

RU

Гиперо-гипонимические отношения в терминологии компьютерной лингвистики

Польщикова О. Н.

Аннотация. Цель исследования - раскрыть особенности отношений «род - вид» в терминологии компьютерной лингвистики на основе проведения их комплексного анализа. В статье изучена структурная организация терминов компьютерной лингвистики на основе процессов гиперо-гипонимии, показаны связи между терминами в рамках привативных и эквивалентных оппозиций. Научная новизна работы состоит в том, что впервые составлена классификация родо-видовых отношений в терминологии компьютерной лингвистики, сформированных по различным признакам. В результате исследования выявлены типы иерархических структур, отражающие закономерности гиперо-гипонимических связей между элементами терминосистемы компьютерной лингвистики.

EN

Hyper-Hyponymic Relations in Computer Linguistics Terminology

Polshchikova O. N.

Abstract. The aim of the study is to reveal the features of gender-aspect relations in the terminology of computer linguistics on the basis of their comprehensive analysis. The article researches the structural organization of computer linguistics terms based on the processes of hyper-hyponymy, shows the relationship between the terms within the framework of privative and equivalent oppositions. The scientific novelty of the work lies in the fact that a classification of gender-aspect relations in the terminology of computer linguistics formed according to various characteristics has been compiled for the first time. As a result of the study, the types of hierarchical structures have been identified representing the patterns of hyper-hyponymic relationships between the elements of the terminological system of computer linguistics.

Введение

Гиперо-гипонимические отношения причисляются к самым важным парадигматическим отношениям с точки зрения семантики (Кронгауз, 2005). Они «существуют между родовым, широким по объему понятием и входящими в него более узкими, видовыми понятиями» (Гринев-Гриневиц, 2008, с. 82), определяются «между двумя дескрипторами, когда объем понятия нижестоящего дескриптора укладывается в объем понятия вышестоящего дескриптора» (Доржиева, 2018, с. 102), характеризуются «привативными оппозициями единиц, включенной дистрибуцией, совместимыми подчиненными понятиями» на основе «логико-семантической субординации» (Новиков, 1982, с. 241). При этом «слова-гипонимы выражают видовые понятия, а слова-гиперонимы – родовые» (Головин, Кобрин, 1987, с. 59). Главные функции гипонимов, к которым относятся «систематизация терминов и толкование значений», реализуются «путем обобщения и конкретизации посредством признака видового отличия» (Коновалова, 2001, с. 51). Отношения «род – вид» представляются «значимыми для научно-профессионального знания, поскольку с их помощью выстраивается иерархия специальных понятий» (Асадова, Акаева, 2020, с. 22). Посредством данного типа отношений «может быть структурирован как словарный состав языка в целом, так и его специальная лексика» (Коновалова, 2001, с. 51).

Отношения гиперо-гипонимии проявляются «в научно-технических текстах и приводят к появлению в языковой системе полисемичных терминов, терминов-синонимов, терминов-антонимов» (Головин, Кобрин, 1987, с. 60). В выражении родо-видовых связей «проявляется системный характер терминологии», без их выявления невозможно «описание системной организации терминологии» и проведение «лексикографической работы, потому что принцип организации словаря любого типа основывается на определении слова через родовое и видовое отличие» (Польщикова, Польщикова, 2018, с. 156). Изучение иерархической организации терминов «позволяет не только составить представление о содержании предметной области, но и лучше понять принципы формирования отраслевой терминологии» (Сарангова, 2015, с. 57).

Гиперо-гипонимическими семантическими процессами «пронизаны терминологические системы различных областей знаний» (Головин, Кобрин, 1987, с. 59). Родо-видовые отношения выявлены в терминологии строительства (Доржеева, 2018) и терминологии финансовой сферы (Сарангова, 2015), гипонимические связи рассмотрены в зоологической терминологии (Галимова, 2018), отношения между гиперонимами и гипонимами представлены для систематизации научных понятий политической терминологии (Лату, 2019). Изучена структурная организация инфокоммуникационных терминов на основе процессов гиперо-гипонимии (Польщикова, Польщикова, 2018), показаны особенности отношений «род – вид», присущие юридической терминологии (Батюшкина, 2020) и экономической терминологии (Коновалова, 2001). Системные отношения между гиперонимами и гипонимами исследованы в терминологии цифровой фотографии (Асадова, Акаева, 2020), родо-видовые связи выявлены в терминологии фонетики (Лукина, 2019), выделены гиперо-гипонимические модели многокомпонентных горнозаводских терминов (Юдина, 2013).

Следует отметить, что в научной литературе недостаточно освещены вопросы функционирования отношений «род – вид» в терминологии компьютерной лингвистики, являющейся сравнительно новой профессиональной сферой. Научное направление компьютерной лингвистики связано с проведением междисциплинарных исследований в области автоматической обработки естественного языка (Агузумцян, Великанова, Польщиков и др., 2021; Польщиков, Лазарев, Константинов и др., 2020; Польщиков, Польщикова, Игитян и др., 2019; Хобсон, Ханнес, Коул, 2020; Fathi, Maleki, 2018; Johri, Khatri, Al-Taani et al., 2021; Polshchikov, Lazarev, Konstantinov et al., 2020; Polshchikov, Lazarev, Polshchikova et al., 2019; Velikanova, Polshchikov, Likhosherstov et al., 2021). Терминам компьютерной лингвистики присущи характерные черты, обусловленные особенностями формирования соответствующей совокупности специальных наименований. Быстрое развитие программно-аппаратных средств автоматизации, технологий искусственного интеллекта и моделирования естественного языка усложняет адекватное пополнение терминологии компьютерной лингвистики, оно происходит «во многом стихийно, неуправляемо, в результате чего проявляются негативные явления в виде, например, нежелательной синонимии, чрезмерного употребления англицизмов» (Прохорова, Польщикова, Польщикова и др., 2021, с. 31). В этой связи исследования терминологии компьютерной лингвистики, в частности имеющихся в ней гиперо-гипонимических отношений, способствуют её упорядочению, дальнейшей стандартизации и являются актуальными.

Для достижения поставленной цели исследования в статье решены следующие задачи: 1) представить формирование родо-видовых отношений по различным классификационным признакам в терминологии компьютерной лингвистики; 2) выявить типы структур группирования терминов на основе гиперо-гипонимических связей; 3) показать связи между терминами в рамках привативных и эквиолентных оппозиций.

Анализ отношений «род – вид» в терминологии компьютерной лингвистики выполнен на основе применения различных методов исследования. Методы описания и лингвокогнитивного моделирования применены для систематизации специальных названий. Использование дефинитивного анализа фактического материала позволило изучить семантические особенности рассмотренных терминов. Метод количественного анализа применен при подсчете гипонимов в составе родо-видовых групп.

Решение поставленных задач стало возможно благодаря теоретической базе, представленной трудами, которые посвящены описанию системных отношений между терминологическими единицами (Асадова, Акаева, 2020; Большакова, Воронцов, Ефремова и др., 2017; Головин, Кобрин, 1987; Гринев-Гриневич, 2008; Кронгауз, 2005; Кузнецова, 1989; Новиков, 1982; Трубецкой, 2000). Источниками фактического материала стали тексты, размещенные в научных журналах и книгах по компьютерной лингвистике, словари и специализированный тезаурус (РАТКЛ, 2022).

Практическая значимость работы определяется возможностью использования её результатов в лексикографической деятельности и учебном процессе для подготовки специалистов по профилю компьютерной лингвистики.

Основная часть

Структурированность терминологии компьютерной лингвистики основана на различных лексико-семантических, в частности гиперо-гипонимических отношениях. В состав родо-видовых групп входят термины, которые «выражают один и тот же семантический признак с разной степенью конкретизации» (Польщикова, Польщикова, 2018, с. 156). Так, слово *перевод* является гиперонимом к гипонимам *машинный перевод*, *человеко-машинный перевод*, *статистический машинный перевод*, *нейронный машинный перевод*. Термин *корпус* является родовым по отношению ко многим видовым терминам: *национальный корпус*, *лингвистический корпус*, *диалектный корпус*, *диахронический корпус*, *синхронический корпус*, *одноязычный корпус*, *многоязычный корпус*, *открытый корпус*, *мониторный корпус*, *параллельный корпус*, *сравнительный корпус*.

В тексте переход «от гиперонима к гипониму (при условии, если они обозначают одно и то же) связан с получением дополнительной информации, с конкретизацией обозначаемого» (Новиков, 1982, с. 242), например: *Смысл есть всюду, где есть знаковые единицы языка (морфемы, слова, предложения)* (Большакова, Воронцов, Ефремова и др., 2017, с. 15); *Самые распространенные типы лексической неоднозначности – это омонимия и полисемия* (Прикладная и компьютерная лингвистика, 2016, с. 18); *Более сложными видами лексических ресурсов являются тезаурусы и онтологии* (Большакова, Воронцов, Ефремова и др., 2017, с. 22); *Кроме того, части слов можно разделить на еще более мелкие смысловые единицы. Буквы или графемы обладают тональностью и смыслом* (Хобсон, Ханнес, Коул, 2020, с. 69).

В терминологии компьютерной лингвистики выявлены родо-видовые группы, в которые входят пары гипонимов, противопоставленных по определенному признаку, например: *ориентированное дерево – неориентированное дерево, многоцелевой парсер – одноцелевой парсер, ручное реферирование – автоматическое реферирование*. Значение гиперонима тем шире, «чем больше существует терминов, конкретизирующих значение родового понятия» (Польщикова, Польщикова, 2018, с. 156). Гипероним *разметка* охватывает не менее семи гипонимов: *морфологическая разметка, синтаксическая разметка, семантическая разметка, анафорическая разметка, просодическая разметка, дискурсная разметка, ручная разметка*. Более широкое значение имеет термин-гипероним *анализ*, охватывающий более 20 гипонимов: *референциальный анализ, анализ мнений, анализ речи, анализ спектра, анализ тональности, анализ тональности текста, анализ Фурье, графематический анализ, глубинный анализ текстов, интеллектуальный анализ текста, интеллектуальный анализ текстовой информации, авторегрессионный анализ, синтаксический анализ, лексико-семантический анализ, автоматический синтаксический анализ, морфологический анализ, автоматический морфологический анализ, семантический анализ, латентно-семантический анализ, сентимент-анализ, слуховой анализ, фонетический анализ* и др.

Термины компьютерной лингвистики часто обозначают понятия, классифицируемые по различным признакам, образуя несколько гипонимических групп. Так, гипероним *онтология* имеет гипонимические группы, различаемые по типу наполнения (*общая онтология, онтология задач, предметная онтология*) и по цели практического применения (*онтология представления, прикладная онтология*). Термин *система распознавания речи* образует три гипонимические группы по следующим признакам: по объему словаря (*система распознавания речи с малым словарем, система распознавания речи со средним словарем, система распознавания речи с большим словарем, система распознавания речи со сверхбольшим словарем*), по привязке к конкретному диктору (*дикторозависимая система распознавания речи, дикторонезависимая система распознавания речи*), по типу объектов распознавания (*система распознавания коротких фраз речи, система распознавания слитной речи, система распознавания спонтанной разговорной речи*).

В зависимости от особенностей признаков, конкретизирующих гипероним, в терминологии компьютерной лингвистики можно выделить родо-видовые группы с тремя типами иерархических структур: 1) простое дерево; 2) разветвленное дерево; 3) дерево смешанного типа.

Структура первого типа состоит только из двух терминологических уровней: на верхнем уровне находится гипероним, а на нижнем – согипонимы. Структура такого типа на примере гиперо-гипонимической группы термина *морфология* представлена на Рисунке 1.

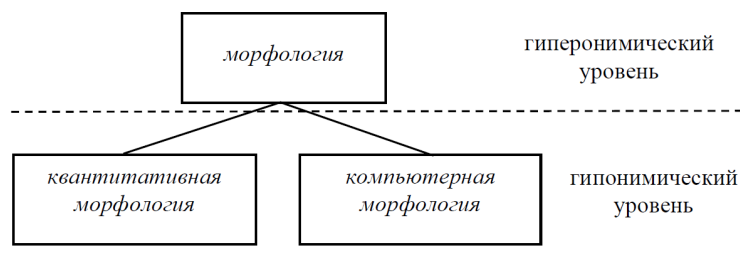


Рисунок 1. Простое дерево

Второй структурный тип между уровнями гиперонимов и гипонимов содержит промежуточные уровни признаков. Так, иерархическая структура родо-видовой группы термина *онтология* между верхним (гиперонимическим) и нижним (гипонимическим) уровнями содержит промежуточный уровень признаков (Рисунок 2). Наиболее сложным является третий тип иерархической структуры, представляющий собой сочетание простых и разветвленных деревьев. В такой структуре смешанного типа между одними парами гиперонимических и гипонимических уровней имеются промежуточные уровни признаков, а между другими – не имеются.

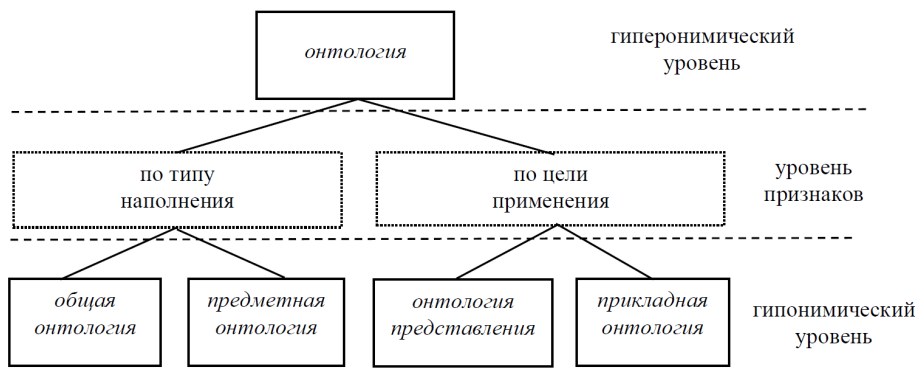


Рисунок 2. Разветвленное дерево

Гиперо-гипонимические отношения в терминологии компьютерной лингвистики можно рассматривать с использованием логической классификации смысловых различительных оппозиций (Трубецкой, 2000). Терминологические оппозиции привативного типа «представляют собой очень важный для структуры лексической системы тип связей», в котором «находят свое выражение отношения рода и вида, представленные соответственно основным и маркированным членами» (Кузнецова, 1989, с. 66). В рамках привативных оппозиций «один из членов обладает указанным свойством (маркированный член оппозиции), а у другого этот признак отсутствует» (Соловьева, 2015, с. 180). В этом смысле «гипонимы выступают в качестве маркированных элементов, а гиперонимы представляют основу такой оппозиции» (Кузнецова, 1989, с. 45), например: *направление компьютерной лингвистики* (гипероним) – *машинный перевод* (гипоним); *процедура автоматического морфологического анализа* (гипероним) – *лемматизация* (гипоним); *задача извлечения информации из текстов* (гипероним) – *распознавание именованных сущностей* (гипоним); *характеристика системы информационного поиска* (гипероним) – *релевантность* (гипоним). Термин «может входить в такие оппозиции как в качестве основного, так и в качестве маркированного, уточняющего члена», уточняясь «как бы в двух направлениях: за счет соотнесенности с более общим словом и за счет соотнесенности с менее общим, конкретизирующим его словом» (Кузнецова, 1989, с. 66). Так, в сфере компьютерной лингвистики та или иная *модель* зачастую представляется в виде *графа*, разновидностью которого является *дерево*, поэтому в терминологии слово *граф* является гипонимом по отношению к гиперониму *модель* и гиперонимом по отношению к гипониму *дерево*.

В терминологии компьютерной лингвистики распространены эквивалентные (равнозначные) оппозиции как отношения, возникающие между логически равноправными согипонимами, например: *абсолютная частота леммы* – *относительная частота леммы*; *двумерный вектор признаков* – *многомерный вектор признаков*; *семантическая валентность* – *синтаксическая валентность*; *квантитативная лингвистика* – *корпусная лингвистика* – *прикладная лингвистика*.

Заключение

Проведенное исследование позволяет сформулировать изложенные ниже выводы.

Формирование родо-видовых отношений между терминами, обозначающими понятия сферы компьютерной лингвистики, осуществляется по различным классификационным признакам. Выявленные группы «род – вид» в составе исследуемой терминологии отличаются также числом охваченных гипонимов (от двух до двадцати и более).

Распространенными типами структур группирования терминов компьютерной лингвистики на основе гиперо-гипонимических связей являются простые и разветвленные деревья, а также деревья смешанного типа.

В терминологии компьютерной лингвистики отношения между гиперонимами и гипонимами выступают в виде привативных оппозиций, а между логически равноправными согипонимами имеются эквивалентные оппозиции.

Перспективы дальнейшего исследования видятся в изучении процессов полисемии и антонимии в терминологии компьютерной лингвистики, а также в проведении этимологического анализа специальных названий этой предметной области.

Источники | References

1. Агузумян Р. В., Великанова А. С., Польщиков К. А., Игитян Е. В., Лихошерстов Р. В. 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.
2. Асадова А. Ю., Акаева Х. А. Опыт репрезентации прикладной терминологической системы посредством терминологической сети // Актуальные проблемы филологии и педагогической лингвистики. 2020. № 2.
3. Батюшкина М. В. О выражении юридическими терминами отношений «род - вид» и «целое - часть» // Вестник Кемеровского государственного университета. 2020. № 4 (84).
4. Большакова Е. И., Воронцов К. В., Ефремова Н. Э., Клышинский Э. С., Лукашевич Н. В., Сапин А. С. Автоматическая обработка текстов на естественном языке и анализ данных М.: Изд-во НИУ ВШЭ, 2017.
5. Галимова О. Н. Некоторые особенности гипонимических связей (на примере зоологической терминологии татарского языка) // Филологические науки. Вопросы теории и практики. 2018. № 9-1 (87).
6. Головин Б. Н., Кобрин Р. Ю. Лингвистические основы учения о терминах. М.: Высшая школа, 1987.
7. Гринев-Гриневич С. В. Терминоведение. М.: Академия, 2008.
8. Доржиева О. А. Отношения «род - вид», «целое - частное» в терминологии строительства // Интеллектуальный потенциал XXI века: сб. ст. Междунар. науч.-практ. конф. (г. Казань, 29 января 2018 г.). Уфа, 2018.
9. Коновалова Е. А. Деривационный потенциал и парадигматические отношения современной русской экономической терминологии // Ученые записки Таврического национального университета им. В. И. Вернадского. 2001. № 1.
10. Кронгауз М. А. Семантика. М.: Академия, 2005.
11. Кузнецова Э. В. Лексикология русского языка. М.: Высшая школа, 1989.

12. Лату М. Н. Языковая репрезентация системной организации политических понятий в дефинициях // Политическая лингвистика. 2019. № 1.
13. Лукина О. И. Гипонимия терминов фонетики во французском языке // Мир науки, культуры, образования. 2019. № 1 (74).
14. Новиков Л. А. Семантика русского языка. М.: Высшая школа, 1982.
15. Польщикова К. А., Лазарев С. А., Константинов И. С., Польщикова О. Н., Свойкина Л. Ф., Игитян Е. В., Балакшин М. С. Модель для оценки эффективности выполнения робототехнической системой коммуникативных функций // СТИН (СТанки ИНструмент). 2020. № 6.
16. Польщиков К. А., Польщикова О. Н., Игитян Е. В., Балакшин М. С. Алгоритм поддержки принятия решений по выбору средств обработки больших массивов естественно-языковых данных // Научные ведомости Белгородского государственного университета. Серия «Экономика. Информатика». 2019. Т. 46. № 3.
17. Польщикова О. Н., Польщикова А. К. Структурная организация инфокоммуникационных терминов на основе гиперо-гипонимических отношений // Филологические науки. Вопросы теории и практики. 2018. № 11-1 (89).
18. Прикладная и компьютерная лингвистика / под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. М.: Ленанд, 2016.
19. Прохорова О. Н., Польщикова О. Н., Польщикова А. К., Деев А. В. Системность терминологии компьютерной лингвистики // Известия Юго-Западного государственного университета. Серия «Лингвистика и педагогика». 2021. Т. 11. № 1.
20. Русско-английский тезаурус по компьютерной лингвистике (ПАТКЛ). 2022. URL: <https://uniserv.iis.nsk.su/thes/search.php>
21. Сарангова Т. А. Гиперо-гипонимические отношения терминологии, обозначающей связи с инвесторами (на материале тематической группы «Номинация лиц, связанных с Investor Relations») // Вестник Российского университета дружбы народов. Серия «Лингвистика». 2015. № 2.
22. Соловьева Н. В. Оппозиция как объект теоретического исследования (на материале английского языка) // Филологические науки. Вопросы теории и практики. 2015. № 4-2 (46).
23. Трубецкой Н. С. Основы фонологии. М.: Аспект Пресс, 2000.
24. Хобсон Л., Ханнес Х., Коул Х. Обработка естественного языка в действии. СПб.: Питер, 2020.
25. Юдина Т. М. Структурные типы и системные отношения терминов в ЛСГ «Названия лиц по профессии» в горнозаводской терминологии начала XVIII века // Ученые записки Петрозаводского государственного университета. 2013. № 7-1 (136).
26. Fathi E., Maleki S. B. Deep Neural Networks for Natural Language Processing // Handbook of Statistics. 2018. Vol. 38.
27. Johri P., Khatri S. K., Al-Taani A. T., Sabharwal M., Suvanov S., Kumar A. Natural Language Processing: History, Evolution, Application, and Future Work // Lecture Notes in Networks and Systems. 2021. Vol. 167.
28. Polshchikov K. A., Lazarev S. A., Konstantinov I. S., Polshchikova O. N., Svoikina L. F., Igityan E. V., Balakshin M. S. Assessing the Efficiency of Robot Communication // Russian Engineering Research. 2020. Vol. 40. No. 11.
29. Polshchikov K., Lazarev S., Polshchikova O., Igityan E. The Algorithm for Decision-Making Supporting on the Selection of Processing Means for Big Arrays of Natural Language Data // Lobachevskii Journal of Mathematics. 2019. Vol. 40. No. 11.
30. Velikanova A. S., Polshchikov K. A., Likhosherstov R. V., Polshchikova A. K. The Use of Virtual Reality and Fuzzy Neural Network Tools to Identify the Focus on Achieving Project Results // Journal of Physics. 2021. Vol. 2060.

Информация об авторах | Author information



Польщикова Ольга Николаевна¹, к. филол. н.

¹ Белгородский государственный национальный исследовательский университет



Polshchikova Olga Nikolaevna¹, PhD

¹ Belgorod National Research University

¹ polshchikova@bsu.edu.ru

Информация о статье | About this article

Дата поступления рукописи (received): 10.03.2022; опубликовано (published): 30.06.2022.

Ключевые слова (keywords): компьютерная лингвистика; терминология; родо-видовые отношения; гипероним; гипоним; computer linguistics; terminology; gender-aspect relations; hyperonym; hyponym.