

СЛОВАРНЫЕ ПРОЕКТЫ И ТРУДЫ DICTIONARY PROJECTS AND WORKS

Научная статья
УДК 81`374
doi: 10.17223/22274200/34/1

Корпусный подход в создании тезауруса идиом татарского языка

Гузель Казбековна Гизатова¹

¹ *Казанский федеральный университет, Казань, Россия, guzelgizat52@gmail.com*

Аннотация. Представлен проект «Словаря-тезауруса идиом татарского языка». Описываются основные принципы, лежащие в основе его построения. Необходимость создания такого словаря связана с тем, что в настоящее время в отечественной лексикографической практике отсутствует полное идеографическое описание идиом татарского языка. Исследование базируется на основных постулатах когнитивной лингвистики и, в первую очередь, на системной организации структурирования семантических полей. Результаты анализа показали, что применение корпусов текстов представляет новую информацию о семантике татарских идиом.

Ключевые слова: тезаурус, идиомы, татарский язык, семантические поля, корпус

Для цитирования: Гизатова Г.К. Корпусный подход в создании тезауруса идиом татарского языка // Вопросы лексикографии. 2024. № 34. С. 5–23. doi: 10.17223/22274200/34/1

Original article

doi: 10.17223/22274200/34/1

A corpus-based approach to a thesaurus of Tatar idioms

Guzel K. Gizatova¹

¹ *Kazan Federal University, Kazan, Russian Federation, guzelgizat52@gmail.com*

Abstract. The main aim of the project is an ideographic description of modern Tatar idioms, i.e., organization of idioms according to semantic fields which are defined as a set of lexical units connected with each other by some salient semantic features. The totality of semantic fields along with the conceptual links between them construct the thesaurus of a given language, which can be represented in the form of a semantic network. The need for a new type of dictionary is motivated by the fact that, at present, there is no a comprehensive dictionary of Tatar idioms built on a thesaural principle. The idioms are explored along two revealing lines of inquiry. The first line focuses on a semantic description of idioms and their placement into appropriate taxa according to their meanings. The second line illustrates advantages of organizing a dictionary based on an analysis of corpus data. The idioms were collected from monolingual and bilingual Tatar phraseological dictionaries by continuous sampling. The illustrative materials were sampled from the Written Tatar Corpus containing 620 million words; the Tatar National Corpus (260 million words); and corpus query system Sketch Engine, its **Tatar Web** Corpus (102 million words). This made it possible to find all occurrences of idioms under analysis and obtain statistically representative data. The findings were processed manually. The Thesaurus is based on an inductive method, that is, from idioms to semantic fields and not from an abstract logical outline to idioms. The aim of the first stage was to describe the semantic features of each idiom and assign a descriptor or several descriptors to the idiom under consideration. At the second stage, the idioms were grouped into semantic fields/taxa in accordance with their meanings. The third stage of classification is connected with checking the correctness of the conceptual marking of idioms, which is of primary importance for regrouping the existing idiom stock. At this point, the frequency of each idiom was checked in the corpus. After idioms were classified on the basis of their dictionary definitions, it was necessary to clarify the information about their meaning and usage. For this purpose, idioms were searched in text corpora. In our case, a corpus-illustrated approach is used as a source of illustrative material. Its purpose is the selection of contexts that could be a source of paradigmatic examples. This approach is relevant in lexicographic work. There is no doubt that a dictionary in which material is drawn from the corpus is more objective, than one with illustrative examples based on linguistic intuition. The results of the research have proved that the use of illustrative materials based on

corpus data presents new information on the semantic features of idioms, which contributed to the achievement of the aim.

Keywords: thesaurus, idioms, Tatar language, semantic fields, corpus

For citation: Gizatova, G.K. (2024) A corpus-based approach to a thesaurus of Tatar idioms. *Voprosy leksikografii – Russian Journal of Lexicography*. 34. pp. 5–23. (In Russian). doi: 10.17223/22274200/34/1

Введение

Цель данной статьи – представить концепцию «Словаря-тезауруса идиом татарского языка» (далее – Тезаурус), которая описывается впервые в отечественной лексикографии. Для достижения основной цели статьи были поставлены следующие задачи: 1) представить принципы, по которым идиомы классифицируются и распределяются по таксонам словника, 2) проиллюстрировать преимущества использования корпусных данных для составления словаря.

Отличие Тезауруса (идеографического словаря) от других лексикографических источников заключается в способе структурирования языкового материала. Идиомы в словаре организованы на основе их значений, а не в алфавитном порядке. Необходимость создания нового типа фразеологического словаря мотивирована тем, что в настоящее время в отечественной лексикографии Тезауруса татарских идиом нет. Новизна подхода к составлению словаря проявляется в отражении семантических связей между таксонами – антонимических, родовидовых и отношений включения, которые охватываются системой парадигматических отсылок.

В зависимости от типа словаря варьируются как область фиксируемого знания, так и методы его представления. «В наиболее четком и ясном виде идея представления знаний о мире связывается с так называемыми идеографическими словарями, поскольку в качестве входа словарных статей фигурируют не выражения естественного языка, а выражения семантического метаязыка, непосредственно соотносимые со сферой знаний» [2. С. 364].

В работе мы опираемся на следующее определение термина «тезаурус» (от греч. *θησαυρός* – сокровище, сокровищница) – это «словарь, который в явном виде фиксирует семантические отношения между составляющими его единицами» [3. С. 148]. Тезаурус Роже (*Roget's Thesaurus of English Words and Phrases* – первое издание в 1852 г.)

является первым и наиболее известным из имеющихся тезаурусов в мировой лексикографической практике [4]. Словарь имеет индекс концептов и составлен согласно семантическим категориям. Тезаурус Роже – первый источник теоретически обоснованной модели логической структуры английской лексики. Этот труд был мотивирующим фактором становления идеографической лексикографии во многих языках: французском [5], немецком [6], испанском [7].

Наибольший интерес в отечественной лингвистике к идеографической лексикографии в лексической системе языка проявился в 70–80-е гг. прошлого столетия. В 1976 г. издан «Минимальный идеографический словарь» Ю.Н. Караулова [8], в 1995 вышел в свет «Идеографический словарь русского языка», составленный О.С. Барановым [9]. В тюркологии в 1975 г. опубликован первый двуязычный словарь, организованный по тематическому принципу «Русско-узбекский тематический словарь» [10]. Слова размещены по 16 темам, которые объединены в 3 цикла: «Природа», «Жизнь и деятельность человека» и «Пространство, время, количество, счёт». После значительного затишья в три десятилетия интерес к тематическим и идеографическим словарям в тюркологии пробуждается вновь и ряд двуязычных словарей выходит в свет: «Русско-азербайджанский тематический словарь» [11], «Русско-турецкий тематический словарь» [12], «Русско-якутский тематический словарь» [13]. Все названные словари, созданные с учетом достижений лексикографии на уровне своего периода и степени исследованности языка, могут служить базой для последующих еще более плодотворных лексикографических изысканий.

В отличие от лексических тематических словарей, замыслом нашего проекта является построение фразеологического «Словаря-тезауруса современного татарского языка». Объект лексикографирования – фразеологические единицы (ФЕ) современного татарского литературного языка. В работе понятия «идиомы» и «фразеологические единицы» рассматриваются как синонимичные, соответственно, взаимозаменяемые, и за основу взято их определение, данное А.В. Куниным: «Фразеологические единицы, или идиомы, – это раздельнооформленные единицы языка с полностью или частично пересмысленными значениями» [14. С. 7].

У истоков разработки теории фразеологической идеографии в отечественной лингвистике стоят А.М. Эмирова [15], В.М. Мокиенко [16], Р.И. Яранцев [17], А.Н. Баранов, Д.О. Добровольский [18],

Д.О. Добровольский, Ю.Н. Караулов [19]. Практическое воплощение «Словарь-тезаурус современной русской идиоматики» получил в 2007 г., авторами словника являются А.Н. Баранов и Д.О. Добровольский [20]. Здесь следует подчеркнуть, что Словарь-тезаурус является первым и самым полным собранием идиоматических выражений современного русского языка в отечественной и зарубежной лексикографии. Этот труд представляет значительную веху в истории лексикографии и фразеологической идеографии, в частности.

Нужно отметить, что интерес к лексикографии, организованной по понятийным рубрикам во фразеологии татарского языка, пробудился гораздо раньше. Первые шаги к созданию фразеологического словаря татарского языка были предприняты еще в 50-е гг. прошлого столетия. Словарь «Фразеология, пословицы и поговорки татарского языка» вышел в свет в 1957 г. [21]. Он состоит из 6 000 единиц и включает более 60 семантических полей, таких как «богатство–бедность», «дружба–вражда», «щедрость–жадность» и т.д. Трехтомный словарь пословиц татарского языка «Татарские народные пословицы» Н. Исанбета (1957–1969) насчитывает около 8 000 единиц и также построен по тематико-идеографическому принципу [22]. В 1991 г. был издан «Учебный тематический русско-татарский фразеологический словарь» Л.К. Байрамовой [23]. Его достоинством является то, что это первый двуязычный фразеологический тематический словарь, организованный по темам (14 разделов и 95 тем). В целом следует признать, что выход в свет указанных словарей явился значительным вкладом в лексикографию татарского языка, их новизна бесспорна.

В последующий двадцатилетний период фразеологические идеографические словари татарского языка не издавались, но шла подготовительная работа по сбору фактического материала. Вдохновляющим толчком к началу работы в этом направлении послужил выход в свет «Словаря-тезауруса современной русской идиоматики» А.Н. Баранова и Д.О. Добровольского [20], который на сегодняшний день является единственным в отечественной и мировой лексикографии словарем, претендующим на полный охват идиоматики русского языка за последние 60 лет [20. С. 3].

Так, в 2010 г. вышел в свет «Идеографический фразеологический словарь (на материале татарского, русского и английского языков)» под редакцией автора настоящего исследования [24]. Это самое полное в отечественной лексикографии собрание фразеологических вы-

ражений татарского, русского и английского языков, распределенных по смысловым категориям (таксонам). Словарь насчитывает около 3 000 ФЕ в каждом из представленных в источнике языков и включает 80 семантических полей.

С 2013 г. на кафедре европейских языков и культур Казанского федерального университета ведется работа над «Словарем-тезаурусом идиом современного татарского языка», речь о котором пойдет в настоящей статье.

Татарский язык и его текущий статус

Прежде чем перейти непосредственно к предмету исследования – представлению семантики идиом в идеографическом словаре следует сказать несколько слов о татарском языке и его сегодняшнем статусе. Татарский язык относится к тюркской ветви алтайской языковой семьи; в мире проживает около 7 млн татар, из них 4.71 млн человек – в России. Татары – коренное население Российской Федерации, второй основной по численности этнос и самый многочисленный народ в РФ после русских (по данным Всероссийской переписи населения 2020–2021 гг.). Около одного миллиона татар живет в Москве, татары проживают также на Урале, в Сибири и Средней Азии, значительные диаспоры представлены в США, Австралии, Финляндии, Канаде. Вследствие географической рассредоточенности, а также ряда политических и социальных факторов татарский язык подвержен серьезным трансформациям, и есть все более убедительные доказательства того, что существует угроза его исчезновения и в недалеком будущем язык не сможет быть благополучно передан последующим поколениям и бесследно растворится. Родной язык каждого народа – это его самая главная ценность, богатство, которое надо сохранить и передать последующим поколениям. Подробно о сохранении татарского языка, об исчезновении образных средств языка см. статью автора [25].

Концепция Словаря-тезауруса идиом татарского языка

В качестве теоретической базы построения Тезауруса используются труды, выполненные в рамках когнитивной лингвистики [26–31] и, прежде всего, системной организации структурирования семантических полей [1; 18–20]. Для составления словаря применялись как спе-

циальные лингвистические методы (семантический анализ), так и общенаучные методы, такие как описание, классификация, анализ, синтез и индукция. В основе построения словника лежит индуктивный подход – не от понятийной схемы к идиомам, как в случае с лексическими тезаурусами, а от идиом к семантическим полям. Для идеографической классификации лексических единиц дедуктивный подход – от понятийной схемы – вполне оправдан, так как лексика, в отличие от фразеологии, полностью покрывает понятийное пространство языка. Как отмечают А.Н. Баранов и Д.О. Добровольский, «...многие априорно постулируемые семантические поля могут вообще отсутствовать во фразеологии. Тем самым, исходя из чисто логической схемы, невозможно предсказать в ней существование тех или иных полей» [20. С. 1112].

Словарь описывает единицы, которые представляют собой центральную часть фразеологической системы татарского языка – идиомы. Основной структурной единицей Тезауруса является семантическое поле, или таксон, представляющий собой группу идиом, объединенных семантическим сходством. Например, в таксон ЭМОЦИИ в словнике включаются идиомы, выражающие разные эмоциональные состояния:

түбәсе күккә тигән ‘макушкой коснулся неба’ в знач.: ‘сверх меры рад’; жиденче кат күккә ашу ‘быть на седьмом небе’ в знач.: ‘быть безгранично счастливым’;

(2) чәчләре үрә тору ‘волосы встали дыбом’; кот очу ‘душа улете-ла’; жан бугазга килү ‘душа подступила к горлу’, все ФЕ в знач.: ‘сильно испугаться’;

(3) балтасы суга төшкән (кеше) кебек ‘будто топор в воду уронил’ в знач.: ‘морально подавленный, упавший духом’; борын салындыру ‘повесить нос’, в знач.: ‘приходить в уныние, отчаяние’; әвене янган мыни ‘будто у него овин сгорел’ в знач.: ‘быть в мрачном, унылом состоянии’.

Крупные таксоны, как правило, подразделяются на подтаксоны, которые могут дробиться и дальше до нескольких уровней. В частности, каждая из представленных выше идиом входит в разные подтаксоны семантического поля ЭМОЦИИ. Идиомы (1) входят в подтаксон РАДОСТЬ, УДОВОЛЬСТВИЕ, СЧАСТЬЕ¹, идиомы (2) относятся к

¹ Названия таксонов и подтаксонов заимствованы из «Словаря-тезауруса современной русской идиоматики» [20].

подтаксону СТРАХ, ИСПУГ, УЖАС, а в подтаксон ГОРЕ, ОТЧАЯНИЕ, СТРАДАНИЕ входят идиомы (3).

Тезаурус отличается от тематических словарей татарского языка [21–23] ориентацией на современный аутентичный языковой материал, а не на примеры из лексикографических источников. Иллюстративный материал для Тезауруса собран из двух корпусов: «Письменного корпуса татарского языка», насчитывающего более 620 млн слов [32], и корпуса «Туган тел» (в переводе – «родной язык») [33], включающего 260 млн словоупотреблений. Кроме того, была использована поисковая система Sketch Engine, ее подкорпус татарского языка [34].

Классификация идиом и её корпусная верификация

Работа над Тезаурусом состояла из нескольких этапов. Сначала создавалась база данных словаря, и на предварительной стадии (2013) эмпирический материал составлял 2 900 идиом татарского языка. Во время первого этапа сбора материала единицы были отобраны методом сплошной выборки из одноязычных и двуязычных лексикографических источников татарского языка [21–23; 35–36].

Единицы словаря формируются в ходе исследования, а не задаются *apriori*. Следуя стратегии Д.О. Добровольского [37. Р. 264], на первом этапе систематизации материала к значению каждой идиомы приписывается маркер (дескриптор). Важным свойством фразеологических единиц является семантический синкретизм, т.е. способность совмещения нескольких значений в одной форме. Ср. примеры (4) и (5):

(4) Хөрмәтле геройлар, без план тулсын дип, салкында, яңгырда көн дими, төн дими, фидакарь **бил бөгеп**, тир түгеп, йөкләмәләрне үтәп йөргәндә, кайберәүләр жылы өйдә чэй эчеп утырдылар... – *Уважаемые герои, мы, для выполнения плана, и в холод, и в снег и в дождь, днем и ночью самоотверженно гнем спину; однако, в то время как мы делаем все для выполнения наших обязательств, некоторые сидят у себя дома в тепле, распивая чай...* (перевод наш. – Г.Г.).

(5) Меңбаш, белеп торсаң иде, мин беркайчан да **җаһилләргә** баш биреп, **бил бөгеп** яшәмәдем. – *Если хочешь знать, Минбаш, я никогда не подчинялся невеждам и не унижался перед ними* (перевод наш. – Г.Г.).

Судя по контекстам, ФЕ *бил бегү* – ‘гнуть спину’ имеет значение не только ‘нести непосильный труд’, но также означает и ‘выражать

покорность, раболепствовать, заискивать, унижаться'. Каждому из значений идиомы был приписан семантический дескриптор, который отражает наиболее существенные элементы плана содержания этой идиомы. В частности, идиома *бил бегу* имеет два дескриптора: 'тяжелый труд' и 'покорность, унижение'.

На втором этапе идиомы группировались согласно их значению по таксонам или семантическим полям. Идиома *бил бегу* была распределена в два семантических поля (таксона) Тезауруса – ТРУД и ПОКОРНОСТЬ, УНИЖЕНИЕ.

Третий этап построения словаря включает работу, связанную с проверкой правильности концептуальной маркировки идиом. Этот этап является важнейшим в перегруппировке существующего массива эмпирического материала исследования.

Заключительным этапом работы является поиск аутентичного материала в корпусах татарского языка. Целью этого этапа является верификация использования идиом в современном дискурсе. В ряде случаев информация в словарях отличалась от реальных контекстных условий использования той или иной единицы в аутентичных текстах, представленных в корпусах. В результате выявления этих различий (в частности, квазисинонимии или полисемии идиом, не представленных во многих словарях татарских идиом), вносились изменения, связанные с включением идиом в другие таксоны и подтаксоны словаря.

Например, согласно информации во фразеологических словарях татарского языка, ФЕ *авызны чөйгә элү* 'повесить рот на вешалку' имеет значение 'голодать' (ср. *положить зубы на полку*). Поэтому эта ФЕ была определена в таксон ЕДА, ПИЩА, который дробится на более мелкие подтаксоны, одним из которых является – НЕДОСТАТОК ЕДЫ, ГОЛОД. От этого подтаксона идут парадигматические отсылки (→) к семантическим полям БЕДНОСТЬ и ТРУДНОСТИ.

Далее, поиск в «Письменном корпусе татарского языка» выдал еще одно вхождение ФЕ в значении 'обижаться'. В этом случае ФЕ была определена в таксон ЭМОЦИОНАЛЬНЫЕ СОСТОЯНИЯ, в его подтаксон ОБИДА, который имеет отсылку к подтаксону этого же поля (→) ПЛАЧ. См. контексты (6) и (7) из корпуса.

(6) *Заманында бәрәңге кәлҗемәсе ашадык, алабута белән арыш онын катнаштырып икмәк пешердек, кычыткан, балтырган ашлары белән тукландык, чәй урынына яфрак, кишер суы эчтек, авызны*

чөйгә элгән чаклар да күп булды. – В свое время мы ели котлеты из картофеля, пекли хлеб из ржаной муки, смешанной с лебедой, вместо чая пили напиток из листьев, из моркови, порой просто **голодали** (перевод наш. – Г.Г.).

(7) *Егеткә сүз әйтеп булмый, әйттеңме авызын чөйгә элп куя.* – Парню слова не скажи, стоит сказать – сразу **обижается** (перевод наш. – Г.Г.).

В итоге, ФЕ *авызын чөйгә элү* попадает в два семантических поля Тезауруса: 1) НЕДОСТАТОК ЕДЫ, ГОЛОД → ТРУДНОСТИ... → БЕДНОСТЬ и 2) ОБИДА → ПЛАЧ.

В результате классификации материала было выделено 86 семантических полей. Наиболее продуктивными в татарской фразеологии являются семантические поля ЭМОЦИИ, МЫШЛЕНИЕ, РЕЧЕВЫЕ АКТЫ, ПРОБЛЕМЫ – ОТСУТСТВИЕ ПРОБЛЕМ.

На данном этапе работы Словарь-тезаурус насчитывает около 6 000 ФЕ и более 22 000 контекстов их употребления.

Тезаурус идиом татарского языка как семантическая сеть

При формировании семантической сети словника авторы следуют главным принципам, лежащим в основе организации «Словаря-тезауруса современной русской идиоматики» [20]. Базовой единицей словника является таксон, который состоит из семантического или концептуального маркера (дескриптора) и группы лингвистических знаков (слов и/или идиом), отражающих суть соответствующего концепта. К тому же Тезаурус включает систему дополнительных средств, указывающих на взаимосвязь концептов [37. Р. 264].

В табл. 1 на примере семантического поля ИСТИННОЕ – ЛОЖНОЕ рассмотрим внутренние связи и взаимодействие концептов внутри таксона.

Как видно из табл. 1, крупные таксоны делятся на подтаксоны, например, таксон ИСТИННОЕ–ЛОЖНОЕ имеет один подтаксон ИСТИНА, ПРАВДА – ЛОЖЬ, ОБМАН (1). Последний имеет три подтаксона: ИСТИНА, ПРАВДА (1.1), ЛОЖЬ, ОБМАН (1.2) и ЧЕСТНОСТЬ – НЕЧЕСТНОСТЬ (1.3). Таксоны 1.2 и 1.3 также подразделяются на более мелкие подтаксоны. В некоторых случаях глубина такого деления в Тезаурусе может достигать до пяти-шести уровней. Парадигматические отсылки указывают на «зоны сгущения значений» (таксоны или подтак-

соны словника), которые семантически близки друг к другу [2]. Например, идиомы таксона ПРИТВОРСТВО семантически связаны с идеей НЕЧЕСТНОСТЬ / НЕИСКРЕННОСТЬ / ЛИЦЕМЕРИЕ и поэтому от таксона ПРИТВОРСТВО идет отсылка → к таксонам НЕЧЕСТНОСТЬ / НЕИСКРЕННОСТЬ / ЛИЦЕМЕРИЕ.

Таблица 1

Фрагмент Тезауруса: «ИСТИННОЕ–ЛОЖНОЕ»

ИСТИННОЕ– ЛОЖНОЕ 1. ИСТИНА, ПРАВДА – ЛОЖЬ, ОБМАН 1.1. ИСТИНА, ПРАВДА <i>битлеген салдыру</i> ‘снять маску’; <i>дореслек-кэ туры карау</i> ‘смотреть правде в глаза’ 1.2. ЛОЖЬ, ОБМАН → ЛОЖНЫЙ СТАТУС → ПРЕДАТЕЛЬСТВО 1.2.1. УЧАСТНИКИ, ИНСТРУМЕНТЫ СИТУАЦИИ ЛЖИ, ОБМАНА <i>күз буяучы</i> ‘замазывающий глаза’, в знач.: ‘очковтиратель’; <i>кэкре куллы</i> ‘криворукый’ в знач.: «вор». 1.2.2. НЕВЕРБАЛЬНЫЙ ОБМАН → ВОРОВСТВО, КРАЖА <i>агач атка атландыру</i> ‘посадить на деревянную лошадь’, в знач.: ‘обвести вокруг пальца’; <i>кэкре каенга терэту</i> ‘прислонить к кривой березе, в знач.: ‘обвести вокруг пальца’; <i>куык очыру</i> ‘пускать пузыри’, в знач.: ‘обманывать; создавать о себе ложное впечатление; водит за нос’ 1.2.2.1. ПРИТВОРСТВО → НЕЧЕСТНОСТЬ / НЕИСКРЕННОСТЬ/ ЛИЦЕМЕРИЕ <i>башны тилегэ салу</i> ‘притворяться дурачком’ 1.2.3. ЛОЖЬ, ВЕРБАЛЬНЫЙ ОБМАН, ВВЕДЕНИЕ В ЗАБЛУЖДЕНИЕ → ХВАСТОВСТВО <i>борчак сибү</i> ‘сыпать горох’, в знач.: ‘обманывать’, ср.: ‘врать как сивый мерин’;	<i>куык очыру</i> ‘пускать пузыри’, в знач.: ‘обманывать; создавать о себе ложное впечатление’; <i>сабын күбеге очыру</i> ‘пускать мыльные пузыри’ в знач.: ‘создавать о себе ложное впечатление’, ср.: ‘пускать пыль в глаза’ 1.2.4. САМООБМАН, ИЛЛЮЗИИ <i>хыялга бату</i> ‘утонуть в грезах’ в знач.: ‘предаваться бесплодным фантазиям’ ср.: ‘витать в облаках’ 1.3. ЧЕСТНОСТЬ – НЕЧЕСТНОСТЬ 1.3.1. ЧЕСТНОСТЬ, ОТКРОВЕННОСТЬ, ИСКРЕННОСТЬ → КЛЯТВА <i>турыдан туры айту</i> ‘сказать напрямую’, в знач.: ‘сказать правду’; <i>исеме белэн атау</i> ‘назвать по имени’, в знач.: ‘называть вещи своими именами’. 1.3.2. НЕЧЕСТНОСТЬ, НЕИСКРЕННОСТЬ, ЛИЦЕМЕРИЕ → ХИТРОСТЬ... → ПРИТВОРСТВО → БЕЗНРАВСТВЕННОСТЬ <i>көйрык болгау</i> ‘вилять хвостом’, в знач.: ‘увиливать, прибегая к хитрости и уловкам’; <i>ике яклы пычак</i> ‘обоюдоострый нож’, в знач.: ‘и нашим, и вашим’
--	---

В табл. 1 в целях экономии места представлено только по несколько ФЕ семантического поля ИСТИННОЕ – ЛОЖНОЕ. В целом в нашем материале на данный момент имеется 366 ФЕ данного поля. Следует также подчеркнуть, что каждая ФЕ в Тезаурусе вводится несколькими аутентичными контекстами из корпусов татарского языка. В табл. 2 представлен фрагмент таксона ИСТИННОЕ–ЛОЖНОЕ, в котором идиома *куык очыру* иллюстрируется примерами из корпуса. В данной таблице приведен пример использования лишь одной ФЕ, так как формат статьи не дает нам возможность представить большее количество контекстов из корпуса. Как видно, в табл. 1 выделяются подтаксоны НЕВЕРБАЛЬНЫЙ ОБМАН (1.2.2) и ВЕРБАЛЬНЫЙ ОБМАН (1.2.3). Идиома *куык очыру* ‘пускать пузыри’, в знач.: ‘обманывать; создавать о себе ложное впечатление’ имеет как вербальный, так и невербальный характер употребления, поэтому она включена в оба подтаксона.

В табл. 2 представлен образец словарной статьи Словаря-тезауруса без перевода на русский язык. Перевод дан ниже и выведен за пределы табл. 2. Лемма набирается полужирным шрифтом, за леммой идут примеры из корпусов, которые вводятся знаком . Знак • отделяет примеры друг от друга. Каждый пример имеет указание на источник: ПКТЯ – «Письменный корпус татарского языка» или НКТЯ – «Национальный корпус татарского языка».

Таблица 2

Словарная статья идиомы *куык очыру*

1.2.2. НЕВЕРБАЛЬНЫЙ ОБМАН → ВОРОВСТВО, КРАЖА Куык очыру  Син, Идмас, куык очырма эле. Энис масаючан кешелэрне яратмый иде (НКТЯ) • Э Мөгаллимэ яшел күлмәк аулак өйләргә йөри, килешлеитеп сугышка эле китеп бетмәгән егетләрне үзенә карата, кәпрәя, <i>куык очыра</i> (ПКТЯ) • Ул Гиндстанда укып кайткан, соңын Хвалыньск шәһәрәндәге татарлар өчен – гиндистан мәдрәсәләре тибында мәдрәсәләр салдырам, дип <i>куык очырып</i> йөргән, шул максат белән кешедән байтак кы-	1.2.3. ЛОЖЬ, ВЕРБАЛЬНЫЙ ОБМАН, ВВЕДЕНИЕ В ЗАБЛУЖДЕНИЕ Куык очыру  Миңлегата <i>куык очыра</i>. Әй мактана, әй мактана: ‘Курыкмыйм, - ди Миңлегата, - бернидән дә, куркмыйм, - ди, - аюдан да, бүредән дә...’ (ПКТЯ) • Яушев мәжлесләрдә мактаньрга ярата, ә Муса әче телле, теге <i>куык очыра</i> башлагач, бер ничә сүз белән авызын каплар (НКТЯ) • Нәби кайчакта үзе дә ялганына ышанып куя. Бөрвакыт ни сөйләсә дә ышанмый башладылар аңа. ‘Һи, <i>куык очыра</i> башлады инде’, диләр
---	---

на акча жыеп алган һәм, шул акчаны кесәсенә салып, беркем алдындада хисап бирмичә, кыяклаган аферист... (ПКТЯ)	малайлар. Шуннан соң Нәби иптәшләре алдында күп сөйләмәс булды (ПКТЯ)
--	---

Ниже, в контекстах (8–13) приводится перевод высказываний с идиомой *куык очыру* согласно порядку их употребления в табл. 2.

1.2.2. НЕВЕРБАЛЬНЫЙ ОБМАН → ВОРОВСТВО, КРАЖА

(8) *‘Тебе, Идмас, нечего тут **пускать** нам **пыль в глаза**’ – Анис не любил хвастливых людей и наконец-то высказал все Идмасу в лицо (перевод наш. – Г.Г.).*

(9) *А Маглима как наденет свое зеленое платье, так и ходит по укромным вечеринкам, ладненько танцует, чтобы на нее заглядывались ребята, еще не ушедшие на фронт, умеет выставять себя и **водить их за нос** (перевод наш. – Г.Г.).*

(10) *Он учился в Индии, затем вернулся домой, поселился в Хвалынске; там он убедил местных татар, что будет строить для них мусульманские школы, как в Индии. В принципе он **водил их за нос**, собрал с людей крупную сумму денег, забил деньгами карманы и был таков... исчез бесследно, ни перед кем не отчитавшись (перевод наш. – Г.Г.).*

1.2.3. ЛОЖЬ, ВЕРБАЛЬНЫЙ ОБМАН, ВВЕДЕНИЕ В ЗАБЛУЖДЕНИЕ

(11) *Минлегата **пускает пыль в глаза**, как хвастается, ой как хвастается. ‘Не боюсь, – говорит Минлегата, – ничего не боюсь – ни медведя, ни волка...’ (перевод наш. – Г.Г.).*

(12) *Яушев любит хвастаться на всяких мероприятиях, застольях. А Муса очень злой на язык, когда тот начинает **пускать пыль в глаза**, он буквально несколькими словами умеет заставить хвастуна замолчать (перевод наш. – Г.Г.).*

(13) *Наби порой даже сам начинал верить своим выдумкам. В конце концов, наступил такой момент, когда люди перестали ему верить, что бы он ни сказал. ‘Фу, – говорили ребята, – опять начал **мозги пудрить**’. В конце концов Наби решил больше при друзьях много не говорить (перевод наш. – Г.Г.).*

Как видно из приведенных контекстов (8–10), значение идиомы *куык очыру* носит невербальный характер. В контексте (8) Идмас, судя по всему, постоянно пытается оказать психологическое воздействие на окружающих, создавая ложное впечатление о себе, представляя свое

положение, способности, возможности выше, чем они есть на самом деле. Пример (9) наглядно демонстрирует манипулятивное воздействие Маглиммы, героини повествования, на остальных его участников. В приведенном контексте идиома *куык очыру* выражает лживое поведение Маглиммы, которая пытается выставить себя в выгодном свете, поднять свой имидж с целью соблазнения молодых мужчин. В контексте (10) также очевидно манипулятивное воздействие основного участника повествования на большое количество людей. В данном примере весьма наглядно представлено использование манипулятором религиозных чувств людей. Он создает ложные иллюзии у населения для управления их поведением. В данном контексте идиома *куык очыру* вполне соответствует русской идиоме *водить за нос*.

Основные семантические связи, представленные в Словаре-тезаурусе: антонимические, родовидовые и отношения включения [2. С. 393].

Рассмотрим способы их отражения в структуре словника.

1. Антонимичные отношения представлены в названиях таксонов: РАДОСТЬ – ПЕЧАЛЬ, СВОБОДА – ОТСУТСТВИЕ СВОБОДЫ, ДРУЖБА – ВРАЖДА. В каждом из таксонов выделяются подтаксоны, содержащие соответствующие антонимичные семы. Например, таксон РАДОСТЬ – ПЕЧАЛЬ имеет три подтаксона: 1) РАДОСТЬ, УДОВОЛЬСТВИЕ, СЧАСТЬЕ, 2) СМЕХ – СЛЕЗЫ и 3) ПЕЧАЛЬ, СОЖАЛЕНИЕ, УНЫНИЕ. Каждый из приведенных подтаксонов имеет свои более мелкие подтаксоны и т.д.

2. Родовидовые отношения отличаются тем, что таксон выступает в родовой функции по отношению к другим таксонам, обозначающим виды, выделяемые в рамках этого рода. Как видно из табл. 1, от подтаксона 1.2.2. НЕВЕРБАЛЬНЫЙ ОБМАН делается отсылка (→) к таксонам словника ВОРОВСТВО, КРАЖА, являющимися видами НЕВЕРБАЛЬНОГО ОБМАНА.

3. Отношения включения характеризуют ситуацию, «когда идиомы таксона *X* содержат сему *a*, как и идиомы таксонов *Y*, ..., *Z*. Но при этом значение идиом, входящих в таксоны *Y*, ..., *Z*, оказывается более специфированным, чем значение идиом таксона *X*, но не наоборот» [2. С. 393]. Например, в словнике в таксоне УСПЕХ делается отсылка к таксонам ОТСУТСТВИЕ ПРОБЛЕМ и ПРЕВОСХОДСТВО, но не наоборот, так как идеи «удача в достижении чего-либо» и «общественное признание» содержатся во всех идиомах таксонов ОТСУТСТВИЕ ПРОБЛЕМ и ПРЕВОСХОДСТВО.

Заключение

Проект «Словаря-тезауруса идиом современного татарского языка» является важной вехой и шагом вперед в татарской лексикографической практике. Он представляет новый подход к исследованию фразеологии – тезаурусный принцип структурирования семантических полей идиом татарского языка. Собранный эмпирический материал позволил нам составить представление о фразеологической системе современного татарского языка, структуре семантических полей и внутренних связях между ними.

На данном этапе работы Словарь-тезаурус насчитывает около 6 000 ФЕ и более 22 000 контекстов их употребления. В результате классификации материала было выделено 86 семантических полей, самыми продуктивными в татарской фразеологии являются семантические поля ЭМОЦИИ, МЫШЛЕНИЕ, РЕЧЕВЫЕ АКТЫ, ПРОБЛЕМЫ – ОТСУТСТВИЕ ПРОБЛЕМ.

Словник Тезауруса является оригинальным: в его основе лежит База данных по современной идиоматике татарского литературного языка. Корпусный анализ идиом татарского языка позволил уточнить традиционные представления о семантике идиом, выявить их трансформации и модификации. Так, для ряда идиом татарского языка, которые представлены в словарях как моносемичные, выявлены дополнительные значения.

Новизна словаря заключается также в том, что в нём представлены разные типы семантических отношений: антонимические (отражены в названиях таксонов), а также родовидовые (отражены в системе парадигматических отсылок).

Список источников

1. Баранов А.Н., Добровольский Д.О. Семантические отношения во фразеологии // Компьютерная лингвистика и интеллектуальные технологии : материалы ежегодной Международной конференции «Диалог-2011». № 10 (17). М. : РГТУ, 2011. С. 51–60.
2. Баранов А.Н., Добровольский Д.О. Аспекты теории фразеологии. М. : Знак, 2008. 656 с.
3. Караулов Ю.Н. Лингвистическое конструирование и тезаурус литературного языка. М. : Наука, 1981. 366 с.
4. Roget P.M. Roget's Thesaurus of English Words and Phrases. London : Penguin, 1998. 1440 p.
5. Boissière P. Dictionnaire analogique de la langue française. Paris : Edité par Larousse et Boyer, 1862. 1456 p.

6. Dornseiff, F. Der deutsche Wortschatz nach Sachgruppen. 5. Aufl. Berlin : de Gruyter, 1959. 964 p.
7. Casares J. Diccionario ideológico de la lengua española. Barcelona : Gredos, 1959. 1456 p.
8. Караулов Ю.Н. Минимальный идеографический словарь // Общая и русская идеография. М. : Наука, 1976. 354 с.
9. Баранов О.С. Идеографический словарь русского языка. М. : Изд-во ЭТС. 1995. 814 с.
10. А.Н. Тихонов. Русско-узбекский тематический словарь. Ташкент : Ўқитувчи, 1975. 519 с.
11. Русско-азербайджанский тематический словарь. М. : T&P Books Publishing, 2019. 256 с.
12. Русско-турецкий тематический словарь. М. : T&P Books Publishing, 2019. 258 с.
13. Тарабукина М.В. Русско-якутский тематический словарь. Якутск : Издательский дом СВФУ, 2012. 415 с.
14. Кунин А.В. Курс фразеологии современного английского языка : учеб. пособие для ин-тов и фак. иностр. яз. Дубна : Феникс+. 2005. 488 с.
15. Эмирова А.М. К теоретическим основам фразеологической идеографии // Вопросы фразеологии. 1979. № 14. С. 50–61.
16. Мокиенко В.М. О тематико-идеографической классификации фразеологизмов // Словари и лингвострановедение / под ред. Е.М. Верещагина. М. : Русский язык, 1982. С. 108–121.
17. Яранцев Р.И. Словарь-справочник по русской фразеологии. М. : Русский язык, 1985. 304 с.
18. Баранов А.Н., Добровольский Д.О. К проблеме построения тезауруса русских идиом // ИАН СЛЯ. 1992. Т. 51, № 5. С. 60–67.
19. Добровольский Д.О., Караулов Ю.Н. Идиоматика в тезаурусе языковой личности // Вопросы языкознания. 1993. № 2. С. 5–15.
20. Словарь-тезаурус современной русской идиоматики / под ред. А.Н. Баранова, Д.О. Добровольского. М. : Мир энциклопедий Аванта+. 2007. 1135 с.
21. Заляй Л.З. Фразеология, пословицы и поговорки татарского языка. Казань : Татар. кн. изд-во, 1957. 244 с.
22. Исанбет И. Татарские народные пословицы : в 3 т. Казань : Татар. кн. изд-во. Т. 1. 1959. 916 с.; Т. 2. 1963. 960 с.; Т. 3. 1967. 1014 с.
23. Байрамова Л.К. Учебный тематический русско-татарский фразеологический словарь. Казань : Татаркнигоиздат, 1991. 158 с.
24. Гизатова Г.К. Галайиева Л.Ш., Люстиг З.А. Идеографический фразеологический словарь (на материале русского, татарского и английского языков). Казань : Изд-во Казанского ун-та. 2010. 420 с.
25. Gizatova G. A nation without a language is a nation without heart // Language Endangerment. Disappearing metaphors and shifting conceptualizations. Amsterdam/Philadelphia : John Benjamins Publishing House, 2015. P. 176–199.
26. Кубрякова Е.С. Об установках когнитивной науки и актуальных проблемах когнитивной лингвистики // Вопросы когнитивной лингвистики. 2004. № 3. С. 6–17.

27. Croft W., Cruse D.A. Cognitive Linguistics. Cambridge : Cambridge University Press, 2004. 356 p.
28. Geeraerts D., Cuypkens H. (eds.). The Oxford Handbook of Cognitive Linguistics. Oxford : Oxford University Press, 2010. 1368 p.
29. Lakoff G. Ten Lectures on Cognitive Linguistics. Leiden; Boston: Brill, 2017. 226 p.
30. Wierzbicka A. Back to definitions: Cognition, semantics, and lexicography // Lexicographica. 1992. № 8. P. 10–25.
31. Татарско-русский словарь : в 2 т. Казань : Магариф, 2007.
32. Письменный корпус татарского языка (ПКТЯ). URL: <https://www.corpus.tatar>
33. Национальный корпус татарского языка «Туган тел» (НКТЯ). URL: <https://tugantel.tatar>
34. Sketch Engine Tatar Corpus. URL: <https://www.sketchengine.eu/tatar-corpus-from-the-web/>
35. Сафиуллина Ф.С. Татарско-русский фразеологический словарь. Казань : Магариф, 2001. 335 с.
36. Исанбет Н. Фразеологический словарь татарского языка : в 2 т. Казань : Татар. кн. изд-во. Т. 1. 1989. 495 с.; Т. 2. 1990. 365 с.
37. Dobrovolskij D. Idioms in a Semantic Network: Towards a New Dictionary-Type // Proceedings of the 6th EURALEX International Congress, Amsterdam. 1994. P. 263–270.

References

1. Baranov, A.N. & Dobrovolskij, D.O. (2011) [Semantic relations in phraseology]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii* [Computer Linguistics and Intellectual Technologies]. Proceedings of the International Conference. Dialog [Dialogue]. Vol. 10 (17). Moscow. 25–29 May 2011. Moscow: Russian State University for the Humanities. pp. 51–60. (In Russian).
2. Baranov, A.N. & Dobrovolskij, D.O. (2008) *Aspekty teorii frazeologii* [Aspects of the Theory of Phraseology]. Moscow: Znak.
3. Karaulov, Yu.N. (1981) *Lingvisticheskoe konstruirovanie i tezaurus literaturnogo yazyka* [Linguistic Construction and Thesaurus of the Literary Language]. Moscow: Nauka.
4. Roget, P.M. (1998) *Roget's Thesaurus of English Words and Phrases*. London: Penguin.
5. Boissière, P. (1862) *Dictionnaire analogique de la langue française*. Paris: Edité par Larousse et Boyer.
6. Dornseiff, F. (1959) *Der deutsche Wortschatz nach Sachgruppen*. 5th ed. Berlin: de Gruyter.
7. Casares, J. (1959) *Diccionario ideológico de la lengua española*. Barcelona: Gredos.
8. Karaulov, Yu.N. (1976) *Obshchaya i russkaya ideografiya* [General and Russian Ideography]. Moscow: Nauka.
9. Baranov, O.S. (1995) *Ideograficheskiy slovar' russkogo yazyka* [Ideographic Dictionary of the Russian Language]. Moscow: Izd-vo ETS.
10. Tikhonov, A.N. (1975) *Russko-uzbekskiy tematicheskii slovar'* [Russian-Uzbek Thematic Dictionary]. Tashkent: Ўқитувчи.

11. Taranov, A. (2019) *Russko-azerbaydzhanskii tematicheskii slovar'* [Russian-Azerbaijani Thematic Dictionary]. Moscow: T&P Books Publishing.
12. Taranov, A. (2019) *Russko-turetskii tematicheskii slovar'* [Russian-Turkish Thematic Dictionary]. Moscow: T&P Books Publishing.
13. Tarabukina, M.V. (2012) *Russko-yakutskii tematicheskii slovar'* [Russian-Yakut Thematic Dictionary]. Yakutsk: M. K. Ammosov North-Eastern Federal University.
14. Kunin, A.V. (2005) *Kurs frazeologii sovremennogo angliiskogo yazyka* [Course of Phraseology of the Modern English Language]. Dubna: Feniks+.
15. Emirova, A.M. (1979) K teoreticheskim osnovam frazeologicheskoy ideografii [On the theoretical foundations of phraseological ideography]. *Voprosy frazeologii*. 14. pp. 50–61.
16. Mokienko, V.M. (1982) O tematiko-ideograficheskoi klassifikatsii frazeologizmov [On the thematic-ideographic classification of phraseological units]. In: Vereshchagin, E.M. (ed.) *Slovari i lingvostranovedenie* [Dictionaries and Linguacultural Studies]. Moscow: Russkii yazyk. pp. 108–121.
17. Yarantsev, R.I. (1985) *Slovar'-spravochnik po russkoy frazeologii* [Dictionary and Reference Book on Russian Phraseology]. Moscow: Russkii yazyk.
18. Baranov, A.N. & Dobrovol'skiy, D.O. (1992) K probleme postroeniya tezaurusa russkikh idiom [On the problem of constructing a thesaurus of Russian idioms]. *IAN SLYa*. 5 (51). pp. 60–67.
19. Dobrovol'skiy, D.O. & Karaulov, Yu.N. (1993) Idiomatika v tezauruse yazykovoy lichnosti [Idioms in the thesaurus of a linguistic personality]. *Voprosy yazykoznaniiya*. 2. pp. 5–15.
20. Baranov, A.N. & Dobrovol'skiy, D.O. (eds) (2007) *Slovar'-tezaurus sovremennoy russkoy idiomatiki* [Dictionary-Thesaurus of Modern Russian Idioms]. Moscow: Mir entsiklopediy Avanta+.
21. Zalyay, L.Z. (1957) *Frazeologiya, poslovitsy i pogovorki tatarskogo yazyka* [Phraseology, Proverbs and Sayings of the Tatar Language]. Kazan: Tatarskoe knizhnoe izdatel'stvo.
22. Isanbet, I. (1959–1967) *Tatarskie narodnye poslovitsy* [Tatar Folk Proverbs]. Vols 1–3. Kazan: Tatarskoe knizhnoe izdatel'stvo.
23. Bayramova, L.K. (1991) *Uchebnyy tematicheskii russko-tatarskii frazeologicheskii slovar'* [Educational Thematic Russian-Tatar Phraseological Dictionary]. Kazan: Tatknigoizdat.
24. Gizatova, G.K. Galyavieva, L.Sh. & Lyustig, Z.A. (2010) *Ideograficheskii frazeologicheskii slovar' (na materiale russkogo, tatarskogo i angliiskogo yazykov)* [Ideographic Phraseological Dictionary (Based on the Russian, Tatar and English languages)]. Kazan: Kazan State University.
25. Gizatova, G. (2015) A nation without a language is a nation without heart. In: *Language Endangerment. Disappearing metaphors and shifting conceptualizations*. Amsterdam/Philadelphia: John Benjamins Publishing House. pp. 176–199.
26. Kubryakova, E.S. (2004) Ob ustanovkakh kognitivnoy nauki i aktual'nykh problemakh kognitivnoy lingvistiki [On the attitudes of cognitive science and current problems of cognitive linguistics]. *Voprosy kognitivnoy lingvistiki*. 3. pp. 6–17.
27. Croft, W. & Cruse, D.A. (2004) *Cognitive Linguistics*. Cambridge: Cambridge University Press.

28. Geeraerts, D. & Cuyckens, H. (eds) (2010) *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.
29. Lakoff, G. (2017) *Ten Lectures on Cognitive Linguistics*. Leiden; Boston: Brill.
30. Wierzbicka, A. (1992) Back to definitions: Cognition, semantics, and lexicography. *Lexicographica*. 8. pp. 10–25.
31. Asylgaev, Sh.N. et al. (eds) (2007) *Tatarsko-russkiy slovar'* [Tatar-Russian Dictionary]. Kazan: Magarif.
32. *Pis'mennyy korpus tatarskogo yazyka (PKTYa)* [Written Corpus of the Tatar Language (PKTYa)]. (n.d.) [Online] Available from: <https://www.corpus.tatar>.
33. *Natsional'nyy korpus tatarskogo yazyka "Tugan tel" (NKTYa)* [Tugan Tel. National Corpus of the Tatar language (NKTYa)]. (n.d.) [Online] Available from: <https://tugantel.tatar>
34. *Sketch Engine Tatar Corpus*. (n.d.) [Online] Available from: <https://www.sketchengine.eu/tatar-corpus-from-the-web/>
35. Safiullina, F.S. (2001) *Tatarsko-russkiy frazeologicheskiy slovar'* [Tatar-Russian Phraseological Dictionary]. Kazan: Magarif.
36. Isanbet, N. (1989–1990) *Frazeologicheskiy slovar' tatarskogo yazyka* [Phraseological Dictionary of the Tatar Language]. Vols 1–2. Kazan: Tatarskoe knizhnoe izdatel'stvo.
37. Dobrovol'skiy, D. (1994) Idioms in a Semantic Network: Towards a New Dictionary-Type. *Euralex 1994*. Proceedings of the 6th EURALEX International Congress. Amsterdam. 30 August – 3 September 1994. Amsterdam: [s.n.]. pp. 263–270.

Сведения об авторе:

Гизатова Гузель Казбековна – д-р филол. наук, профессор кафедры европейских языков и культур Казанского федерального университета (Казань, Россия). E-mail: guzelgizat52@gmail.com

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Guzel K. Gizatova, Dr. Sci (Philology), professor, Kazan Federal University (Kazan, Russian Federation). E-mail: guzelgizat52@gmail.com

The author declares no conflicts of interests.

*Статья поступила в редакцию 01.10.2024;
одобрена после рецензирования 24.11.2024; принята к публикации 10.12.2024*

*The article was submitted 01.10.2024;
approved after reviewing 24.11.2024; accepted for publication 10.12.2024*