properties of word senses is their frequency in a language, as the different senses are not distributed evenly. However, this information is not represented in dictionaries. We cannot rely on the ordering of word senses in a dictionary to obtain this information, as it is not always consistent with real sense distribution in a language. In the Russian lexicographic tradition the ordering of senses follows etymological principles: the first sense of a polysemous word is usually the original, non-figurative meaning (Kruglikova, 2012). For example, the Russian word *veha* can be described as having two distinctly different senses: (1) ‘boundary-mark’ and (2) ‘a milestone in smb's life’ (Apresjan, 2014). Although native speakers might agree that the first sense of the word *veha* is rare, we cannot quickly check this assumption; instead, relative frequency is assessed subjectively by intuition.

The lack of word sense frequency information becomes a problem in language learning and teaching. Nesi and Haill (2002) stress the problem of learners being satisfied with the first sense listed in a dictionary, even if the meaning does not fit the context, which often leads to incorrect interpretations. The information about sense frequency is especially necessary if a dictionary is going to be used for text production (Lew, 2013). Discussing the question of word lists for teaching a language, Beck et al. (2013) state that there is no way to obtain the relative frequency of one meaning or sense of a word from the general frequency of this word. It evokes the problem of selecting the appropriate meaning that should be studied first. The same problem can be illustrated for Russian. For example, the first dictionary sense of the Russian word *bremya*—'heavy load'—is perceived as rare in comparison with its second sense—'burden' (according to the *Large Explanatory Dictionary of Russian* (Kuznetsov, 2014)). So, the information about word sense frequency could help students prioritise learning the most relevant sense of a word.

Word sense frequencies can be useful for theoretical studies of the meaning structures of polysemous words. The information about relative sense frequencies can be a basis for comparing the cross-linguistic meaning structure of cognate words in two languages (like *base*—*basa*, *clay*—*klej* in English and Russian) and translation equivalents (like *thing*—*veshch'* in English and Russian). Iomdin and colleagues (2016) described three cases of cognates in Russian and English whose meaning structures are dissimilar: words with senses that have no match in the other language (*vagon*—*wagon*, *gradus*—*grade*); words with one or more matching senses for which the most frequent senses drastically differ (*avtoritet*—*authority*, *artist*—*artist*); and words in which several senses match but others do not (*blok*—*block*). The authors discovered that people tend to transfer meaning structures of cognates from their own language to the other language. Thus, information about common mistakes in cognate usage and sense frequencies can be important for language learners as well as for linguists.

The question of word sense frequencies is studied as a practical application to automated word sense disambiguation tasks (Navigli, 2009). The most frequent sense detection is widely studied (Mohammad & Hirst, 2006; McCarthy et al., 2007; Loukachevitch & Chetviorkin, 2015) and is known to be an important baseline, and difficult to overcome for many word sense disambiguation systems (Agirre et al., 2007; Navigli, 2009). Furthermore, psycholinguistic experiments with homonyms and polysemes use information about sense frequency as a factor. Several studies (Klein & Murphy, 2001; Pylkkänen et al., 2006; Foraker & Murphy, 2012) showed that sense frequencies and sense dominance influence processing speed.

In this paper, we present an approach to word sense frequency estimation that is based on corpora and explanatory dictionaries. It allows us to automatically obtain sense frequency distributions from raw corpora and uses dictionary information for training. We extend previously reported works (Lopukhina et al., 2016; Lopukhina et al., in print) in a number of different directions: (1) We apply the method to verbs and adjectives, while previous studies included only nouns. We get sense frequencies from academic and web-based corpora and compare distributions. (2) We experiment with sense granularity for nouns and evaluate our method on coarse-grained and fine-grained sense inventories. (3) We compare the *Active Dictionary of Russian* (Apresjan, 2014), that was used for sense inventory and training data, to two other dictionaries: the *Large Explanatory Dictionary of Russian* (Kuznetsov, 2014) and the *Russian Language Dictionary* (Evgenyeva, 1981–1984). Thus we aim to study whether our approach can be generalized to any explanatory dictionary. We conduct our research on the Russian language.

2. Word Sense Frequency Estimation

For the purpose of word sense frequency estimation, for each word we perform automated word sense disambiguation on contexts sampled from corpora, and then calculate relative sense frequencies in the sample. We need a word sense inventory, a source of word contexts (a corpus), and a word sense disambiguation technique. We use only existing linguistic resources, without any additional annotation except for evaluation.

2.1 Word Sense Inventories

As a source of word senses we chose an explanatory dictionary—this type of sense inventory is the most natural for our task and, besides, many languages have dictionaries, but not all possess WordNet-like resources.

For our research, we principally used the *Active Dictionary of Russian* (Apresjan, 2014). This dictionary has three major advantages: first, it is the most developed explanatory dictionary of Russian which reflects contemporary language; second, it uses a consistent and systematic approach to polysemy—each word sense is identified by a set of its unique properties and similar words are described similarly; and third, for each word sense it provides many examples and collocations. They are used by our word sense disambiguation technique for training. We have already presented the results of sense frequency estimation for 440 polysemous and homonymous nouns from the *Active Dictionary of Russian* (Lopukhina et al., 2016; Lopukhina et al., in press). Our current research is focused on verbs and adjectives from the first issue of the dictionary.

In order to answer the question of whether more coarse-grained sense distinction can boost performance (Navigli, 2006), we experimented with sense granularity of nouns, verbs and adjectives from the *Active Dictionary of Russian.* All senses that were described as components of one block and have indexes (like 1.1, 1.2, 1.3) were merged and considered as one sense. This clustering of senses inevitably leads to the loss of details: such as the loss of scope for the verb *brodit'*: 1.1 'to travel from place to place on foot, usually without a particular direction or purpose' and 1.2 'to travel around the world with no particular purpose'; or the loss of specificity for the adjective *belyj*: 7.1 'good' (*white magic*) and 7.2 'legal' (*reported salary*). Nevertheless, coarse-grained senses are distinct, interpretable and different from other senses of a word. We aim to test whether a more coarse-grained sense inventory will provide better results in our task.

Despite its advantages, the *Active Dictionary of Russian* has one important drawback—it is an ongoing project: only 17%of the dictionary vocabulary has been described and edited (approximately 1960 words out of 11,150). Therefore, in this study, we also tested two more explanatory dictionaries: the academic *Russian Language Dictionary* (Evgenyeva, 1981–1984) and the *Large Explanatory Dictionary of Russian* (Kuznetsov, 2014). Both have electronic versions that we used in our research. These dictionaries have the most similar definitions among all the explanatory dictionaries of Russian and have similar distributions of entries by the number of senses (Kiselev et al., 2015). The major disadvantage of these dictionaries that prevented us from using them from the very beginning is the lack of collocations and illustrative sentences, which are crucial for our technique. The average number of examples and collocations for 14 nouns in the *Russian Language Dictionary* is 4.5, in the *Large Explanatory Dictionary of Russian* is 7.5, and in the *Active Dictionary of Russian* is 20. For the purpose of the current study we selected 14 polysemous nouns, extracted all the collocations and illustrative sentences in their entries from the dictionaries and compared the performance of our method for these three sense inventories.

2.2 Corpus

The corpus is a source of contexts for disambiguation. The choice of corpus influences sense frequency, because word sense distributions vary from corpus to corpus. For nouns it was found that 67 out of 440 words have different most frequent senses in the academic and in the web-based corpora (Lopukhina et al., in print). The difference was explained by the difference in content of the corpora. For purposes of the current study, we also used the contexts from the same two corpora: the Russian National Corpus (RNC, http://ruscorpora.ru/en, 230 million tokens in the main corpus), a resource created by a consortium of linguists and software developers; and the ruTenTen11 web-based corpus, the largest Russian internet corpus, consisting of 18 billion tokens integrated into the Sketch Engine system (Kilgarriff et al., 2004). Web corpora are known for having more recent data and for providing relevant and comparable linguistic evidence for lexicographic purposes (Ferraresi et al., 2010). Therefore, we expect to find differences in sense frequency distributions for verbs and adjectives in these two corpora. To estimate word sense frequency we sample 1,000 random contexts for each word in both corpora. Sample sizes yield a statistical error below 3.1%.

2.3 Word Sense Disambiguation Method

In this study we use the word sense disambiguation (WSD) method based on semantic vectors that is described in detail in Lopukhina et al. (in press). This method can achieve good disambiguation accuracy even on a small number of examples available in the dictionary, and is very robust to overfitting. The basis of the method is a vector representation of context or a dictionary example, which is obtained as a weighted sum of semantic vectors for words: this representation aims to capture the sense of a context. Context vectors for all illustrative examples, collocations, synonyms, etc. for a particular sense are averaged to form a single sense vector. Such vectors are built for all dictionary senses. When disambiguating a new context, its vector is calculated in the same way (as a weighted sum of word vectors), and the method assigns this context to the sense with the closest sense vector. In Lopukhin & Lopukhina (2016) we studied several variations of the method, and have decided to use the most simple and robust variant in this paper.

Word vectors were trained using word2vec skip-gram algorithm on a 2 billion lemmatized corpus (combined RuWaC, lib.ru and Russian Wikipedia) with vector dimension 1024, window size 5 and negative sampling. Word weights were estimated on the same corpus. Implementation of the method is available online on https://github.com/lopuhin/sensefreq.

3. Evaluation

Quantitative evaluation is comprised of three parts: evaluating WSD accuracy for different parts of speech, coarse-grained vs. fine-grained senses, and different dictionaries. In the evaluation for different parts of speech we focus on verbs and adjectives, and also include results on nouns for comparison—evaluated in more detail in Lopukhina et al. (in press). In the coarse-grained sense evaluation we compare WSD accuracy when using coarse and fine-grained senses from the *Active Dictionary of Russian* for nouns, verbs and adjectives. For the evaluation of the different dictionaries we compare WSD accuracy obtained when training on entries from the *Active Dictionary of Russian* and when training on entries from two other dictionaries.

3.1 Word Sense Disambiguation for Verbs and Adjectives

We evaluated word sense disambiguation accuracy and sense frequency estimation error of our method for words of three different parts of speech: nouns, verbs and adjectives. We used two different kinds of training data: full contexts from the corpus and entries from the *Active Dictionary of Russian* (AD). For this study at least 100 contexts were labelled for each word, and 50 random contexts were used for training, while the rest were used for evaluation in a fivefold cross-validation scheme. When training on dictionary entries, all labelled contexts from the corpus were used for training. Frequency error was measured as maximum absolute error in sense frequency estimation averaged across all words.

Results are presented in Table 1. We provide two baselines: the first dictionary sense baseline and the MFS (most frequent sense) baseline. MFS is a powerful baseline that assigns all contexts to the most frequent sense and is often hard to beat (Navigli, 2009). The first dictionary sense baseline assigns all contexts to the first dictionary sense and is more relevant for methods trained on dictionary entries. This baseline is more powerful than a random one, because the first sense is often the most frequent.

Part of speech	**Nouns**	**Verbs**	**Adjectives**
Number of words	17	20	14
Avg. number of senses	3.82	5.00	5.93
First sense baseline	0.50	0.59	0.55
MFS baseline	0.67	0.63	0.62
Accuracy training on contexts	0.80	0.72	0.69
Accuracy training on AD entries	**0.76**	**0.69**	**0.68**
Frequency error (AD entries)	**0.10**	**0.14**	**0.14**

Table 1: WSD accuracy for nouns, verbs and adjectives

We see that training on 50 contexts from the corpus gives more accurate predictions than training on dictionary entries, although the difference for adjectives is very small. Nouns have the highest accuracy while also having the lowest number of senses, and adjectives have the lowest accuracy and the highest number of senses. Verbs have significant negative Pearson correlation between number of senses and accuracy: −0.7, while the correlation between nouns and adjectives is more moderate, at −0.3. The average number of senses given in Table 1 is for words used for evaluation, but it is similar across all polysemous words in the *Active Dictionary of Russian*: 3.33 for nouns, 5.17 for verbs and 3.79 for adjectives—only adjectives display a significant difference.

Figure 1 shows a distribution of WSD accuracy when training on AD entries. We see that verbs have a more diverse distribution, with some scoring as low as 0.2 but also many having scores above 0.9, while adjectives have few words with accuracy higher than 0.8.

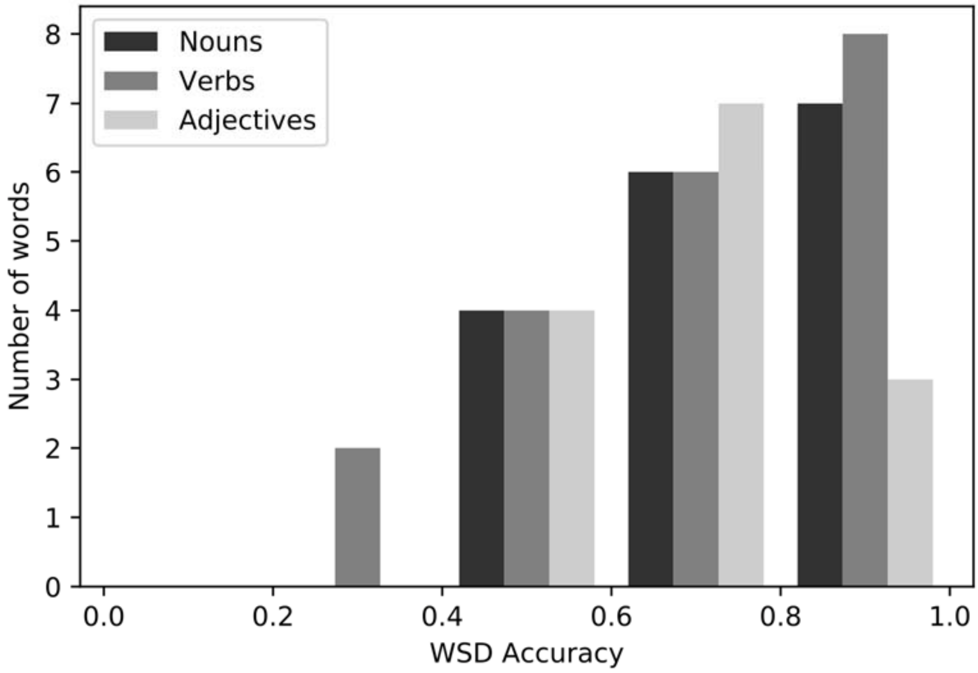

Figure 1: Distribution of WSD accuracy for nouns, verbs and adjectives

Sense frequency estimation error for verbs and adjectives is higher than for nouns, but is still low in an absolute sense, lower than 15% for all parts of speech. This means that our method gives reliable sense frequency estimation for all parts of speech.

3.2 Coarse-grained Sense Inventory

The *Active Dictionary of Russian* provides a two-level hierarchical sense inventory: senses are numbered as *x.y* (e.g. *2.1*), making it possible to evaluate word sense disambiguation on coarse-grained senses, formed by lumping together fine-grained components of one semantic block (e.g. *2.1, 2.1 → 2*). As a result, most words have fewer senses, and most senses have more training examples. Results of this evaluation are presented in Table 2. We see that all parts of speech have significantly fewer coarse-grained senses on average, and accuracy for coarse-grained senses increases for nouns and especially verbs, and is almost the same for adjectives. For verbs, the result can be explained by a general tendency to obtain a higher accuracy for fewer senses. We suppose that the lower accuracy gain for adjectives may be explained as follows: adjectives get different senses in contexts with nouns while verbs and nouns have more diverse contexts. More limited contexts for adjectives can be the reason for the results we obtained.

Part of speech		Nouns	Verbs	Adjectives
Number of senses	Fine	3.82	5.00	5.93
	Coarse	2.77	3.15	4.07
First sense baseline	Fine	0.50	0.59	0.55
	Coarse	0.56	0.66	0.60
Accuracy	Fine	**0.76**	**0.69**	**0.68**
	Coarse	**0.80**	**0.79**	**0.79**

Table 2: Coarse and fine sense inventories for the *Active Dictionary of Russian*

3.3 Other Dictionaries

The *Active Dictionary of Russian* is a very attractive resource for computational linguistics methods due to its very comprehensive and systematic descriptions. However, its wordlist is small compared to other dictionaries, and only the first volume has been published at the time of writing. Thus, it is interesting to check how our method works on other dictionaries with larger wordlists, namely the *Russian Language Dictionary* (Evgenyeva, 1981–1984), denoted as MAS, and the *Large Explanatory Dictionary of Russian* (Kuznetsov, 2014), denoted as BTS. Since all these dictionaries have different sense inventories, we had to perform sense mapping: each sense in MAS or BTS was mapped to one or more senses in AD. If some AD sense did not have any corresponding sense in MAS/BTS, contexts with this sense were removed from test data. Words where only one sense was left or where one AD sense corresponded to several MAS/BTS senses were discarded. Evaluation was performed only on nouns: we selected 11 nouns for MAS and 14 nouns for BTS. Results are presented in Tables 3 and 4. In order to compare the quality of training data in MAS/BTS to the *Active Dictionary of Russian*, we also measured word sense disambiguation accuracy with mapped senses but AD training data (denoted as AD* in the table).

Sense inventory	Training data	
	BTS/MAS	AD*
MAS	**0.66**	0.75
BTS	**0.65**	0.72

Table 3: WSD accuracy for other dictionaries (MAS and BTS) compared to AD

Sense inventory	Training data	
	BTS/MAS	**AD***
MAS	**0.20**	0.13
BTS	**0.21**	0.15

Table 4: Sense frequency estimation error for MAS and BTS compared to AD

We see that both MAS and BTS perform significantly worse than AD, and that BTS performs better than MAS when compared with the *Active Dictionary of Russian*. Sense frequency estimation error for MAS and BTS is also larger but could still be useful for some tasks. In Table 5 we compare average number of examples per sense and average number of words per sense: BTS has a larger number of examples than MAS, which might explain differences in WSD performance (relative to AD) between MAS and BTS.

	MAS	**BTS**	**AD**
Number of examples per sense	4.5	7.3	20
Number of words per sense	62	49	216

Table 5: Average number of examples and words per sense in training data

4. Results and Discussion

We obtained sense frequencies for Russian verbs, adjectives (in this study) and nouns (Lopukhina et al., in press) in the academic Russian National Corpus and web-based ruTenTen11. All data are available online: http://sensefreq.ruslang.ru/. Word sense frequency distributions differ depending on the part of speech and on the corpora used. In Lopukhina et al. (in press) we reported on sense frequencies for 440 nouns. In this study, we applied our method to all homonymous and polysemous verbs and adjectives from the first issue of the *Active Dictionary of Russian* and obtained word sense frequencies for 329 Russian verbs and 256 adjectives.

First, we compared the first sense in the *Active Dictionary of Russian* with the most frequent sense in the RNC and ruTenTen11. The ratio of verbs where the first dictionary sense is the most frequent (excluding homonyms) is 50% in the RNC and 48% in ruTenTen11. For adjectives, the first dictionary sense coincides with the most frequent sense in 61% of cases in the RNC and 59% in ruTenTen11. This means that, for verbs and adjectives, the meaning described first in a dictionary differs from the most common sense of the word in contemporary language in about half of cases.

The discrepancy between the first sense of verbs in the Active Dictionary of Russian and the most frequent sense in the RNC can be observed in the following examples. The first dictionary sense of the verb *gladit'* is 'to iron', while in 83% of cases in the RNC it is used in the other sense—'to gently move your hand over skin, hair, or fur'. The first literal sense of the verb *bolet'* is 'to be ill'. In the RNC, this sense is the third most frequent (20%); the most frequent is 'to feel pain somewhere in your body' (46%) and the second most frequent, 'to be a fan, to encourage somebody's favourite sportsman or team' (31%). For several verbs, the most frequent meaning is a metaphorical one; it is normally described after a literal one in the dictionary, e.g. *vykroit'* ('to succeed in getting enough of something, especially time and money, by making a lot of effort', 87%), *vyputat's'a* ('to get yourself out of a situation that you no longer want to be involved in', 88%), *galdet'* ('to make noise (about people)', 92%), *vkluchit's'a* ('to start to take part in a particular activity that has started before', 71%), *votsarit's'a* ('something starts to happen and have an effect, and is not likely to stop for a long time', 75%), *vsplyt'* ('to appear in somebody's mind without special reason', 53%).

For adjectives, the discrepancy between the first dictionary sense and the most frequent sense in the Russian National Corpus can be illustrated by the following examples. The word *gluhoj* in the RNC is used in 30% of cases in collocations with *sound,* in the sense of 'a low sound made when one hard heavy object hits another', while its first dictionary sense is 'not able to hear anything' (12%). In some cases, a collocation may be very frequent and thus increases the frequency of an adjective. A good illustration for this observation is the word *vishn'ovyj*: its most frequent sense in the RNC is 'related with a tree that produces cherries' (57%), evidently because of the spread of the name of the Anton Chekhov play 'The Cherry Orchard', in the texts of the academic corpus. For the adjective *burnyj,* the distribution of sense frequencies is completely opposite to the ordering of senses in the dictionary: *stormy weather* (4%), *stormy wind or sea* (15%), *rapid growth* (34%) and *wild passion, stormy romance* (47%). As for verbs, for some adjectives the most frequent sense has undergone a semantic shift and is metonymical, as in the examples *bir'uzovyj* (*turquoise color*, 80%), *antikvarnyj* (*antique shop*, 59%), *belokuryj* (*fair-haired boy*, 55%), *golovnoj* (*head*, 41%).

We think that including the information about the most frequent sense and overall sense frequency distribution in explanatory dictionaries is relevant for dictionary users. Robert Lew (2013) suggested that the information about the most frequent sense would be necessary for text production (such as essay writing) but not for comprehension, as dictionary users usually do not look up a frequent sense of a word. We advocate the need for these conclusions to be tested as soon as the information about sense frequencies of words in dictionaries becomes available. Moreover, it may help to include dictionaries in natural language processing tasks like word sense disambiguation, as necessary information regarding the most frequent sense will become available in explanatory dictionaries and connected with their sense inventories.

We compared the most frequent senses for verbs and adjectives in the Russian National Corpus that contains more literary contexts, with the most frequent senses in the up-to-date web-based ruTenTen11. The corpora have a high degree of overlap: the ratio of the same most frequent sense is 80% for verbs and 82% for adjectives. The difference can be explained by the content of the corpora. The RNC provides quite literary most-frequent senses: as in the examples *close relative* for the word *blizhnij*, *boulevard bench* for *bulvarnyj* and *bitter laugh, bitter irony* for *gor'kij*, as compared with the colloquial uses *the nearest place*, *tabloid novels* and *bitter taste,* respectively. For words such as *anglijskij* and *almaznyj* the most frequent senses in ruTenTen11 are narrower and more specific than in the RNC: 'the English language'/'related to England' and 'produced using cutting diamond'/'related to a diamond' (ruTenTen11/RNC in both examples). These observations are also relevant for verbs. Moreover, we observed that for some verbs the most frequent senses in ruTenTen11 are metaphorical, while in the RNC they are literal. For example, *bazirovat's'a* 'to base a decision or idea on particular information'/'to be based somewhere', *bredit'* 'to talk nonsense'/'to be delirious' and *vooruzhit'* 'to provide yourself or other people with useful information or equipment to achieve the goal'/'to provide yourself or other people with weapons'.

Our aim was to study whether our approach to word sense frequency estimation can be generalized to any explanatory dictionary and therefore we compared the accuracy of our method for three dictionaries: the *Active Dictionary of Russian* (AD), the *Large Explanatory Dictionary of Russian* (BTS) and the *Russian Language Dictionary* (MAS). The comparison was performed on nouns, because nouns normally have more distinct senses (compared to other parts of speech), as many of them refer to objects existing in the real world (Iomdin et al., 2014). In BTS and MAS, the number of collocations and illustrative sentences is much less than in the AD. The lack of examples prevented our method from building solid sense vectors and thus the accuracy of the method trained on BTS and MAS is worse compared to that on the AD. The difference in sense inventories also influenced the results: the word *al'bom* has three senses in the AD—'a book with blank pages, used for drawing', 'a book in which you can collect things such as photographs or stamps' and 'a collection of several songs or pieces of music recorded as an MP3 file, on a CD etc'. The last is rather frequent in the Russian National Corpus (33%) and the most frequent in ruRenTen11 (73%), but is absent in both BTS and MAS. This implies that many contexts are not covered by senses described in these dictionaries. To ensure a good performance, our method requires an up-to-date sense inventory with several typical illustrative sentences and collocations for each sense used for training.

5. Conclusion

This paper continues the study of the automated word sense frequency estimation for Russian words. We applied the method based on semantic vectors and trained on collocations and illustrative sentences from the *Active Dictionary of Russian* to ambiguous verbs and adjectives from the first issue of the dictionary. As a result, we obtained sense frequencies for 329 verbs and 256 adjectives. All the data are available on http://sensefreq.ruslang.ru. Subsequently, the word sense frequency database now contains frequency distributions for nouns, verbs and adjectives in the academic Russian National Corpus and the web-based corpus ruTenTen11 (1025 ambiguous words in total). We evaluated frequency estimation error for verbs and adjectives and found that it is slightly worse than for nouns but still below 15%.

We experimented with sense granularity in the *Active Dictionary of Russian* and found that using more coarse-grained senses improves disambiguation accuracy, and a hierarchical approach to sense description can be very helpful when fine-grained distinctions between senses are not important for the task at hand.

In order to test our approach on other dictionaries we compared word sense disambiguation accuracy obtained when training on the *Active Dictionary of Russian* to the *Large Explanatory Dictionary of Russian* and the *Russian Language Dictionary.* We found out that although the accuracy on the other two dictionaries is above the baseline, it is substantially lower than on the *Active Dictionary of Russian.* Many collocations and illustrative examples for each sense are important for achieving good disambiguation accuracy.

The information about word sense frequency may have several applications: for lexicography and language learning, for the theoretical and experimental study of polysemy, and for different NLP tasks. The method presented in this paper can be applied to any language with a sufficiently large corpus and a dictionary with contemporary vocabulary that provides several examples of each sense.

6. Acknowledgements

This research was supported by RSF (project no.16–18–02054: “Semantic, statistic, and psycholinguistic analysis of lexical polysemy as a component of a Russian linguistic worldview”). The authors would also like to thank the two anonymous reviewers for their valuable comments.

7. References

Agirre, E., Marquez, L. & Wicentowski, R. (eds.). (2007). *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007).* Association for Computational Linguistics, Prague, Czech Republic.

Apresjan, J. (2000). *Systematic Lexicography.* Oxford.

Apresjan, J. (ed.). (2014). *Active Dictionary of Russian.* A-G. JSK, Moscow.

Beck, I., McKeown, M. G. & Kucan, L. (2013). *Bringing Words to Life: Robust Vocabulary Instruction.* Guilford Press.

Evgenyeva, A. (ed.). (1981–1984). *Russian Language Dictionary.* Russian language, Moscow.

Ferraresi, A., Bernardini, S., Picci, G. & Baroni, M. (2010). Web corpora for bilingual lexicography: A pilot study of English/French collocation extraction and translation. *Using Corpora in Contrastive and Translation Studies.* Newcastle: Cambridge Scholars Publishing, pp. 337–359.

Foraker, S. & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of meaning dominance. *Journal of memory and language* 67.4, pp. 407-425.

Iomdin, B. (2014). Polysemous words in and out of the context. *Voprosy jazykoznanija.* Vol. 4. Moscow.

Iomdin, B., Lopukhina, A. & Nosyrev, G. (2014). Towards a word sense frequency dictionary. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2014".* Bekasovo, Moscow, pp. 204–229.

Iomdin, B., Lopukhin, K., Lopukhina, A. & Nosyrev, G. (2016). Word sense frequency of similar polysemous words in different languages. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2016".* Moscow, pp. 201–211.

Kilgarriff, A., Rychly, P., Smrz, P. & Tugwell, D. (2004). The Sketch Engine. *Euralex 2004. Proceedings.* Lorient, France, pp. 105–116.

Kiselev, Y., Krizhanovsky, A., Braslavski, P., Menshikov, I., Mukhin, M. & Krizhanovskaya, N. (2015). Russian Lexicographic Landscape: a Tale of 12 Dictionaries. *Proceedings of the International Conference "Dialog 2015*", pp. 254-272.

Klein, D. & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language* 45.2, pp. 259-282.

Kruglikova, L. (2012). The big academic dictionary of Russian as a successor of Russian academic lexicography traditions. *Cuadernos de Rusistica Espanola*, 8, pp. 177-198.

Kuznetsov, S. (ed.). (2014). *Large Explanatory Dictionary of Russian.* Norint, St. Petersburg.

Lew, R. (2013). Identifying, ordering and defining senses. In H. Jackson (ed.) *The Bloomsbury Companion to Lexicography.* London: Bloomsbury Publishing, pp. 284–302.

Lopukhin, K. & Lopukhina, A. (2016). Word sense disambiguation for Russian verbs using semantic vectors and dictionary entries. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2016"*, pp. 393-405.

Lopukhina, A., Lopukhin, K., Iomdin, B. & Nosyrev, G. (2016). The Taming of the Polysemy: Automated Word Sense Frequency Estimation for Lexicographic Purposes. *Proceedings of the XVII EURALEX International Congress. Lexicography and Linguistic Diversity* (6–10 September, 2016). Tbilisi: Ivane Javakhishvili Tbilisi State University, pp. 249-257.

Lopukhina, A., Lopukhin, K. & Nosyrev, G. (in press). Automated word sense frequency estimation for Russian nouns. In M. Kopotev, O. Lyashevskaya, A.

Mustajoki (eds), *Quantitative Approaches to the Russian Language*. Routledge.

Loukachevitch, N. & Chetviorkin, I. (2015). Determining the most frequent senses using Russian linguistic ontology RuThes. *Proceedings of the Workshop on Semantic Resources and Semantic Annotation for Natural Language Processing and the Digital Humanities at NODALIDA*, pp. 21–27

McCarthy, D., Koeling, R., Weeds, J. & Carroll, J. (2007). Unsupervised acquisition of predominant word senses. *Computational Linguistics* 33:4, pp. 553–590.

Mohammad, S. & Hirst, G. (2006). Determining word sense dominance using a thesaurus. *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL-2006)*, pp. 121–128.

Navigli, R. (2006). Meaningful clustering of senses helps boost word sense disambiguation performance. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of ACL*. Association for Computational Linguistics, USA, pp. 105–112.

Navigli, R. (2009). Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR)* 41:2, pp. 1–69, Article 10.

Nesi, H. & Haill, R. (2002). A study of dictionary use by international students at a British university. *International Journal of Lexicography* 15.4, pp. 277–306.

Pustejovsky, J. (1996). *Lexical semantics: The problem of polysemy*. Oxford.

Pylkkänen, L., Llinás, R. & Murphy, G. L. (2006). The representation of polysemy: MEG evidence." *Journal of cognitive neuroscience* 18.1, pp. 97-109.

ruscorpora.ru/en. Accessed at: http://ruscorpora.ru/en. (10 July 2017)

github.com/lopuhin/sensefreq. Accessed at: https://github.com/lopuhin/sensefreq. (10 July 2017)

sensefreq.ruslang.ru. Accessed at: http://sensefreq.ruslang.ru/. (10 July 2017)