МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Филологический факультет
Кафедра русского языка
Лаборатория общей и сибирской лексикографии
ИНСТИТУТ РУССКОГО ЯЗЫКА ИМ. В.В. ВИНОГРАДОВА РАН
ГОСУДАРСТВЕННЫЙ ИНСТИТУТ РУССКОГО ЯЗЫКА им. А.С. ПУШКИНА
ПОД ЭГИДОЙ РОССИЙСКОГО ОБЩЕСТВА ПРЕПОДАВАТЕЛЕЙ
РУССКОГО ЯЗЫКА И ЛИТЕРАТУРЫ (РОПРЯЛ)

ЛЕКСИКОГРАФИЯ ЦИФРОВОЙ ЭПОХИ

Сборник материалов Международного симпозиума
(24–25 сентября 2021 г.)

Томск
Издательство Томского государственного университета
2021

DOI: 10.17223/978-5-907442-19-1-2021-115

Н.Г. Кантышева, О.И. Девяткова
Nadezhda G. Kantysheva, Olga I. Devyatkova

**Проект терминологического словаря таможенной сферы:
корпусный подход**
**Project of a terminological dictionary of the customs sphere:
corpus approach**

Тюменский государственный университет, Тюмень –
University of Tyumen, Tyumen
nkantyscheva_@mail.ru

Аннотация. Доклад освещает тематические аспекты современных исследований по корпусной лексикографии; сложности отбора языкового материала и спектр задач корпусных инструментов в процессе составления терминографических ресурсов. Рассмотрен алгоритм построения терминологического словаря таможенной сферы в рамках Евразийского экономического союза.

Summary. The paper deals with general overviews of thematic aspects research within the field of the modern corpus lexicography; difficulties in selecting language material and a wide range of tasks for corpus tools in the process of compiling terminographic resources. The method of constructing a terminological dictionary of the customs sphere within the framework of the Eurasian Economic Union is considered.

Ключевые слова: обработка текстов на естественном языке, корпусная лексикография, словари на основе корпусов, терминологический словарь

Keywords: natural language processing, corpus lexicography, corpus-based dictionary, terminological dictionary

Повсеместная цифровизация и интеграция наук меняют статус словарей нового поколения, которые в настоящее время расширяются за счет включения электронных баз данных. Лексикографическое направление становится одним из приоритетных в современной лингвистике, происходит обновление технологий и методик лексикографирования, которые позволяют анализировать тенденции развития современной корпусной лексикографии. Доказательством этого являются многочисленные публикации, выполненные по тематике словарных проектов и появление разнообразных форм словарей. Такие предпосылки обуславливают необходимость формирования навыков изучения естественного языка на основе

корпусов и составления электронных лексикографических источников, соответствующих принципам организации информационной среды.

В целях соблюдения этого баланса к задаче составления любого словаря на основе корпуса нужно подходить комплексно, учитывая, с одной стороны, исследовательские и дидактические потребности ученых-лингвистов, с другой стороны, современные инструменты моделирования лексикографических ресурсов. Заявленный в классических источниках по лексикографии тезис комплексного подхода при составлении словарей постулируется и в современных работах [1–5]. На практике предлагаемые в научной литературе методики и инструменты не столь активно применяются в процессе создания электронных лексикографических ресурсов [6], либо их применение не всегда позволяет в полном объеме решить поставленные задачи.

Несмотря на систематическую работу в области исследования терминологических систем, отдельные терминологии все еще не описаны в полной мере лингвистами и лексикографами. В частности это касается терминологии такой интегративной сферы, как таможенное дело в рамках Евразийского экономического союза (ЕАЭС). Данный тип институционального дискурса является молодым и интенсивно развивающимся. Таможенный кодекс ЕАЭС вышел в редакции 2018 года и поэтому требует, прежде всего, ясности в определении основных понятий изучаемой области. Это особенно релевантно в свете последних событий, поскольку таможенное право регламентирует международные отношения, в частности, в условиях политического и экономического кризиса. В ЕАЭС обеспечивается свобода движения товаров, услуг, капитала и рабочей силы и проведение скоординированной, согласованной политики в отраслях экономики.

В процессе моделирования терминологического словаря таможенной сферы применяется комплексный подход, сочетающий корпусные технологии, статистические методы, качественную и количественную методики. Таможенная терминология изучается в естественных условиях ее функционирования, в реальных условиях профессиональной коммуникации. В связи с этим в исследовании решается вопрос о количественной и качественной оптимизации отбора текстов. Выборочная совокупность текстов представляет усредненную схему вероятностного построения словарей, обобщающую наборы ситуаций, в которых написаны тексты. Адресатом словаря могут стать специалисты, работающие в таможенной сфере, а также преподаватели и студенты таможенных и экономических факультетов, лингвисты и переводчики.

Подобные терминографические продукты являются вспомогательной дидактической базой для педагогов современного университета, поскольку позволяют применять и изучать необходимый терминологический аппарат в рамках образовательного пространства, анализировать системные отношения термина, строить логическую дефиницию специального понятия. Обучающие приёмы овладения терминосистемой помогают студенту упорядочивать научную терминологию, уточнять взаимосвязи и взаимоотношения между понятиями, формировать навык самостоятельного конструирования краткого определения любого термина.

Литература

1. *Bubenhofen N., Ptashnyk S.* Korpora, Datenbanken und das Web: State of the Art computergestützter Forschung in der Phraseologie und Lexikographie // *Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie.* Hohengehren, 2010. S. 7–20.
2. *Mederake N, Klosa A., Müller-Spitzer C.* Internetlexikografie. Ein Kompendium // *International Journal of Lexicography.* Vol. 31, is. 4. P. 515–518.
3. *Карпова О.М.* Английская лексикография. М.: Академия, 2010.
4. *Копотев М.В.* Введение в корпусную лингвистику. Praha: Animedia, 2014.
5. *Саженин И.И.* Корпусные методы в лексикографии: опыт создания модели словарного корпуса: дис. ... канд. филол. наук. Новосибирск, 2013. 192.
6. *Селегей В.П.* Электронные словари и компьютерная лексикография. URL: <https://docplayer.ru/36764723-Elektronnye-slovary-i-kompyuternaya-leksikografiya.html> (дата обращения: 25.07.2020).

DOI: 10.17223/978-5-907442-19-1-2021-116

А.Б. Коконова
Anna B. Kokonova

**Фонотека кабинета диалектологии МГУ:
от архива к базе данных
Record library of the Moscow State University dialectology room:
from archive to database**

Московский государственный университет имени М.В. Ломоносова, Москва –
Moscow State University, Moscow, Russia
annakokonova@gmail.com

Аннотация. Фонотека кабинета диалектологии МГУ является собранием уникальных записей культурно-исторического характера. Задача нашего