

ДЕРИВАЦИОННЫЙ АНАЛИЗ ТЕРМИНОВ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ

©Эрметова Ж. И., ORCID: 0000-0003-3300-4238, SPIN-код: 2118-2034,

Ph.D., Ургенчский государственный университет,

г. Ургенч, Узбекистан, jamilaermetovam@gmail.com

©Матякубова Н. М., ORCID: 0000-0002-0518-3630, SPIN-код: 8636-6543,

Ургенчский государственный университет, г. Ургенч, Узбекистан, pristinegirl@mail.ru

DERIVATIONAL ANALYSIS OF TERMS OF COMPUTATIONAL LINGUISTICS

©Ermetova J., ORCID: 0000-0003-3300-4238, SPIN-code: 2118-2034, Ph.D.,

Urgench State University, Urgench, Uzbekistan, jamilaermetovam@gmail.com

©Matyakubova N., ORCID: 0000-0002-0518-3630, SPIN code: 8636-6543,

Urgench State University, Urgench, Uzbekistan, pristinegirl@mail.ru

Аннотация. Описывается словообразовательный анализ терминологии компьютерной лингвистики на английском языке. Анализируются термины компьютерной лингвистики общим объемом в 706 единиц, основанные на материале словарей и справочной литературы по компьютерной лингвистике. Анализ проведен на основе классификации Л. Б. Ткачевой. Выделены наиболее продуктивные способы терминообразования, а также их структурные типы. Данное исследование может представлять интерес для исследований в области компьютерной лингвистики и переводческой деятельности.

Abstract. Describes the derivational analysis of the English terminology of computational linguistics. The study is conducted on a compilation of computer terms with a total volume of 744 units, compiled on the basis of computer and linguistic glossaries and dictionaries. The analysis was done based on the classification of L. Tkacheva. Made it possible to identify the most productive methods of term formation, as well as their structural types. The study may be of interest to researchers in the field of computational linguistics and translation activities.

Ключевые слова: деривационный анализ, терминология, компьютерная лингвистика, способы словообразования.

Keywords: derivational analysis, terminology, computational linguistics, types of wordformation.

В наше время едва ли можно представить сферу жизни, не включающую компьютерные технологии. Компьютеры стали неотъемлемой частью человеческого существования. В то же время специальный язык и специальная терминология данной области стремительно развиваются. Компьютерная терминология превратилась в современную предметную область лингвистических знаний, о которой можно судить по ее распространенности среди людей разных социальных групп и возрастов. И естественно, что в лингвистике произошел так называемый «терминологический взрыв». Однако в результате этого прогресса многим специалистам и любителям компьютерных технологий стало намного сложнее освоить новшества в электронно-вычислительных машинах (ЭВМ), так как они все создаются на базе английского языка. В нашей письменной и устной речи возникла необходимость в переводе



терминов на узбекский язык. В настоящий момент во всем мире ведутся работы по переводу постоянно растущей компьютерной терминологии, и теперь они возобновляются и на узбекском языке. Направление компьютерной лингвистики является новым в Узбекистане и еще требует кропотливой работы в сравнительно-сопоставительном плане.

В свою очередь, компьютерная терминология узбекского языка расширяется и дополняет систему терминологических ресурсов в результате появления новой терминологии на иностранных языках, в частности на английском языке. Таким образом, терминология языка тесно связана с социально-экономическими явлениями и процессами, которые происходят в мире и затрагивают нашу страну. Наблюдения показывают, что большая часть лексического слоя, который широко используется в компьютерной терминологической системе узбекского языка, сформировалась под влиянием английского языка.

Основная причина, по которой была определена тема исследования — это массовое внедрение компьютерной терминологии в язык. Поскольку компьютерная терминология не ограничивается сферами ее использования, язык активно затрагивается как система и как средство общения. Исследование описывает словообразовательный анализ терминологии компьютерной лингвистики на английском языке.

Слово «термин» подразумевает любое слово или словосочетание, которое употребляется для точной номинации научных или технических понятий или явлений, свойственное определенной области деятельности, и обслуживающее коммуникативные потребности данной сферы [1].

Следует провести грань между такими понятиями, как «компьютерная терминология» и «терминология компьютерной лингвистики». Несмотря на явную схожесть понятий, они вбирают в себя терминологические единицы разных областей. Говоря о «компьютерной терминологии» в основном используются термины, связанные с компьютером, его составными частями, компьютерными программами, компьютерными процессами и т.д.. В свою очередь терминология компьютерной лингвистики включает как и абстрактные понятия, связанные с теоретической лингвистикой и теорией познания, так и понятия прикладного характера, как обработка естественного языка и ее практические результаты. Выборка для исследования происходила на основе литературы по компьютерной лингвистике по частоте использования. Именно поэтому данная выборка (состоящая из 706 лексических единиц), содержит, как и компьютерную терминологию, так и терминологию компьютерной лингвистики.

Целью исследования является осуществление деривационного и структурного анализа терминологии компьютерной лингвистики и ее классификация.

Деривационный анализ исследуемой терминологии

На сегодняшний день вопросы терминологий, относящихся к различным отраслям науки и техники, номинации становятся значимыми. Узбекские исследователи провели анализ различных лингвистических аспектов терминоведения, а именно словообразование терминов различных отраслей: в автодорожной сфере, автомобильной сфере, в сфере международных отношений, медицинской сфере, экологической сфере и др. [2–6].

В свою очередь, терминология компьютерной лингвистики представляет собой совокупность различных систем терминов, так как она связана с рядом наук: лингвистикой, информатикой, искусственным интеллектом, математикой, логикой, философией, психологией, когнитивной психологией, психолингвистикой, антропологией и другими. Конкретная группа терминов была выбрана на основе тщательного анализа двуязычных и одноязычных словарей и глоссариев и справочной литературы. Специфика вычислительной

терминологии включает в себя термины вышеупомянутых наук. Терминология компьютерной лингвистики — это конгломерат различных терминов, описывающих процессы обработки естественного языка и моделирования естественного языка с вычислительной точки зрения.

Методы для исследования определены особенностями материала, поставленной целью и задачами. В ходе работы были использованы следующие методы:

1. прием деривационного анализа как основной метод;
2. когнитивный метод;
3. методы классификации и систематизации;

Структурный анализ образования данных терминологических единиц проводился на основе классификации Л. Б. Ткачевой [7], согласно которой выделяются простые, сложные и многокомпонентные терминологические сочетания.

Объектом исследования является термины компьютерной лингвистики общим объемом в 706 единиц. Тут следует отметить, что эти данные были выявлены на анализе лишь части терминов относящихся к компьютерам и компьютерной лингвистике, в то время как их общее количество насчитывает более 10000 единиц. Отбор терминов происходил по их частоте использования в литературе, указанной в списке литературы [8–11]. Мы считаем, что данное исследование может представлять интерес для исследований в области компьютерной лингвистики и переводческой деятельности.

Результаты исследования

Процесс выборки показал наличие ряда однокомпонентных терминов. Они делятся на те, что описывают:

–компьютерные процессы: access, bridge, bug, command, crash, down, environment, fault, font;

–части компьютера: cable, segment, chip, display, hub, mouse, screen, etc.

–лингвистические термины: stem, split, tense, word, etc.

Аффиксация (включая суффиксы и префиксы) является одной из основных форм образования терминов компьютерной лингвистики. Ниже приводится список суффиксов со спадающей активностью:

–tion (application, authentication, collision, configuration, encapsulation, instantiation, instruction, lemmatization, lexicalization, localization);

–ing (booting, encoding, modeling, nesting, parsing, programming, rendering, routing, surfing, telecomputing);

–ics (graphematics, graphemics, linguistics, phonetics, phonotactics, pragmatics, psycholinguistics, semantics, semiotics, sociolinguistics);

–ogy (grammatology, methodology, morphology, ontology, phonology, terminology);

–er/or (editor, header, mailer, programmer);

–able (executable, portable, variable);

–ence/ance (concordance, inheritance).

А теперь приводим список префиксов (приставок) со спадающей активностью:

inter– (interactive, internet, interrupt);

de– (debug, default, decoding);

re– (recall, re-engineering, retiming);

hyper– (hyperlink, hypermedia, hypertext);

micro– (microcomputer, microprocessor, microwave);

in- (inbox, input);
meta- (metacharacter, metadata);
allo- (allomorph, allophone).

Помимо приведенных выше префиксов имеется ряд других префиксов, с помощью которых образуются слова компьютерной лингвистики (con-, dis-, ex-, im-, mini-, non-, pre-, self-, un-), однако на каждый такой префикс приходится незначительное количество слов.

С этимологической точки зрения термины компьютерной лингвистики содержат суффиксы римского происхождения -tion (наиболее производительный), а также греко-латинские префиксы и суффиксы: re-, de-, inter-, in-, hyper-, meta-, micro-, -ing, -ogy, -able, -ence / -ance, -er / -or.

Исследование показало, что наиболее продуктивным способом словообразования в терминах компьютерной лингвистики является формирование двухкомпонентных фраз. Наиболее часто встречающиеся структурные модели двухкомпонентных терминов включают в себя следующие структуры:

N+N: Application layer, base character, character code, command shell, data mining, error checking, file format, hand shaking, information extraction, laser printer;

A+N: Autonomous System, computational semantics, digital image, hard disk, hierarchical file, linguistic corpora, natural language, open platform, phonetic transcription;

V+ed+N: Applied linguistics, bitmapped terminal, compiled language, distributed database, linked data;

V+ing+N: collating sequence, scripting language.

Также были отмечены трехкомпонентные фразы, имеющие следующую структуру:

N+N+N: data link layer, database management system;

A+N+N: "black box" evaluation, hierarchical file structure;

A+A+N: reverse alphabetical order;

V+ed+N+ N: distributed file system;

A+ N +V+ing: statistical language modeling.

Другой продуктивный метод словообразования терминологии компьютерной лингвистики — морфологический и синтаксический. Сокращенная терминология составляет около 18% данной подборки терминологии, которая представлена акронимами: ACL (Access Control List), Application Programming Interface / API, Artificial Intelligence/AI, BIOS (Basic Input / Output System), CLI (Command-Line Interface), CAT (Computer-Assisted Translation), FAQ (Frequently Asked Questions), GUI (Graphical User Interface), HMM (Hidden Markov Model), IPA (International Phonetic Alphabet), TCP (Transmission Control Protocol).

Анализируемая терминология включает в себя многочисленные сложные слова, состоящие из различных компонентов, а именно прилагательных (freeware), существительных (keyboard, software, cyberspace, gateway, mailbox, network, backbone, database), предлогов (backup, output)

Синтаксический тип словообразования представлен следующими примерами: on-line, object-based, user-friendly, dial-up, multi-processing, object-oriented.

На Рисунке отражен приблизительный состав терминологии компьютерной лингвистики на английском языке. Самыми продуктивными способами деривации терминов компьютерной лингвистики являются однокомпонентные и двухкомпонентные слова и

выражения с показателями в 25% и 28% соответственно. В свою очередь, трехкомпонентные выражения (1%) составили меньшинство.

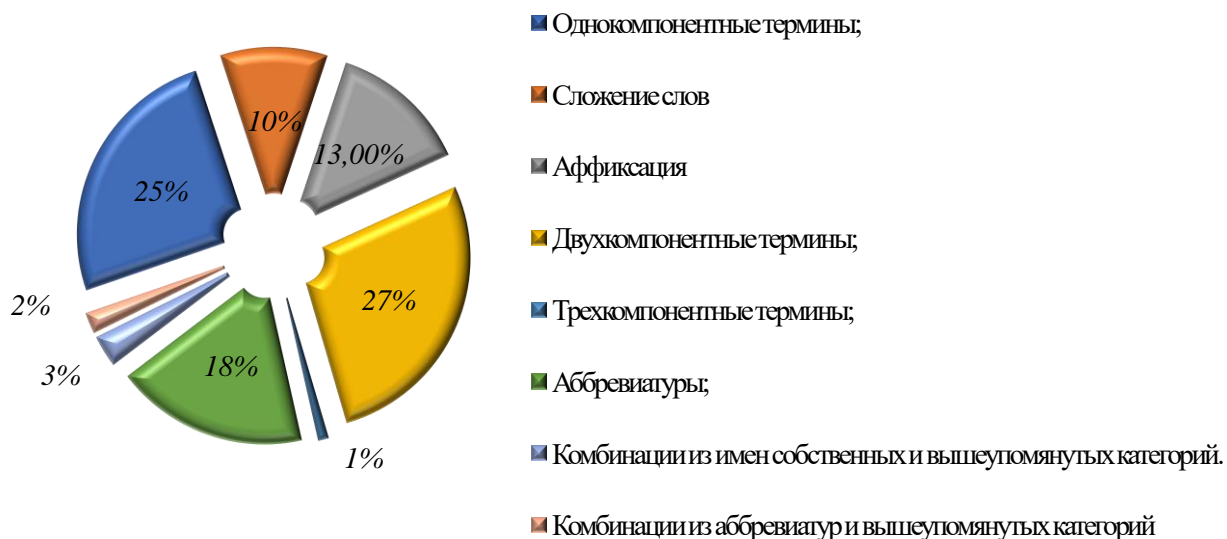


Рисунок. Способы словообразования терминов английской компьютерной лингвистики.

Заключение

Таким образом, выделены следующие структуры терминов: простые термины, сложные термины и терминологические сочетания. Анализ простых терминов с точки зрения аффиксального способа образования показал, что самыми продуктивными суффиксами являются: *-tion*, *-ing*, *-ics*, и префиксами: *inter-*, *re-*, *de-*.

Кроме аффиксации, композитное словосложение как разновидность морфологического способа также является продуктивным в образовании компьютерных терминов.

Проанализировав структурные характеристики наших терминов, мы пришли к выводу, что существует тенденция образования терминов с более чем одной основой.

Среди двух компонентных терминологических сочетаний наиболее продуктивными являются: *N+N* и *A+N*, среди трехкомпонентных: *N+N+N*, *A+N+N* и *A+A+N*.

В сфере образования терминологических единиц продуктивным также является синтаксическое словообразование, широко представленное графическим вариантом.

Список литературы:

1. Ярцева В. Н. Лингвистический энциклопедический словарь. М., 1990.
2. Валиев Т. Автойулсозлик терминларининг шаклий структурасига кура турлари // КарДУ Хабарлари. 2017. №3 (33). С. 120-125.
3. Ermetova J., Hojayev O. Uzbek-English, English-Uzbek explanatory dictionary of automotive terms. Tashkent, 2019.
4. Boltayeva D. Sh. Xalqaro aloqalarni bildiruvchi terminlarning lingvomadaniy hususiyati: magistrlik dissert, Tashkent, 2016.
5. Косимов А. Тиббий терминларнинг изоҳли луғати. Ташкент, 2003.
6. Палуанова Х. Д.. Экологик терминларнинг деривацион таснифи. Ташкент, 2016.
7. Ткачева Л. Б. Основные закономерности развития английской терминологии. Томск, 1987.

8. The handbook of computational linguistics and natural language processing / ed. by A. Clark, C. Fox, S. Lappin. John Wiley & Sons, 2013.
9. Rahimov A. Kompyuter lingvistikasi fanidan o'quv-uslubiy qo'llanma. Andijan, 2012.
10. Butterfield A., Ngondi G. E., Kerr A. (ed.). A dictionary of computer science. Oxford University Press, 2016.
11. Bolshakov I. A., Gelbukh A. Computational linguistics: models, resources, applications // Computational Linguistics. 2006. V. 32. №3. P. 443-444.
12. Ermetova J. Comparative analysis of the use of punctuation marks-dash and hyphen in English and Russia // Central Asian Problems of Modern Science and Education. 2019. V. 3. №4. P. 253-258.

References:

1. Yartseva, V. N. (1990). Lingvisticheskii entsiklopedicheskii slovar'. Moscow. (in Russian).
2. Valiev, T. (2017). Avtoilsozlik terminlarining shaklii strukturasi qura turlari. *KarDU Khabarlari*, (3), 120-125. (in Uzbek).
3. Ermetova, J., & Hojayev, O. (2019). Uzbek-English, English-Uzbek explanatory dictionary of automotive terms. Tashkent. (in Uzbek).
4. Boltayeva, D. Sh. (2016). Halqaro aloqalarni bildiruvchi terminlarning lingvomadaniy hususiyati: magistrlik dissert, Tashkent. (in Uzbek).
5. Kosimov, A. Tibbii terminlarning izokhli lugati. Tashkent, 2003.
6. Paluanova Kh. D. Ekologik terminlarning derivatsion tasnifi. Toshkent, 2016. (in Uzbek)
7. Tkacheva L. B. Osnovnye zakonomernosti razvitiya angliiskoi terminologii. Tomsk, 1987. (in Russian).
8. Clark, A., Fox, C., & Lappin, S. (eds.). (2013). The handbook of computational linguistics and natural language processing. John Wiley & Sons.
9. Rahimov, A. (2012). Kompyuter lingvistikasi fanidan o'quv-uslubiy qo'llanma. Andijan, (in Uzbek)
10. Butterfield, A., Ngondi, G. E., & Kerr, A. (eds.). (2016). A dictionary of computer science. Oxford University Press.
11. Bolshakov, I. A., & Gelbukh, A. (2006). Computational linguistics: models, resources, applications. *Computational Linguistics*, 32(3), 443-444.
12. Ermetova, J. (2019). Comparative analysis of the use of punctuation marks-dash and hyphen in English and Russia. *Central Asian Problems of Modern Science and Education*, 3(4), 253-258.

*Работа поступила
в редакцию 08.03.2020 г.*

*Принята к публикации
11.03.2020 г.*

Ссылка для цитирования:

Эрметова Ж. И., Матякубова Н. М. Деривационный анализ терминов компьютерной лингвистики // Бюллетень науки и практики. 2020. Т. 6. №4. С. 581-586. <https://doi.org/10.33619/2414-2948/53/70>

Cite as (APA):

Ermetova, Zh., & Matyakubova, N. (2020). Derivational Analysis of Terms of Computational Linguistics. *Bulletin of Science and Practice*, 6(4), 581-586. <https://doi.org/10.33619/2414-2948/53/70> (in Russian).

