

УДК 81'33

ТЕНДЕНЦИИ РАЗВИТИЯ КОМПЬЮТЕРНО-ОПОСРЕДОВАННОГО ДИСКУРСА: МЕТАЛИНГВИСТИЧЕСКИЙ АСПЕКТ

А. А. Баркович

Белорусский государственный университет, Минск

Статья посвящена проблемной области компьютерно-опосредованного дискурса, речевой практике, занимающей всё больше коммуникационного пространства. Очевидная аутентичность сферы компьютерно-опосредованной коммуникации предполагает не только эмпирический характер её исследования, но и востребованность метаязыковых обобщений. Такого рода осмысление данных аналитического и синтетического характера целесообразно представить, в первую очередь, как тенденции развития.

Ключевые слова: лингвистическая парадигма, компьютерно-опосредованная коммуникация, компьютерно-опосредованный дискурс, метаязык, металингвистика, тенденция развития.

Металингвистическая интерпретация сведений об особенностях функционирования языковых данных обладает несомненной ценностью в общенаучном контексте. На современном этапе развития высокотехнологичной коммуникации уже накоплена значительная когнитивно-семиологическая совокупность сведений обобщающего свойства, что обуславливает их систематизацию и категоризацию. Подобная лингвистическая репрезентация актуальна: проблематика сферы компьютерно-опосредованной коммуникации, в том числе, создание искусственного интеллекта, информационной реальности и другие амбициозные стратегии, нуждаются в многоуровневом научном сопровождении. Такая задача может быть решена средствами лингвистической парадигматики.

На фоне универсализма и фундаментальности классических лингвистических канонов в условиях динамичного развития коммуникационной сферы компьютерно-опосредованного типа ярко проявляются аутентичные характеристики развития современного дискурса – речевой деятельности определённой социальной, личностной и прагматической идентичности. Научное сопровождение компьютерно-опосредованного дискурса (далее – КОД) в рамках лингвистической парадигматики обусловлено не только языковым опосредованием таких фундаментальных аспектов коммуникации, как социальная и ментальная практика, но и языковой сущностью данных, функционирующих в компьютерно-опосредованной коммуникации (далее – КОК). Несомненной ценностью в методологическом плане обладает практика выявления и мониторинга существенных и долгосрочных векторов развития сферы – тенденций. Систематизация и метаязыковое структурирование КОД не только позволяют учитывать потенциал сферы, но и делают динамику развития сферы предсказуемой и понятной [1].

Тенденции развития такой динамичной практики как КОД по определению не могут составлять закрытую, раз и навсегда сформировавшуюся совокупность. Широкое рассмотрение проблемной области КОД предполагает поиски методологического консенсуса в рамках множества дисциплин лингвис-

тического характера и пространственно-временную экспликацию релевантных данных. В рамках компактного лингвистического анализа актуальны парадигматическая определённость и характеристика подтверждённых практикой репрезентативных данных сферы.

Лингвистическая парадигматика сферы КОК обусловлена как широкими эмпирическими возможностями сбора данных, так и существенными сложностями её *метаязыкового* и *металингвистического* охвата. Пожалуй, определяющим сущностным фактором функционирования КОК является её поликодовый характер, через призму которого КОД неизбежно экстраполируется как многоаспектный феномен. При помощи информационного обеспечения современной компьютеризированной коммуникации структурирование и агрегация как отдельных информационно-текстовых фрагментов КОД, так и гипертекстов – вполне решаемые задачи. Высокотехнологичные возможности обработки данных в сфере КОК позволяют не только получать структурированные сведения о характере релевантного дискурса, но и осуществлять мониторинг языковых отношений.

В данном контексте среди тенденций развития компьютерно-опосредованного типа дискурсивной практики необходимо отметить: во-первых, глобализацию КОД; во-вторых, массовый перевод дискурсивной практики в электронный формат; в-третьих, унификацию языковых стандартов сферы; в-четвёртых, научную кооперацию при специализированном сопровождении такой деятельности; в-пятых, развитие узкой специализации в деятельности по организации и сопровождению КОК. Данный ряд тенденций является открытым «нечётким» множеством элементов, что, в принципе, характерно практике метаописания поликодовых систем [2].

1. Глобализация в коммуникационной сфере, при всей неоднозначности тенденции, является благоприятным фактором, создающим единое пространство разработки локальных научных проблем с опорой на мировой опыт: многое уже сделано, и, разработанные для презентации отдельных фрагментов речевой практики алгоритмы, как правило, можно адаптировать к исследовательской специфике подобного ряда задач.

Использование уже протестированных технологий и приобретение определённых лицензионных продуктов сопровождения КОД позволяют научиться эффективному использованию возможностей коммуникационной сферы и постепенно вырабатывать собственную стратегию на оригинальном языковом материале. Коммуникационная сфера, которая раньше разъединяла носителей разных языков, в условиях быстрого технического прогресса и интернационализации науки постепенно становится сферой интеграции не только научных данных, но и культурных достижений. Степень адаптивности КОД при реализации глобальных коммуникационных проектов явно превосходит соответствующие возможности уже привычных каналов обмена информацией – газет, радио, телевидения и т.д. В тоже время, инерция использования дифференцированных языковых ресурсов, которые традиционно препятствуют объединению разрозненного национально-специфичного культурного «содержания», не преодолена в КОД и вряд ли будет когда-нибудь полностью исчерпана – в силу субъективизма речевой практики.

Обратимся к примерам. Использование большинством европейских языков латинской графики несомненно способствует успешным стереотипным исследованиям с использованием универсальных программ-менеджеров, таких как *MonoConc*, *WordSmyth Tools* и др. [11; 13]. Кстати, некоторые из них работают и с кириллицей (*MonoConc*). Симптоматично, что работоспособные алгоритмы часто используются в более поздних исследовательских проектах, в том числе иноязычных. Использование русским языком кириллицы накладывает определённые ограничения на использование в её контексте достижений романо-германской филологии, но на примере русских корпусов можно видеть, что проблемы программного обеспечения могут успешно решаться и при использовании кириллической графики [3; 4 и др.].

Показательной является набирающая популярность технология сбора доступных для скачивания текстов из Интернета – «веб-краулинг» (англ. – *web crawling*). С помощью этой технологии для создания корпусов используются материалы, уже выложенные в Интернете. Примером миллиардного лингвистического корпуса, созданного по интернет-технологиям, стал проект *WaCky project (Web as Corpus kool yinitiative)*, в виде *ukWaC*, *deWaC* и *itWaC* – соответственно английскому, немецкому и итальянскому языкам текстов [5]. Очевидно, что подключение к подобному проекту любого следующего из обеспеченных КОД языков не должно оказаться нерешаемой проблемой.

2. Массовый перевод дискурсивной практики в электронный формат постепенно становится доминирующим коммуникационным стандартом, не только определяющим аутентичность новых артефактов, но и поглощающим достижения докомпьютерной эпохи. Дискурсивная практика КОК демонстрирует повышенную вариативность и разнообразие текстового наполнения сферы. Объём информации в виртуальной реальности уже увеличился до такого уровня, каким даже гипотетически не смогла бы удивить сфера печатного книгоиздания или звукозаписи. Компьютерные средства хранения и модификации информации постоянно совершенствуются и демонстрируют «всеядность», с которой никогда уже не смогут соперничать издательства бумажных книг и компакт-дисков. Значительную часть интернет-контента представляют копии уже изданных традиционными способами текстов: в КОД широко распространены копии копий, варианты, кавер-версии и т. д.

При этом, «компетенции» компьютерных программ непостижима эвристика человеческого разума и широчайшая контекстная референтность традиционной коммуникации: языковой материал – в огромном количестве – на серверах Всемирной паутины пока лишь накапливается. Однако принципиальное понимание путей эффективного решения проблем его интерпретации, реферирования, систематизации до сих пор отсутствует. Пожалуй, наиболее существенной лингвистической проблемой лавинообразного расширения КОД является его недостаточная обеспеченность языковыми эталонами: словарями, специализированными корпусами, электронными библиотеками. Так, например, для более полного использования преимуществ «корпусного формата» требуется наличие «академических» онлайн-словарей, максимально упрощённый доступ к ним (в том числе, через Интернет) и интерактивный режим их презентации. При этом, функционирование самодостаточного в плане языковой однородности фрагмента КОД обусловлено накоплением «критической массы»

релевантного той или иной кодовой / поликодовой общности текстового материала.

Опосредование языковой практики сферы КОК электронными (цифровыми) семиотическими системами уже привело к формированию целых квазинаучных направлений исследования сферы, таких как, например, «Гуглология» (от *Google*) и «Яндексология» (от *Яндекс*). Перспективность таких векторов исследования сферы зависит от успешности и стабильности соответствующих денотатов. Называть такую поисковую идеологию наукой, очевидно, ещё рано, но она уже стала объектом метаязыковой практики [10].

3. Развитие глобального языкового пространства в сфере КОК протекает активно, сопровождаясь тенденцией к языковой унификации и КОД. Как и другие тенденции развития сферы КОК, данная тенденция имеет сложный, противоречивый и неоднозначный характер. Специфика языковой унификации КОК – в интерактивном характере взаимодействия языковых систем КОД. Очевидно, что все языки в большей или меньшей степени испытывают интерференцию активно формируемой поликодовой общности КОК; меньше других, – возможно, английский, который закрепился в качестве основного интеркоммуникационного средства в КОК.

Приоритет английского языка в КОД чем-то напоминает «Пиррову победу»: исторически и так «многослойный», английский язык создаёт и перерабатывает большое количество интернациональной лексики, часто основанной на греческих и латинских корнях. Такое разнонаправленное «донорство» может привести к тому, что и без того давно далёкий от традиций Шекспира язык может остаться английским только номинально, превратившись в современный аналог эсперанто.

Дискурс КОК (как может быть никакой иной дискурс) обусловлен межъязыковыми отношениями сферы. Данные, в частности, интернет-дискурса, объективно свидетельствуют о существовании языковых *приоритетов* КОК: на первом месте с показателем в 54,7 % – английский язык, на втором с 5,9 % – русский, на третьем – немецкий язык с аналогичным показателем [14]. Вместе с тем, за один год, с начала 2012 г. по март 2013 г., русскоязычная аудитория Интернета выросла относительно других языков больше, чем на один процент, подняв русский язык с третьей позиции на вторую и продемонстрировав уступающие только китайскому языку темпы роста [там же]. Понятно, господству английского языка в Интернете пока ничто не угрожает. По данным другого источника, *Internet World Stats – Usage Population and Statistics*, примерно в это же время, по количеству пользователей русский язык, с приблизительно 60 млн. подключенного к Интернету населения России, занял только седьмое место в рейтинге [9]. Первое место по количеству интернет-пользователей занял Китай, хотя показатели китайского интернет-контента достаточно скромны – шестое место [9]. Русский язык, между тем, преобладает в сфере КОК Беларуси, Украины и других республик *ex-СССР*. Согласно приведённым данным, русский язык имеет абсолютный приоритет не только в Интернете России: 79,0 % – на Украине, 86,9% – в Беларуси, 84,0 % – в Казахстане, 79,6 % – в Узбекистане, 75,9 % – в Киргизстане и 81,8 % – в Таджикистане [14].

4. Кооперация научной деятельности в условиях компьютерно-опосредованной практики является одной из характерных тенденций лингвистического сопровождения КОД, которая обеспечивает значимые результаты как в контексте решения глобальных задач, так и в обеспечении успешного изучения проблематики отдельных языков.

Примеров успешного международного сотрудничества в сфере КОК уже достаточно много. Например, в проекте *Национальный корпус русского языка* (НКРЯ) русский язык стал центром притяжения языковых данных широкого спектра. Среди аффилированного в НКРЯ языкового материала – тексты на английском, немецком, итальянском, польском, украинском, белорусском и других языках [4]. Расширение круга включённых в научную орбиту иноязычных данных – не субъективизм авторов корпуса, а эффективная методика исследования КОД. Симптоматично, что НКРЯ смог стать лидирующим проектом в контексте восточнославянских языков – русского, украинского и белорусского. Например, русско-белорусский модуль в НКРЯ – *Параллельный корпус (белорусский)* – появился одновременно (в 2012 г.) с первым доступным в Интернете белорусскоязычным корпусом *Corpus Albaruthenicum*, превзойдя последний количественно в десять раз [6].

Есть сведения о привлечении широкого круга славянских языков и при реализации ряда других проектов, например *«InterCorp: projekt paralelních korpusů Filozofické fakulty Univerzity Karlovy v Praze»*, где заявлено наличие около 7 588 000 русских, 1 493 000 – украинских, 1 308 000 – белорусских, 26 879 000 – болгарских, 12 625 000 – хорватских, 2 664 000 – македонских, 47 640 000 – польских, 40 108 000 – словацких, 33 741 000 – словенских, 6 972 – сербских и, естественно, 99 547 000 – чешских единиц из общего количества в 867 287 000 словоупотреблений [8]. Подобная статистика во многом отражает уровень развития лингвистики КОК в каждой отдельно взятой славянской стране.

Многоязычный корпусный проект *ParaSol: A Parallel Corpus of Slavic and other languages* был создан Рупрехтом фон Вальденфельсом (*Ruprecht von Waldenfels*) и Роландом Майером (*Roland Meyer*) на базе Института славянских языков и литератур Бернского университета и Института славистики Регенсбургского университета в Швейцарии [12]. На фоне 18 085 532 словоупотреблений в рубрике «другие языки» показатель в 7 646 832 токенов представленности славянских языков – очень высокий. В этом ресурсе, наряду с другими славянскими языками, представлены русские тексты в виде 3 637 357 токенов и 78 997 лемм; украинские тексты – 1 017 057 токенов и 33 563 леммы; белорусские тексты – 482 467 токенов и 24 131 леммы. Здесь белорусские и украинские тексты обрабатывались благодаря помощи специалистов из НКРЯ [12].

5. Тенденцию развития узкой специализации в деятельности по организации и сопровождению КОК, как и другие описанные выше тенденции, целесообразно рассматривать в комплексном аспекте. Она не является изолированной и самодостаточной, развивается во взаимодействии с тенденциями «глобализации», «кооперации», «унификации», «электронного форматирования». Преимущества использования специализированного инструментария и его опытных «операторов» в исследованиях КОД контрастны и целесообразны.

Присущая сфере КОК специфика не только усложняет её программными средствами, но и позволяет в сравнительно короткий срок добиваться высоких результатов в проведении исследований даже с нуля. Сужение спектра компетенции позволяет улучшить структуру узкоспециальных умений и навыков. Множественность источников конкретной направленности позволяет повысить степень верификации компактно представленных данных.

Тенденция к узкой специализации в науке стала очевидной с середины XX в. Такое положение вещей обусловлено не только расширением объективной стороны научно-исследовательской деятельности, но и необходимостью выполнения исследований на эмпирическом материале, изучения локальных объектов в их взаимосвязях и перспективе. С другой стороны, узкая специализация приводит к появлению целых лакун в научной сфере: прорывные узкие направления активно и плодотворно разрабатываются, систематическая и комплексная работа остаётся в тени.

Компьютерно-опосредованная дискурсивная практика обладает и другими особенностями развития, характеризовать которые как тенденции феномена КОД, скорее всего, преждевременно. В металингвистическом аспекте достаточно широкий круг интра- и экстралингвистических факторов может оказывать влияние разной степени интенсивности на языковые системы одного или нескольких языков. Результаты такого воздействия могут проявляться, например, в активном создании «контаминаций» из элементов английского и других языков, расширении сфер сленга, жаргона и т.д. Стихийный, на первый взгляд, интернет-поиск с помощью *Google* еженедельно анализируют на *Euronews*, что тоже является показательным. Кстати, сам телевизионный канал *Euronews* выходит в эфир параллельно на четырнадцати языках, что является прекрасной иллюстрацией противоречивости тенденции языковой унификации КОК [7]. Тем не менее, под тенденциями развития КОД есть основания понимать все же некие «универсалии» – присущие компьютерно-опосредованной дискурсивной практике закономерности в глобальном масштабе.

Языковая картина КОК полна и стихийных, и умело организованных предрассудков и мифов, являясь пока малокодифицированной сферой [1: 90]. Названный ряд тенденций имеет комплексный характер, объединяя наиболее влиятельные тенденции развития КОД. Характерной особенностью тенденций КОД является их внутренняя противоречивость и интерактивность, обусловленные языковой спецификой сферы. При нарастающей «синхронизации» пока разобщенных языковых сред КОК есть прекрасные возможности для анализа динамики сферы, использования преимуществ синкретичного существования и развития коммуникационных феноменов и артефактов. В современных условиях именно сфера КОК, при насыщенности разного рода кодовыми системами и поликодовыми конгломератами, успешно выполняет роль экспериментальной площадки для широкого спектра лингвистических исследований. КОД уже давно не только отражает, но и формирует мейнстрим речевой практики, оказывая влияние и на отдельные языковые системы, и на универсальную языковую общность сферы КОК. Отсутствие единого выразительного тренда развития КОД свидетельствует об активности и незавершённости процесса кристаллизации метаязыковой структуры КОД.

Список литературы

1. Баркович А. А. Интернет-дискурс: компьютерно-опосредованная коммуникация: учеб. пособие. М.: Флинта: Наука, 2015. 288 с.
2. Кофман А. Введение в теорию нечётких множеств. М.: Радио и связь, 1982. 432 с.
3. МГУ–ЛОКЛЛ: Компьютерный корпус текстов русских газет конца XX-ого века [Электронный ресурс]. URL: http://www.philol.msu.ru/~lex/corpus/corpus_descr.html (дата обращения: 20.12.2014).
4. Национальный корпус русского языка [Электронный ресурс]. URL: <http://www.ruscorpora.ru> (дата обращения: 20.12.2014).
5. Baroni M. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora [Electronic resource]. URL: http://wacky.sslmit.unibo.it/lib/exe/fetch.php?media=papers:wacky_2008.pdf (accessed at: 20.12.2014).
6. Corpus Albaruthenicum [Электронный ресурс]. URL: <http://grid.bntu.by/corpus> (дата обращения: 20.12.2014).
7. Euronews [Electronic resource]. URL: <http://ru.euronews.com/news> (accessed at: 20.12.2014).
8. InterCorp: projekt paralelních korpusů Filozofické fakulty Univerzity Karlovy v Praze [Electronic resource]. URL: <http://www.korpus.cz/intercorp/?req=page:info> (accessed at: 20.12.2014).
9. Internet World Stats – Usage and Population Statistics [Electronic resource]. URL: <http://www.internetworldstats.com> (accessed at: 20.12.2014).
10. Kilgarriff A. Googleology is bad science [Electronic resource]. URL: <http://www.kilgarriff.co.uk/~Publications/2007-K-CL-oogleology.pdf> (accessed at: 20.12.2014).
11. Monoconc: Text Searching Software [Electronic resource]. URL: <http://www.monoconc.com> (accessed at: 20.12.2014).
12. ParaSol: A Parallel Corpus of Slavic and other languages [Electronic resource]. URL: <http://parasol.unibe.ch> (accessed at: 20.12.2014).
13. WordSmith Tools [Electronic resource]. URL: <http://www.lexically.net/wordsmith> (accessed at: 20.12.2014).
14. W3Techs – extensive and reliable web technology surveys [Electronic resource]. URL: <http://w3techs.com> (accessed at: 20.12.2014).

DEVELOPMENT TRENDS IN COMPUTER-MEDIATED DISCOURSE: METALINGUISTIC ASPECT

A.A. Barkovich

Belarusian State University, Minsk

The article deals with the problem domain of computer-mediated discourse, speech practice, that takes more and more communication space. The apparent authenticity of computer-mediated communication means not only the empirical nature of research, but requires metalinguistic generalizations as well. It is necessary to introduce such understanding of analytic and synthetic data first of all as development trends.

Keywords: *linguistic paradigm, computer-mediated communication, computer-mediated discourse, metalanguage, metalinguistic, development trend.*

Об авторе:

БАРКОВИЧ Александр Аркадьевич – кандидат филологических наук, доцент кафедры прикладной лингвистики Белорусского государственного университета, e-mail: albark@tut.by.

