

*Вестник Ивановского государственного университета.*

*Серия: Гуманитарные науки. 2023. Вып. 3. С. 64—73.*

*Ivanovo State University Bulletin. Series: Humanities. 2023. Iss. 3. P. 64—73.*

Научная статья

УДК 81'26

DOI: 10.46726/И.2023.3.8

## **ЭВОЛЮЦИЯ МЕТОДОВ КОНСТРУИРОВАНИЯ ТЕЗАУРУСОВ: ОТ ИНТУИТИВНОЙ КОМПИЛЯЦИИ К ОНТОЛОГИЧЕСКОМУ ИНЖИНИРИНГУ (На материале английского языка)**

***Инна Алексеевна Воронцова***

Ярославский государственный педагогический университет им. К.Д. Ушинского,  
г. Ярославль, Россия, aginnal@yandex.ru

**Аннотация.** Статья содержит исторический обзор и анализ методов тезаурусного моделирования в англоязычной лексикографии. Основной задачей исследования является систематизация семантически-ориентированных подходов, применявшихся для описания лексики английского языка на разных этапах развития лексикографической практики (X—XXI вв.). В ходе исследования автор приходит к выводу, что отправной точкой англоязычной тезаурусной лексикографии являются интуитивно компилируемые тематические глоссарии, основанные на применении парадигматического или сочетании парадигматического и синтагматического подходов к организации и описанию лексики в словаре и репрезентирующие фрагменты «осязаемой» картины мира. Формализация и специализация научного знания обеспечивает постепенный перенос научных методов на создание словарей с понятийной организацией словника, в результате чего происходит размежевание иерархического и аналогического (алфавитно-понятийного) подходов, с одной стороны, и разделение общей и специальной идеографии, с другой. Развитие технологий, в первую очередь корпусных, и интеллектуального анализа текста выводит общие и специальные тезаурусы на уровень продукта и одновременно инструмента инженерии знаний. Тезаурус, фокусирующийся на лингвистических данных, и онтология, систематизирующая экстралингвистические знания, позволяют моделировать предметные области и (в потенции) целостную картину мира во всей полноте взаимосвязей, устанавливаемых между идеями и обозначающими их лексическими единицами.

**Ключевые слова:** тезаурус, онтология, корпус, инженерия знаний, парадигматический подход, синтагматический подход

**Для цитирования:** Воронцова И.А. Эволюция методов конструирования тезаурусов: от интуитивной компиляции к онтологическому инжинирингу: (На материале английского языка) // Вестник Ивановского государственного университета. Серия: Гуманитарные науки. 2023. Вып. 3. С. 64—73.

Original article

## EVOLUTION OF THESAURUS CONSTRUCTION METHODS: FROM INTUITIVE COMPILATION TO ONTOLOGICAL ENGINEERING (With special reference to English)

*Inna A. Vorontsova*

Yaroslavl State Pedagogical University named after K.D. Ushinsky,  
Yaroslavl, Russia, arinnal@yandex.ru

**Abstract.** The article contains a historical survey and analysis of thesaurus modeling methods in English lexicography. The research objective is to systematize the semantically oriented approaches adopted to describe English vocabulary at different stages of lexicography development (X—XXI centuries). The author concludes that intuitively compiled thematic glossaries, based on the application of paradigmatic or a combination of paradigmatic and syntagmatic approaches to the organization and description of vocabulary and representing fragments of a “tangible” world picture serve as the starting point of English thesaurus lexicography. The formalization and specialization of scientific knowledge provides a gradual transfer of scientific methods on thesaurus construction, which results in the departure of the hierarchical approach from the analogical (alphabetic and concept-based) method of vocabulary organization, on the one hand, and the separation of general- and specific-purpose ideography, on the other. The development of technologies, primarily corpus and text mining, has brought general- and specific purpose thesauri to the level of a product and tool of knowledge engineering. Thesaurus focusing on linguistic data and ontology systematizing extralinguistic knowledge allow for the modeling of subject areas and (in potency) of the world picture in the entirety of the relationships established between ideas and the lexical units denoting them.

**Keywords:** thesaurus, ontology, corpus, knowledge engineering, paradigmatic approach, syntagmatic approach

**For citation:** Vorontsova I.A. Evolution of thesaurus construction methods: from intuitive compilation to ontological engineering: (With special reference to English), *Ivanovo State University Bulletin, Series: Humanities*, 2023, iss. 3, pp. 64—73.

Тезаурус как лексикографическая форма имеет богатую историю и известен со времен античной Греции. Первый англоязычный (латинско-английский) словарь (глоссарий) с идеографо-тематической организацией лексики — глоссарий аббата Эльфрика (*Abbot Aelfric's Glossary*) — появился в X в. Порядок следования тематических групп в глоссарии не был лишен логики, однако обнаруживал некоторые противоречия: так, блок тем, представленный лексическими множествами «человек, родственные отношения, части тела, болезни, постройки, пища, напитки», не содержал наименований предметов одежды, которые, на наш взгляд, весьма органично смотрелись бы в данном ряду; последние, однако, выносились в блок, объединявший наименования игр и развлечений, оружия и др. [Воронцова 2004: 80]. Тем не менее неоценимой заслугой данного тематического справочника является заданная в нем дедуктивная логика представления лингвистического материала: Бог — планеты — земля — человек — окружение человека (нерукотворное и рукотворное). Именно такая структурная организация будет заимствована в словари тезаурусного типа последующих поколений [Stames, Noyes: 199].

В начале XIII в. британские лексикографы апробировали новый формат идеографо-тематической презентации лексики: парадигматический подход реализовывался на этапе отбора лексики и конструирования словника, словарная статья (рубрика) являла собой связный текст, демонстрировавший употребление отобранной лексики в контексте. Первым словарем, построенным по такому принципу, был латинско-французско-английский словарь «О названиях предметов утвари» (*De Nominibus Utensilium*) А. Неккама. Словарь описывал кухонную утварь, методы приготовления пищи, дом, домашнюю обстановку, сельскохозяйственные орудия и т. п. в составе текста на латинском языке с его подстрочным переводом на французский и английский языки. В связи с этим интересен и опыт У. Биббсуорта, автора поэтического «Трактата» (*Le Tretiz*, конец XIII в.) на французском языке, построенного по тематическому принципу и содержавшего подстрочный перевод на английский язык и комментарии грамматического характера. Тематические группы, представленные в «Трактате», отличались практикоориентированностью и прагматизмом: строение человеческого тела, процесс одевания и наименования предметов одежды, хорошие манеры и правильное питание, производство тканей, рыболовство, сельское хозяйство, погода и времена года, наименования цветов, плодов и деревьев, птиц и животных, дом и домашняя обстановка, приготовление пищи и сервирование стола и др. [Воронцова 2004: 83]. «Трактат» был написан по заказу знатной особы для обучения ее детей французскому языку, однако впоследствии приобрел более широкую популярность и был включен в позднесредневековый учебник французского языка *Femina Nova* [Wright: x - xi]. Словари подобного рода не отличались широтой тематического охвата и репрезентировали фрагмент «осязаемой» картины мира, формируемой преимущественно материальными объектами. Это, вероятно, было обусловлено учебной направленностью словарей, а также гибридным (парадигматико-синтагматическим) подходом к организации лексики, усложнявшим как структуру словаря, так и его содержание — запоминанию подлежала не только лексика, но и грамматика языка.

Опыт сочетания парадигматического и синтагматического подходов был повторен в XVI в. У. Хорманом («Повседневный» (*Vulgaria*), 1519) и Р. Уиттинтоном («Повседневный» (*Vulgaria*), 1520) и в XVII в. Яном А. Коменским (*Janua Linguarum Reserata*, 1631 и *Orbis Sensualium Pictus*, 1657), затем тематическая традиция практически полностью перешла в парадигматическую (парадигматико-эпидигматическую) плоскость. Возрождение идеи применения синтагматического подхода к описанию лексики в англоязычном тезаурусе произошло в конце XX в. Такой подход представляется весьма обоснованным в свете современной парадигмы лексикографического знания: в научной литературе все чаще звучит мысль о необходимости рассмотрения тезауруса как семантической сети, строящейся с учетом традиции употребления слов носителями языка, ибо значение слова формируется только в результате накопления знаний о множестве контекстов употребления данного слова в речи [Осокина].

Признавая смелость и новизну лексикографических опытов X—XIV вв., отметим, однако, что фрагментарность картины мира, отражаемой в ранних тезаурусах английского языка, а также субъективность и, возможно, мировоззренческая ангажированность, присущая авторским попыткам классифицировать лексику, указывает на донаучность методов конструирования тезаурусов в этот период. Тематические глоссарии X—XIV вв. были лишены возможности опираться на системные знания о мире и составлялись во многом интуитивно.

Создаваемые в XV—XVIII вв. тезаурусы («Слова» (*Vocabula*, 1496) и «Повседневный» (*Vulgaria*, 1508) Дж. Стэнбриджа; «Краткий словарь для начинающих» (*Shorte Dictionarie for Yonge Begynners*, 1556) Дж. Виталса; тематический словарь Дж. Рея (*Dictionarium Trilingue*, 1675); «Лондонский тематический словарь» (*The London Vocabulary*, ок. 1700) Дж. Гринвуда и др. обнаруживали тенденцию к расширению тематического диапазона и увеличению словника за счет вновь появляющихся в языке слов, например, названий цеховых специальностей (сапожник, золотых дел мастер, кузнец и т. п.), а также некоторому углублению идеографических связей в рубрике, то есть иерархизации смысловых отношений в микроструктуре словаря, однако принципиального развития лексикографическая форма тезауруса в этот период не получила.

Конец XVIII в., однако, открывает эру синонимических словарей английского языка: это труд Дж. Барклея, изданный в 1774 г. под названием «Полный и универсальный словарь английского языка» (*Complete and Universal English Dictionary*), словари Дж. Траслера *Difference between Words Esteemed Synonymous in the English Language* (1776), Эстер Л. Пийоцци *British Synonymy, or an attempt at regulating the choice of words in familiar conversation* (1794), У. Тейлора *English Synonyms Discriminated* (1813), Г. Крэбба *English Synonymes Explained* (1816), Р. Уатлея *Selection of English Synonyms* (1854) и др. [Воронцова 2004: 106]. Отметим, что англоязычные словари американского языкового варианта, содержащие последовательную информацию о синонимах, связаны с великими именами Н. Уэбстера и Дж. Вустера: *An American Dictionary of the English Language. Unabridged, revised, illustrated edition, with added synonymy* (1859) и *A Pronouncing, Explanatory, and Synonymous Dictionary of the English Language* (1855) соответственно [Burkett: 176—177, 207—208].

Значительная часть синонимических словарей XVIII — 1-й пол. XIX века имела дефинитивный характер, предоставляя пользователю толкование заглавных единиц словаря и энциклопедическую справку об обозначаемых ими понятиях. Отдельные словари (словарь Эстер Л. Пийоцци — см. выше) делали особый акцент на включение синонимов в коллокации и привычные сентенции. Издавались и словари кумулятивного (списочного) плана.

В целом английские словари конца XVIII — 1-й пол. XIX в. отличаются более высокой степенью научности методов по сравнению со своими предшественниками. Достаточно вспомнить принципы лексикографического описания языкового материала, сформулированные С. Джонсоном: разграничение значений в структуре многозначных слов, документирование значений и слов посредством привлечения цитат из авторитетных источников, маркировка лексических единиц с помощью функциональной системы помет и т. д., однако тезаурусы (словари синонимов) этого периода не всегда опираются на первоисточники (адресация в словарях практически отсутствует), толкования и комментарии по-прежнему носят субъективный характер (авторы словарей считают достаточным апеллировать к собственному языковому чутью), язык описания еще не в полной мере подчиняется требованиям к научному стилю, комментарии, данные к коллокациям и узуальным сентенциям, часто имеют ироничную тональность [Воронцова 2004: 107—110]. К неоспоримым достижениям синонимической лексикографии этого периода относится собранный в словарях богатый языковой материал и большой практический опыт построения и расположения словарных статей, устройства алфавитных указателей, использования различных типов синонимических характеристик [Мельник: 4].

Разработка синонимических словарей активно продолжилась во второй половине XIX в., когда их число достигло более 300 [Там же: 7]. Синонимические словари постепенно «обрастали» чертами аналогического тезауруса, фиксируя в рубриках антонимическую, гипонимическую и гиперонимическую лексику. Отдельные издания имели классификационную схему, основанную на выявлении смысловых связей между заглавными словами (например, словарь Г. Крэбба — см. выше), иными словами, обнаруживали черты идеографического тезауруса.

Поистине исторической вехой в развитии идеографической лексикографии — как национальной, так и мировой — стал тезаурус Питера М. Роже (*Roget's Thesaurus of English Words and Phrases*, 1852 г.). В контексте проводимого исследования особую значимость приобретает системность подхода автора тезауруса к классификации и рубрикации лексики: в предисловии к словарю П.М. Роже прямо указывает на гомологичность собственных методов и принципов, используемых при классификации «особей в различных областях естественной истории». По мнению П.М. Роже, убедительная классификация слов по их смыслам невозможна до тех пор, пока должным образом не изучены и не организованы объекты действительности, называемые этими словами [цит. по: Морковкин: 41]. Идеи научной рубрикации были положены в основу синоптической карты словаря, соответствующим (иерархическим) образом лексика организовывалась и в корпусе тезауруса (классы, категории, секции, рубрики). Таким образом, тезаурус П.М. Роже по праву считается первой научно обоснованной попыткой создания некоторого макета логически упорядоченной лексики языка.

Развитие информационных технологий позволило реализовать идеи логической рубрикации понятий в информационно-поисковых системах. Постепенно сформировалось устойчивое понятие тезауруса как инструмента информационного поиска, дескрипторной статьи как основной структурной единицы тезауруса, дескриптора как единицы, разрешенной к использованию при поиске информационных документов [Гетман: 7—8].

Последняя четверть XX и XXI век стали свидетелями становления специальной (терминологической) идеографии. *Thesaurus of Psychological Index Terms*, 1974; *Multilingual Thesaurus of Geosciences*, 1988; *Thesaurus for Graphic Materials*, 1995; *Unified Astronomy Thesaurus* на *astrothesaurus.org* и др. — ничтожно малая часть примеров специальных тезаурусов, созданных за это время. Отраслевые идеографические словари, а затем и справочные ресурсы, и банки данных электронного формата появились в ответ на запрос научных и производственных сообществ и информационных служб. Большой опыт, накопленный научными школами в типологизации терминов и описании таксонов, относительная предельность лексического (терминологического) фонда, подлежащего включению в тезаурус, и достижения специалистов в области создания инструментов информационного поиска способствовали интенсивному развитию терминологической идеографии для многих отраслей знания.

В это же время выходит достаточно большое количество тезаурусов для общих целей — идеографических, ассоциативных, аналогических (с алфавитной организацией слов-наименований рубрик) — в печатной (*The New American Roget's College Thesaurus*, 1958; *The Structure of Associations in Language and Thought*, 1965; *Collins Paperback Thesaurus*, 1986; *Longman Language Activator*, 1993; *Oxford Learner's Wordfinder Dictionary*, 1997 и др.), а затем и электронной форме (*Collins Desktop Wordfinder English Thesaurus*

on CD-ROM, 1999; *Roget's Thesaurus V (0.59)*, thesaurus.com и др.). Электронные тезаурусы встраиваются в структуру гибридных или комплексных лексикографических ресурсов. Это чаще всего означает, что тезаурусный фасет «вложен» в толковый словарь (*Cambridge Dictionaries, Collins Dictionary and Thesaurus, Longman Dictionary of Contemporary English, Macmillan Dictionary, Oxford Dictionary, Thesaurus and Translator on Lexico.com, Merriam-Webster's Online Dictionary and Thesaurus* и др.). Интернет-пространство изобилует и тезаурусами, функционирующими автономно или как часть словарных агрегаторов: *Thesaurus.com, Freethesaurus.com, Powerthesaurus.org, Thesaurus.plus* и др. [Воронцова 2022b: 189]. Появляется сетевой тезаурус *Princeton WordNet*, по его подобию создаются и другие ресурсы (*Verbnet, Imagenet*).

XXI столетие — это эпоха цифровизации и интеллектуализации словарей. Электронно-цифровые технологии, изменившие саму модель лексикографического процесса [Воронцова 2022b: 188], оказали сильнейшее влияние и на развитие тезаурусной лексикографии. Потенциал лексикографической формы тезауруса как «онтологической сетевой базы знаний, содержащей лексическую основу для всеобщей модели мироздания» [Пацкин] был раскрыт в новом научно-прикладном направлении, получившем название «Инженерия знаний». Существенную, если не определяющую, роль здесь сыграли развитие методов корпусной лингвистики, совершенствование систем корпусного менеджмента, а также разработки в области интеллектуального анализа текста (*text mining*). Тезаурусное моделирование, основанное на методах интеллектуальной обработки больших данных, позволило учесть и отразить многообразие связей и отношений, в которые идеи, облеченные в словесную форму, вступают друг с другом в экстралингвистической действительности и воссоздающем ее множестве текстов.

Компетентный корпус текстов становится источником и «регулятором» тезауруса. Корпусно-ориентированный подход к построению тезаурусов и онтологий предложен в одноименной статье В.П. Захарова. Этапы конструирования тезауруса включают экстракцию лексики (терминов) на основе дистрибутивного анализа, кластеризацию отобранных лексем в микрополя на основании значений статистической меры, показывающей степень семантической близости слов к ключевому, насыщение тезауруса коллокациями и коллигациями (лексико-синтаксическими шаблонами). Пользователю предоставляется возможность перехода от каждого слова-кластера ко всем контекстам его употребления в корпусе и уточнения релевантности его нахождения в тезаурусе [Захаров: 7]. На наш взгляд, принцип, предложенный и реализованный В.П. Захаровым с помощью системы *Sketch Engine*, демонстрирует новый канон конструирования тезаурусов (прежде всего, специальных или отраслевых), позволяющий ускорить разработку тезауруса, увеличить глубину и степень детализации представления языкового материала, выявить в совокупности связи, устанавливаемые между членами кластера и полем (синонимические, гипонимические, гиперонимические, ассоциативные и др.), повысить объективность представляемых данных.

В начале XXI в. концепция тезауруса соотносится с концепцией онтологии — артефакта, структуры, описывающей значения элементов некоторой системы [Соловьев В.Д., Добров Б.В.], преимущественно специального знания. При всей неоднозначности и разноречивости подходов к определению и разграничению понятий «тезаурус» и «онтология», представляется продуктивным взгляд на тезаурус как на лексическую систему в проекции

на семантику, а на онтологию — как систему семантики и прагматики, до известной степени в проекции на язык. Иными словами, тезаурус — это система, организующая лингвистическую информацию о соответствующей проекции естественного языка на данную предметную область, онтология — система знаний об этой предметной области, в идеале — ее формальная модель (МПО, модель предметной области). По мнению А.С. Нариньяни, ядром интеллектуальных систем с активным использованием компьютерной лингвистики становится тандем «тезаурус» плюс «онтология», высокая интегрированность которого позволяет считать его системой нового типа. Для ее обозначения автор вводит термин «ТЕОН» [Нариньяни]. Вслед за А.С. Нариньяни, В.Э. Клюкин, добавляя к существующему термину усеченную основу термина «глоссарий», вводит понятие и описывает инструмент «ГЛОТЕОН», где тезаурус отвечает за состав лексики, описывая родовидовые отношения с помощью дескрипторных групп, онтология содержит представление знаний об описании предметной области на естественном языке со ссылкой на факты, представленные в глоссарии и тезаурусе, а глоссарий предназначен для описания лексем и семантических сущностей [Клюкин: 27].

На современном этапе появилось (и продолжает создаваться) множество специальных онтологий, описывающих некоторую область знаний или систематизирующих лексику определенного (часто специализированного) корпуса текстов. Онтологии строятся и «поверх» существующих тезаурусов. В работе “*From Thesaurus to Ontology*” авторы описывают опыт создания онтологии на базе «Тезауруса искусства и архитектуры» (*Art and Architecture Thesaurus*) посредством дополнительной иерархизации понятий (фреймирование с выделением дополнительных слотов), установления связей между понятиями, формулировки дескрипторов, насыщения статей дополнительной экстралингвистической («знаниевой») информацией, загрузки визуальных рядов (фото, иллюстраций и др.) [Wielinga et al.]. Рассуждая о рациональных подходах к созданию онтологий на базе существующих тезаурусов, А.Ю. Филиппович упоминает о таких проектах, как *OntoWordNet*, *MikroKosmos* и *OntoSem*, участники которых высказывают мысль о том, что построение онтологии на базе существующего тезауруса — это не просто «склейка» ресурсов, но их реструктуризация и переработка, баланс систем с четко простроенными взаимосвязями [Филиппович].

Тезаурус рассматривается и как методологическая основа построения дву- и многоязычных онтологий предметных областей. Е.Е. Котова и И.А. Писарев описывают онтологическую модель предметной области «Онтологический инжиниринг» для русского и английского языков. Конструирование модели включает следующие шаги: анализ запросов пользователей, формирование корпуса текстов, определение формата тезауруса терминологической структуры предметной области, формирование частотного словаря, удаление из словаря незначимых слов, пополнение тезауруса из машинно-читаемых словарей (*WordNet*, *Oxford Dictionary*, *Wikipedia*), формализация классификационной структуры, рецензирование прототипа онтологии, расширение корпуса текстов, эксплуатация онтологии [Котова, Писарев]. Содержательно онтология включает термин (лемму), дефиницию, гипероним и синонимы к ключевым терминам.

Несмотря на то, что концепция онтологии чаще ассоциируется со специальной предметной областью, разрабатываются и онтологии для общих целей. К таковым, на наш взгляд, можно отнести онлайн-ресурс *BabelNet*, созданный на базе *Princeton Wordnet* и отмеченный в нем в разделе *Related Projects*.

*BabelNet* интегрирует в единую многоязычную сеть информацию из различных источников на разных языках (более 50 источников, более 500 языков) [Navigli et al.; Qi et al.; Al-Sayed, Hassan...]. Ресурс фактически дает комплексное лингво-энциклопедическое описание апеллятивной лексики и именованных сущностей (апеллятивов и онимов), организуя их в синсеты (наборы синонимической лексики и смысловых коррелятов в разных языках), агрегируя знания о словах и сущностях, полученных из таких ресурсов, как *WordNet*, *Wikipedia*, *Wikidata*, *Wiktionary*, *OmegaWiki* и др. [Воронцова 2022а: 149].

Визитной карточкой онтологий является особый способ визуализации информации с помощью ментальных (концептуальных) карт или концептуальных графов. Так, ментальная карта в *BabelNet* визуализирует синсет во всем многообразии типов смысловых отношений, устанавливаемых между членами синсета, лежащих как в лингвистических, так и в экстралингвистических плоскостях [Воронцова 2022а: 149]. Отметим, что ментальные карты могут быть доступны в качестве дополнительного сервиса в идеографических тезаурусах, например, в *Roget's Hyperlinked Thesaurus*, или служить основой макроструктуры тезауруса — *Thinkmap Visual Thesaurus*.

Таким образом, за одиннадцать столетий лексикографическая форма тезауруса эволюционировала от интуитивных тематических компилятов до сложного инструмента онтологической инженерии. Опираясь на системные научные знания о мире и языке, тезаурус сегодня выходит за пределы словарной формы, трансформируясь в фундамент и каркас модели отдельных отраслей знания и научной, и наивной картин мира в целом. Тезаурус является источником, результатом и методологией работы с языковыми и внеязыковыми данными.

#### Список литературы / References

- Воронцова И.А. Историко-типологическое исследование тезаурусов английского языка: дис. ... канд. филол. наук. Иваново, 2004. 276 с.  
(Vorontsova I.A. Historical and typological study of thesauruses of the English language: diss. ... Cand. of Sciences (Philology), Ivanovo, 2004, 276 p. — In Russ.)
- Воронцова И.А. Актуальные подходы к конструированию ономазиологических словарей (на материале английского языка) // Верхневолжский филологический вестник. 2022а. № 3 (30). С. 146—154.  
(Vorontsova I.A. Modern approaches to onomasiological dictionary construction in the english language, *Verhnevolzhski philological bulletin*, 2022, no. 3 (30), pp. 146—154. — In Russ.)
- Воронцова И.А. Современные практики англоязычной онлайн-лексикографии // Верхневолжский филологический вестник. 2022b. № 1 (28). С. 186—194.  
(Vorontsova I.A. Modern practices of English online lexicography, *Verhnevolzhski philological bulletin*, 2022, no. 1 (28), pp. 186—194. — In Russ.)
- Гетман И.М. Тезаурус как инструмент современного языкознания (на материале русистики): дис. д-ра. филол. наук. Киев, 1991. 365 с.  
(Getman I.M. Thesaurus as a tool of modern linguistics (on the material of Russian studies): diss. Doctor of Sciences (Philology), Kyiv, 1991, 365 p. — In Russ.)
- Захаров В.П. Корпусно-ориентированный подход к построению тезаурусов и онтологий // Структурная и прикладная лингвистика. 2015. № 11. С. 123—141.  
(Zakharov V.P. Corpus-based approach to thesaurus and ontology construction, *Structural and applied linguistics*, 2015, no. 11, pp. 123—141. — In Russ.)

- Котова Е.Е., Писарев И.А. Построение тематических онтологий с применением метода автоматизированной разработки тезаурусов // Известия СПбГЭТУ «ЛЭТИ». 2016. № 3. С. 37—47.
- (Kotova E.E., Pisarev I.A. Construction of thematic ontologies using the method of automated thesauri development, *Proceedings of Saint Petersburg Electrotechnical University*, 2016, no. 3, pp. 37—47. — In Russ.)
- Клюкин В.Э. Композиция ГЛОТЕОН: глоссарий + тезаурус + онтология // Машиностроение и компьютерные технологии. 2011. № 9. С. 22—31.
- (Klyukin V.E. GLOTEON composition: glossary + thesaurus + ontology, *Engineering and Computer Technology*, 2011, no. 9, pp. 22—31. — In Russ.)
- Мельник О.А. Принципы построения синонимических словарей английского языка: автореф. дис. ... канд. филол. наук. Л., 1954. 15 с.
- (Melnik O.A. Principles of construction of English synonymic dictionaries: abstract of diss. ... Cand. of Sciences (Philology), Leningrad, 1954, 15 p. — In Russ.)
- Морковкин В.В. Идеографические словари. М.: МГУ, 1970. 69 с.
- (Morkovkin V.V. Ideographic dictionaries, Moscow, 1970, 69 p. — In Russ.)
- Нариньяни А.С. Кентавр по имени ТЕОН: тезаурус + онтология // Труды международного семинара ДИАЛОГ'2001. М., 2001. Т. 1: Теоретические проблемы. Аксаково. С. 184.
- (Narinyani A.S. Centaur named THEON: thesaurus + ontology, *Proceedings of the International Seminar Dialogue '2001*, Moscow, 2001, vol. 1: Theoretical problems. Aksakovo, p. 184. — In Russ.)
- Осокина С.А. Сетевая модель языкового тезауруса: особенности построения // Сибирский филологический журнал. 2016. № 3. С. 191—198.
- (Osokina S.A. The network model of the language thesaurus: construction specifics, *Siberian journal of philology*, 2016, no. 3, pp. 191—198. — In Russ.)
- Пацкин А.И. Гиперсловари на базе системы «Абриаль» // Труды международного семинара ДИАЛОГ'2002. М., 2002. URL: <http://patskin.ru/artic/hyperdic.htm> (дата обращения: 14.08.2023).
- (Patskin A.I. Hyperwords based on the Abrial system, *Proceedings of the International Seminar Dialogue '2002*, Moscow, 2002. — In Russ.)
- Соловьев В.Д., Добров Б.В., Иванов В.В., Лукашевич Н.В. Онтологии и тезаурусы: учебное пособие. Казань; Москва, 2006. 157 с.
- (Solovyev V.D., Dobrov B.V., Ivanov V.V. Lukashevich N.V. Ontologies and thesauri: tutorial, Kazan, Moscow, 2006. 157p. — In Russ.)
- Филиппович А.Ю. Тезаурус vs. Онтология. URL: <http://blogs.it-claim.ru/andrey/2011/01/18/тезаурус-vs-онтология/> (дата обращения: 14.08.2023)
- (Filippovich A.Y. Thesaurus vs. Ontology. — In Russ.)
- Al-Sayed M.M., Hassan H.A., Omara F.A. Mapping lexical gaps in cloud ontology using Babelnet and FP-growth, *International Journal of Computer Science & Security (IJCSS)*, 2019, vol. 13, no. 2, pp. 36—52.
- Burkett E. American Dictionaries before 1861, The Scarecrow Press Inc. Metuchen. N.J. and London, 1979, 312 p.
- Navigli R. etc. Ten years of BabelNet: A survey, *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 2021, pp. 4559—4567.
- Qi F. et al. Towards building a multilingual sememe knowledge base: predicting sememes for BabelNet synsets, *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, vol. 34, no. 5, pp. 8624—8631.

Starnes W., Noyes G. The English Dictionary from Cawdrey to Johnson 1604—1755, Chapel Hill, 1946, 433 p.

Wielinga B.J. et al. From thesaurus to ontology, *Proceedings of the 1st international conference on Knowledge capture*, 2001, pp. 194—201.

Wright T. Anglo-Saxon and Old-English Vocabularies, vols. 1, 2, Darmstadt, 1968, 434 p.

*Статья поступила в редакцию 24.04.2023; одобрена после рецензирования 17.05.2023; принята к публикации 09.06.2023.*

*The article was submitted 24.04.2023; approved after reviewing 17.05.2023; accepted for publication 09.06.2023.*

#### ***Информация об авторе / Information about the author***

**Воронцова Инна Алексеевна** — кандидат филологических наук, доцент кафедры теории и практики перевода, Ярославский государственный педагогический университет им. К.Д. Ушинского, г. Ярославль, Россия, arinna1@yandex.ru

**Vorontsova Inna Alekseevna** — Candidate of Sciences (Philology), Associate Professor of the Department of Theory and Practice of Translation, Yaroslavl State Pedagogical University named after K.D. Ushinsky, Yaroslavl, Russian Federation, arinna1@yandex.ru