

ПОЛИСЕМИЯ В ТЕРМИНОЛОГИИ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ

Нематова Нилуфар Гайратуллаевна

Аннотация. Цель исследования - раскрыть особенности процесса полисемии в терминологии компьютерной лингвистики на основе проведения его комплексного анализа. В статье изучены причины терминологической многозначности, показано, что полисеманты могут иметь и очень похожие по смыслу значения, и слабо связанные друг с другом дефиниции. Научная новизна работы состоит в том, что впервые составлена классификация функционирующих в терминологии компьютерной лингвистики полисемантов по признакам эксплицитности (имплицитности) и типам категориальной многозначности. В результате исследования показана распространенность различных типов многозначных терминов компьютерной лингвистики.

POLYSEMY IN COMPUTER LINGUISTICS TERMINOLOGY

Нематова Нилуфар Гайратуллаевна

Abstract. The aim of the study is to reveal the features of the polysemy process in the terminology of computer linguistics based on its comprehensive analysis. The article examines the causes of terminological polysemy, shows that polysemants can have both very similar meanings and definitions that are weakly related to each other. The scientific novelty of the work lies in the fact that for the first time a classification of the polysemants functioning in the terminology of computer linguistics has been compiled according to the features of explicitness (implicitness) and categorial polysemy types. As a result of the study, the occurrence of various types of polysemantic terms in computer linguistics is shown.

Ключевые слова (keywords): компьютерная лингвистика; терминология; многозначность; полисеманты; бисеманты; computer linguistics; terminology; polysemy; polysemants; bisemants.

Введение

Терминология компьютерной лингвистики находится в состоянии непрерывной динамики, что обусловлено стремительным развитием новых технологий, совершенствованием методов, моделей и

программных средств в области автоматической обработки естественного языка (англ. Natural Language Processing, NLP) (Агузумцян, Великанова, Польщиков, 2021; Польщиков, Лазарев, Константинов, 2020; Alzayed, Al-Hunaiyyan, 2021; Goldberg, 2017; Farzindar, Inkpen, 2015; Polshchikov, Lazarev, Konstantinov, 2020; Polshchikov, Lazarev, Polshchikova et al., 2019; Velikanova, Polshchikov, Likhosherstov et al., 2021). В результате этих процессов термины компьютерной лингвистики возникают в некоторой степени стихийно. Отсутствие автохтонных лексем для описания новых понятий в области NLP приводит к широкому использованию в ней заимствованных терминов и переводных эквивалентов.

Потребность во взаимопонимании в профессиональном общении требует рационального регулирования терминологии компьютерной лингвистики, осуществления целенаправленных действий по упорядочению, стандартизации соответствующего понятийного аппарата, гармонизации и унификации как на национальном, так и на международном уровнях. Все это невозможно без глубокого анализа лексических особенностей терминологии компьютерной лингвистики, в частности наличия в ней полисемии.

Явление многозначности изучалось в различных работах, полисемические специальные названия выявлены многими исследователями отраслевых терминологий. Рассмотрена полисемия терминов экономического подязыка (Аверина, 2015), исследована многозначность терминов медицины (Гончарук, Гущина, 2009; Вольфберг, 2017), изучена полисемия в технической терминологии (Чистюхина, 2020). Проведен анализ особенностей полисемии в терминологии нанотехнологий (Иванова, 2010), правоохранительной деятельности (Галкина, 2015), юридической сферы (Кожанов, 2014), банковского дела (Шетле, 2009), трубопроводного транспорта (Горохова, 2015), сельского хозяйства (Щербинина, 2020). Однако вопросы функционирования полисемии в терминологии компьютерной лингвистики на сегодняшний момент освещены недостаточно. Актуальность изучения этого языкового явления обусловлена растущим интересом к терминологии компьютерной лингвистики в условиях динамичного развития современных средств автоматической обработки естественно-языковой информации, сопровождаемого образованием новых отраслевых понятий и терминов.

Для достижения поставленной цели исследования в статье решены следующие задачи: 1) выявить причины возникновения

терминологической полисемии; 2) представить классификацию многозначных терминов компьютерной лингвистики; 3) показать типы категориальной многозначности в терминологии компьютерной лингвистики.

Исследование явления полисемии в терминологии компьютерной лингвистики выполнено на основе применения методов описания, сравнения, дефинитивного анализа. Теоретической базой послужили научные труды, посвященные изучению особенностей специальной лексики (Головин, Кобрин, 1987; Даниленко, 1977; Лотте, 1961; Суперанская, Подольская, Васильева, 2012).

Практическая значимость исследования состоит в возможности применения его материалов и результатов при подготовке и проведении учебных занятий по терминоведению, а также в лексикографической деятельности.

Основная часть

Полисемия считается негативным явлением с точки зрения однозначности трактовки специальных понятий. Под полисемией (многозначностью) понимается «наличие у единицы языка более чем одного значения». Многозначные слова (бисеманты, полисеманты) «служат для обозначения различных предметов и явлений действительности». Наличие определенной семантической связи между значениями многозначного слова «даёт основание считать их значениями одного и того же слова» (РЯЭ, 1997, с. 352).

Понять, в каком значении употребляется полисемант, помогает контекст. Как отметила В. П. Даниленко (1977), «в специальной речи термин всегда находится в контекстном окружении, которое конкретно реализует одно из возможных значений термина. Имеются в виду весьма характерные для терминологической лексики случаи категориальной многозначности, преодоление которой обеспечивается стандартным контекстом» (с. 60). При этом Д. С. Лотте (1961) утверждал, что значение термина «не может зависеть от того предложения, в котором он употреблен, а должно определяться лишь всей системой понятий и, соответственно, терминологией данной дисциплины, области знания» (с. 74).

По мнению Б. Н. Головина и Р. Ю. Кобрина (1987), в гуманитарных, естественных и технических науках имеется значительное количество полисемических терминов. Исследователи, характеризуя полисемию как негативное явление в терминологии, отметили, что многозначность является «одним из факторов, затрудняющих коммуникацию» (с. 48).

Основной причиной возникновения терминологической многозначности является недостаточный объем корневого словарного материала по сравнению с количеством научно-технических понятий, поэтому для обозначения специальных понятий используются общеупотребительные слова. Термины-полисеманты, «восходящие к общеупотребительным словам, конкретизируют свое значение в контексте» (Головин, Кобрин, 1987, с. 49). К причинам возникновения полисемии также относится «обозначение уже существующим в науке термином новых объектов и понятий – в зависимости от авторской исследовательской позиции, принадлежности к научной школе» (Головин, Кобрин, 1987, с. 51). Кроме того, полисемия может быть обусловлена некритическим заимствованием терминов из иностранных языков.

Несмотря на определенное негативное отношение к полисемии в терминосистемах, ее выделяют как объективное явление, при котором значения слов взаимосвязаны и взаимообусловлены. Зачастую «недостаточное количество отличительных признаков не позволяет присваивать двум похожим предметам разные имена» (Суперанская, Подольская, Васильева, 2012, с. 44), поэтому употребление многозначных терминов способствует экономии языковых средств. В профессиональной речи, как известно, важно стремиться к однозначности термина, однако на практике избавиться от полисемии не удастся. Это обстоятельство подтверждает вывод о том, что «однозначность – не свойство термина, а требование, к нему предъявляемое» (Суперанская, Подольская, Васильева, 2012, с. 44).

В пределах терминологии компьютерной лингвистики можно выделить эксплицитную (межсистемную) и имплицитную (внутрисистемную) полисемию. Характерным примером эксплицитных полисемантов выступает термин вес: «Числовое значение, отражающее важность данного термина или связи между двумя понятиями для некоторого текста» (терминология компьютерной лингвистики) (РАТКЛ, 2022); «Сила тяжести» (физическая терминология); «Значение, влияние, авторитет», «Совокупность лиц одной специальности, какого-либо одного официального или служебного положения» (общелитературный язык) (БТСРЯ, 2000, с. 121).

К числу эксплицитных полисемантов относится термин сценарий: «Система связанных между собой фреймов, применяемая для описания сложного события или динамически развивающейся ситуации» (терминология компьютерной лингвистики) (Николаев,

Митренина, Ландо, 2016, с. 66); «Литературно- драматическое произведение, содержащее подробное описание действия и текст речей персонажей, на основе которого создаётся фильм» (кинематографическая терминология); «Сюжетная схема, план театральной пьесы, оперы, балета» (театральная терминология); «Заранее подготовленный детальный план, программа проведения какого-либо зрелища, мероприятия» (общеупотребительная лексика) (БТСРЯ, 2000, с. 1296). Лексема конверсия также имеет несколько значений в разных терминологиях: «Переход слова из одной части речи в другую» (терминология компьютерной лингвистики) (Николаев, Митренина, Ландо, 2016, с. 199); «Изменение условий ранее выпущенного государственного займа, например, понижение процента, изменение срока погашения» (финансовая терминология); «Свёртывание мощностей оборонной промышленности и переориентация её на производство товаров народного потребления» (экономическая терминология) (БТСРЯ, 2000, с. 447).

В терминологии компьютерной лингвистики нередко встречаются эксплицитные бисеманты. Два значения имеет термин валентность: «Способность слова сочетаться в тексте с другой языковой единицей, прежде всего с другим словом» (терминология компьютерной лингвистики) (РАТКЛ, 2022) и «Способность атома соединяться с определённым числом других атомов» (химическая терминология) (БТСРЯ, 2000, с. 110). Термин нормализация также имеет две дефиниции: «Процесс постановки слова и словосочетания в каноническую форму» (терминология компьютерной лингвистики) (Николаев, Митренина, Ландо, 2016, с. 14); «Приближение к норме» (общелитературный язык) (БТСРЯ, 2000, с. 656).

В ряде случаев эксплицитная и имплицитная полисемии очень тесно граничат, как будто частично «накладываются» друг на друга. Одним из примеров подобного явления является термин шаблон, имеющий несколько дефиниций: «Описание языковой конструкции, употребляемой для отражения в тексте искомого факта или объекта» (терминология компьютерной лингвистики); «Набор различных атрибутов текста, присущих целевой ситуации» (терминология компьютерной лингвистики) (РАТКЛ, 2022); «Документ, используемый в качестве образца для создания новых документов» (компьютерная терминология) (Дорот, Новиков, 2004, с. 512); «Приспособление, инструмент для проверки форм готовых изделий» (техническая терминология);

«Принятый образец, которому следуют без размышлений; трафарет, штамп» (общеупотребительная лексика) (БТСРЯ, 2000, с. 1488).

Лексема газеттир в отрасли компьютерной лингвистики обозначает, во-первых, «список специфических для решаемой задачи извлечения именованных сущностей лексем и словосочетаний», во-вторых, «программный ресурс, задачей которого является распознавание во входном тексте слов и словосочетаний, релевантных для заданной предметной области» (РАТКЛ, 2022). В географической терминологии одноименная лексема обозначает «текстовый список, связывающий между собой топонимы географических объектов и их локализацию, т.е. координаты, листы карты» (Лунева, 2002).

Полисеманты могут иметь очень похожие по смыслу значения. Так, термин индекс имеет несколько дефиниций: «Формальное описание документа на информационно-поисковом языке» и «Множество признаков, ключевых слов, тематических элементов, терминов и термов, которые отражают основную тематику текста» (терминология компьютерной лингвистики) (РАТКЛ, 2022); «Число целого типа или арифметическое выражение, принимающее целочисленное значение, приписываемое элементу массива или другой конструкции данных для идентификации этого документа» (компьютерная терминология) (Дорот, Новиков, 2004, с. 207);

«Цифровой показатель, выражающий в процентах последовательные изменения какого-либо экономическое явление» (экономическая терминология); «Числовой или буквенный указатель, помещаемый чаще всего внизу буквы, входящей в математическое выражение» (математическая терминология); «Указатель, список, перечень чего-либо» и «Система условных обозначений (буквенных, цифровых или комбинированных) в системе какой-либо классификации» (общеупотребительная лексика) (БТСРЯ, 2000, с. 391). Ряд похожих значений выявлен у полисеманта разметка: «Процесс приписывания текстам и их компонентам специальных меток» (терминология компьютерной лингвистики) (РАТКЛ, 2022); «Параметры размещения текстового или иллюстративного материала на странице, в число которых входят, например, размеры полей страницы, расстояния между колоннотитулами и основным текстом» (компьютерная терминология); «Режим отображения страницы на экране компьютера, при котором видны все необходимые параметры» (компьютерная терминология) (Дорот, Новиков, 2004, с. 409); «Перенесение на заготовку точек и линий с чертежа для обозначения

раз- мера и мест обработки» (чертежная терминология); «Знак, метка, поставленные с целью обозначения чего- либо, указания на что-либо» (общеупотребительная лексика) (БТСРЯ, 2000, с. 1073).

В исследуемой терминологии часто встречаются полисемические термины, определения которых связаны между собой очень слабо, однако определенное сходство обозначаемых ими понятий все же имеется. Явление такого едва заметного сходства значений можно объяснить особенностями процессов создания терминов, происходящих при поиске удачных лексем для обозначения новых понятий компьютерной лингвистики.

Отрасль автоматической обработки естественного языка в современном обществе является одной из наиболее динамично развивающихся. Это приводит к интенсивному появлению новых понятий, требующих идентификации посредством терминов. Объекты и их свойства, явления и процессы, имеющие непосредственное отношение к тематике компьютерной лингвистики, достаточно специфичны, поэтому для их именованья иногда слишком сложно подобрать функционально подобные общеупотребительные слова или термины других отраслей. Так, термин сцена в сфере компьютерной лингвистики – это «структура памяти, организующая действия и состояния, связанные единой целью и единой обстановкой, заданной в пространстве» (РАТКЛ, 2022). Одноименная лексема имеет другие значения в театральной терминологии: «Специальная площадка, на которой даются представления»; в литературной, художественной и кинематографической терминологиях: «Отдельная часть действия, эпизод в пьесе, литературном произведении, картине»; в общеупотребительной лексике:

«Происшествие, ссора, крупный разговор, объяснение» (БТСРЯ, 2000, с. 1296). В, казалось бы, в совсем разных дефинициях можно найти общие первичные значения – «структура, специально организованное пространство» либо «часть какого-то действия». Термин выброс с общими компонентами значения «отличающийся элемент» представлен следующими дефинициями: «Объект, не характерный для данного класса или находящийся в окружении другого класса» (терминология компьютерной лингвистики) (Николаев, Митренина,

Ландо, 2016, с. 129); «Поступление в окружающую среду загрязняющих веществ от промышленных или сельскохозяйственных

предприятий» (экологическая терминология) (ЭС, 2022); «Всплеск, короткий импульс» (техническая терминология) (СТП, 2022).

Значительно реже в терминологии компьютерной лингвистики используются имплицитные бисеманты. Так, термин форманта имеет следующие значения: «Акустическая характеристика звука речи (преимущественно гласного), связанная с частотой голосового тона и образующая тембр звука» (БТСРЯ, 2000, с. 1429); «Область концентрации энергии в спектре звуков речи» (РАТКЛ, 2022).

Широкое распространение в терминологии получила категориальная многозначность, «состоящая в том, что содержание понятия складывается из признаков, принадлежащих одновременно нескольким категориям» (Даниленко, 1977, с. 68-69). В терминологии компьютерной лингвистики выявлено несколько основных типов категориальной многозначности: действие и результат действия (высказывание – выражение мысли и законченный по смыслу речевой отрезок), свойство и величина (валентность – способность слова сочетаться с другими элементами и число присоединяемых к глаголу актантов), процесс и величина (затухание – процесс уменьшения амплитуды звука и величина потери звуковой энергии).

Заключение

Выполненное исследование позволило сделать следующие выводы:

1. Возникновение многозначных терминов обусловлено недостаточным объемом имеющегося корневого материала для обозначения специальных понятий, наличием категориальной многозначности, чрезмерным заимствованием терминов.

2. Многозначные термины компьютерной лингвистики могут быть классифицированы по признакам эксплицитности (имплицитности) и числу обозначаемых значений. Широко распространены эксплицитные полисеманты и бисеманты, редко встречаются имплицитные бисеманты.

3. К основным типам категориальной многозначности, выявленным в терминологии компьютерной лингвистики, относятся «действие – результат», «свойство – величина», «процесс – величина».

Перспективы дальнейших исследований состоят в изучении омонимических отношений в терминологии компьютерной лингвистики и составлении тематических групп специальных названий этой предметной области.

ИСТОЧНИКИ | REFERENCES:

1. Аверина М. А. К вопросу о полисемии англицизмов в русской экономической терминологии // Актуальные проблемы гуманитарных и естественных наук. 2015. № 1-1.
2. Агузумцян Р. В., Великанова А. С., Польшчиков К. А. О применении интеллектуальных технологий обработки естественного языка и средств виртуальной реальности для поддержки принятия решений при подборе исполнителей проектов // Экономика. Информатика. 2021. Т. 48. № 2.
3. Большой толковый словарь русского языка (БТСРЯ) / сост. и гл. ред. И. А. Кузнецов. СПб.: Норинт, 2000.
4. Вольфберг Д. М. Полисемия и омонимия в английской медицинской терминологии // Филологические науки. Вопросы теории и практики. 2017. № 3-2 (69).
5. Галкина В. В. Особенности полисемии в английской терминологии правоохранительной деятельности // Омский научный вестник. 2015. № 4 (141).
6. Головин Б. Н., Кобрин Р. Ю. Лингвистические основы учения о терминах. М.: Высшая школа, 1987.
7. Гончарук В. В., Гущина Л. Н. Пути развития полисемии и ее роль в медицинской терминологии // Журнал Гродненского государственного медицинского университета. 2009. № 3 (27).
8. Горохова Н. В. Роль полисемии в специальной терминологии (на материале англоязычной терминологии трубопроводного транспорта) // Филологические науки. Вопросы теории и практики. 2015. № 7-1 (49).
9. Даниленко В. П. Русская терминология. Опыт лингвистического описания. М.: Наука, 1977.
10. Дорот В. Л., Новиков Ф. А. Толковый словарь современной компьютерной лексики. СПб.: ПХВ-Петербург, 2004.
11. Иванова О. Б. Полисемия в английской терминологии нанотехнологии // Вестник Московского государственного областного университета. Серия «Лингвистика». 2010. № 3.
12. Кожанов А. А. Неоднозначность: полисемия и омонимия в юридической терминологии (на материале немецкого языка) // Вестник Брянского государственного университета. 2014. № 2.
13. Лотте Д. С. Основы построения научно-технической терминологии: вопросы теории и методики. М.: Изд-во АН СССР, 1961.

14. Лунева Н. В. Традиционные газеттеры и задачи создания электронных геобиблиотек. 2002. URL: <https://www.dialog-21.ru/digest/2002/articles/luneva>
15. Николаев И. С., Митренина О. В., Ландо Т. М. Прикладная и компьютерная лингвистика. М.: Ленанд, 2016.
16. Польщиков К. А., Лазарев С. А., Константинов И. С. Модель для оценки эффективности выполнения робототехнической системой коммуникативных функций // Станки Инструмент. 2020. № 6.
17. Русский язык. Энциклопедия (РЯЭ) / гл. ред. Ю. Н. Караулов. М.: Дрофа, 1997.
18. Русско-английский тезаурус по компьютерной лингвистике (ПАТКА). 2022. URL: <https://uniserv.iis.nsk.su/thes/search.php>