

**Қазақстан Республикасы конституциялық және қылмыстық заңның мәтіндегі
мүмкін емес құқықты қорғау**

С.М.Иманбаев

«Астана» Халықаралық университеті (Астана, Қазақстан)

Аңдатпа

Мақала құқық қорғау органдарының қызметкерлері мен өзге де адамдар тарапынан Азаптауға жол бермеу құқығын қорғауды бұзу туралы өзекті мәселені, лауазымды адамның өзі жасады деп күдіктенген адамдардан тану үшін адамгершілік немесе физикалық азап пен ауыртпалықты қасақана келтіру мақсатымен зерттеуге арналған. Азаматтарға қатысты азаптауды қолдану теориясы мен практикасының проблемалық мәселелері қарастырылады. Қолданыстағы ұлттық заңнаманы, доктриналық тәсілдер мен сот практикасын талдауды ескере отырып, азаптау актілерінің алдын алу бойынша тиімді шараларды әзірлеу және қабылдау қажеттілігі негізделеді. Қадір-қасиет құқығын қорғауды және Азаптауға жол бермеу құқығын қорғау мәселесін шешу үшін заңнаманы жетілдіру жолдары ұсынылды. Бұл мақалада азаптаудың құқықтық құбылысы, оның ішінде анықтамалар, тарихи түсінік, негізгі жағдайлар және туындайтын мәселелер қарастырылады.

Түйін сөздер: қылмыстық құқық, гуманизм, абырой, азаптау, зорлық-зомбылық, физикалық зорлық-зомбылық, азап.

**Protection of the right to inadmissibility of torture in the context of Constitutional
and Criminal law of the Republic of Kazakhstan**

S.M.Imanbayev

Astana International University (Astana, Kazakhstan)

Abstract

The article is devoted to the study of the topical issue of violation of the protection of the right to inadmissibility of torture by law enforcement officers and other persons, with the purpose of intentionally causing moral or physical suffering and pain by an official in order to obtain a confession from persons of whom he is suspected. The problematic issues of the theory and practice of the use of torture against citizens are considered. Taking into account the analysis of current national legislation, doctrinal approaches and judicial practice, the necessity of developing and taking effective measures to prevent acts of torture is justified. The ways of improving legislation to solve the problem of violation of the protection of the right of dignity and the right of protection to the inadmissibility of torture are proposed. This article examines the legal phenomenon of torture, including definitions, historical understanding, key cases and emerging issues.

Keywords: criminal law, humanism, dignity, torture, violence, torment, physical violence.

Поступила в редакцию: 08.02.2024

Одобрена: 21.03.2024

Первая публикация на сайте: 10.05.2025

MPHTI: 16.21.21

ВОЗМОЖНОСТИ ТЕХНОЛОГИИ «BIG DATA» В ЭТИМОЛОГИИ

**¹Б.В.КРАМАРЕНКО^{ID}, ²Ж.А.БАЯНБАЕВА^{ID}*

¹Кокишетауский университет имени Ш.Уалиханова (Кокишетау, Казахстан)

²Казахский национальный университет имени аль-Фараби (Алматы, Казахстан)

**kramareko55bv@gmail.com, bayanbayeva.zhadyra@kaznu.kz*

Аннотация

Технология исследования «big data» (с англ. «большие данные») доказала свою эффективность в разных сферах жизнедеятельности: управлении, медицине, логистике и т.д. Однако в лингвистике она используется фрагментарно. О применении этой технологии в этимологии и компаративистике

(правда, не в автоматизированном виде) нами опубликовано несколько статей. На основе полученных результатов разработан конструктивный подход к этимологизации слов, отличающийся от традиционного. Его суть: в большинстве языков есть корпус слов, относящихся к ядерной лексике, которые кроме традиционных элементов: корня, аффиксов и флексий – содержат иные структурные компоненты: конструкты, детерминативы, форманты отрицания, совокупность которых можно назвать праморфологией. Цель данной статьи заключается в обосновании и демонстрации возможностей использования технологии «big data» в этимологии и компаративистике, а также в разработке теоретических предпосылок для создания этимологического словаря нового типа, который на основе конструктивного подхода будет определять этимологическое значение определенного слова не одного языка, а сразу нескольких. В статье продемонстрировано, как для повышения эффективности этимологического поиска можно использовать ряд техник и методов анализа больших данных, уже сложившихся в ходе исследований в других научных сферах: «Data Mining», «моделирование», «статистика» и др. Данное исследование показывает возможности использования разных методов анализа больших данных в лингвистике.

Ключевые слова: детерминатив, конструкт, конструктивный подход, праморфология, технология «big data», формант отрицания.

Введение

Технология «big data» (Tbd) в условиях углубляющейся цифровизации является одной из самых востребованных. Ее эффект обусловлен тем, что локальное рассмотрение чего-либо не дает исследователям четкого представления о природе происходящего, а применение технологии больших данных позволяет видеть более полную картину и динамику исследуемого процесса. Так, тысячи рентгеновских снимков позволили медикам визуально увидеть процесс развития определенных заболеваний от заражения до летального исхода.

Этимология и компаративистика тоже относятся к явлениям, состоящим из тысяч деталей, которые рассредоточены во времени и пространстве, поэтому применение Tbd является для этих направлений лингвистики вполне оправданным. Однако до недавнего времени у лингвистов-исследователей не было даже возможности посмотреть, как то или иное слово выглядит на письме в большом количестве языков. Сегодня с помощью сервисов онлайн-перевода «Яндекс-переводчик» и «Google-переводчик» в течение нескольких минут можно увидеть, как слово выглядит на 160 языках, а число вновь вводимых языков на этих сервисах постоянно растет. Но привычка использовать локальный фокус зрения вместо общего или полного взгляда на проблему еще дает о себе знать.

Основная часть

Наши исследования представляют собой первые систематические попытки использования новой ситуации в этимологии, связанную с реализацией возможностей, которые предоставляют компьютерная техника и большие данные, создаваемые с помощью широкодоступных сервисов: сопоставлять слова нескольких десятков языков, а не нескольких, как это делалось прежде [Крамаренко Б.В., 2021; Крамаренко Б.В., 2023].

Результаты этих исследований показывают, что в большинстве языков строение слов ядерного лексического фонда, помимо того, что представлено в современной морфологии, имеет скрытую, а точнее, забытую функционально-семантическую структуру, которую условно можно назвать праморфологией. Эта структура включает:

1. «Конструкт» (Kt) – общую часть одного слова в разных языках, основная функция которого заключается в обозначении объекта номинации. Выделяется девять эталонных конструктов, к которым восходят все остальные – девиантные.
2. Четыре «детерминатива» (D), которые наряду с конструктом определяют буквальное значение слова:
 - а) «D-активности или движения» имеют слова, которые обозначают активность предмета номинации или стали результатом чьей-то активности (*гром*, *гроза*; *берлога*, *гнездо*);

- б) «D-внутри» имеют слова, описывающие внутреннее пространство, а также что-то вбирающее в себя (*верх, ворота, портал, пещера; вулкан, озеро, ущелье*);
- в) «D-пространства» характеризует слова, описывающие протяженность или пространственный масштаб чего-либо (*даль, длина, степь, стоянка, простор*);
- г) «D-наличия или присутствия» подчеркивает наличие того, что обозначает семантический ключ слова, его конструкт (*море, мир*).

3. «Внутрисловный формант отрицания», как элемент слова, меняет его общее значение на противоположное (*бегать, белый, ноль*);

4. «Семантический комплемент» представляет две-три разные буквы, имеющие одинаковый функционально-семантический смысл. Они все читаются, но при расшифровке слова имеют смысл как один элемент слова (*скорость, скакать, скрипеть*).

Все прамофологические элементы слова, состоят из разных букв (от 1 до 4) и имеют свое семантическое значение, разную дистрибуцию (размещение), но их материальное воплощение таково, что порою трудно понять, что это: детерминатив, формант отрицания или остаток девиантного конструкта, поэтому важно определить, какую функцию выполняет конкретная буква (или их малая совокупность), чтобы определить, какую семантическую роль этот элемент играет в общем значении слова.

Литературный обзор

К сожалению, статей, посвященных использованию технологии больших данных в этимологии, практически нет. Массу статей, посвященных компьютерной лингвистике, ежегодным конференциям специалистов этого направления, нельзя рассматривать как материал, раскрывающий проблемы этимологического поиска. Компьютерная лингвистика направлена на математическое моделирование интеллектуальных процессов с помощью компьютеров. Одна из ее целей – использование математических моделей для описания естественных языков [Смагин М.Р., Филатова Е.Ю., 2017].

Относительно близки нашему направлению идеи Ноама Хомского о существовании у людей врожденной схемы обработки информации и формирования абстрактных структур языка, представляющих собой лингвистические универсалии, общие для всех языков. Однако считаем, что единство языков или их моногенез проявляется не только в существовании универсальной грамматики, которую разработал Н.Хомский [Корниенко М.А., 2018]. Праморфологическая структура слов, о которой пойдет речь ниже, полагали мы, сложилась в дограмматический период, но она тоже является лингвистической универсалией. Об этом говорят наши статьи.

Методы

К сожалению, накопление больших данных в компаративистике только начинается. Указанные выше сервисы онлайн-перевода представляют собой совокупность более полутора сотен двуязычных автоматизированных словарей. Чтобы увидеть, как выглядит то или иное слово сразу на 50 или 100 языках, не говоря уже о фразеологизмах, необходимо поработать с этими двуязычными автоматизированными словарями в ручном режиме. Пока только так можно составить большие данные по каждому из исследуемых слов.

Использование Tbd в разных областях деятельности привело к тому, что появился целый набор техник и методов анализа больших данных, которые с успехом переносятся в другие сферы при решении задач, совершенно отличных от тех, которым они обязаны своим появлением. Выделяют девять таких методов и техник [Панасюк Д., 2017]. Считаем, что часть из них можно использовать в лингвистике.

Результаты и дискуссия

Сравнительно-исторический метод давно стал основным в этимологии, но он, на наш взгляд,

имеет недостаток: при поиске значения слова используется лишь несколько родственных языков или языки из той же лингвистической семьи. Не всегда оправдан алгоритм поиска на основе звукового соответствия и нахождения фонетической близости искомого слова со словами других языков. При этом часто не обращалось внимание на то, что фонетически близкие слова имели значения, довольно далекие от искомого.

Слово *gorod*, к примеру, в этимологическом словаре описывается так: «первоисточником является корень *ghard* в индоевропейском праязыке, что и объясняет наличие родственных слов в других языках: немецкое *Garten* (*cad*), англ. *garden* (*cad*), албан. *garth* (*ograda*), готск. *grhas* (дом)» [Семенов В.А., 2005]. И в этой словарной статье есть традиционная ссылка на индоевропейский праязык, а далее следует перечисление однокоренных слов, из которых читатель и должен сделать вывод, что *gorod* – это *огороженное место*, как *cad* (англ. *garden*), русск. *огород* и т.д. Мы сочли такую практику недостаточно обоснованной.

Наличие однокоренных слов с модельным словом праязыка еще не гарантия правильной этимологизации. Кроме того, даже при наличии таких слов можно сделать неправильный вывод. При этимологизации *gorod* обычно упоминают и слово *ogorod*, но не соотносят его со словоформами этого же слова из других языков. Приставка *o* в данном случае может иметь значение 'около'. И *ogorod* означает не столько – *огороженный участок*, сколько участок, находящийся *около города*. В противном случае городская стена (как *ograda* и семантическая основа этимологии этого слова) простиралась бы в древних городах на сотни километров, а они, как известно из истории, были достаточно компактны. Кроме того, как совместить *огороженное место* с однокоренными словами других языков, имеющими такие понятия: «дом» или «лит. *gardinỹs* «хлев для мелкого скота»; жем. *gardis*, вин. *gařdi* «решетчатый борт воза», др.-инд. *grhás* «дом» [Фасмер М., 1964]. Но из всей совокупности родственных и фонетических близки слов за основу этимологического вывода взят все же такой признак города как его «огороженность». Хотя очевидно, что речь идет о рукотворности и созданности и *города*, и *ограды*, и *хлева для скота*, и *решетчатого борта воза*.

Нас не устраивал такой подход, поэтому мы стали искать другие. Еще в 80-х годах прошлого века попытку нахождения детерминативов в современном армянском языке предприняла Г.В.Сукиасян. Детерминатив (*D*) – это несамостоятельный и нечитаемый определитель слова, существовавший в логографических языках древности. В условиях высокой омонимичности древних слов детерминатив указывал к какой категории относится слово в том или ином контексте. Автором выделены 5 групп читаемых *DD*, имеющих от одного до трех вариантов, включающих от одной до трех букв. Она считала их асемантичным элементом, нужным только для создания словарного разнообразия [Сукиасян Г.В., 1984].

В позиции этого исследователя мы сочли ошибочной идею асемантичности *DD*. Так же, как и позицию автора статьи о детерминативе в российской энциклопедии, который полагал, что в настоящее время поиск детерминативов в современных языках является с научной точки зрения несостоятельным [БРЭ, 2004].

Дальнейшее знакомство с используемыми техниками и методиками анализа больших данных позволит доказать нашу точку зрения

Методика «*Data Mining*» буквально – «добыча данных» или их интеллектуальный, глубокий анализ, в результате чего появляются такие данные, которые являются, действительно, новыми, нетривиальными, порою трудно объяснимыми. Чем больше таких данных добывается – тем результативнее процесс. Сегодня для того, чтобы использовать эти большие данные, их нужно сначала создать. В нашем случае это были таблицы в, которые мы собрали переводы каждого исследуемого слова на большинстве языков, имеющихся в сервисах. Это делалось для углубленного их анализа в последующем.

Среди принципов этимологического поиска есть – морфологический [Откупщиков Ю.В., 2001]. Исследуя как выглядит слово-корень *gorod*, в разных языках мы обнаружили порадоксальное

явление: в словах была выявлена структурная единица меньше корня, который, как гласит современная морфология, считается неделимым. Но для обнаружения этой единицы мы в соответствие с принципом релятивизма упростили ее параметры, представив ее в виде $v+R+v$ (где v -гласная буква), а учитывая такой глобальный процесс как ротаизм (в данном случае обратный) представили эту единицу как $v+R/L+v$. Поскольку к ней как к несущей конструкции были прикреплены и другие элементы слова, мы назвали ее «конструктом».

Обоснованность нашего утверждения, мы проверили еще одним методом анализа Tbd – «визуализацией», что позволило выделить те новые данные, какие удалось добыть с помощью «Data Mining». Их можно было показать с помощью инфографики (диаграмм, схем, рисунков и т.д.), специальных компьютерных программ, но не менее информативно они представлены в виде таблицы 1, в которой в отдельных столбцах 3А и 3Б разместили найденные нами новые компоненты слова, которые, полагаем, появились в период, предшествовавший морфологии. Визуализированные данные противоречили сложившейся научной парадигме, но отрицать их наличие с позиций релятивизма невозможно.

Примеры в таблице 1 приведены на основе транслитерации, существующей в указанных выше интернет-сервисах.

Таблица 1

Конструкты в слове город в некоторых языках

0	1 А	2А	3А	0	1Б	2Б	3Б
№	язык	слово	конструкт	№	язык	слово	конструкт
1	русс.	<i>g/oro/d</i>	oro	7	казах.	<i>k/ala;</i>	ala
2	венгр.	<i>v/áro/s;</i>	áro	8	греч	<i>p/óli;</i>	óli
3	баск.	<i>h/iri/a</i>	iri	9	грузин.	<i>k'/ala/k'i</i>	ala
4	бенгал.	<i>śah/ara;</i>	ara	10	зулу	<i>id/olo/bha</i>	olo
5	румын	<i>ora/ş</i>	ora	11	франц.	<i>v/ille</i>	ille
6	сингал.	<i>nag/ara/ya</i>	ara	12	шотланд	<i>bha/ile</i>	ile

Столбцы таблицы 1 с конструктами ясно показывают, что условно одинаковые комплексы букв, ставших элементом одного слова, но в разных языках, действительно, существуют. В таблице приведены примеры лишь из некоторых языков, а их с учетом девиаций слова в разных языках значительно больше. Так слово *g/oro/d* (*город*) в ряде славянских языков имеет форму *g/ra/d* (*град*) или *g/ar/d* (*град*), а индийских языков существует более десятка, которые имеют структуру и фонемный ряд близкий к бенгальскому языку.

Исследования показали, что семантически и функционально тождественны друг другу и иные конструкты со звукокомплексами *айа, ала, ара, ана, ама, лон(о), лом(о), рон(о), ром(о)*.

Еще одной из техник Tbd является «пространственный анализ». Он используется в основном в геоинформационных системах. Но в лингвистике такой анализ важен при исследовании структуры слова. Материалы, приведенные в таблице 1, дают информацию о размещении, то есть о дистрибуции конструкта в составе слова. Как видно, он может размещаться в начале, середине или конце слова, как соответственно: в румын. *ora/ş*, в русс. *g/oro/d*, казах. *k/ala*. Налицо свободная дистрибуция.

Не менее важен пространственный анализ ареалов проживания носителей языков, приведенных в вышеуказанной таблице. Они проживают в Европе, Азии, Южной Африке, представляют разные языковые группы и семьи, а это означает, что фонетическая общность конструктов не может быть результатом заимствования слова из языка в язык и лишний раз подтверждает справедливость теории моногенеза языков.

В некоторых языках слово состоит из одного конструкта. Так слово *город* на языке йоруба – *ilu*; на марийском – *ola*; на вымершем шумерском – *iri* [Гаршин И., 2022].

Иногда конструкт может сокращаться до одной буквы или быть нулевым. Сравните: русс. *спать* (s/p/at') и англ. 'to s/lee/p', нем. 'sch/lee/f/en', нидер. 's/la/p/en'. В языках германской группы четко выделяются конструкты **lee**, **la**, восходящее к **ala**. В русском слове его либо нет, либо его функции выполняет буква **a**.

«Статистический анализ» был одним из первых методов, который использовали в языкознании, опираясь на большие данные. К ним можно отнести работу систематиков, которые выявляют сколько в мире языковых семей.

Метод лексикостатистики разработал С.А.Старостин [Коровина Е.Н., 2019]. Он, для его использования, разработал и создал вместе с другими участниками проекта «Вавилонская башня» большие базы данных на разных языках, в которых по словам из списка Сводеша устанавливалось родство языков из разных языковых семей [Вавилонская башня, 1998].

К статистике прибегал и Ноам Хомский, разрабатывая свою универсальную грамматику [Корниенко М.Ф., 2018].

Для целей нашего исследования статистический анализ был модифицирован, главным в нем было выявление различных элементов праморфологической структуры слова: конструкта, детерминативов и т.п. Важно было выявить какие функции выполняет каждая буква (или их небольшая совокупность) в слове и с учетом этого выявить ее семантическое значение. Поэтому эта исследовательская техника получила название «метода функционально-семантической статистики».

Статистка конструктов по слову *город* свидетельствует, что к форме LON(O)//LOM(O) или RON(O)//ROM(O) восходят 8 яз.; к форме АYA – 5 яз.; к ALA – 14 яз.; к ARA – 30 яз.; к ANA – 12 яз.; к AMA – 10 языков. В целом 101 яз. из 109 демонстрировали наличие указанных конструктов.

Подчеркнем, что статистика, проводимая нами при этимологизации, носила поисковый характер. Важно было доказать наличие определенных элементов в слове, поэтому она не всегда соответствовала требованиям, предъявляемым к проведению статистических исследований: формированию выборки, учету родства языков, использованию сложного математического аппарата и т.д.

В слове *г/оро/д* кроме конструкта есть другие элементы слова, которые не соответствуют современной морфологии. Ведь приставки *г*, как и суффикса *д* в русском языке нет.

Чтобы определить, чем является буква **Г**, нами был проанализирован корпус слов, обозначающих соматизмы (части тела) в русском языке. Выяснилось, что там, где есть эта буква, и чередующиеся с ней (К, Х, З, С, Ж, Ш, Щ, Ч, Ц), соматизмы очень подвижны: *нога*, *глаза*, *колени*, *шея*, *хвост*, *плечо*, *сердце*, *живот*, *волос*.

В слове *глаз* в 58 языках из 98 были эти буквы. На этом основании, а также учитывая идею Сукиасян Г.В. о *D* нами был выделен детерминатив «активности или движения».

Статистика по слову *дорога* показала, что буква **Д**, а также чередующаяся с ней **Т** встречаются в 54 языках из 109. Это позволило выделить «детерминатив пространства», так как функционально-семантическая статистика этого слова показывает устойчивую тенденцию использования этих букв в данной роли. Но, с другой стороны, почти в каждом втором языке для номинации пространственных параметров предмета речи в слове либо не используется этот детерминатив, либо он в процессе использования стерся или был заменен на другой звук, (а затем и букву) более удобный при произношении.

Методика «Смешение и интеграция данных» дает самые неожиданные результаты, так как смешиваться могут самые различные компоненты. Так, этимологизация слова *белый* показала, что в ряде языков оно содержало тот же конструкт *lan*, что и в слове *lan/d*

некоторых языков, имеющее значение *земля, земляной (черный)*. Например, в каталан. '*b/lan/c*', испан. '*b/lan/co*', итал. '*bi/an/ca*', а после «ротацизма» в португал. '*branco*'.

Если учесть, что буква **С** – это детерминатив активности, то выходило, что с помощью элементов *b/b+гласная* были созданы антонимы слову *земляной = черный*. И слово *белый* буквально имеет смысл – «не черный». Это позволило выделить *внутрисловный формат отрицания* как элемент праморфологической структуры слова. Подтверждает правильность этимологизации этого колоронима слово *al/ba* из латинского языка. В нем конструкт *al<ala* подвергся девиации, а слово метатезе. Но это не изменило смысла слова: вместо «не черный» стало «черного нет».

Учитывая, что внутрисловный формант отрицания – это тоже универсалия, решено было посмотреть: какие лингвистические средства для его выражения используются. Статистика по слову *белый* дала следующие результаты: в 91 языке из 109 форманты отрицания присутствовали в виде букв: *b, v, w, f, p, d, t, n, m, u, o*.

Однако этимологизация слова *черный* на многих языках показала, что в них тоже есть буквы (*b, v, w, f, p, u, o*), которые вроде бы должны выполнять функции форманта отрицания. Но в этом слове их не могло его быть. Это слово ассоциация. Семантика слов подсказывала, что *черный* на этих языках тоже означает *земляной*, но под этим, очевидно, подразумевалась земля, находящаяся внутри грунта, а не на ее поверхности. На этом основании нами был выделен еще один детерминатив – «*внутри*», раскрывающий семантику слов с внутренней структурой (*полка, шкаф, кабина*), нечто вбирающее что-то в себя (*вода, воздух*), находящегося внутри чего-то (*винт, болт*).

Еще один детерминатив *наличия или присутствия* обозначался буквой *М/М+гласный*. Он содержался в 10 языках в словоформах *слово*, где семантически не могло быть формантов отрицания.

Все элементы структуры слова в праморфологии являются универсалиями, то есть присущи большинству языков. Однако анализ слов, обладающих в русском языке одинаковыми конструктами (*голова, золото, молодой, колода*), показал, что они не могут иметь одинаковую семантику. Данные конструкты, имея разное значение, являются семантическими ключами слов для понимания их смысла. А чтобы выяснить значение каждого, необходимо в этимологизируемом слове подобрать такой семантический ключ, который будет подходить для других структурных компонентов этого же слова в большинстве языков. Поэтому вполне закономерно, что мы использовали ещё один метод анализа в Tbd – «*имитационное моделирование*».

Наш опыт подобного моделирования искомых значений представлен в таблице 2. В нем конструкты имеют семантический ключ *вода*, хотя в рассматриваемом слове предмет номинации – *рыба*.

Таблица 2

Семантический ключ *вода* в слове *рыба* разных языков

Язык	Семанти- ческая структура слова	Конст- рукт: <i>вода</i>	Д- актив- ности	Д- прост- ранства	Д- внутри	Общее значение
рус.	ry+b{a}	ry			ba	воды внутри
кхмер [្រ័]	t+re+i	re	i[й]	t		пространством воды движется
греч. ψάρι	p+s+ári	ári	s		p	внутри воды движется
англ.	f+i+sh	i	h		f	внутри воды движется
каталан.	p+ei+x	ei	x		p	внутри воды движется
аймара	ch+all+wa	all	ch		wa	движется воды внутри
казах.	б+алы+қ	алы	қ		б	внутри воды движется

фин.	к+ала	ала	к			движется водой
тайск. ปลา	P+lā	lā			р	внутри воды
ирлан.	ia+sc	ia	sc			водой движется
непал. माछा	mā+chā	mā	chā			водой движется
фриз.	f+i+sk	I	sk		f	в воде движется

Как видно, в таблице 2 значение гиперонима 'рыба' на разных языках примерно одинаковое, хотя разность мотивом налицо. В русском и тайском важным сочли нахождение денотата *внутри воды*, в германской, романской и тюркской группах подчеркнули, что объект *внутри воды движется*, в других обратили внимание, что объект *водой движется*. Тем самым имитационное моделирование достигло своей цели - семантический ключ подобран верно.

Полагаем, что наиболее эффективен конструктивный подход при выяснении этимологии именно гиперонимов, так как на появление гипонимов большее влияние оказывают другие словообразовательные модели: заимствование, морфологический способ (суффиксальный чаще всего), указания на внешние признаки предмета и их метафорические и метонимические переносы [Карасева Т.В., Цинь Чжэнь, 2009].

Значение слова *рыба* в традиционной этимологии считается неясной. С одной стороны, обращается внимание на фонетически близкие слова других языков: латинское – *rubeta* (*жаба*) [Семенов А.В., 2005]; древнемецкое *ruppe* – «личинка» или *rupa* – «головастик» [Успенский Л.В., 2024]. С другой стороны, сближают со славянскими словами *рябой*, *рыхлый* (с первоначальным значением «быстрый»). С позиций конструктивного подхода указанные фонетически близкие слова, на которые ссылаются этимологи, имеют другие семантически ключи. Например, в латин. слове *rubeta* – «жаба» имеется два детерминатива: «внутри» – *b* и «пространства» – *t*, а конструкт *ru* имеет семантический ключ – «болото», (он, кстати, сохранился в диалекте русского языка в слове *жабьё* – «болото, топь» [Шанский Н.М., 2024].

В казахском языке *бақа* (*baqa*) имеет тот же праморфологический состав слова, что и в русском языке: «*D-внутри*» и «*D-активности*», но только после метатезы. Однако, в турецком *qırbaqa* – «лягушка, жаба» содержится конструкт, который «стерся» в русском и казахском языках. Слово имеет буквальное значение «идет болота внутри идет (находится)».

Все это подтверждает нашу гипотезу расшифровки слова *рыба* не только в русском языке.

Неравномерность забвения праморфологии в разных языках неодинаковая. Такая ситуация позволяет восстановить утраченные элементы в структуре слов, там, где они исчезли. Это делается с помощью аналогии с теми родственными языками, где они сохранились. Например, сравнивая бел. *сонца* (*сонца*) и укр. *sontse* (*сонце*) и русс. *solntse* (*солнце*), можно ясно видеть, что произошла утрата *Л* в конструкте словоформ белорусского и украинского языков, восходящих к конструкту *лон(о)*.

Заключение

Данное исследование еще раз подтвердило широкие возможности использования технологии Tbd в этимологии и компаративистике. Оно показало, что сложившиеся практики анализа больших данных в разных сферах деятельности, эффективно работают и в лингвистике, с учетом, естественно, специфики науки и особенностей исследуемого материала. В частности, были использованы: методики «*Data Mining*»; «*Пространственный анализ*»; «*Статистический анализ*», адаптированный под задачи исследования и ставший «*Методом анализа функционально-семантического статистическим*»; методика «*Смешение и интеграция данных*»; «*Имитационное моделирование*».

В статье были использованы материалы предыдущих наших исследований, в которых в меньшей степени обращалось внимание на описание методико-инструментальных аспектов исследовательской деятельности, но все вышеперечисленные техники и методы анализа больших данных в них, конечно, присутствовали, благодаря чему и были выявлены и

праморфологическая структура слова и ее отдельные элементы: конструкторы, детерминативы, форманты отрицания, семантические компоненты и сложился конструктивный подход к этимологизации слов, дополняющий сравнительно-исторический метод в лингвистике. В данной статье уделено больше внимания технике и алгоритмам анализа больших данных.

В таблице 2 отсутствуют два элемента праморфологии: *внутрисловный формант отрицания* и *детерминатив наличия*, но сама структура таблицы может служить основой для создания этимологического словаря нового типа. В нем, как видно, из приведенного опыта этимологизации слова «рыба», наряду с указанием праморфологических элементов слова в словоформах разных языков, указывается семантический ключ слова и дается общее значение каждой национально-лингвистической словоформе. Такой словарь, считаем, будет зримым доказательством моногенеза языков, а, самое главное, он позволит увидеть сущность национально-лингвистического видения мира у носителей каждого языка, разность и сходство мотивом каждой словоформы, сможет показать их внутреннее единство, что будет способствовать совершенствованию методик изучения языков и развитию полилингвизма.

Отметим, что пределы или границы использования Tbd в этимологизации слов на основе конструктивного подхода уже очерчены. Не все слова можно этимологизировать с его помощью, а только те, которые входят в ядерный лексический фонд национальных языков и, как правило, являются в нем гиперонимами.

Дальнейшие исследования о том, как возникла праморфология, каковы хотя бы приблизительные хронологические рамки ее появления и последующего забвения, можно будет делать только на основе появления значительного корпуса слов, этимологизированных на основе конструктивного подхода,

Будущие исследования могут касаться изучения преемственности праморфологии и морфологии в различных языках. Праморфология не могла не оставить каких-то следов в компонентах современных словоформ. В русском языке, например, детерминатив «Внутри» в некоторых ситуациях превратился в приставку «В» или одноименный предлог. В казахском и некоторых других тюркских языках детерминатив пространства в виде букв «Д-Т» сохранился в окончании направительного падежа *да-де, та-те*.

Элементы праморфологии очень напоминают относительно простой лингвистический конструктор, который можно будет использовать для создания новых искусственных или художественных языков, отдалено напоминающие те, что создал для получения читателями эстетического удовольствия Дж. Р.Р.Толкин – автор «Властелина колец» [Соснин Е., 2011]. Лингвопроектирование или лингвомоделирование как разновидность интеллектуального досуга можно будет использовать и для прагматических целей: привития интереса к изучению лингвистики, этимологии, компаративистики, древних языков, а также усвоения основ конструктивного подхода.

Лингвистический конструктор может быть использован и для реализации, так называемого, «нэйминга» – разработки названий кампаний, брендов, наименования новых товаров и услуг.

Список использованных источников

- Крамаренко Б.В. (2021) *Этимология в эпоху цифровизации: конструктивный подход*. Кокшетау: РИО Кокшетауский университет имени Ш.Уалиханова, 172.
- Крамаренко Б.В. (2023) «Big data» и структура слова. *Научные горизонты*, 1(65), 15-29 [Электронный ресурс]: URL: https://www.sciencehorizon.ru/wp-content/uploads/2017/08/Nauchnyie_gorizontyi_1_65_2023.pdf (дата обращения: 10.01.2024).
- Смагин М.Р., Филатова Е.Ю. (2017) Компьютерная лингвистика: теория и практика. *Успехи в химии и химической технологии*, XXXI, 14, 77-79.

- Корниенко М.А. (2018) Лингвофилософия Ноама Хомского: от картезианской традиции к генеративной грамматике. *Вестник Томского государственного университета Философия. Социология. Политология*, 43, 88-100 [Электронный ресурс]: URL: http://journals.tsu.ru/uploads/import/1723/files/43_088.pdf (дата обращения: 10.01.2023).
- Панасюк Д. (2017) *Что такое Big Data?* [Электронный ресурс]: URL: <https://medium.com/adtech-pioneers/что-такое-big-data-a0e503dc7fce> (дата обращения: 10.01.2024).
- Фасмер М. (1964) *Город. Этимологический словарь русского языка*. М.: Прогресс, 1964-1973.
- Сукиасян Г.В. (1984) Детерминативы в армянском языке. *Автореферат кандидатской диссертации*. Ереван, 19.
- БРЭ (2004) *Большая российская энциклопедия* [Электронный ресурс]: URL: <https://old.bigenc.ru/linguistics/text/1949830> (дата обращения: 12.01.2024).
- Откупщиков Ю.В. (2001) *Очерки по этимологии*. СПб.: Санкт-Петербургский университет, 480.
- Гаршин И. (2022) Шумерская лексика (словарь эме-гир): *Сайт Игоря Гаршина* [Электронный ресурс]: URL: <https://www.lexicons.ru/old/sh/sumerian/index.html> (дата обращения: 10.01.2024).
- Коровина Е.В. (2019) Ранжирование базой лексики С.А.Старостина: материалы к улучшению. *Индоевропейское языкознание и классическая филология*, 23-1 [Электронный ресурс]: URL: <https://cyberleninka.ru/article/n/ranzhirovanie-bazisnoy-leksiki-s-a-starostina-materialy-k-uluchsheniyu> (дата обращения: 10.01.2024).
- Вавилонская башня (1998). Проект: Этимологическая база данных: *Сайт* [Электронный ресурс]: URL: <https://starlingdb.org/indexru.htm> (дата обращения 10.01.2024).
- Карасева Т.В., Цинь Чжэнь (2009) Названия рыб в современном русском языке. *Актуальные вопросы современной филологии и журналистики*, 5, 9-24 [Электронный ресурс]: URL: <https://cyberleninka.ru/article/n/nazvaniya-ryb-v-sovremennom-russkom-yazyke> (дата обращения: 10.01.2024).
- Семенов А.В. (2005) *Город. Этимологический онлайн-словарь* [Электронный ресурс]: URL: <https://gufo.me/dict/semenov/gorod> (дата обращения: 10.01.2024).
- Успенский Л.В. (2024) Почему не иначе?: *Школьный этимологический онлайн словарь* [Электронный ресурс]: URL: <https://lexicography.online/etymology/uspensky/> (дата обращения 10.01.2024).
- Шанский Н.М. (2024) Происхождение слова «рыба». *Этимологический онлайн-словарь* [Электронный ресурс]: URL: <https://lexicography.online/etymology/p/рыба> (дата обращения: 10.01.2024).
- Соснин Е. (2011) Образы древнегерманской мифопоэтической модели мира и их реконструкция в произведениях Дж. Р.Р.Толкина. *Автореф. дисс. ... канд. филол. наук*, Барнаул, 23.

References

- Kramarenko B.V. (2021) *Etimologiya v epohu cifrovizacii: konstruktivnyj podhod*. Kokshetau: RIO Kokshetauskij universitet imeni Sh.Ualihanova, 172.
- Kramarenko B.V. (2023) «Big data» i struktura slova. *Nauchnye gorizonty*, 1(65), 15-29 [Elektronnyj resurs]: URL: https://www.sciencehorizon.ru/wp-content/uploads/2017/08/Nauchnye_gorizontyi_1_65_2023.pdf (data obrashcheniya: 10.01.2024).
- Smagin M.R., Filatova E. Yu. (2017) Komp'yuternaya lingvistika: teoriya i praktika. *Uspekhi v himii i himicheskoy tekhnologii*, XXXI, 14, 77-79.
- Kornienko M.A. (2018) Lingvofilosofiya Noama Homskogo: ot kartezijskoj tradicii k generativnoj grammatike. *Vestnik Tomskogo gosudarstvennogo universiteta Filosophiya. Sociologiya. Politologiya*, 43, 88-100 [Elektronnyj resurs]: URL: http://journals.tsu.ru/uploads/import/1723/files/43_088.pdf (data obrashcheniya: 10.01.2023).

- Panasyuk D. (2017) *Chto takoe Big Data?* [Elektronnyj resurs]: URL: <https://medium.com/adtech-pioneers/chto-takoe-big-data-a0e503dc7fce> (data obrashcheniya: 10.01.2024).
- Fasmer M. (1964) *Gorod. Etimologicheskij slovar' russkogo yazyka*. M.: Progress, 1964-1973.
- Sukiasyan G.V. (1984) *Determinativy v armyanskom yazyke. Avtoreferat kandidatskoj dissertacii*. Erevan, 19.
- BRE (2004) *Bol'shaya rossijskaya enciklopediya* [Elektronnyj resurs]: URL: <https://old.bigenc.ru/linguistics/text/1949830> (data obrashcheniya: 12.01.2024).
- Otkupshchikov Yu.V. (2001) *Ocherki po etimologii*. SPb.: Sankt-Peterburgskij universitet, 480.
- Garshin I. (2022) *Shumerskaya leksika (slovar' eme-gir): Sajt Igorya Garshina* [Elektronnyj resurs]: URL: <https://www.lexicons.ru/old/sh/sumerian/index.html> (data obrashcheniya: 10.01.2024).
- Korovina E.V. (2019) *Ranzhirovanie bazinoj leksiki S.A.Starostina: materialy k uluchsheniyu. Indoevropskoe yazykoznanie i klassicheskaya filologiya*, 23-1 [Elektronnyj resurs]: URL: <https://cyberleninka.ru/article/n/ranzhirovanie-bazisnoj-leksiki-s-a-starostina-materialy-k-uluchsheniyu> (data obrashcheniya: 10.01.2024).
- Vavilonskaya bashnya. Proekt: Etimologicheskaya baza dannyh: *Sajt* [Elektronnyj resurs]: URL: <https://starlingdb.org/indexru.htm> (data obrashcheniya 10.01.2024).
- Karaseva T.V., Cin' Chzhen' (2009) *Nazvaniya ryb v sovremennom russkom yazyke. Aktual'nye voprosy sovremennoj filologii i zhurnalistiki*, 5, 9-24 [Elektronnyj resurs]: URL: <https://cyberleninka.ru/article/n/nazvaniya-ryb-v-sovremennom-russkom-yazyke> (data obrashcheniya: 10.01.2024).
- Semenov A.V. (2005) *Gorod. Etimologicheskij onlajn-slovar'* [Elektronnyj resurs]: URL: <https://gufo.me/dict/semenov/gorod> (data obrashcheniya: 10.01.2024).
- Uspenskij L.V. *Pochemu ne inache?: Shkol'nyj etimologicheskij onlajn slovar'* [Elektronnyj resurs]: URL: <https://lexicography.online/etymology/uspensky/> (data obrashcheniya 10.01.2024).
- Shanskij N.M. *Proiskhozhdenie slova «ryba». Etimologicheskij onlajn-slovar'* [Elektronnyj resurs]: URL: <https://lexicography.online/etymology/r/ryba> (data obrashcheniya: 10.01.2024).
- Sosnin E. (2011) *Obrazy drevnegermanskoj mifopoeticheskoy modeli mira i ih rekonstrukciya v proizvedeniyah Dzh. R.R.Tolkina. Avtoref. Diss. ... kand.filol.nauk*, Barnaul, 23.

Этимологиядағы «Big data» технологиясының мүмкіндіктері

¹Б.В.Крамаренко, ²Ж.А.Баянбаева

¹Ш.Уалиханов атындағы Көкшетау университеті (Көкшетау, Қазақстан)

²Әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан)

Аңдатпа

«Big data» зерттеу технологиясы (ағылш. «үлкен деректер») өмірдің әртүрлі салаларында: менеджментте, медицинада, логистикада және т.б. өзінің тиімділігін дәлелдеді. Алайда, лингвистикада ол фрагментті түрде қолданылады. Бұл технологияны этимология мен салыстыруда қолдану туралы (автоматтандырылған түрде болмаса да) біз бірнеше мақалалар жарияладық. Алынған нәтижелер негізінде дәстүрлі сөздерден ерекшеленетін сөздерді этимологизациялаудың сындарлы тәсілі жасалды. Оның мәні: көптеген тілдерде ядролық лексикаға қатысты сөздер корпусы бар, олар түбірлер, аффикстер және флексиялар сияқты дәстүрлі элементтерден басқа, конструкциялар, детерминативтер, теріске шығару форманттарынан тұратын құрылымдық компоненттерден тұрады, олардың жиынтығын праморфология деп атауға болады. Бұл мақаланың мақсаты - этимология мен салыстымалы тіл білімінде «big data» технологиясын қолдану мүмкіндіктерін негіздеу және көрсету, сонымен қатар конструктивті тәсіл негізінде бір тілдің емес, бірден бірнеше сөздің этимологиялық мағынасын анықтайтын жаңа типтегі этимологиялық сөздікті құрудың теориялық алғышарттарын әзірлеу. Мақалада этимологиялық іздеудің тиімділігін арттыру үшін басқа ғылыми салалардағы зерттеулер барысында қалыптасқан үлкен деректерді талдаудың бірқатар әдістері мен тәсілдерін қалай қолдануға болатындығы көрсетілген: «Data Mining», «модельдеу», «статистика» және т.б. Бұл зерттеу лингвистикада үлкен деректерді талдаудың әртүрлі әдістерін қолдану мүмкіндіктерін көрсетеді. *Түйін сөздер:* детерминатив, конструктивтілік, конструктивтік тәсіл, праморфология, «big data» технологиясы, теріске шығару форматы.

Possibilities of Big data technology in etymology

¹B.V.Kramarenko, ²Zh.A.Bayanbayeva

¹Sh.Ualikhanov Kokshetau University (Kokshetau, Kazakhstan)

²al-Farabi Kazakh National University (Almaty, Kazakhstan)

Abstract

The "big data" research technology has proven its effectiveness in various fields such as management, medicine, logistics, etc. However, its application in linguistics has been fragmented. Several articles have been published on the non-automated use of this technology in etymology and comparative linguistics. Based on the obtained results, a constructive approach to word etymologization has been developed, differing from the traditional approach. Its essence lies in the fact that in most languages, there is a corpus of words related to core vocabulary, which, in addition to traditional elements such as roots, affixes, and flexions, also includes other structural components: constructs, determinatives, negation formants, the combination of which can be referred to as pramorphology. The goal of this article is to justify and demonstrate the possibilities of using "big data" technology in etymology and comparative linguistics. Additionally, it aims to develop theoretical prerequisites for creating a new type of etymological dictionary that, based on the constructive approach, will determine the etymological meaning of a specific word in multiple languages. The article illustrates how various techniques and methods of big data analysis, established in research in other scientific fields such as Data Mining, modeling, statistics, etc., can be used to enhance the efficiency of etymological research. This study demonstrates the possibilities of using different methods of big data analysis in linguistics

Keywords: determinative, construct, constructive approach, pramorphology, big data technology, negation formant.

Поступила в редакцию: 23.01.2024

Одобрена: 25.03.2024

Первая публикация на сайте: 10.05.2025

MPHTI: 16.21.25

БАСТАУЫШ СЫНЫПТА АҒЫЛШЫН ТІЛІ САБАҒЫНДА РӨЛДІК ОЙЫНДАР АРҚЫЛЫ МӘДЕНИЕТАРАЛЫҚ МЕДИАЦИЯЛЫҚ ДАҒДЫЛАРДЫ ДАМУ

*¹A.C.ТОКЖУМАНОВА^{ID}, ¹Б.Т.РАХИМБАЕВА^{ID}

¹академик Е.А.Бөкетов атындағы Қарағанды университеті
(Қарағанды, Қазақстан)

*adelya181002@gmail.com, byan.turabekovna@gmail.com

Аңдатпа

Балалардың оқу іс-әрекетін мақсатты қалыптастырудың маңызды құралы ойын екені белгілі. Ойынның маңызды функцияларының бірі - әлеуметтену функциясы. Ойын - бұл баланы әлеуметтік қатынастар жүйесіне қосудың, оның мәдени байлығын игерудің ең қуатты құралы. Шет тілін үйрену қазіргі әлемде күн сайын танымал болып келеді, бірақ ең көп таралған тіл - ағылшын тілі, өйткені бұл бүкіл әлемдегі негізгі және халықаралық тілдердің бірі. Ағылшын тілін тиімді түсінуді, сөйлеуді және оқуды үйрену туралы сұрақ туындайды. Ағылшын тілін үйрену процесі соншалықты қарапайым емес және көп күш, уақыт және міндетті түрде қызығушылықты қажет етеді. Білім берудің бастапқы кезеңінде - бастауыш мектепте берілетін білім маңызды рөлдердің бірі болып табылады. Бұл жаста адам жаңа ашылуларға және жаңа нәрселерді үйренуге қызығушылық таныта бастайды, сондықтан бұл маңызды кезең. Балаларды ағылшын тілін үйренуге қызықтыру үшін білім беру технологияларын қолдану қажет. Осындай технологиялардың бірі - ойын. Бұл нұсқаулықта тек ойындар ғана емес, рөлдік ойындар да бар.

Түйін сөздер: ойын, рөлдік ойын, халықаралық, бастауыш мектеп.