

<https://doi.org/10.30853/filnauki.2019.9.51>

Кузякин Александр Сергеевич

ВАРИАТИВНОСТЬ ПРОИЗНОШЕНИЯ СЛОЖНОСОКРАЩЕННЫХ ТЕРМИНОВ ИЗ ОБЛАСТИ ДОЛГОСРОЧНОГО ХРАНЕНИЯ ДАННЫХ В НЕМЕЦКОМ ЯЗЫКЕ

Статья посвящена анализу особенностей звуковой репрезентации сложносокращенных компьютерных терминов в немецком языке на примере узкой предметной области долгосрочного хранения данных. Подчеркиваются отличия компьютерного синтеза специального текста в речь от живого воспроизведения коренным носителем языка. Выясняются границы нормы и причины вариации в произношении терминологических аббревиатур и акронимов, зависящие как от алгоритма и словарной базы искусственного интеллекта компьютера, так и от индивидуальных особенностей говорящего.

Адрес статьи: www.gramota.net/materials/2/2019/9/51.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2019. Том 12. Выпуск 9. С. 251-256. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2019/9/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: phil@gramota.net

УДК 81'373:811.11

Дата поступления рукописи: 13.07.2019

<https://doi.org/10.30853/filnauki.2019.9.51>

Статья посвящена анализу особенностей звуковой репрезентации сложносокращенных компьютерных терминов в немецком языке на примере узкой предметной области долгосрочного хранения данных. Подчеркиваются отличия компьютерного синтеза специального текста в речь от живого воспроизведения коренным носителем языка. Выясняются границы нормы и причины вариации в произношении терминологических аббревиатур и акронимов, зависящие как от алгоритма и словарной базы искусственного интеллекта компьютера, так и от индивидуальных особенностей говорящего.

Ключевые слова и фразы: сложносокращенное слово; графическая аббревиатура; акроним; омоним; фонетическая вариация; компьютерный термин; англицизм.

Кузякин Александр Сергеевич, к. филол. н., доцент
Российский университет транспорта, г. Москва
a_kuzyakin@mail.ru

ВАРИАТИВНОСТЬ ПРОИЗНОШЕНИЯ СЛОЖНОСОКРАЩЕННЫХ ТЕРМИНОВ ИЗ ОБЛАСТИ ДОЛГОСРОЧНОГО ХРАНЕНИЯ ДАННЫХ В НЕМЕЦКОМ ЯЗЫКЕ

Под сложносокращенными словами (сложными аббревиатурами) здесь понимаются слова, образованные действием аббревиации и основосложения, которые состоят из начальных букв или слогов слов и основ или из сочетания их с полными основами. В современном немецком языкознании эта категория слов в разной степени соотносится с понятиями “Abkürzung” («сокращение») [17, S. 2], “Kurzwort” («сложносокращенное слово»), “Kunstwort” («искусственно образованное слово»), “Kürzel” («графическая аббревиатура») и др. [10, S. 9-11], например: “z.B.” (zum Beispiel), “MA” (Master of Arts) («Магистр искусств»), “EDV” (elektronische Datenverarbeitung) («электронная обработка данных»), “LIFE” (Lifecycle Information for E-Literature) («жизненный цикл электронной литературы»), “UV” (Ultraviolett) («ультрафиолетовый»), P (Pixel < picture element) («пиксель, минимальный элемент изображения»).

Под термином, вслед за О. С. Ахмановой, мы понимаем слово или словосочетание специального (научного, технического и т.п.) языка, создаваемое (принимается, заимствуемое и т.п.) для точного выражения специальных понятий или обозначения специальных предметов [1, с. 474]. Идеальный термин, по мнению С. В. Гриневы-Гриневича, в пределах данной системы понятий должен быть однозначным, систематичным, стилистически нейтральным [4, с. 27-31].

Общественный запрос на повышение компьютерной осведомленности и доступность интернет-текстов для различных категорий населения, в том числе слабовидящих, приводит к тому, что сложносокращенные компьютерные термины из области специальных текстов переходят в аудитории, звучат в презентациях и радиопередачах, озвучиваются в программах-читалках, и то, в какой звуковой форме они будут представлены конечному пользователю, может сыграть важную роль в удовлетворении этого запроса. **Актуальность** работы состоит в том, что она исследует пласт лексики из области информационных технологий, которые в эпоху цифровизации и Интернета вещей становятся определяющими для жизни современного общества и развития немецкого языка.

Целью настоящей работы является выяснение тенденций, влияющих на произношение сложносокращенных слов, взятых из одной узкой предметной области компьютерных технологий – долгосрочного хранения данных.

Задачи, которые ставит перед собой статья, следующие: во-первых, определить место сложносокращенных терминов области долгосрочного хранения данных в массиве сложносокращенных компьютерных терминов немецкого языка; во-вторых, проследить, как лингвистическое разграничение акронимов и аббревиатур отражается в алгоритмах их представления в гипертекстах; в-третьих, сравнить алгоритмы современных программ-читалок между собой; в-четвертых, выяснить причины, определяющие выбор живым носителем языка варианта чтения сложносокращенного слова, совпадающего или отличающегося от компьютерного чтения.

Материалом для данного исследования послужил словарь акронимов и аббревиатур, размещенный в приложении к малой энциклопедии долгосрочного цифрового хранения данных, в объеме 331 лексических единиц [18]. Выбор пал на этот словарь, потому что, во-первых, технология долгосрочного хранения данных, хоть и заимствована из США, но уже долгое время развивается на германской почве и имеет свои национальные теоретические наработки, а во-вторых, потому что процент собственно немецких языковых образований здесь более заметен, чем где-либо. Термины, подобные рассматриваемым далее сложносокращенным словам, лингвисты часто называют англо-немецкими гибридами с использованием английских компонентов [6, с. 97], не имеющими немецких соответствий и используемых вследствие этого в немецкой литературе, преимущественно в сфере информационных технологий, экономике и менеджменте [3, с. 53].

Научная новизна настоящего исследования, помимо фокуса на узкой предметной терминологической области долгосрочного хранения данных, состоит в совмещении лингвистического анализа формы и содержания терминов с анализом психолингвистического механизма речевой деятельности говорящего, опирающегося на ассоциации и интенции, а также в оценке возможности репликации этого механизма в алгоритмах искусственного интеллекта.

В нашей выборке нашли отражение такие тематические разделы, как: 1) операционные системы: “DOS” (Disk Operating System) («дисковая операционная система, ДОС»), “Mac OS” (Macintosh Operating System) («операционная система фирмы “Apple Computer” для Power Macintosh»); 2) системы обработки и хранения данных: “CRIG” (Common Repository Interfaces Group) («группа интерфейсов общего банка хранения данных»), “PCM” (Puls-Code-Modulation) («Импульсно-кодированная модуляция, ИКМ для оцифровки аналоговых сигналов»); 3) системы обработки графики: “GIF” (Graphics Interchange Format) («формат графического изображения»), “PNG” (Portable Network Graphics) («формат графических файлов для работы в сети»); 4) аппаратное обеспечение: “RAID” (Redundant Array of Inexpensive (Independent) Disks) («дисковый массив»), “SCSI” (Small Computer System Interface) («интерфейс малых компьютерных систем»); 5) Интернет: “S/MIME” (Secure/Multipurpose Internet Mail Extensions) («безопасный протокол передачи электронной почты»), “IMAP” (Internet Message Access Protocol) («протокол доступа к интернет-сообщениям»); 6) коммуникация и системы передачи данных: “GIS” (Geoinformationssystem) («географическая информационная система»); 7) мультимедиа: “DVD” (Digital Versatile Disc) («цифровой видеодиск»), “CD-ROM” (Compact Disc Read-Only Memory) («компакт-диск без возможности перезаписи»); 8) программное обеспечение: “CAD” (Computer Aided Design) («автоматизированное проектирование»); 9) программирование: “J2EE” (Java Platform, Enterprise Edition) («среда и технология разработки корпоративных приложений на языке Java»); 10) прочее: названия немецких и зарубежных учреждений, библиотек, университетов, законодательных инициатив, комитетов, формальных процедур, необходимых для осуществления долгосрочного цифрового хранения данных. Именно в этом разделе количество немецких сложносокращенных слов наибольшее: “GRUR” (Deutsche Vereinigung für gewerblichen Rechtsschutz und Urheberrecht) («Немецкий союз правовой охраны результатов умственного труда в экономике и авторского права»), “HUB” (Humboldt-Universität zu Berlin) («Университет А. Гумбольдта в Берлине»), “PflAV” (Pflichtablieferungsverordnung) («требование обязательного экземпляра»), “RSWK” (Regeln für den Schlagwortkatalog) («Правила составления предметного каталога») и др.

Сравнение тематики терминов-аббревиатур узкой предметной области долгосрочного хранения данных с категориями немецкого компьютерного лексикона [14] показывает, во-первых, что в целом рубрикация совпадает, однако количество сокращений во второй и третьей значительно превышает количество сокращений в остальных восьми группах. Во-вторых, существенным отличием изучаемых сложносокращенных слов является практическое отсутствие многозначности. Этот факт увеличивает лексикографическую ценность узко предметных терминологических словарей, соответствующих идеальному определению термина.

Методика анализа предполагала озвучивание списка терминологических акронимов и аббревиатур немецкими синтезаторами речи, озвучивание того же списка носителем языка, сравнение живого чтения с синтезированным и, где это возможно, с данными словарей, определение коммуникативных интенций и спонтанных ассоциаций носителя со звуковыми образами сложносокращенных слов, сравнение этих ассоциаций с тем смыслом, который, возможно, закладывал в них создатель при первичной номинации. Такой подход напоминает упрощенный тест Тьюринга, позволяющий отличить искусственный интеллект компьютера от интеллекта человека применительно к акронимам и аббревиатурам. Упрощение в том, что заранее известно, где компьютер, а где человек.

Первым вопросом такого теста является умение компьютера и человека отличать чтение акронима от чтения аббревиатуры. Под акронимом, согласно словарю “Duden”, понимается произносимое как отдельное слово сокращение, состоящее из инициальных букв или слогов, составных частей композита, или слов, образующих словосочетание [11], например: “PC” [pe`tze] (англ. “personal computer”), “PIN” [pin] (persönliche Identifikationsnummer) («персональный идентификационный номер»), “LAN” [lan] (local area network) («локальная сеть»). Под аббревиатурами в узком смысле, называемыми здесь графическими аббревиатурами, понимаются усеченные слова (Wortkürzungen), которые, согласно М. Д. Степановой, представляют собой более или менее произвольные сокращения слов с различной словообразовательной структурой [7, с. 152], которые произносятся не так, как пишутся [10, S. 9], например: “Abk.” [’ap, kvɪʃtsʊŋ] (Abkürzung) («сокращение»), “z.B.” [zum`baiʃpʲl] («например»).

Человек, ориентируясь на точки, скобки, расположение буквы или буквосочетания перед цифрой или после цифры, перед именем или после имени и другие элементы контекста, относительно легко понимает лакунарность сокращенных графических образов слов или сочетаний и разворачивает их в целые слова и сочетания. Поскольку графические аббревиатуры не всегда отделяются на письме точкой и могут писаться слитно, раздельно, с пропуском, с варьированием строчных и прописных букв, например: z.B. = zB = z B, то их легко смешать с акронимами с вариативным написанием типа “PKW” и “Pkw” [’pe:ka:ve:] (Personalkraftwagen) («легковой автомобиль»). Кроме того, графические аббревиатуры (неполные аббревиатуры) могут на правах элементов входить в сложносокращенные термины, например: “e-Learning” [i:’lə:nɪŋ] («обучение при помощи электронных программ»). Казалось бы, такое положение вещей играет против компьютера, в пользу человека, поскольку вопрос о разграничении двух типов чтения становится проблематичным.

Однако это не так. В первых версиях компьютерного «языка гипертекстовой разметки» “HTML” (HyperText Markup Language) слова, написанные прописными, читались только по буквам, позже программисты, чтобы придать воспроизводимому тексту больше естественности, стали сопровождать сложносокращенные слова тегами – <abbr> («аббревиатура») или <acronym> («акроним»). Тэги не были видны в самом тексте, но сопровождали его в файле метаданных. Если графическая аббревиатура ‘B’ была маркирована тэгом <abbr>, то в программе, конвертирующей текст в речь, буква читалась как слово, которое она замещает, т.е. ‘Byte’ («байт»). В последних версиях “HTML5” тэг <acronym> уже не поддерживается. Компьютеры постоянно совершенствуются, и алгоритмы искусственного интеллекта сейчас могут определить в слове, написанном

прописными, признаки акронима (аналога обычного слова), а не аббревиатуры, которую следует разворачивать в полные слова. С другой стороны, акронимы и аббревиатуры, согласно Яну Хелбушу, сейчас снабжаются при помощи атрибута “title”, используя “Neudeutsch”, “Tooltips” («тултипы, контекстные меню»), которые могут открываться или по предварительной установке в программе-читалке, или по требованию пользователя при наведении курсора на сложносокращенный термин [12].

Компьютерные алгоритмы синтеза речи – это коммерческие ноу-хау компаний, разные по цене и качеству, опирающиеся на разные базы словарей. Маловероятно, что хорошие алгоритмы спутают устойчивую графическую аббревиатуру с акронимом, однако один и тот же акроним в разных программах может быть прочитан по-разному. Например, акронимы “DAT” (Digital Audio Tape) («цифровая аудиолента») и “COM” (Computer Output on Microfilm/-fiche) («Автоматизированный вывод информации из ЭВМ непосредственно на микрофильм») были прочитаны синтезатором *Translate.Google* соответственно как [dat] и [kom] [18], а синтезатором *ispeech* как [de_a_te] и [tze_o_em] [15]. Словарь “Duden” поддержал *ispeech* в случае с DAT [de_a_te] [11] (или послужил источником для немецких транскрипций этой программы), однако для акронима “COM” в онлайн-версиях “Duden” и “Langenscheidt” места не нашлось. “Duden” регистрирует “COM” в указанном значении в своем словаре сокращений [10], однако там нет транскрибированных произношений, а в словаре произношений [9] нет слова “COM”.

Ни программы, ни словари не объясняют причины своих предпочтений, и в этом их сходство. Искусственный интеллект может хорошо различать акронимы и аббревиатуры, потому что опирается на огромную базу словарей, но он не может объяснить предпочтений в их звуковой форме. Очевидно, только живые люди могут решить, какой вариант в каких случаях для них лучше и почему. Вариативность произношения сложносокращенных слов опирается на норму.

Формальными факторами, определяющими нормативное чтение сокращения, являются: длина слова, количество согласных и гласных, место гласных в слове, удобопроизносимость сочетаний звуков. Чем длиннее сложносокращенное слово и чем равномернее гласные распределены среди согласных, тем больше вероятность, что оно будет прочитано и человеком, и компьютером как цельное слово, состоящее из одного или более фонетических слогов. В нашей выборке представлены сокращения длиной от 1 до 9 букв. Большинство сложносокращенных слов трехбуквенные – 41% и четырехбуквенные – 31,5%. Трех-, четырех- и пятибуквенные сложносокращенные слова вместе составляют почти более 4/5 выборки, т.е. явно преобладают. Только пятая часть приходится на оставшиеся 1-, 2-, 6-, 7-, 8- и 9-буквенные аббревиатуры и акронимы, вместе взятые.

Короткие сокращения из 2-3-4 букв, состоящие только из гласных, или только из согласных, или начинающиеся с гласных и заканчивающиеся на согласный, или начинающиеся с гласного и заканчивающиеся на гласный, озвучиваются чаще как инициальные акронимы, в которых каждая буква произносится отдельно, как в алфавите, например: “CF” [tse_ef] (Compact Flash) («карта флеш-памяти»), “EU” [e_u] (Europäische Union) («Европейский Союз»), “AP” [a_pe] (Arbeitspaket) («рабочий пакет»), “HSM” [ha_es_em] (Hierarchisches Speichermanagement) («управление иерархической памятью»), “AAD” [a_a_de] (Access to Archival Databases) («доступ к архивным базам данных»), “ADR” [a_de_er] (Advanced Digital Recording) («усовершенствованная цифровая запись»), “AP” [a_ef_pe] (Zeitschrift für Medien- und Kommunikationsrecht < Archiv für Presserecht) («Журнал о вопросах права в СМИ»), “ATA” [a_te_a] (Advanced Technology Attachment) («интерфейс для подключения внешних устройств в АТ-совместимых компьютерах, интерфейс (стандарта) АТА»), “AVI” [a_fau_i] (Audio Video Interleave) («файловый формат программного сжатия чередующихся аудио- и видеоданных, формат AVI»).

Если в четырехбуквенном сокращении нет гласных, а есть только согласные, или только один гласный и три согласных, или один согласный и три гласных, оно будет прочитано по буквам, например: “BDSG” [be_de_es_ge] (Bundesdatenschutzgesetz) («федеральный закон о защите информации»), “AHDS” [a_ha_de_es] (Arts and Humanities Data Service) («информационная служба в сфере изобразительного искусства и гуманитарных наук»), “AKEA” [a_ka_e_a_] (Arbeitskreis Elektronische Archivierung des Verbands der Wirtschaftsarchive) («Ассоциация электронного архивирования Союза хозяйственных архивов»), “BABS” [be_a_be_es] (Bibliothekarisches Archivierungs- und Bereitstellungssystem der Bayerischen Staatsbibliothek) («библиотечная система архивирования и выдачи Баварской государственной библиотеки»).

Как цельные слова произносятся чаще трех-, четырех- и пятибуквенные сокращения с промежуточными гласными. Например: “DIN” [din] (Deutsches Institut für Normung) («Германский институт стандартизации»), “BAM” [bam] (Bibliotheken, Archive und Museen) («Библиотеки, архивы и музеи»), “CAD” [kxd] (Computer Aided Design) («автоматизированное проектирование»), “DOI” [doJ] (Digital Object Identifier) («цифровой идентификатор объекта»), “AOLA” [a`ola] (Austrian On-Line Archive) («Австрийский онлайн архив»), “CERN” [tjern] (Conseil Européen pour la Recherche Nucléaire) («Европейская организация по ядерным исследованиям»), “ASCII” [ˈaskJ] (American Standard Code for Information Interchange) («Американский стандартный код обмена информацией»), “CD-ROM” [tzede`rom] (Compact Disc Read-Only Memory) («постоянное запоминающее устройство на компакт-диске»), “CLOCKSS” [kloks] (Controlled Lots of Copies Keep Stuff Save) («создание достаточного количества копий способствует сохранению данных») и др.

Если в сокращении много букв, то в целях экономии времени произнесения предпочтительным будет вариант чтения сокращения как цельного слова, составленного из: 1) букв или слогов компонентов: “GEVER” [ge`fer] (Geschäftsverwaltung) («управление предприятием»); 2) букв или слогов одного компонента и несокращенной формы другого компонента производящего словосочетания: “DigiTool” [ˈdigiHI] (Digital Asset

Management Tool) («инструменты цифрового управления активами»), “e-Learning” [J_IWniN] (electronic learning) («обучение с использованием интернет-технологий»); 3) частей компонентов: “DigCCurr” [ˈdɪgkjuːr] (Digital Curation Curriculum) («план цифрового оздоровления»), “DIMAG” [ˈdɪmag] (Digitales Magazin des Landesarchivs Baden-Württemberg) («Цифровой журнал федеральной земли Баден-Вюртемберг»); 4) искусственного сращения фонетического акронима с алфавитным акронимом: “CIDOC-CRM” [ˈsɪdɔk_tze_er_em] (International Committee for Documentation – Conceptual Reference Model) («Международный комитет по документации – модель понятийного каталога»), DIN-NABD [dɪn_en_a_be_de] (Deutsches Institut für Normung – Normenausschuss Bibliotheks- und Dokumentationswesen) («Институт стандартизации ФРГ – комитет по библиотечному делу и документоведению»).

Некоторые сокращения произнести вообще нельзя. В нашем материале встретилось сокращение настолько искусственное, что его произнесение вслух у коренной немки вызвало затруднение. Написать его можно, а вот прочесть нельзя: “GenTAufzV” (Gentechnikaufzeichnungsverordnung) («предписание о документировании генной инженерии»). Близко к “GenTAufzV” по неестественности наш информант отнесла также “GEVER” (Geschäftsverwaltung) («управление предприятием») и “OAI-ORE” (Open Archives Initiative Protocol – Object Exchange and Reuse) («Протокол инициативы открытых архивов – обмен объектами хранения и их многократное использование»), отметив, что они скорее будут произнесены соответственно как целый композит или словосочетание, нежели как сокращения. Очевидно, что подобные искусственные построения носят глубоко отсылочный характер и в специальном тексте играют роль квазиместоимений, которые авторы заготавливают на всякий случай, а потом забывают ими пользоваться. В исследованной нами энциклопедии эти сокращения встретились только один раз.

Вариации в произношении анализировались с психолингвистических позиций в русле методов прямого (свободного) ассоциативного эксперимента, когда испытуемый сам называет ассоциации с тем или иным словом-акронимом [5, с. 81-108], и метода интен-анализа, когда внимание исследователя направлено на то, чем вызвано данное высказывание, какова его целевая направленность, зачем оно нужно говорящему [Там же, с. 114-117].

Вариации в произношении могут быть обусловлены рядом причин. Во-первых, привычность слова для говорящего на немецком языке как родном. Привычность непосредственно связана с возрастом, родом деятельности, образовательным уровнем. Нашим информантом была 22-летняя немецкая студентка из Мангейма, изучающая прикладную информатику, знающая английский язык на продвинутом уровне. Слова типа “DIN”, “BAM”, “DOI”, “CERN”, “ASCII”, “CD-ROM” не вызвали сомнений, они воспринимаются как устоявшиеся. Однако для близкого им по структуре “COM” (Computer Output on Microfilm/-fiche) «Автоматизированный вывод информации из ЭВМ непосредственно на микрофильм» ею было предложено чтение [tse_o_em], потому что “COM” как [kom] ассоциативно закреплено для нее как представительницы немецкой молодежи, часами сидящей в Интернете, за доменным именем “com.” (commercial) («коммерческий»). То же слово “COM” [kom] (communications port) («последовательный порт ПК») ей близко и понятно как программистке. Информант намеренно выбирает алфавитное чтение для привычного слова в непривычном значении, чтобы избежать омонимии, которая для терминологии в рамках одной предметной сферы в идеале неприемлема. По условию эксперимента, протекавшего в форме вопросно-ответного диалога с записью на цифровой носитель, информант читал вслух только левую часть словаря, т.е. акронимы и аббревиатуры в количестве 331 единицы. Она видела перед собой определения сложносокращенных терминов, но не озвучивала их. В такой коммуникативной ситуации основной интенцией говорящей было не столько «ответить», сколько «уточнить» и «подтвердить» соответствие фонетической формы термина его когнитивному содержанию, опираясь свой языковой опыт и фоновые знания, причем сделать это максимально корректно, т.е. избегая нежелательных оттенков смыслов. Эту интенцию можно назвать ассертивной.

Другие случаи ухода от омонимии: 1) “DIDL” (Digital Item Declaration Language) («язык описания мультимедийного файла формата MPEG-21») произносится как [de_i_de_el], чтобы избежать ассоциации с “Diddl” [didl] («мышонок из детских комиксов немецкого художника Томаса Голетца»); 2) “ATA” (Advanced Technology Attachment) произносится как [a_te_a], чтобы избежать ассоциации с “ata” (gehen) [ˈata] («идти гулять (в разговоре с детьми)»); 3) “DIP” (Dissemination Information Package) («информация о возможности скачивания файла в одном из доступных форматов») произносится [de_i_pe], чтобы не смешиваться с немецким существительным “Dip” [dip] («соус»); 4) “DIAS” (Digital Information Archiving System) произносится как [daiqs], чтобы уйти от немецкого омофона “Dias” [dias] (мн.ч. от Dia < Diapositiv) («диапозитив»); 5) “GRATE” (Global Remote Access to Emulation Services) («глобальный удаленный доступ к программам эмуляции») произносится [greit], чтобы уйти от немецкого “Grate” [ˈgrate] (мн.ч. от “Grat” («кромка»)), и др. Как видно из примеров, уход от омонимии осуществляется как алфавитным чтением акронима (DIDL, ATA, DIP), так и его англоязычным произношением (DIAS, GRATE).

Случайные совпадения терминологических сокращений с английскими словами для немцев в целом не играют роли, хотя для людей хорошо знакомых с английским языком они могут придавать акронимам в немецких текстах некоторую стилистическую пикантность и полифонизм, подсознательно заставляя их участвовать в диалоге культур в широком смысле, в понимании М. М. Бахтина [2]. Акроним “CRiB” (Conversion and Recommendation of Digital Object Formats) («сервис по конвертированию и выбору цифрового формата объекта») произносится [tze_er_i_be] не для того, чтобы избежать аналогии с английским неформальным существительным “crib” [krib] («шпаргалка») (хотя аналогия тут вполне может быть, т.к. в обоих значениях есть ассоциация с подсказкой), а просто для того, чтобы быть корректным, говорить наверняка. Мы поступаем также в русском языке, когда есть необходимость сослаться на маргинальное или непонятное

слово: берем слово в кавычки и называем каждую букву в отдельности и этим как бы дистанцируемся от возможных обвинений в некорректности.

Однако в ряде случаев очевидная омонимия с немецким или английским словом не вызывает отторжения, более того, омоним служит готовой языковой заготовкой, по которой говорящий формирует звуковой образ нового неизвестного еще сокращения, например: чтение акронима “DACHS” [daks] (Digital Archive for Chinese Studies) («цифровой архив китаистики») опирается на немецкое сокращение-омофон “DAX” [daks] (Deutscher Aktienindex) («немецкий биржевой индекс»), а акроним “DRIVER” [ˈdraivɪq] (Digital Repository Infrastructure Vision for European Research) («Проект инфраструктуры цифрового хранилища для европейских научных исследований») опирается на англицизм “der Driver” [ˈdraivɪq] («драйвер (программа, управляющая устройством)»).

При фонетическом разграничении акронимов один и тот же компонент производящего сочетания может быть представлен разным количеством букв и, следовательно, разными фонемами, например: компонент “Europäische” представлен фонемой [e] в “EU” [e_u] (Europäische Union) («Европейский Союз») и фонемой [oi] в “EuGH” [oi_ge_ha] (Europäische Gerichtshof) («Европейский суд»).

Другим направлением фонетической вариации является возможность произнесения некоторых согласных и гласных в акронимах в германизированном и/или англизированном виде, например: “GIF” (Graphics Interchange Format) («формат обмена графическими данными») допускает чтение [gif] и [dʒif], “HP” (Hewlett Packard) «компания Hewlett Packard» допускает [ha_pe] и [eitS_pJ]. Причем даже когда речь идет о принтерах этой компании, эти варианты идут на равных, одна половина сокурсников нашей информантки называют большой факультетский принтер [ha_pe], другая – [eitS_pJ]. Очевидно, при выборе англизированного чтения к ассертивной интенции добавляется интенция «продемонстрировать знание английского языка». С другой стороны, в целой группе терминологических сокращений, связанных с технологией гипертекста, “HTML” (Hypertext Markup Language) («гипертекстовый язык»), “HTTP” (Hypertext Transfer Protocol) («протокол передачи гипертекста»), “HTTPS” (Hypertext Transfer Protocol Secure) («протокол защищенной передачи гипертекста»), имеющих явно английские корни, среди немецких IT-специалистов предпочтение отдается именно германизированным вариантам произношения, соответственно: [ha_te_em_el], [ha_te_te_pe], [ha_te_te_pe_es].

Наметившаяся тенденция к германизированному произношению компьютерных терминов, возможно, связана с активностью инициативных групп, выступающих за использование родного языка и терминологии во всех областях науки, включая информатику [8], и предлагающих к использованию списки компьютерных терминов с немецкими корнями вместо становящихся привычными англицизмов [19]. Возможно, тут играет роль и национальная гордость немцев, считающих Германию родиной первого в мире компьютера. Немец Конрад Цузе произвел и поставил на рынок первую коммерческую версию работающего в двоичном коде программируемого счетного устройства (Rechner Z4) еще в 1950 году [16]. Во всяком случае, при выборе говорящим германизированного произношения термина-англицизма к ассертивной интенции добавляется интенция «быть или выглядеть патриотичным».

Еще одно направление вариации можно связать с местными особенностями произношения. Например, акроним “HUB” (Humboldt-Universität zu Berlin) («Университет Гумбольдта в Берлине») студенты в Мангейме скорее произнесут нейтрально как [ha_u_be], в то время как для студентов Гумбольдт Университета в Берлине он будет неформально [hub]. Обе группы, по-видимому, знают о существовании в немецком языке омографа-англицизма “der Hub” [hAb] («хаб, устройство соединения ветвей многоточечной сети»), однако предпочитают уходить от омонимии с интернет-термином в произношении. При выборе информантом произношения “HUB” (Humboldt-Universität) [ha_u_be] к ассертивной интенции добавляется интенция «указать на свое место рождения», возможно, даже подсознательно, т.е. «я не из Берлина» или «я местная».

Местный узуз может влиять на ассоциативное поле сложносокращенного слова. Акроним MA [em_a] (Master) «магистр, магистерская программа обучения» произносится также как аббревиатура города Мангейм – “MA” (Mannheim). Совпадение звучаний обоих слов приводит не к многозначности термина, полностью исключенной контекстом, но к его полифоничности, обогащению его ассоциативного семантического поля. Выбор информантом варианта чтения “MA” (Master) [em_a], а не рекомендованного словарем произношений “Duden” англизированного [em_ei], свидетельствует о сращении ассертивной интенции с подсознательной интенцией «указать на свое место рождения».

Некоторые графические аббревиатуры международных единиц измерения могут молодежью произноситься как акронимы. Например, “MB” (Megabyte) [ˈme:gabaʏt] может произноситься неформально как [em_be], хотя аналогичный по структуре “GB” (Giga-Byte) [ˈgi:gabaʏt] такой опции пока не имеет. Это происходит потому, что в современных популярных немецких тарифных планах для смартфонов учет потребленного интернет-трафика в целом и в отдельных телефонных приложениях идет по понедельно и помесечно в мегабайтах, а не в гигабайтах, поэтому чтобы не выйти за границы тарифа, надо считать мегабайты – [em_be]. К ассертивной интенции здесь прибавляется подсознательная интенция «декларировать свою принадлежность к молодежи».

Таким образом, можно утверждать, во-первых, что сложносокращенные термины из области долгосрочного хранения данных тематически отражают компьютерный лексикон в целом, причем статистически преобладающими являются трех- и четырехбуквенные акронимы, произносимые по буквам, как в алфавите, и не имеющие омонимов в обычном языке. Во-вторых, нами было выяснено, что программисты посредством тегирования переводят важную для понимания специальных текстов теоретическую лингвистическую проблему разграничения акронимов и аббревиатур в область языковой практики, предоставляя возможность информированному пользователю самому произвольно вызывать у сложносокращенных слов в гипертексте контекстные меню, поясняющие их деривационную структуру и смысл. В-третьих, было обнаружено,

что современные программы, преобразующие текст в речь, имеют как достоинства, так и недостатки. К первым относятся опора на нормативное немецкое произношение, широкую базу данных словарей, возможность настраивать темп речи. Ко вторым относятся коммерческая ориентированность, закрытость алгоритмов, неполнота и противоречивость используемых баз словарей, в которых отсутствуют новейшие терминологические аббревиатуры из области долгосрочного хранения данных. Главный недостаток – механистический характер синтезированной речи, поскольку звуковая форма слов подается в отрыве от их понятийного и ассоциативного содержания, а коммуникация слушателя с искусственным интеллектом программы носит очень ограниченный характер. В-четвертых, в результате психолингвистического эксперимента были выявлены причины фонетического варьирования, которые делятся на ассоциативные и интенциональные. Если информант обнаруживает в значении узуального омонима терминологического акронима «некорректные» или «неуместные» ассоциации, то термин произносится как сочетание букв в алфавитном чтении. К «неуместным» в нашем материале были отнесены ассоциации «детский», «смешной», «пища», «архитектура», «предмет школьного обихода» и др. «Уместные» ассоциации, например, с «биржевой активностью» (DACHS – DAX) или «программным обеспечением» (DRIVER – DRIVER) на произношение не влияют. Коммуникативные интенции определялись как условиями эксперимента, так и личными особенностями личности информанта. Интенции не всегда очевидны и осознаны. Основной инвариантной интенцией явилась «ассертивность», т.е. подтверждение наличия связи между определенной словесной формой и когнитивным содержанием. Она базируется на языковом опыте и фоновых знаниях говорящего. Дополнительными интенциями, вызывающими фонетическое варьирование, явились: «продемонстрировать знание английского языка», «быть или выглядеть патриотичным», «указать на свое место рождения», «декларировать свою принадлежность к молодежи».

Список источников

1. **Ахманова О. С.** Словарь лингвистических терминов. М.: Советская энциклопедия, 1969. 608 с.
2. **Бахтин Михаил** [Электронный ресурс] // Философская энциклопедия. URL: https://dic.academic.ru/dic.nsf/enc_philosophy/7986/БАХТИН (дата обращения: 12.07.2019).
3. **Голубева Т. И.** Особенности функционирования англо-немецких гибридов в аббревиации // Филологические науки. Вопросы теории и практики. 2014. № 8 (38). Ч. 1. С. 53-56.
4. **Гринёв-Гриневиц С. В.** Терминоведение. М.: Академия, 2008. 304 с.
5. **Денисенко В. Н., Чеботарева Е. Ю.** Современные психолингвистические методы анализа речевой коммуникации: учеб. пособие. М.: РУДН, 2008. 258 с.
6. **Зиновьева А. Ф., Миляева Н. Н.** Немецко-русский словарь сокращений. Экономика и менеджмент. М.: Р.Валент, 2007. 296 с.
7. **Степанова М. Д., Чернышева И. И.** Лексикология современного немецкого языка. М.: Высшая школа, 1962. 310 с.
8. **Däßler K.** Wissen, Intelligenz und Muttersprache [Электронный ресурс]. URL: <https://www.mathint.com/grundlagen/muttersprache.pdf> (дата обращения: 12.07.2019).
9. **Duden.** Das Aussprachewörterbuch. Siebente Auflage. Berlin: Duden Verlag, 2015. 926 S.
10. **Duden.** Das Wörterbuch der Abkürzungen. 5. Auflage. Mannheim: Dudenverlag, 2005. 480 S.
11. **Duden Online-Wörterbuch** [Электронный ресурс]. URL: <https://www.duden.de/woerterbuch> (дата обращения: 08.07.2019).
12. **Hellbusch J. E.** Zett Punkt Bee Punkt [Электронный ресурс]. URL: <https://www.barrierefreies-webdesign.de/knowhow/acronym-und-abbr/erlaeuterung.html> (дата обращения: 11.07.2019).
13. <https://translate.google.com/> (дата обращения: 10.07.2019).
14. <https://www.computerlexikon.com> (дата обращения: 10.07.2019).
15. **Ispeech** [Электронный ресурс]. URL: <https://www.ispeech.org/text.to.speech> (дата обращения: 10.07.2019).
16. **Konrad Zuse** [Электронный ресурс] // Wikipedia. URL: https://de.wikipedia.org/wiki/Konrad_Zuse (дата обращения: 10.07.2019).
17. **Metzler Lexikon Sprache.** 3. Auflage / hrsg. von H. Glück unter Mitarbeit von F. Schmöe. Stuttgart – Weimar: Verlag Metzler, 2005. 782 S.
18. **NESTOR Handbuch** [Электронный ресурс]: eine kleine Enzyklopädie der digitalen Langzeitarchivierung / hrsg. von H. Neuroth, A. Obwald, R. Scheffel, S. Strathmann, K. Huth. URL: https://www.langzeitarchivierung.de/Webs/nestor/DE/Publikationen/nestor_Handbuecher/nestor_handbuecher_node.html (дата обращения: 10.07.2019).
19. **Wörterliste englisch/deutsch für häufig benutzte Begriffe der Alltagsinformatik** [Электронный ресурс] // GMI: Gesellschaft für Mathematische Intelligenz. URL: <https://www.mathint.com/grundlagen/inferm.htm> (дата обращения: 12.07.2019).

PHONETIC VARIABILITY OF COMPOUND ABBREVIATIONS IN THE SPHERE OF THE GERMAN LONG-TERM DATA STORAGE TERMINOLOGY

Kuzyakin Aleksandr Sergeevich, Ph. D. in Philology, Associate Professor
Russian University of Transport (MIIT), Moscow
a_kuzyakin@mail.ru

The article examines the phonetic peculiarities of the German compound abbreviations of computer terms by the example of the narrow subject area of long-term data storage. The paper emphasizes differences between text-to-speech synthesis of computer terms and their live presentation by a native speaker. The author considers the limits of standard pronunciation and the causes of phonetic variations of terminological abbreviations and acronyms conditioned by both computer artificial intelligence algorithms and terminology database and a speaker's individual features.

Key words and phrases: compound abbreviation; graphic abbreviation; acronym; homonym; phonetic variation; computer term; Anglicism.