

Эрик Гайкович АЗГАЛЬДОВ

Москва, Российская Федерация

Редактор журнала «Studies on Russian Economic Development» (английская версия академического журнала «Вопросы прогнозирования»)

Переводчик, синхронный переводчик (английский язык)

Единый язык vs. межъязыковой перевод

На протяжении всей письменной истории каждый народ ощущал свою самобытность и одновременно мог чувствовать себя частью единого человечества, которое почему-то оказалось расколотым на носителей разных языков. (Я употребляю *язык* в значении «нация», потому что язык – это первый признак нации. Неслучайно самоназвание большинства народов совпадает с названием их языка: русский, italiano, français, Deutsch).

В более широком смысле язык является, может быть, определяющим свойством человека. Ведь человек – «общественное животное» (по Аристотелю), а человеческое общество невозможно без коммуникации, т.е. языка. (Впрочем, доказано, что дельфины могут общаться, называя друг друга по именам, а некоторые человекообразные обезьяны способны выучить несколько сотен слов. Но языки общения животных не имеют ни грамматики, ни синтаксиса – обязательного признака настоящего языка.)

Возникновение языка и языкового многообразия – одна из самых непостижимых загадок. В мифологиях и религиях происхождение речи и языка обычно приписывается богам (или Богу). Вот что говорит Библия: «Господь Бог образовал из земли всех животных полевых и всех птиц небесных и привел к человеку, чтобы видеть, как он назовет их, и чтобы, как наречет человек всякую душу живую, так и было имя ей» (Быт. 2,19). В иудео-христианской традиции первоначальным, единым языком считается язык, на котором говорили Ной и его сыновья: «На всей земле был один язык и одно наречие» (Быт. 11,1). Многообразие языков – это следствие самонадеянности людей, вознамерившихся построить Вавилонскую башню. Бог «смешал язык людей», нарушив функцию коммуникации и сделав невозможным их совместный труд. В результате народы рассеялись по всей земле.

Просвещенные и любознательные люди издавна пытались раскрыть тайну происхождения языка. Геродот рассказывает о фараоне Псамметихе (VII в.

до н. э.), который поручил воспитание двух детей глухонемым, чтобы узнать, на каком языке они заговорят. Когда их привели к нему, один якобы произнес нечто, звучащее как «хлеб» по-фригийски. Из этого фараон поспешил заключить, что первым языком в мире был фригийский. В аналогичном эксперименте, поставленном шотландским королем Яковом V (XVI в.), дети будто бы заговорили на иврите. Лично я считаю более достоверными результаты, полученные императором Священной Римской империи Фридрихом II (XIII в.), а также индийским падишахом Акбаром Великим (XVI в.): в их случаях дети вообще не заговорили. В самом деле, теперь мы уже знаем, что человек не может овладеть языком вне общения с другими людьми; косвенным доказательством этого служат истории «маугли» – младенцев, оказавшихся и росших в природной среде вне человеческого социума.

Уже в новое время, в 1866 г., Парижское лингвистическое общество постановило закрыть дискуссию по этому вопросу, сочтя его неразрешимым. Интерес к данной проблематике вновь возник после Второй мировой войны, когда активно развивались разные школы антропологии, нейролингвистики, психолингвистики и др. Вопросы, на которые пытается ответить сегодняшняя наука: почему (из чего) возник язык; был ли это постепенный или относительно одномоментный процесс; возник ли язык на раннем этапе развития человека (150–200 тыс. лет назад) или сравнительно поздно (40–50 тыс. лет назад). И наконец, ведутся дискуссии о том, возникли ли современные языки на разных континентах независимо друг от друга или вначале появился один протоязык, который затем распространился по всей Земле?

В пользу первого предположения как будто говорит теория «универсальной грамматики», согласно которой в мозгу каждого человека с рождения имеется особый механизм, который охватывает любые грамматические системы, позволяя ребенку овладевать любым языком. Если эта гипотеза верна, тогда, действительно, представители рода человеческого, жившие на разных континентах, могли бы выработать собственные языки независимо друг от друга.

Однако больше сторонников имеет гипотеза моногенеза языка, которая исходит из того, что все человечество – это потомки так называемой «Митохондриальной Евы», жившей в Восточной Африке примерно 150 тыс. лет назад. Предполагается, что тогда же возник и праязык, впоследствии разделившийся на две или несколько ветвей. Одно из соображений в пользу теории моногенеза: в любом языке наличествуют корни и грамматические показатели, имеющие параллели в самых разных языках мира. Так, в русском и ирландском языках находят около 30 % неслучайных совпадений. Некоторые отечествен-

ные лингвисты развивают теорию, согласно которой русский язык – прямой потомок праязыка и египетские надписи лучше всего читаются по-русски.

Трудно подсчитать, сколько живых языков сейчас в мире, потому что не всегда возможно разграничить язык и диалект. Для английского языка уже давно не существует понятия нормативного языка, и принято говорить об английском языке Великобритании, США, Новой Зеландии и т.д. Появилось даже название *Ebonics*, обозначающее язык чернокожего населения США. Вот другие примеры. Во времена СФРЮ сербский, хорватский и боснийский считались одним языком, а теперь, когда их носители живут в разных государствах, это уже отдельные языки. В молодой Республике Молдове, напротив, молдавский язык признан тождественным румынскому и переведен на латинскую письменность. Корсиканские националисты стремятся доказать отдельность корсиканского языка от итальянского на основании того, что в нем есть две мягкие гласные, которых нет в итальянском языке. Неудивительно, что существуют большие расхождения в оценках численности живых языков: от трех до почти семи тысяч.

Что же на самом деле многоязычие – проклятие или благо? Должны ли мы рассматривать его как наказание людям за своеволие или как магистральный путь развития всего сущего – принципиальную установку на многообразие?

Во все века передовые умы человечества сознавали глубинное единство рода человеческого и мечтали о вненациональности, т.е. о преодолении национальной и языковой разобщенности. Платон говорил, что боги облагодетельствовали бы человечество, даровав ему общий язык. Вольтер как-то сказал: «Различие языков – одно из самых больших бедствий общественной жизни».

Многих мыслителей занимала проблема универсального, или философского, языка – идеального языка, свободного от несовершенства естественных национальных языков. Они видели в таком языке инструмент правильного мышления и классификации понятий, а также основу для построения справедливого общества. Во многих утопических проектах единый язык выступает одной из главных составляющих счастья и гармонии в обществе.

В исторической ретроспективе мы видим наивность подобных проектов. Первой документированной попыткой такого рода был город Уранополис («Небесный город»), основанный в IV в. до н. э. греческим утопистом Александархом на горе Афон, в Македонии. Жители этого города должны были говорить на совершенно искусственном, «придуманном», языке. Мы не знаем, как и когда закончился этот проект, но городок под таким названием существует близ горы Афон и поныне.

Вопросы единого языка и перевода интересовали и утопистов последующих поколений, прежде всего Томаса Мора и Томмазо Кампанеллу. Даже у Мишеля Нострадамуса в его «Центуриях» имеются рассуждения о некоем новом языке – посреднике между латынью и восточными языками. Разработкой универсального языка занимались такие мыслители, как Ян Амос Коменский, Ньютон, Декарт и Лейбниц. Наша просвещенная императрица Екатерина II наказала специальной комиссии ученых провести сопоставительное изучение языков разных народов с целью выработать всемирный язык. Из этой попытки, разумеется, ничего не вышло, но был создан сравнительный словарь 200 западных и восточных языков.

Пока ученые занимались поиском идеального языка, жизненная практика диктовала необходимость какого-то языка-посредника. (Точнее сказать – языков, потому что разные формы межъязыкового взаимодействия, коммерции и интеллектуального общения приводили к унификации систем письменности, счета, номенклатур, нотного письма, а сегодня уже и языков программирования.) Средневековая латынь с V по XVIII вв. являлась профессиональным языком Европы; кстати, и Ломоносов большинство своих работ писал на этом языке. В XIX в. вся образованная Европа говорила по-французски. В коммерции и межкультурном общении были в ходу языки типа пиджин, или лингва-франка, т.е. смеси из трех и более языков с упрощенными словарем и грамматикой.

Постепенно на смену идее замены этнических языков неким всеобщим языком пришла идея искусственного вспомогательного языка, используемого лишь для целей международного общения. Питательной средой была Европа периода, который принято называть *Belle époque* (с конца XIX в. до Первой мировой войны). Сейчас мы охарактеризовали бы этот период как первую волну глобализации или начало формирования мирового рынка. С одной стороны – успехи естественных наук и техники, рост промышленности, всемирные выставки, расцвет Британской империи и формирование других колониальных империй. С другой – стремление людей выйти за рамки узконациональных интересов, объединиться с «братьями по классу», чтобы создать новое, более справедливое общество. Подсчитано, что в XVII в. появился 41 проект искусственных языков, в XVIII в. – 50, а в XIX в. – уже 246!

Среди искусственных языков международного общения наиболее известны волапук, эсперанто, интерлингва и идо. Все они построены на основе слов и элементов естественных (национальных) языков и имеют более или менее упрощенную грамматику. Каждый такой язык – это как бы попытка воссоединить рассыпавшийся «послевавилонский» язык. Однако в основном они

строятся на лексике европейских языков: немецкого, английского, французского, итальянского, русского и испанского. Между тем в новых проектах международных языков делаются попытки преодолеть европоцентричность. Новейшая разработка этого рода – язык *Lingwa de Planeta*, создаваемый группой петербургских ученых; в его словарь наряду с корнями европейских языков включены слова из восточных языков – китайского и хинди.

Но в реальности почти треть населения земного шара говорит на английском языке, хотя он является родным только для 380 миллионов людей. Большая часть содержания Интернета – на английском языке, хотя 44 % пользователей говорят на другом языке. Ни один язык в истории человечества не был настолько распространен и популярен. В чем же причина такой распространенности английского языка?

Во-первых, английский язык – это язык общения величайшей империи всех времен – Британской, а также сменившей ее после Второй мировой войны новой квази-империи – Соединенных Штатов. Не случайно послевоенный период иногда называют *Pax Americana* по аналогии с *Pax Romana* (буквально «римский мир»; *мур* в смысле «мирное существование»). Если экспансия Римской империи привела к распространению в Европе латыни, то англо-саксонская экономическая и культурная экспансия сделала английский язык международным языком общения, коммерции, науки и (массовой) культуры. К достоинствам английского языка относится и то, что данный язык германской группы содержит значительное число латинских корней. Это сближает его с романскими языками, число говорящих на которых тоже исключительно велико. К тому же специально создавались упрощенные варианты английского языка: *BASIC* (British, American, Scientific, International, Commercial) со словарем в 850 слов, а совсем недавно – *Globish*, словарь которого насчитывает 1500 слов. Разумеется, это не полноценные языки, а лишь средство общения.

Что же дальше? Учитывая, что в нынешнем столетии Китай выходит в мировые экономические лидеры, не станет ли китайский язык, при всей его сложности, лидером и в международной коммуникации?

До сих пор я говорил, что многоязычие традиционно рассматривалось как беда, как проклятие для человечества и что мыслители и ученые всех времен искали способ снять это проклятие. Возможно, первым, кто исповедовал другие взгляды, был великий философ и лингвист Вильгельм фон Гумбольдт. Его брат Александр, столь же знаменитый естествоиспытатель и географ, писал, что Вильгельму было дано «глубже проникнуть в структуру большего числа языков, чем когда-либо удавалось постичь одному человеческому разуму». Для Гумбольдта многоязычие – это движущая сила в развитии человеческого духа. При всем раз-

нообразии языков каждый из них благодаря своей структуре и строению способен отражать определенный взгляд на мир. Согласно одной из доктрин современной лингвистики (разделяемой, впрочем, далеко не всеми) все известные языки – от языков австралийских аборигенов до современного английского – в равной мере «продвинуты» и в принципе переводимы друг в друга.

Но и в наше время проблема межъязыкового общения отнюдь не потеряла остроту. Конечно, было бы хорошо, если бы каждый владел несколькими языками, но полиглоты всегда были большой редкостью. Сегодня почти половина граждан Европейского Союза может говорить на языке, не являющемся для них родным, а примерно четверть европейцев может общаться на двух иностранных языках, и это естественно. В Великобритании же лишь треть населения владеет иным языком, кроме родного английского. Что тоже вполне понятно.

На протяжении всей истории человечества необходимую функцию посредников в общении и понимании других культур выполняли переводчики и толмачи (под переводчиками здесь понимаются лица, занимающиеся переводом письменных документов, а под толмачами – устные переводчики). В Первом послании к Коринфянам апостол Павел наставляет: «Если кто говорит на незнакомом языке, говорите двое, или много трое, и то порознь, а один изъясняй» (Кор. 14, 27). Если бы не переводчики, человечество в массе своей не знало бы Библии и других священных книг, не было бы обмена научной информацией, не возникло бы понятие «мировая литература».

Впрочем, переводческая деятельность (как и всякая другая) дает примеры как блестящей, так и весьма посредственной работы. Не зря у итальянцев есть полушутливое выражение – *traduttore traditore* («переводчик – предатель»). В XX в. стала остро ощущаться потребность в рационализации перевода, в первую очередь письменного. В новом индустриальном, а затем постиндустриальном обществе переводчики, эти «почтовые лошади просвещения», явно становились тормозом в развитии. Жизнь требовала их замены «железным конем», а еще лучше – умной электронной машиной.

Эти чаяния, казалось, начали воплощаться в жизнь с середины XX в., когда произошла компьютерная революция. В числе прочего она породила цикл кибернетических, информационных и когнитивных наук и дала толчок развитию прикладной лингвистики. Начались интенсивные работы по моделированию человеческого мозга, семантическому и синтаксическому анализу языка и синтезу речи. Ставилась цель создать так называемый искусственный интеллект и в конечном счете человекоподобных роботов. После первых десятилетий «бури и натиска» казалось, что мы уже близки к построению систем автоматического (или машинного) перевода. Но где-то в 80-е гг. энтузиазм увял, и ра-

боты почти остановились. Почему? Оказалось, что построить выполняющего команды робота куда легче, чем наделить его «разумом». Да и первые компьютеры не обладали достаточной мощностью и быстродействием, что тормозило исследования и разработки.

Но в последние два десятилетия работы в этой области активизировались в разных странах мира, и в них участвуют тысячи ученых и специалистов. Соответственно, сюда направляются значительные денежные средства. Все это обусловлено несколькими факторами, которые на самом деле тесно взаимосвязаны.

Первый из них – глобализация в экономике, политике и культуре, расширение зон свободной торговли и движения труда и капитала. Все это сопровождается ростом транснациональных корпораций, что вызывает необходимость локализации товаров. Речь идет о том, что мир потребляет все больше одинаковых товаров и их описания и технические инструкции необходимо переводить на местные языки. Это предполагает большие материальные затраты и стимулирует инвестиции в машинный перевод.

Одновременно наблюдается сокращение полномочий национальных государств и усиление международных организаций. Возникновение ООН, а затем Европейского Союза с их принципиальной установкой на поддержку многоязычия содействовало развитию синхронного перевода. Сейчас в Европейском Союзе представлено более 20 официальных языков. Если говорить о межъязыковом переводе, это означает более 200 языковых пар, или почти 400 направлений перевода. На обслуживание перевода эта организация тратит 1,1 млрд евро в год, или 1 % всего бюджета. А Европейский парламент тратит на перевод дебатов и документов целых 30 % своего бюджета.

Еще один важнейший фактор – стремительное развитие информационно-коммуникационных технологий. Начиная с первых машин на вакуумных лампах до сегодняшних цифровых оптических компьютеров (т.е. за период чуть более полувека) мощность компьютеров удваивалась каждые полтора года. А в ближайшие годы нас, по-видимому, ждет переход на биокомпьютеры, которые будут работать уже на клеточном уровне и обладать мощностью, во много тысяч раз превышающей сегодняшнюю. Для иллюстрации: если раньше синтаксический анализ предложения требовал нескольких часов, то теперь – долю секунды. И конечно, появление и распространение Интернета действительно превратило весь цивилизованный мир в «глобальную деревню», о которой когда-то писал Маклуэн. Люди в разных странах, говорящие на разных языках и еще недавно даже не подозревавшие о существовании друг друга, получили возможность общаться почти напрямую и практически в реальном времени!

Наконец, в последние годы на этом поле появился новый крупный игрок. Речь идет о США. До недавних пор в этой практически одноязычной стране уделялось сравнительно мало внимания вопросам автоматического перевода. Но сейчас в контексте борьбы с международным терроризмом Министерство обороны США выделяет громадные средства на разработки в этой сфере: соответствующие ассигнования на 2008 г. составили 75 млн долларов. Поставлена задача создать систему автоматического устного перевода с английского на разные иностранные языки, которая могла бы работать в боевой обстановке. Уже разработана версия одностороннего перевода, которая использовалась в Ираке и Афганистане. В перспективе планируется создание более совершенных систем двустороннего устного перевода.

Сегодня перевод – самая быстрорастущая отрасль. В мире работает около 5 тыс. переводческих компаний (если считать организации со штатом не меньше 5 человек). В 2005 г. доходы переводческой отрасли составили, по разным оценкам, от 8 млрд долларов до 30 млрд евро. Столь большой разброс в цифрах объясняется тем, что в переводческой отрасли широко практикуется субподряд (до четырех уровней!), и часто бывает трудно установить фактическую сумму затрат на перевод по всей цепочке подрядов.

Но традиционный, «ручной», перевод дорог и не оперативен. Посмотрим же, что предлагают нам наука и технологии – сегодня и в перспективе – в области обработки естественного языка, куда входит и межъязыковой перевод. (Здесь мы не говорим о художественном переводе, который в принципе не поддается автоматизации. Мы не касаемся и так называемых «карманных переводчиков», которые на самом деле не более чем обычные разговорники, только с дисплеем и менее удобные.)

Хорошее представление о современном уровне работ в этой области дает большой проект *TC-STAR* (Technology and Corpora for Speech to Speech Translation), финансируемый Европейским Союзом. Вот основные направления и достижения этого проекта.

Автоматическое распознавание речи. Это очень сложная научная задача – намного сложнее, чем, например, игра в шахматы. К настоящему времени уже созданы системы, способные воспринимать связную речь (причем разных людей) и переводить ее в печатный текст. Сейчас решается задача построения системы, способной распознавать не диктовку, а естественную, непринужденную речь.

Преобразование текста в речь, иначе говоря, синтез речи. Уже существуют достаточно эффективные системы этого рода, причем в последние годы достигнуты большие успехи в воспроизведении естественного голоса. Ведутся работы по синтезированию разных голосов.

Машинный перевод, или письменный перевод с одного языка на другой. На первом этапе задачи машинного перевода пытались решать на основе лингвистики и детального кодирования человеческих знаний. Это оказалось очень сложным делом, и сейчас чаще используются статистические методы, которые уже хорошо зарекомендовали себя в системах распознавания речи. Речь идет о том, что в систему загружается очень большой словарь лексических единиц, с которыми и сравнивается переводимый текст. Процент ошибок и в том, и в другом случае довольно высок. Фактически машина выдает сырой текст, который требует «ручного» редактирования. Но главное преимущество таких систем – оперативность и дешевизна по сравнению с переводом, выполняемым человеком. В Интернете такие системы широко применяются для перевода веб-сайтов и чатов.

Перевод устной речи на одном языке в текст на другом языке. В этих системах два компонента: распознавание устной речи и машинный перевод. Трудность создания подобных систем состоит в том, что они должны работать в таких условиях, где говорящие в принципе не думают о том, чтобы быть понятыми машиной, например на телевидении, в ходе совещаний и т.д.

Перевод «речь в речь». В такой системе объединены все вышеуказанные методы. Она должна позволять людям разговаривать друг с другом на разных языках, используя в качестве переводчика компьютер. Крайне сложная задача – сохранить индивидуальный стиль говорящего. Уже имеются опытные системы, обеспечивающие довольно высокое качество перевода, и они применялись для передачи сути выступлений во время дебатов в Европейском парламенте.

На сегодняшний день самым практичным является «машиноподобный» перевод на основе так называемой *переводческой памяти*. В отличие от систем машинного перевода здесь базовый словарь лексических единиц не закладывается в систему заранее, а формируется самим переводчиком на протяжении всей его карьеры. Компьютер же лишь сохраняет эту информацию. В ходе работы над очередным переводом компьютер выступает как «переводческая память», подсказывая варианты перевода слов или фраз на основе ранее переведенных текстов. В подобной системе сам переводчик, а не машина, решает, насколько ему подходит предложенный вариант. Понятно, что система совершенствуется по мере роста ее памяти. Чаще всего такие системы применяются в больших переводческих проектах, когда параллельно работают несколько человек и их тексты сводятся в общую базу данных. Особенно удобны они при переводе больших юридических текстов или технических описаний. Среди наиболее известных систем этого рода можно назвать *Wordfast*, *Déjà Vu* и *SDL Trados*.

Мы лишь в самом начале пути, и разглядеть горизонты будущего не дано никому. Но уже появилась надежда, что в новом, глобализующемся мире нам удастся сохранить большинство живых языков, установив между ними электронные мосты перевода.