

GERMAN-RUSSIAN IDIOMS ONLINE: ON A NEW CORPUS-BASED DICTIONARY¹

Dobrovol'skij D. O.

(dm-dbrv@yandex.ru, dobrovol'skij@gmail.com)

V. V. Vinogradov Russian Language Institute RAS,
Moscow, Russia

The paper focuses on the structure and principles for constructing a new German-Russian phraseological dictionary based on corpus data. Fragments of this dictionary are available on the website of the German Language Institute in Mannheim: "Deutsch-russische Idiome online" http://wvonline.ids-mannheim.de/idiome_russ/index.htm. Relevant information is also made available via the Europhras homepage at <http://www.europhras.org>. In section 1, I formulate certain general principles of modern bilingual phraseology. Section 2 discusses the state of the art of German-Russian phraseography and explains the need for a new German-Russian phraseological dictionary. In Section 3, key features of the new corpus-based dictionary are considered. The basic difference between the present dictionary and traditional ones is that all examples of idiom usage are taken from text corpora DeReKo and DWDS, and in individual cases from the German-language Internet. Parallel texts from the Russian National Corpus (RNC) are also used. The use of authentic examples based on text corpora is a new approach in bilingual lexicography. Traditional dictionaries were based on a limited body of randomly selected examples, and the use of the idioms was often not even exemplified. The advantages of using corpora consist not only in more detailed and well thought-out illustrations of the expressions being described, but also in additional possibilities that the corpus provides for compiling the idiom-list and structuring entries.

Key words: dictionary, corpus, bilingual lexicography, phraseology, idiom, German, Russian

1. Corpus-based bilingual phraseography and cross-linguistic equivalence

The lexicographic treatment of the notion of equivalent in dictionaries based on corpus data encounters certain problems. Not infrequently, the generally accepted equivalent of an idiom cannot always be used to translate authentic texts.

Let us take an example. The German idiom *jmdn. an der Nase herumführen* has a "standard" equivalent in Russian, namely the idiom *водить за нос кого-л.* It would

¹ This paper is based on work supported by the RGNF under Grants 11-04-00105a, 12-04-12041, 12-34-10413 and by the Basic Research Program "Corpus Linguistics" of the Presidium of the Russian Academy of Sciences.

be somewhat odd to doubt that these idioms are basically equivalent, since they are identical with respect to both their lexicalized meaning and image component. Nevertheless, it turns out that it is far from always possible to translate the expression *jmdn. an der Nase herumführen* with the Russian idiom *водить за нос кого-л.* Numerous contexts with the idiom *jmdn. an der Nase herumführen* can be found in text corpora in which this idiom has to be translated into Russian either by the verbs *надуть* and *одурачить* or by the idiom *обвести вокруг пальца*.

- (1) Die Aktionäre fühlen sich vom größten deutschen Industriekonzern *an der Nase herumgeführt*. (Mannheimer Morgen, 08.08.1995)
У акционеров такое чувство, что самый большой промышленный концерн Германии *обвел их вокруг пальца*.
In Wahrheit hatte er [Wolfgang Schäuble] aber 100.000 Mark <...> bekommen <...>. Und das hat er im Deutschen Bundestag <...> verschwiegen und hat das erst später, vier Wochen später in einem Fernsehinterview aufgedeckt und da haben viele gesagt, <...> der hat den Deutschen Bundestag *an der Nase herumgeführt*. [www.stroebele-online.de/themen/spendenaafaere/29273.html]
На самом деле он [Вольфганг Шойбле] получил 100 000 марок. Причем он скрыл это от бундестага и только позднее, спустя четыре недели, признался в этом во время телеинтервью. И многие сказали тогда: он просто *одурачил* немецкий парламент.

Consequently, despite the intuitively felt equivalence of the expressions *jmdn. an der Nase herumführen* and *водить за нос кого-л.*, this equivalence cannot be considered complete. For the lexicographer interested in a maximally precise description of the material, such instances are problematical. Either we acknowledge that *jmdn. an der Nase herumführen* and *водить за нос кого-л.* are equivalent, in which case it is necessary to explain why the “standard” equivalent is unacceptable in a number of contexts, or we deny that a relationship of bilingual equivalence obtains between *jmdn. an der Nase herumführen* and *водить за нос кого-л.*, and focus exclusively on translating specific contexts. Such a solution, however, is counterintuitive.

There are at least two ways out of this cul-de-sac. Either we refrain from giving equivalents and replace them with an explanation (here permissible target-language correspondences can be given in a special field in the entry — cf. Lubensky 1995), or we provide the given equivalents with a commentary indicating relevant limitations.

In our dictionary we have followed the second path. Thus for the German idiom *jmdn. an der Nase herumführen* we give the Russian equivalent *водить за нос кого-л.* and explain divergences in the use of the idioms in the commentary, where we point to the fact that the Russian idiom *водить за нос кого-л.* is an imperfectiva tantum, i.e. it cannot normally be used in the perfective aspect. Contexts such as *a народ не дурак, за нос его так просто не проведешь* or *за нос такого провести нетрудно* are encountered quite rarely. The use of this idiom in the perfective aspect is licensed only in non-veridical contexts. For more detail see (Dobrovolskij 2013).

A question that arises from the perspective of phraseological theory (especially its comparative aspects) concerns the essence of cross-linguistic equivalence of idioms. Does it really exist?

At first glance, this question seems to be quite simple. Those not involved in idiom research would immediately give a positive answer to this question. Really, how can one doubt the existence of cross-linguistic equivalence in the field of phraseology when there are so many bilingual, and even some multilingual dictionaries of idioms? The aim of such dictionaries is, above all, to provide the user with the knowledge of cross-linguistic idiom-equivalents, therefore such equivalents must exist. And besides, there is rather a long tradition of contrastive idiom research (compare, e.g., review articles Dobrovol'skij 2002 and Korhonen 2007).

1.1. Types of phraseological equivalence

Within this tradition some well-known types of phraseological equivalence are discriminated. For the sake of simplicity, these types have been reduced here to the following four main classes:

- (i) "full equivalents",
- (ii) "partial equivalents",
- (iii) "phraseological parallels", and
- (iv) "non-equivalents".

(i) "Full equivalents" (or "absolute equivalents") are idioms of L1 and L2 which are identical with regard to meaning, syntactic and lexical structure, and imagery basis. Compare German *seine Hand ins Feuer legen für etw.* and English *to put one's hand into the fire for sth.*; English *to rest on one's laurels*, German *auf seinen Lorbeeren ausrufen* and Russian *почивать на лаврах*. Some "full equivalents" allow for morphological or certain lexical alternations, cf. the singular-plural alternation in the following idioms: German *von Kopf bis Fuß* (lit.: "from head to foot") and Russian *с головы до ног* (lit.: "from head to feet"), or German *ganz Ohr sein* (singular), English *to be all ears* (plural) and French *être tout oreilles* (plural). It seems more natural to regard equivalents of this kind as (i) rather than (ii), though in the tradition of contrastive idiom research also quite different views on this issue can be found.

(ii) "Partial equivalents" are idioms of L1 and L2 which have identical or near-identical meanings, but do not fully correspond in syntactic and lexical structure, or imagery basis. Compare English *to get out of bed on the wrong side* and Russian *встать не с той ноги* (lit.: "to get out [of bed] with the wrong foot"), German *aus einer Mücke einen Elefanten machen* (lit.: "to make an elephant out of a mosquito") and Russian *делать из мухи слона* (lit.: "to make an elephant out of a fly"), or German *die Hände über dem Kopf zusammenschlagen* (lit.: "to strike the hands over the head") and Russian *схватиться за голову* (lit.: "to grip one's head").

(iii) "Phraseological parallels" are different idioms of L1 and L2 which correspond to each other in the core meaning, but not with regard to the image component. Cf. English *hot potato* and German *heißes Eisen*, English *to be like a cat on hot bricks* and German *wie auf glühenden Kohlen sitzen*, English *to buy a pig in a poke* and

German *die Katze im Sack kaufen*, German *jmd. hat nicht alle Tassen im Schrank* and Russian *у кого-л. не все дома* (lit.: “sb. does not have them all at home”), English *to take a sledgehammer to crack a nut* and Russian *стрелять из пушек по воробьям* (lit.: “to use cannons to shoot at sparrows”), or English *to play a dirty trick on sb.* and Russian *подложить свинью кому-л.* (lit.: “to put a pig on sb.”), English *spic and span* and Russian *одетый с иголки* (lit.: “dressed from the needle”). “Phraseological parallels” are semantically similar, but their planes of expression mostly do not have much in common.

(iv) “Non-equivalents”, i.e. a given L1-idiom has no idiomatic correspondences in L2; compare the Russian idiom *объяснить на пальцах что-л.* (lit.: “to explain sth. on fingers”) meaning ‘to explain sth. as simply as possible’ which has no idiomatic counterpart in English, or German *etw. nicht übers Herz bringen können* (lit.: “not to be able to bring sth. over the heart”) which can only be translated into Russian using free word combinations such as *не мочь себя заставить (сделать что-л.)* ‘not to be able to force oneself (to do sth.)’.

So, at least in category (i) we seemingly are dealing with real cross-linguistic equivalents, also most members of category (ii) and even some members of category (iii) can be considered good candidates for equivalence. Also from the point of view of text translation there seems to be rather a clear case for the existence of equivalents. It is obvious that almost every literary text analyzed so far contains idioms, sometimes a lot of them. While translating these texts into other languages, translators have to tackle the issue of cross-linguistic equivalence of idioms encountered in the original.

Nevertheless, I doubt that the question as to whether cross-linguistic idiom-equivalents really exist can be answered positively without a serious linguistic discussion. It appears that there are too many stumbling blocks on the way to finding real cross-linguistic equivalents. So, it seems to be an exciting task to discover them, to discuss their nature and, above all, to investigate the reasons why idioms and other semantically corresponding lexical items which look like full equivalents do not always function as such. Cf. example (1) above.

1.2. Aspects of equivalence

It seems expedient to distinguish two different aspects of equivalence:

- a) equivalence in translation; that is, the relationship between an idiom of language L1 and its translation into language L2 in a particular text, and
- b) equivalence in the language system; that is, the relationship between the compared idioms of L1 and L2 on the systemic level.

One of the most important differences between translational and systemic equivalence (besides the fact that the former has to do with a concrete text and the latter with the lexical system) consists in the circumstance that equivalence in translation is a unilateral relationship, whereas equivalence in the language system is defined as bilateral. In other words, if an idiom of language L1 is equivalent to an idiom in language L2, this means that the L2 idiom is also equivalent to the corresponding L1 expression. With

respect to equivalence in translation, all that is being said is that an expression in language L2 is being used in the translation of some specific text in language L1 in such a way that between the L1 idiom from this particular text and the L2 expression there is a relationship of semantic correspondence. The fact that the translation of some L1 idiom into language L2 is its equivalent (at least with respect to this particular context) does not, of course, mean that the relationship can be reversed. That is, the L1 idiom should not be regarded as an equivalent of the expression used in the translation of this idiom into language L2 (even if this expression is an idiom, which is not at all obligatory). Obviously, the study of equivalence in translation broadens our notions about the possibilities of cross-linguistic paraphrasing and about the role of contextual conditions in the selection of adequate correspondences, and it contributes to the development of both translation theory and comparative phraseology.

As for equivalence in the language system, its study has both theoretical and practical significance for phraseology. Deserving of special attention from the theoretical point of view is the question of why one and the same concept is expressed by means of an idiom in one language but not in another. Another (no less important) problem concerns the fact that between basically similar idioms in language L1 and language L2, there are practically always certain semantic, pragmatic, and collocational differences that must be discovered and described. This is especially important in cases where a traditional description postulates a relationship of “full equivalence” but ignores the absence of functional interchangeability between the idioms. The practical aspect of systemic equivalence is what is reflected in bilingual dictionaries, where the entry consists of an idiom of language L1 (in the lemma) and its idiomatic (to the extent this is possible) correlates in L2. Can these correlates be regarded as equivalents of the L1 idiom? Yes and no. On the one hand, they must be at least “partial equivalents” or “phraseological parallels”, for otherwise they could not be placed in the corresponding dictionary entry. On the other, often they cannot be used in the translation of specific texts. The reason, as a rule, is that the idioms of L1 and L2 display certain differences in their semantic, pragmatic, and collocational features. They can be considered cross-linguistic equivalents only in a rather approximate comparison of the idioms of the given languages, and are the starting point of a thorough contrastive analysis that attempts to discover the unique properties of each idiom and thereby improve the lexicological and lexicographical description of phraseology.

Obviously, aspects (a) and (b) are, as it were, two sides of the same phenomenon or two approaches to studying it. We assume that one of the principal goals of contrastive phraseology is to discover genuine equivalents — that is, those that are as close as possible with respect to their actual meanings and — ideally — with respect to the inner form of the expressions (i.e. with respect to images underlying their lexicalized meanings), and that function equally well in analogous types of situations, which does not at all imply an obligatory “idiom — idiom” relationship.

What is important for cross-linguistic correspondence, after all, is not “phraseologicalness,” but **functional equivalence**. It is this type of equivalence that is most interesting from the perspective of bilingual lexicography.

Functional equivalents, how I understand them, are counterparts which can be used in the same concrete situations without any informational loss. To find them out we have

to simultaneously go two ways: from text to language system and from language system to text. As we have seen, not all systemic equivalents can function as counterparts in authentic texts, and on the other hand, not all translational equivalents can be included in the dictionary as typical parallels suitable for using in neutral contexts.

In contrast to a conception that is wide-spread within traditional phraseology, I claim that lexical units of any kind in L2 which have the identical meaning and, in the ideal case, near-identical metaphorical basis as the L1-idioms from the source text are excellent functional equivalents, so they have to be considered not only more or less appropriate translational solutions, but also real functional equivalents, i.e. parallels in the lexicons of L1 and L2, which have to be fixed lexicographically.

2. German-Russian phraseography: state of the art

The need for a new German-Russian phraseological dictionary is motivated by the fact that existing such dictionaries do not meet present requirements. Both the vocabulary and the examples in (Binovich, Grishin 1975) are out of date, and the work fails to satisfy current needs with respect to a number of other parameters as well. Although (Dobrovolskij 1997) is on the whole more up to date, it also has certain shortcomings. Its idiom-list is rather limited, and illustrative examples are often arbitrary and unpersuasive, which may be because it was written back in the “pre-corpus era.” Actually, one of the basic goals of our new lexicographical project is to eliminate all the shortcomings of this dictionary and to significantly expand its idiom-list.

Yet another dictionary of this type has appeared recently: “The New German-Russian Phraseological Dictionary” (Shekasjuk 2010). Its phraseme-list is fairly large and up to date, but the work is difficult to use, primarily because the illustrative examples are not translated into Russian, and the division of entries into meanings and selected equivalents often appears hasty and arbitrary.

Thus there is an unquestionable need for a new dictionary containing the most widely used contemporary German idioms together with carefully selected Russian equivalents, explanations facilitating the correct use of these idioms, and good, authentic examples translated into Russian. It is also important that such a dictionary exist not only in print, but also in an online version, which will not only provide easier access to the information but will also ensure continuous revision and improvement.

3. Key features of the new dictionary

The basic difference between the present dictionary and traditional ones is that all examples of idiom usage in it are taken from the text corpora DeReKo and DWDS, and in individual cases from the German-language Internet. Parallel texts from the Russian National Corpus (RNC) are also used. These examples are especially valuable because they have been translated by professional translators rather than by the

authors and editors of the dictionary. Since this part of the parallel corpus of the RNC is still rather modest in size, however, examples needed for the dictionary were rarely encountered.

The use of authentic examples based on text corpora is a new approach in bilingual lexicography. Traditional dictionaries were based on a limited body of generally randomly selected examples, and the use of the idioms was often not even exemplified. The advantages of using corpora consist not only in more detailed and well thought-out illustrations of the expressions being described, but also in the additional possibilities that the corpus materials provide for compiling the idiom-list and structuring entries. For further detail see (Dobrovol'skij 2013).

Yet another advantage of using corpora is that it increases our ability to determine the peculiarities of the formal and semantic structure of idioms, particularly in the description of the ambiguity and variation of a form. Although an analysis of examples of use clearly indicates that polysemy in phraseology is an extremely widespread phenomenon, traditional dictionaries rarely distinguish the different meanings of idioms, and seldom reflect the full diversity of variants actually represented in texts. Dictionaries often register only a single "canonized" form of an idiom that in many cases proves to be not the most frequent one.

In a number of instances text corpora allow us not only to determine the form of a lemma and a selection of its most frequent variants, but also to establish whether a given expression belongs to the sphere of phraseology. For example, Duden 11 (2002) cites four idioms with the noun *Mundwerk*: *jmds. Mundwerk steht nicht still* (ugs.) 'jmd. redet ununterbrochen'; *ein böses/lockeres/loses/frechtes o.ä. Mundwerk haben* (ugs.) 'gehässig/vorlaut/frech o.ä. reden'; *ein gutes/flinkes Mundwerk haben* (ugs.) 'sehr gewandt reden'; *ein großes Mundwerk haben* (ugs.) 'großsprecherisch reden'. Corpus analysis has shown that the noun *Mundwerk* has a much broader combinatorial profile. Compare, e.g., *flottes, vorlautes, geschliffenes Mundwerk*. This noun can also be used without any adjectives, combining with verbs of various meanings. Cf.

- (2) Manchmal wäre es vielleicht sinnvoller, mein *Mundwerk* etwas zu zügeln, nach dem Motto «Reden ist Silber, Schweigen ist Gold». (St. Galler Tagblatt, 08.04.1999)

Consequently, what we have to do with here is not an idiom but a series of relatively free collocations.

References

1. Binovich L. È., Grishin N. N. (1975), German-Russian Phraseological Dictionary [Nemecko-russkij frazeologičeskij slovar'], Russkij jazyk, Moscow.
2. Dobrovol'skij D. O. (1997), German-Russian Dictionary of Current Idioms [Nemecko-russkij slovar' zhivyh idiom], Metatext, Moscow.
3. Dobrovol'skij D. O. (2002), Phrasemes in cross-linguistic perspective [Phraseologismen in kontrastiver Sicht], Lexikology: An International Handbook on the

- Nature and Structure of Words and Vocabularies. Volume 1. de Gruyter, Berlin/ New York, pp. 442–451.
4. *Dobrovol'skij D.* (2013), German-Russian phraseography: on a new dictionary of modern idiomatics, *Phraseodidactic Studies on German as a Foreign Language*. Verlag Dr. Kovač, Hamburg, pp. 121–138.
 5. *Duden 11* (2002) = Duden — Dictionary of German Idiomatics [Duden — Redewendungen. Wörterbuch der deutschen Idiomatik (=Der Duden, Band 11)]. Dudenverlag, Mannheim etc.
 6. *Korhonen J.* (2007), Issues of cross-linguistic phraseology [Probleme der kontrastiven Phraseologie], *Phraseology: An International Handbook of Contemporary Research*. Volume 1. de Gruyter, Berlin/New York, pp. 574–589.
 7. *Lubensky S.* (1995), *Random House Russian-English Dictionary of Idioms*. Random House, New York.
 8. *Shekasjuk B. P.* (2010), *The New German-Russian Phraseological Dictionary* [Novyj nemecko-russkij frazeologičeskij slovar'], Librokom, Moscow.

Digital resources

1. *DeReKo* — Das Deutsche Referenzkorpus des IDS Mannheim im Portal COSMAS II (Corpus Search, Management and Analysis System): <https://cosmas2.ids-mannheim.de/cosmas2-web>
2. *DWDS* — Corpora des Digitalen Wörterbuchs der deutschen Sprache des 20. Jahrhunderts: <http://www.dwds.de>
3. *Online dictionary* “Deutsch-russische Idiome online“: http://wvonline.ids-mannheim.de/idiome_russ/index.htm
4. *RNC* — Russian National Corpus [Nacional'nyj korpus russkogo jazyka]: <http://www.ruscorpora.ru>