

RU

Принципы построения татарско-английского онлайн-словаря проекта Glosbe на уровне макро- и микроструктуры

Сафиуллина Г. Р.

Аннотация. Цель исследования – определить особенности построения электронного татарско-английского словаря. В статье исследуются особенности подачи лексического материала в онлайн-словарях. Данное исследование проводится на материале татарско-английского онлайн-словаря, размещенного на площадке www.glosbe.com. Научная новизна исследования заключается в выявлении мультимодальных аспектов построения словарной статьи англо-татарского словаря Glosbe. В результате выявлено, что словарный проект Glosbe является уникальным в плане подачи грамматических конструкций и иллюстративных примеров благодаря правильной организации макро- и микроструктуры и активному взаимодействию с пользовательской аудиторией, регулярно пополняющей фонд иллюстративных примеров.

EN

Principles of Compiling the Tatar-English Online Dictionary of the Glosbe Project at Macroand Microstructure Levels

Safiullina G. R.

Abstract. The study aims to determine features of compiling an electronic Tatar-English dictionary. The article examines peculiarities of presenting lexical material in online dictionaries. The study is conducted on the material of the Tatar-English online dictionary posted on the website www.glosbe.com. Scientific novelty of the work lies in identifying multimodal aspects of how a dictionary entry is arranged in the English-Tatar Glosbe dictionary. As a result, it was found that the Glosbe dictionary project is unique in the way it presents grammatical constructions and illustrative examples due to proper organisation of its macro- and microstructure and active interaction with the user base, members of which regularly contribute to the repository of illustrative examples.

Введение

В данной статье нами сделана попытка выявить особенности построения татарско-английского онлайн-словаря проекта Glosbe, размещенного на сайте www.glosbe.com, на уровне микроструктуры и макроструктуры [8]. Достижения информационных технологий и всеобщая доступность сети Интернет упростили обработку языковой информации, позволили оперировать объемами информации как в семантическом, так и в структурно-грамматическом плане при формировании корпуса словаря и отборе иллюстративных примеров. Мы являемся свидетелями появления новой формы лексикографической коммуникации, фиксирующей народную лексикографию XXI века, а именно онлайн-словаря, пополнить который может любой зарегистрированный пользователь.

Актуальность данной статьи обусловлена огромной популярностью и востребованностью электронных онлайн-словарей в современную эпоху открытого доступа к информации, размещенной в сети Интернет. Двухязычный татарско-английский словарь проекта Glosbe как один из наиболее полных и систематизированных онлайн-словарей данной пары языков вызывает повышенный интерес лексикографов с точки зрения построения на уровне макроструктуры и микроструктуры.

Цель статьи ставит перед автором следующие задачи: краткий обзор теории вопроса, рассмотрение макроструктуры и микроструктуры татарско-английского словаря Glosbe, изучение аспекта мультимодальности информации в словарной статье [Ibidem].

В статье мы применяем следующие методы исследования: метод наблюдения, сравнения, анализа, синтеза и классификации, а также метод сплошной выборки, сопоставительный метод и метод статистического анализа.

Теоретической базой нашего исследования послужили работы лингвистов Педро А. Фуэртес-Оливера, Х. Бергенхольца [4], С. Тарпа [5], отметивших большой прогресс в использовании новых технологий в онлайн-словарях нашего времени, подтверждая позицию Р. Х. Гоуса [6; 7], говорящего о недостаточно высоком уровне ранних онлайн-словарей, и Т. Ботма, Р. Н. Гоуса, Д. Ж. Принслоу [3], заявляющих о значимости онлайн-лексикографии на современном этапе развития словарного дела. Х. Неси [11] акцентирует внимание на дидактической функции онлайн-словарей. Все типы модальности, разработанные Р. Лью [9] и Х. Лиу [10], могут с успехом применяться в учебной деятельности. Российские ученые В. А. Козырев, В. Д. Черняк [1], Е. Г. Лукашанец [2] отмечают важные особенности современной онлайн-лексикографии, которая не представляется возможной, по мнению Г. Р. Сафиуллиной, Л. С. Сиразовой, Е. Лелаковой, без мультимодальной презентации значения в онлайн-словарях [12].

Практическая значимость работы заключается в возможности применения результатов при разработке электронных онлайн-словарных статей для других тюркских языков и при дальнейшем совершенствовании имеющихся татарско-английских лексикографических онлайн-источников.

Краткий обзор теории вопроса

По нашему мнению, на ранней стадии зарождения электронные словари были составлены по форме, продиктованной традиционным печатным форматом, возможности онлайн-ресурсов не были до конца осознаны авторами, следовательно, потенциал электронной версии поначалу не был раскрыт полностью. Ситуация качественно изменилась с ростом числа онлайн-словарей. Р. Х. Гоус отмечает, что на заре электронной лексикографии слишком много электронных продуктов было разработано без какого-либо влияния новаторских теоретических предложений, следовательно, некоторые электронные словари не оправдывали ожиданий их пользователей [6; 7]. П. Фуэртес-Оливера и С. Тарп отмечали большой прогресс в использовании новых мощных технологий в специализированных онлайн-словарях в последнее время [5]. Х. Неси подчеркивала дидактическую функцию онлайн-словарей: на современном этапе развития электронной лексикографии словарь часто используется изучающими язык в качестве средства овладения словарным запасом, пользователи словаря опираются на семантику при освоении новых слов [11]. Таким образом, одной из ключевых задач актуальной электронной лексикографии является представление значений в мультимодальном тексте.

На наш взгляд, на сегодняшнем этапе развития электронной лексикографии мы можем уверенно заявить, что она проходит через все стадии развития, свойственные печатной лексикографии. Аналогично бумажным изданиям самые первые словари были составлены не лексикографами, так как практика создания словарей опережала теорию. Низкое качество первых электронных словарей можно объяснить отсутствием теоретических моделей. В том случае, если составители онлайн-словаря опирались на лексикографическую теорию, они руководствовались постулатами, разработанными для печатных словарей. Таким образом, по мнению Р. Х. Гоуса, в дальнейшем необходимо разработать теоретические модели для электронных словарей, которые фокусируются на проблемных вопросах, таких как данные, которые должны быть включены в эти словари, структуры для представления и размещения информации, а также то, как они должны реагировать на потребности своих целевых пользователей [7].

Р. Лью предложил структуру мультимодальной лексикографии, проиллюстрировав различные способы указания значения в электронных словарях и продемонстрировав, как они могут дополнять словесные определения. В трихотомии мультимодальных устройств, предложенной Лью, есть три категории: письменный язык, аудиопрезентация вербальных элементов, включая записи человеческого голоса и синтезированную речь, и другие устройства. Третью категорию, предложенную Р. Лью, можно разделить на пять типов: 1) аудио-запись неязыковых звуков; 2) визуальные средства (статические графические иллюстрации, фотографии, графики, немые анимации, таблицы, значки, типографика и т.д.); 3) аудиовизуальные средства (анимации и видеоклипы); 4) пространственные средства (кадрирование, гиперссылки, плавающие подсказки, сенсорные меню, всплывающие окна и т.д.) и 5) другие возможности, которые можно реализовать (например, тактильные или вкусовые режимы) [9].

На данном этапе развития по типологическим характеристикам электронный словарь может быть одноязычным, двуязычным или многоязычным. Словарное приложение может предназначаться для установки на различные девайсы или гаджеты: от компьютера до телефона и др. Целевая аудитория словаря весьма разнообразна – дети или взрослые, изучающие иностранные языки, или носители языка, следовательно, словарь может быть кодирующего и декодирующего типа. По жанровым характеристикам словари могут быть учебными, энциклопедическими, терминологическими, академическими, словарями сленга и неологизмов и др. Все онлайн-словари объединены возможностями мультимодального представления заглавных слов, которыми снабжен Интернет.

Е. Г. Лукашанец в своей статье описывает вариативность в русскоязычных онлайн-словарях сленга и разговорной лексики, постоянно пополняемых пользователями Интернета. Вариативность как главное свойство онлайн-словарей проявляется в варьированности лексикографического описания и в значительно развитой вариантности описываемых в словарях единиц. Как специфическое проявление вариативности лексикографического описания отмечено дублирование словарных статей, в основе которого лежит стремление уточнить толкование значения и выразить отношение к обозначаемому словом понятию [2].

Рассмотрение макроструктуры и микроструктуры татарско-английского словаря Glosbe

Данный проект размещает на своем сайте словари практически на всех языках земного шара. Самыми объемными словарями являются, в порядке уменьшения переводческих соответствий, словари французского (1 819 368 переводческих соответствий), немецкого, испанского, польского, японского, итальянского, китайского, русского, португальского, нидерландского, финского, арабского, шведского, чешского, венгерского, турецкого, хорватского, хинди и вьетнамского языков (113 842 переводческих соответствия). В открытой для пользователей бесплатной базе переводов предложений с одного языка на другой размещается более одного миллиарда примеров. В базу зарегистрированный пользователь может добавить предложение на любом из заявленных языков с его переводом на другой язык, расширяя имеющийся словарь по своему усмотрению. Интерфейс главной страницы сайта весьма удобен, в центре располагается окно для ввода искомого слова и таблица для выбора языка, ниже располагается информация о самых объемных по количеству статей словарях, демонстрируются данные о последних действиях пользователей, есть отсылающее окно на страницу проекта в Facebook.

На уровне макроструктуры словаря при выборе языка (в рамках данного исследования изучаемая пара языков – это татарский и английский языки) пользователь попадает на страницу с развернутым описанием данных о переводческих соответствиях конкретной пары языков. Указывается количество переведенных фраз и предложений, на сегодняшний день это 20 895 лексических единиц и около шести миллионов предложений. В текстовом поле пользователь может ввести запрашиваемое для перевода слово, перевод слова сопровождается всеми имеющимися примерами [8]. Новые заглавные слова и примеры пользователь может добавлять в неограниченном количестве.

На уровне микроструктуры словаря словарная статья состоит из: 1) графического изображения заглавного слова; 2) грамматической пометы частеречной принадлежности; 3) мультимодального аудиосопровождения в виде записи мужского, женского или синтезированного голоса при отсутствии фонетической транскрипции; 4) переводческого соответствия заглавного слова, как правило, в единичном варианте; 5) полного списка иллюстративных примеров, взятых из открытых корпусов переводов татарских текстов на английский язык.

Изучение мультимодальности информации в словарной статье

Под мультимодальностью, вслед за Х. Лиу, мы понимаем набор различных семиотических моделей (текстуальных, визуальных, аудиальных, пространственных и т.д.), от которых зависит восприятие информации пользователями словаря [10, р. 211]. Обратимся к мультимодальному аспекту словарных статей татарско-английского словаря проекта Glosbe. Из всех типов модальности, предложенных Р. Лью, в данном словаре мы наблюдаем письменный язык и аудиопрезентацию вербальных элементов, включая записи мужского и женского голосов, а также синтезированную компьютерную речь, пространственные элементы – гиперссылки на социальную сеть, другие словари проекта Glosbe, всплывающие окна регистрации, смены языков и всплывающее окно добавления собственных примеров. В словарной статье татарско-английского словаря проекта Glosbe отсутствуют аудиозаписи неязыковых звуков, визуальная модальность, анимация и видеоклипы, тактильные или вкусовые режимы.

Что касается базы примеров в данном словаре, то можем выделить три группы. В первую группу входят переводческие соответствия в виде аналогов или эквивалентов, но с отсутствующими примерами перевода предложений. Рассмотрим примеры данных словарных статей татарско-английского словаря проекта Glosbe:

1) -га бирелүчән to be inclined to (быть склонным к чему-либо)

Примеры отсутствуют [8];

2) -гач ук as soon as (после того как) {conjunction adposition}

Примеры отсутствуют [ibidem].

Во вторую группу мы выделили переводческие соответствия в виде эквивалента или аналога, но уже с иллюстрирующими предложениями, взятыми из английского текста Нового Завета и его перевода на татарский язык, из базы wikimatrix и других источников открытого доступа;

3) -дан алып from (от) {adposition}

Пример предложения с “-дан алып”.

Приведем примеры перевода предложения с английского на татарский язык, которые являются эквивалентными.

Алар алтын һәм ладан алып килер, Gold and frankincense they will carry. / Они принесут золото и ладан;

4) -га күрә in view of (ввиду того что) {adposition}

Пример предложения с “-га күрә”.

Хакимнар 6:11-16 га нигезләнен, ни өчен Йәһвә үз хезмәтчеләренәң яшерен сәләтләрен күрә дип әйтпән була? How does the account at Judges 6:11-16 show that Jehovah sees the potential in his servants [ibidem]? / Как запись Судей 6:11-16 показывает, что Иегова видит потенциал в своих слугах?

Рассмотрим пример перевода при помощи аналога:

5) -дан алып from (от) {adposition}

Ул йөрәгендә дан ярату яисә мул тормыш алып бару теләге үстерә башлагандыр. / Он, скорее всего, начал возвращать в своем сердце желание славы и богатой жизни.

Perhaps he began to nurture personal ambitions or a desire for material prosperity. / Возможно, он начал взращивать в своем сердце желание славы и богатой жизни.

В следующем примере, взятом из базы wikimatrix, наблюдается ошибка семантического плана:

6) -дан алып from (от) {adposition}

Предложение «Келлог монополияларга каршы эшләрне алып барып дан казана». / букв.: Зарабатывает славу тем, что противостоит монополии Келлог в словаре ошибочно переведено как “Allowing the citizens to sue the banks”. / букв.: Разрешая гражданам подавать в суд на банк. Мы наблюдаем полное расхождение плана содержания и выражения предложения, что, к сожалению, свойственно народной лексикографии на этом этапе развития, и относим это к недочетам [Ibidem].

В третью группу мы включили словарные статьи и иллюстративные переводы, предложенные самими пользователями словаря.

Данный пример демонстрирует перевод предложения текста научно-популярного характера:

7) алоэ aloe {noun}

Алоэ, кассия һәм дарчин бүгенге Кытай, Һиндстан һәм Шри-Ланка территориясендә үскән.

Aloes, cassia, and cinnamon were found in what is today China, India, and Sri Lanka 12 [Ibidem]. / Алоэ, кассия и корица росли на территории современного Китая, Индии, Шри Ланки (Author: JW_2017_12 JW_2017_).

Проведя выборочный анализ примеров, мы полагаем, что основным достоинством онлайн-словарей является возможность мультимодальной презентации материала и добавления примеров зарегистрированными пользователями. Оба эти нововведения были с успехом применены в составлении татарско-английского словаря проекта Glosbe, что привело к видоизменениям на уровне макроструктуры и микроструктуры. К основным изменениям мы причисляем: включение в макроструктуру словаря гиперссылок на остальные словари проекта Glosbe, выплывающие окна для запрашиваемых для перевода слов, окно для добавления примеров зарегистрированными пользователями. На уровне микроструктуры в словарную статью были включены иллюстративные примеры пользователей словаря. Включение примеров пользователями требует редакторской правки, которая необходима на следующем этапе существования данного проекта. Также на уровне микроструктуры изменения затронули модальность представления заглавного слова. Аудиозапись, которая является хорошим дополнением словесного определения, сопровождает каждую заглавную единицу татарско-английского словаря.

Заключение

В результате проведенного анализа словарных статей, отобранных из корпуса татарско-английского онлайн-словаря проекта Glosbe, мы приходим к выводам о том, что принципы построения статьи татарско-английского онлайн-словаря на уровне микроструктуры и макроструктуры перешли на качественно новый этап, что соответствует изученным нами теоретическим работам. Рассмотренные макро- и микроструктуры соответствуют общепринятым стандартам построения словаря: словарные статьи, следующие за заглавным словом или предшествующие ему согласно алфавитному порядку, помещаются на страницу статьи заголовочного слова. Мультимодальный аспект подачи информации представлен в виде ссылок на аудиосопровождение статьи в нескольких вариантах, на все примеры использования данного заголовочного слова с их переводом на английский язык. Данная подача материала в словарной онлайн-статье оправданна, так как представляет наиболее полную информацию для пользователя в аспекте актуального использования и мультимодальности и может быть с успехом применена в двуязычных словарях других тюркских языков. К перспективам дальнейшего исследования можно отнести углублённое изучение аспекта мультимодальности в онлайн-словарях различных пар языков.

Список источников

1. Козырев В. А., Черняк В. Д. Современные ориентации отечественной лексикографии // Вопросы лексикографии. 2014. № 1 (5). С. 5-16.
2. Лукашанец Е. Г. Вариативность как специфическая черта онлайн-словарей сленга и разговорной лексики // Вопросы лексикографии. 2014. № 1 (5). С. 16-32.
3. Bothma T., Gouws R. H., Prinsloo D. J. The Role of E-lexicography in the Confirmation of Lexicography as an Independent and Multidisciplinary Field // Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity / ed. by T. Margalitadze, G. Meladze. Tbilisi: Tbilisi State University, 2016. P. 111-118.
4. Fuertes-Olivera P. A., Bergenholtz H. E-Lexicography: The Internet, Digital Initiatives and Lexicography. L. - N. Y.: Bloomsbury Academic, 2011. 368 p.
5. Fuertes-Olivera P. A., Tarp S. Theory and practice of specialised online dictionaries: Lexicography versus terminography. Berlin - Boston: Walter de Gruyter GmbH & Co KG, 2014. 269 p.
6. Gouws R. H. Internet Lexicography in the 21st Century // Wortschatz: Theorie, Empirie, Dokumentation / hrsg. von S. Engelberg et al. Berlin: De Gruyter, 2018. P. 215-236.
7. Gouws R. H. Metalexicographic models for multilingual online dictionaries in emerging e-societies // Studies in multilingual lexicography / ed. by M. J. Dominguez Vázquez et al. Berlin: De Gruyter, 2020. P. 29-46.

8. <https://glosbe.com/tt/en> (дата обращения: 15.02.2021).
9. Lew R. Multimodal lexicography: The representation of meaning in electronic dictionaries // *Lexikos*. 2010. Vol. 20. P. 290-306.
10. Liu X. Multimodal Definition: The Multiplication of Meaning in Electronic Dictionaries // *Lexikos*. 2015. Vol. 25. P. 210-232.
11. Nesi H. Dictionaries in electronic form // *The Oxford History of English Lexicography* / ed. by A. P. Cowie. Oxford: Oxford University Press, 2009. P. 458-478.
12. Safiullina G. R., Sirazova L. S., Lelakova E. Multimodal Representation of Meaning in Merriam-Webster Online Dictionaries // *Applied Linguistics Research Journal*. 2020. № 4 (8). P. 47-53.

Информация об авторах | Author information



Сафиуллина Гульшат Рафаиловна¹, к. филол. н., доц.

¹ Казанский федеральный университет



Safiullina Gulshat Rafailevna¹, PhD

¹ Kazan Federal University

¹ safiullina_gulshat@mail.ru

Информация о статье | About this article

Дата поступления рукописи (received): 24.02.2021; опубликовано (published): 30.04.2021.

Ключевые слова (keywords): татарский язык; онлайн-словарь; макроструктура; микроструктура; иллюстративный пример; Tatar language; online dictionary; macrostructure; microstructure; illustrative example.