ЯЗЫКИ НАРОДОВ ЗАРУБЕЖНЫХ СТРАН (С УКАЗАНИЕМ КОНКРЕТНОГО ЯЗЫКА ИЛИ ГРУППЫ ЯЗЫКОВ)/LANGUAGES OF PEOPLES OF FOREIGN COUNTRIES (INDICATING A SPECIFIC LANGUAGE OR GROUP OF LANGUAGES)

MODELLING CONCEPTUALIZATION BASED ON CORPUS STUDIES (A CASE STUDY OF ENGLISH IDIOMS WITH THE COMPONENT "PIG")

Research article

Ivanova E.V.[1], *
[1] ORCID : 0000-0002-1990-3061;
[1] Saint Petersburg State University, Saint Petersburg, Russian Federation

* Corresponding author (e.v.ivanova[at]spbu.ru)

Abstract

The article is aimed at examining the advantages gained from the corpus analysis for exploring and constructing conceptualization models, based on linguistic signs. The goal of the article resides in the consistent description of idioms, representing the conceptualization of the domestic animal "pig", against the background of their frequency in the corpus, as well as the frequency and combining features of the words correlating with their components. The achievement of this goal results in supplementing the delineation of the conceptualization model with the dynamic perspective and introducing additional details into the interpretation of regularities involved in the idiom formation. Further research along these lines will contribute to both the more detailed and complete reconstruction of conceptual fragments in the linguistic picture of the world and more extensive and dynamic models of the processes underlying the formation of idioms.

Keywords: idiom, corpus, corpus analysis, conceptualization, frequency.

МОДЕЛИРОВАНИЕ КОНЦЕПТУАЛИЗАЦИИ НА ОСНОВЕ КОРПУСНОГО ПОДХОДА (НА МАТЕРИАЛЕ АНГЛИЙСКИХ ИДИОМ С КОМПОНЕНТОМ "PIG")

Научная статья

Иванова Е.В.[1], *
[1] ORCID : 0000-0002-1990-3061;
[1] Санкт-Петербургский государственный университет, Санкт-Петербург, Российская Федерация

* Корреспондирующий автор (e.v.ivanova[at]spbu.ru)

Аннотация

В статье рассматриваются преимущества использования корпусного анализа для исследования и конструирования моделей концептуализации, вербализованных языковыми знаками. Цель статьи заключается в последовательном описании идиом, репрезентирующих концептуализацию домашнего животного, обозначаемого существительным "pig", с учетом частотности в корпусе как этих идиом, так и словесных коррелятов их компонентов. Реализация цели дает возможность внести динамический аспект в конструируемую модель концептуализации и добавить уточнения в трактовку закономерностей формирования идиом. Дальнейшие исследования в подобном направлении будут способствовать как более детализированной и полной реконструкции отдельных фрагментов концептуализации в языковой картине мире, так и построению более подробных и динамических моделей процессов, лежащих в основе образования идиом.

Ключевые слова: идиома, корпус, корпусный анализ; концептуализация, частотность.

Introduction

Corpus approach is widely used at the moment in many fields of linguistics, allowing researchers to obtain valuable data on the functioning of words and various collocations in speech. Phraseology has not been neglected in this respect, for corpus data provide indispensable information for introducing corrections into the actual meaning of phraseological units [1], revealing their modifications in speech and specifying their typical context [2]. Corpus studies also make it possible to identify the most frequent phraseological units in order to optimize the selection of educational material when teaching a foreign language [4, P. 47], as well as significantly improve the quality of translation [3], [7], thus solving the problems of pragmatic nature. Corpus data can also be used for reconstructing the perception of the world inherent to native speakers over the period of time to which the concordances of the corpus belong. This perspective of research helps to add a dynamic aspect to the model of conceptualization based on the analysis of world interpretation, reflected in phraseology. It is necessary to admit that this line of research is only beginning to develop in modern linguistics, which allows us to conclude that the subject of the article is relevant and up-to-date.

Research methods and principles

Corpus analysis in the article is based on the Corpus of Contemporary American English (COCA) [10], considered by some scholars the most widely used and reliable corpus of the English language [4], [9]. The research material for the study comprises idioms with the component "pig". Idioms are regarded, following in A.V. Kunin's steps, as phraseological units with totally or partly reinterpreted meaning [5]. Idioms are collected from the NTC's American English Dictionary [11] and the Oxford Dictionary of Idioms [13]. Both dictionaries contain a restricted number of entries, belonging to the type of "smaller"

dictionaries. This fact made it possible to assume that their compilers had chosen the major and the most frequently used idioms, worthy of being in the focus of research. The second dictionary was added because it contains idioms of the main English areal variants, including American, and thus supplements the rather limited material of the first one. The article revolves around the idioms with the component "pig", numbering 8. Though this number looks insignificant, it nevertheless is enough to realize the goal, set in the article, i.e. to demonstrate the advantages of modelling phraseological interpretation of the world against the background of corpus data. Thus, conceptual analysis is employed along with the structural, semantic and corpus analysis.

At the first stage of corpus analysis basically the first 30 entries are taken into consideration, while the first 5, which reflect the most frequent cases, are considered to be prototypical. At the second stage all the entries are examined.

Main results and discussion

The study provides the following results:

1. The comparison of the idiom component composition with the frequency of the words corresponding to those components shows that not all frequent combinations of the noun "pig" are reflected in the idioms. This means that the prototypical features of the conceptualized animal are not necessarily directly employed in the formation of the idioms involving the name of this animal.

2. The frequency of the idioms with the component "pig" reveals the focus and periphery of the phraseological segment of the conceptualization, adding the dynamic aspect to the conceptualization model.

3. Corpus data allows researchers to model not only conceptualization represented by lexical and phraseological units, but also conceptualization connected with the titles and characters of literature and cinematography, forming a special segment of conceptualization model, thus making the model more complete.

The name of the domestic animal "pig" does not belong to the most frequent names of domestic animals encountered in the COCA concordances. As we can see from the list given below, the first position is taken by the noun "dog", followed by the nouns "horse" and "cat":

- dog — 98847;
- horse — 44525;
- cat — 41763;
- bull — 16969;
- pig — 13823;
- cow — 13181.

These figures should be taken with a grain of salt, as it were. As the analysis of the first 30 entries shows, the noun "dog" is predominantly used in its first literal meaning of a domestic animal:

As over-the-top as it may sound, my dog makes me want to be a better person.

The noun "cat" is used both as a name of the animal and in the figurative, secondary meaning:

Rory's two best friends are a cat and a pit bull.

"I don't like to get political -- but I would vote for that cat," she said.

It is also encountered as a component of an idiom:

… long time friend and admirer author and investor Douglass Kass has now let the cat out of the bag.

The data on the noun "bull" happen to be quite misleading though, because this noun is widely used as the second component of the name of the dog breed (14 out of 30 entries):

The pit bull was lying in a kennel with raw, red tissue on his back…

According to the Oxford Advanced Learner's dictionary [12], the noun "pig" has the following meanings:

1. An animal with pink, black or brown skin, short legs, a broad nose and a short curly tail. Pigs are kept on farms for their meat or live in the wild.

2. (Informal, disapproving) an unpleasant or offensive person; a person who is dirty or greedy.

The functioning of the noun "pig" in the concordances also varies. In the first 30 entries it is used as the name of the domestic animal:

Or, is it more like a farmer can have many pigs but a pig can be owned by only one farmer.

In the figurative meaning:

I would have to admit that having a pig as the leading man was quite interesting.

As a component of an idiom:

I am as happy as a pig in mud to know that I am going back to that world

As a component in the name of a totally different animal:

… we even spot an outrageously cute dassie (a myrax that resembles a guinea pig) scuttering across grey rocks.

Hence, it is necessary to conclude that manual processing of the collected data is unavoidable, when analysing the key lexical component in the model of a domestic animal conceptualization.

Now let us look at the combining features of the key word "pig".

The combination "Verb + (a) pig" displays the following results:

- buy — 19;
- kill — 16;
- make a pig — 12;
- killed a pig — 11;
- roast a pig — 10.

In the first combination 12 out of 19 cases include an idiom:

… what the prosecutors will tell you is they don't want to buy a pig in a poke. They want to know exactly what they're going to get

The noun is also used as a direct name of an animal:

So we have to say, give me money to buy a pig," said Berta Arnst, 43, who works as a milkmaid for Matantsev.

In 3 cases out of 12 "make a pig" is part of the idiom "make a pig of oneself":

Though Edith fussed at him not to make a pig of himself, he gobbled up three cookies with cheerful relish.

In all other cases the combination is part of a bigger syntactic structure:

I want you to (bleep) that mud pit' till I make a pig noise.

… and pulled her chin down to make a pig's snout.

With the verbs "to kill", "killed" and "roast" the noun is used in its direct meaning:

When I ask how often they kill a pig, John says every week.

Tonight we're having a feast. We've killed a pig and we've got meat.

…today is my comeback meal, and I decided to keep it simple and roast a pig.

The combination "Verb + the pig" is represented by the following verbs:

- kill — 15;
- put — 11;
- get — 10;
- thank — 10;
- got — 8.

The verbs "kill", "put" and "get/got" denote a physical actions directed at the domestic animal.

You didn't touch it, did you? - I did not kill the pig, Patty. You got a bad record. You got to admit it.

So I make a hole in a box and put the pig inside with its snout sticking out through the hole.

There's still time to get food. I'll go get the pig and you can call everybody.

The combination "thank the pig" comes from the film "The Pig in the City" in 9 cases out of 10. As for the combination "got + the pig" in 7 entries out of 8 a physical action directed at the animal is described, in one entry the noun is used in a figurative meaning, referring to a person.

"His little sister got the pig! " " Is that right. " "And she cried about it too".

Now let us turn to the combination "Adj + pig", or rather "attribute + pig", because not only frequent adjectives in the true sense of the term come up in the search. This combination is by far more widely represented in the corpus than the combination with the verb:

- little — 208;
- fat — 131;
- stuck — 73;
- f***ing — 70;
- porky — 69.

The adjective "little" is used with the noun "pig" to name a young pig, a piglet, to refer to the character of the fairy-tale and to express a derogative attitude to a person.

…a slaughter house, of all places, and they wanted to bring a little pig into their classroom. So he got this little pig from the slaughter house and…

The third little pig was smart and decided that she would take her time and build a strong house.

You are an animal! A disgusting little pig! Hey! Look what he did!

The second most frequently used adjective "fat" combines with the noun to refer both to an animal and person.

I was looking in the local paper, and her father had a fat pig for sale

I mean, look, if you're a fat pig, you can imagine what I am. Listen. You are all gorgeous.

As the Oxford Dictionary of Idioms explains, "a stuck pig is one that is being butchered by having its throat cut" [13, P. 219]. In corpus concordances the collocation "stuck pig" is basically used as part of two idioms: bleed like a pig (14 entries out of the first 30); squeal/scream like a pig (11 entries out of the first 30).

The attribute "porky" describes the physical feature of the animal, while the intensifier "f***ing" is used to enhance the derogative attitude to a person in question.

To sum up, it is possible to conclude that in the concordances of COCA the pig is seen basically as an animal with certain physical features (little, fat, stuck, porky), at which some physical actions can be directed (to kill, to put, to roast, to get, to buy). There also exists a metaphorical side to this conceptualization, revealing the association of the pig with a person and reflecting a derogative attitude both to the person in question and the animal. The usage of the idioms encountered in the first 5 entries provides a phraseological segment in the linguistic interpretation of this domestic animal. Another segment is connected with the association with literary and film characters, whose names are based on this noun. We can also see that by far not all features, characterizing the pig according to the analysis of the first 5 entries, are encountered in the idioms represcnted in the two dictionaries, which means that the prototypical vision of the animal is not directly involved into the formation of the idioms with its name,

Now let us have a closer look at the phraseological segment, without restricting its representation by 30 entries, but taking into account the whole list of concordances with the selected idioms with the component "pig". Their frequency in the corpus is not high, which is, on the one hand, in accordance with the statements of a number of scholars, who mention the low frequency of idioms in speech in general [10, P. 39], [14, P. 5]. On the other hand, it proves the importance of corpus data for lexicography, mentioned in scientific literature [6], as e.g., "in a pig's eye", characterized in the NTC's American English Dictionary as "chiefly North American", has only one entry in COCA, which questions the rationale of including it in the dictionary.

The 8 idioms selected for the analysis display the following frequencies:

- sweat like a pig (sweat profusely) — 71;

- a pig in a poke (something that is bought or accepted without knowing its value or seeing it first) — 45;
- bleed like a (stuck) pig (bleed copiously) — 36;
- squeal (scream, howl) like a stuck pig (squeal or yell loudly and shrilly) — 21;
- make a pig of yourself (overeat) — 12;
- pig in the middle (a person who is placed in an awkward situation between two others) — 6;
- fat as a pig (exceptionally fat) — 4;
- in a pig's eye (expressing scornful disbelief at a statement) — 1.

The first idiom exhibiting the highest frequency (71 entries) has in fact nothing to do with a pig. Pigs sweat very little, or practically not at all due to the size of their sweat glands. The component "pig" in this idiom goes back to pig iron (a form of iron that is not pure [12]). During the iron smelting process hot iron was poured on sand and as the surrounding air reached its dew point, beads of moisture formed on the surface of the pieces. So the idiom creates a very misleading image of this domestic animal. It could be included into the conceptualization model only with some reservations and indications of the false perception of the pig formed in it at the synchronic level of language analysis. Here we will abstain from discussing the desirability of doing this.

The next idiom, frequently combined with the verbs "to buy" and "to sell", reflects the old trick of selling a cat in a sack (a poke) instead of a piglet. The general situation of deceit associated with the particular situation of selling or buying a "false pig" comes up in the concordances of the corpus.

The image of a stuck pig is shared by two idioms: "bleed like a (stuck) pig" and "squeal (scream, howl) like a stuck pig". In both idioms comparison with the situation involving the pig serves to intensify the character of the action, designated by the verbs. (It is necessary to mention in passing that the former idiom is used without the component "stuck" only twice).

Consequently, looking at the frequencies of the listed idioms in the process of modelling the phraseological segment of the conceptualization of the pig, we can conclude that the images of a bleeding pig and a squealing pig associated with the similar situations of bleeding and screaming involving a human being, form the focus of this conceptualization if taken together. The situation of having or rather not having a pig in a sack (poke) is very close to this focus. Other features ascribed to the pig and projected onto the human, such as greediness in eating and fatness, are not in great demand on the part of native speakers and form the periphery of the phraseological segment. Thus, the phraseological segment of the conceptualization model acquires its dynamic configuration.

Conclusion

Corpus data provide researchers with valuable information for the conceptual analysis and construction of conceptualization models, allowing them to add the dynamic perspective to those models. Corpus data also make it possible to trace what frequent features of a mental construct, regarded as prototypical, find their way into the compositional structure of idioms.

Список литературы / References

1. Баранов А.Н. Аспекты теории фразеологии / А.Н. Баранов, Д.О. Добровольский. — Москва : Знак, 2008. — 656 с.

2. Добровольский Д.О. Корпусы текстов и двуязычная фразеография / Д.О. Добровольский // Вестник Новосибирского государственного педагогического университета. — 2015. — № 5 (27). — С. 23–37.

3. Добровольский Д.О. Корпусный подход к исследованию фразеологии: новые результаты по данным параллельных корпусов / Д.О. Добровольский // Вестник Санкт-Петербургского университета. Язык и литература. — 2020. — Т. 17. — № 3. — С. 398–411.

4. Комарова И.А. Исследование английской фразеологии с помощью подходов корпусной лингвистики / И.А. Комарова, М.С. Коган // Компьютерная лингвистика и вычислительные онтологии. — 2019. — Вып. 3. — С. 40–49.

5. Кунин А.В. Курс фразеологии современного английского языка / А.В. Кунин. — Москва : Высшая школа, 1986. — 336 с.

6. Пивоварова Е.В. Метод корпусного анализа в изучении фразеологии немецкого языка. Теоретический обзор / Е.В. Пивоварова // Филологические науки. Вопросы теории и практики. — 2019. — Т. 12. — № 12. — С. 263–268.

7. Пивоварова Е.В. Корпусный анализ как инструмент для выявления семантических изменений у фразеологизмов — «ложных друзей» переводчика / Е.В. Пивоварова // Филологические науки. Вопросы теории и практики. — 2020. — Т. 13. — Вып. 9. — С. 312–320.

8. Corpus of Contemporary American English. — URL: https://www.english-corpora.org/coca/ (accessed: 04.05.2025).

9. Lindquist H. Corpus linguistics and the description of English / H. Lindquist. — Edinburgh : Edinburgh University Press, 2009. — 219 p.

10. Minugh D. Studies in Corpora and Idioms: Getting the Cat Out of the Bag / D. Minugh. — Stockholm : University of Stockholm, 2014. — 217 p.

11. NTC's American English Dictionary / R.A. Spears (ed.). — 3rd edition. — Lincolnwood : NTC Publishing Group, 2000. — 625 p.

12. Oxford Advanced Learner's Dictionary / S. Wehmeier (ed.). — 6th edition. — Oxford : Oxford University Press, 2000. — 1540 p.

13. Oxford Dictionary of Idioms / edited by J. Siefring. — 2nd edition. — Oxford : Oxford University Press, 2004. — 340 p.

14. Rafatbakhsh E. A thematic corpus-based study of idioms in the Corpus of Contemporary American English / E. Rafatbakhsh, A. Ahmadi // Asian-Pacific Journal of Second and Foreign Language Education. — 2019. — Vol. 4. — № 11. — P. 1–21.

Список литературы на английском языке / References in English

1. Baranov A.N. Aspekty teorii frazeologii [Aspects of Phraseology Theory] / A.N. Baranov, D.O.Dobrovolsky. — Moscow : Znak, 2008. — 656 p. [in Russian]

2. DobrovolskyD.O. Korpusy tekstov i dvuyazychnaya frazeografiya [Text Corpora and Bilingual Phraseography] / D.O.Dobrovolsky// Vestnik Novosibirskogo gosudarstvennogo pedagogicheskogo universiteta [Bulletin of Novosibirsk State Pedagogical University]. — 2015. — № 5 (27). — P. 23–37. [in Russian]

3. DobrovolskyD.O. Korpusnyy podkhod k issledovaniyu frazeologii: novye rezul'taty po dannym parallel'nykh korpusov [Corpus Approach to Phraseology Research: New Results from Parallel Corpora] / D.O.Dobrovolsky// Vestnik Sankt-Peterburgskogo universiteta. Yazyk i literatura [Bulletin of Saint Petersburg University. Language and Literature]. — 2020. — Vol. 17. —№ 3. — P. 398–411. [in Russian]

4. Komarova I.A. Issledovanie angliyskoy frazeologii s pomoshch'yu podkhodov korpusnoy lingvistiki [Study of English Phraseology Using Corpus Linguistics Approaches] / I.A. Komarova, M.S. Kogan // Komp'yuternaya lingvistika i vychislitel'nye ontologii [Computational Linguistics and Computational Ontologies]. — 2019. — Issue 3. — P. 40–49. [in Russian]

5. Kunin A.V. Kurs frazeologii sovremennogo angliyskogo yazyka [A Course in Modern English Phraseology] / A.V. Kunin. — Moscow : Higher School, 1986. — 336 p. [in Russian]

6. Pivovarova E.V. Metod korpusnogo analiza v izuchenii frazeologii nemetskogo yazyka. Teoreticheskiy obzor [Corpus Analysis Method in the Study of German Phraseology. Theoretical Review] / E.V. Pivovarova // Filologicheskie nauki. Voprosy teorii i praktiki [Philological Sciences. Issues of Theory and Practice]. — 2019. — Vol. 12. — № 12. — P. 263–268. [in Russian]

7. Pivovarova E.V. Korpusnyy analiz kak instrument dlya vyyavleniya semanticheskikh izmeneniy u frazeologizmov —"lozhnykh druzey"perevodchika [Corpus Analysis as a Tool for Identifying Semantic Changes in Phraseological Units — "False Friends" of Translators] / E.V. Pivovarova // Filologicheskie nauki. Voprosy teorii i praktiki [Philological Sciences. Issues of Theory and Practice]. — 2020. — Vol. 13. — Issue 9. — P. 312–320. [in Russian]

8. Corpus of Contemporary American English. — URL: https://www.english-corpora.org/coca/ (accessed: 04.05.2025).

9. Lindquist H. Corpus linguistics and the description of English / H. Lindquist. — Edinburgh : Edinburgh University Press, 2009. — 219 p.

10. Minugh D. Studies in Corpora and Idioms: Getting the Cat Out of the Bag / D. Minugh. — Stockholm : University of Stockholm, 2014. — 217 p.

11. NTC's American English Dictionary / R.A. Spears (ed.). — 3rd edition. — Lincolnwood : NTC Publishing Group, 2000. — 625 p.

12. Oxford Advanced Learner's Dictionary / S. Wehmeier (ed.). — 6th edition. — Oxford : Oxford University Press, 2000. — 1540 p.

13. Oxford Dictionary of Idioms / edited by J. Siefring. — 2nd edition. — Oxford : Oxford University Press, 2004. — 340 p.

14. Rafatbakhsh E. A thematic corpus-based study of idioms in the Corpus of Contemporary American English / E. Rafatbakhsh, A. Ahmadi // Asian-Pacific Journal of Second and Foreign Language Education. — 2019. — Vol. 4. — № 11. — P. 1–21.