ВОПРОСЫ ЯЗЫКОЗНАНИЯ

№ 2 1973

# ДИСКУССИИ И ОБСУЖДЕНИЯ

К. Б. БЕКТАЕВ, С. К. КЕНЕСБАЕВ, Р. Г. ПИОТРОВСКИЙ

## ОБ ИНЖЕНЕРНОЙ ЛИНГВИСТИКЕ

Двадцать лет отделяет нас от появления первых публикаций, посвященных решению инженерно-лингвистических задач — задач перевода текста с помощью ЭВМ [1]. Трудно было бы найти такие отрасли науки, которые за первые два десятилетия своего существования перенесли столько драматических коллизий, сколько их пережила инженерная лингвистика. Достаточно вспомнить могучий всплеск машинно-переводческих идей, захвативших в конце 50-х годов не только молодых, но и опытных языковедов среднего и даже старшего поколения,— всплеск, сменившийся в середине 60-х годов глубоким разочарованием, в результате которого большинство отечественных и зарубежных пионеров машинного перевода ушло из инженерной лингвистики. Вместе с тем, говоря о неудачах 50-х и первой половины 60-х годов, нельзя забывать о том, что эта «романтическая эпоха» инженерной лингвистики дала начало многим направлениям современного языкознания. Достаточно назвать здесь статистико-комбинаторное моделирование, глубинный синтаксис, исследования по формализации семантики (смысл — текст).

В небольшой статье трудно оценить все многообразие лингвистических идей, родившихся в связи с включением в сферу функционирования естественного языка электронно-вычислительной машины. Мы ставим перед собой более скромные задачи: во-первых, рассмотреть перспективы использования ЭВМ при решении лингвистических задач, во-вторых, обсудить некоторые вопросы лингвистической теории, возникающие в связи с функционированием языка в новой для него коммуникационной системе «человек — машина — человек».

Рассматривая эти вопросы, мы будем ориентироваться лишь на работающие в экспериментальном и промышленном режиме лингвистические программы, оставляя в стороне все нереализованные на ЭВМ «бумажные» алгоритмы. Такое ограничение, как мы увидим дальше, имеет принципиальное значение с точки зрения теории и практики инженерной лингвистики.

Хотя возможности машины и человека неоднократно обсуждались в научной литературе среди языковедов, в том числе и тех, кто занимается вопросами автоматизации, до сих пор не сложилось единого мнения по поводу использования ЭВМ: одни полны скепсиса по поводу применения

[1] См. сб. «Машинный перевод», М., 1957, стр. 305—306, В. П. Берков, Б. А. Ершов, О попытках машинного перевода, ВЯ, 1955, 6; N. Macdonald, Language translation by machine. A report of the first successful trial, «Computers and automata», III, 2, 1954; И. К. Бельская, Л. Н. Королев, И. С. Мухин, Д. Ю. Панов, С. Н. Разумовский, Автоматический перевод с одного языка на другой на электронно-счетной машине, сб. «Труды Всесоюзного математического съезда (Москва, июнь — июль 1956)», 4, М., 1959, стр. 93.

машин в лингвистике, другие считают вполне возможным решать на ЭВМ сложные синтаксические и даже стилистические задачи. Поэтому, говоря об итогах и перспективах применения ЭВМ в лингвистике, напомним еще раз о реальных возможностях человека и машины при решении разного вида лингвистических задач.

Известно, что скорость переработки информации внутри ЭВМ в несколько сотен и даже тысяч раз превосходит скорость преобразования информации в мозгу человека (память ЭВМ БЭСМ-6 перерабатывает информацию в 20 тыс. раз быстрее, чем человеческий мозг). Хотя эта скорость ЭВМ несколько обесценивается более медленным функционированием вводящих и выводящих информацию устройств, скорость ЭВМ при переработке текста намного превосходит возможности человека [2].

Машина обладает большей по сравнению с человеком помехоустойчивостью и надежностью. Если функционирование мозга при переработке текста (например, при его реферировании или переводе) зависит от физиологического и психического состояния человека, а также быстро нарушается в результате утомляемости, то работа ЭВМ характеризуется большой устойчивостью и продолжительностью (например, среднее время безотказной работы ЭВМ «Минск-22» находится в пределах 50 часов).

Для машины характерно устойчивое хранение информации. Напротив, человек при отсутствии тренировки постепенно забывает усвоенный материал. Это преимущество ЭВМ имеет принципиальное значение: однажды составив программу переработки текста и записав текст на ленту, мы получаем возможность обращаться к его переработке через много месяцев и лет.

Вместе с тем ЭВМ имеет два существенных недостатка, ограничивающих ее возможности при решении лингвистических задач.

Во-первых, память современных ЭВМ на несколько порядков меньше памяти человека. Память ЭВМ «Минск-22» равняется примерно $6 \cdot 10^7$ двоичных единиц (дв. ед.) информации, для «Минска-32» этот объем достигает $6 \cdot 10^8$ дв. ед., для ЕС-1020 он составляет $2 \cdot 10^8$ дв. ед., а суммарный объем памяти БЭСМ-6 приближается к $1.7 \cdot 10^9$ дв. ед. Емкость же памяти человека находится, очевидно, в пределах $10^{10}$—$10^{13}$ дв. ед. информации. Иными словами, объем памяти наиболее мощных машин в пять-шесть раз меньше нижнего порога памяти человека. Что же касается средних ЭВМ, к которым в настоящее время имеют более или менее постоянный доступ языковеды, то их память в несколько сотен и даже тысяч раз меньше памяти человека. Малые объемы памяти ЭВМ явно не соответствуют тому количеству информации, которая содержится в естественном языке. Построенная в группе «Статистика речи» действующая система лингвистического обслуживания, включающая автоматический англо-русский словарь по одной отрасли знания, программы автоматического индексирования, аннотирования — перевода и реферирования, а также вспомогательные программы, полностью занимает оперативную и внешнюю память ЭВМ «Минск-22». Если же нужно получить не просто пословный перевод-подстрочник и информационный образ текста, которые выдает только что упомянутая система, но связный и грамматически правильный перевод, то выполнение такой задачи потребует памяти порядка $2 \cdot 10^9 \div 8 \cdot 10^9$ дв.

[2] Машина «читает» текст с перфокарт или с перфоленты со скоростью примерно 1500 символов/сек., в то время как скорость чтения у человека не превышает 25 символов/сек. Стандартное печатающее АЦПУ-128 выводит лингвистическую информацию из ЭВМ со скоростью около 900 символов/сек., новое печатающее устройство ЕС-7030 выдает машинные результаты со скоростью 1400 символов/сек. Скорость выдачи информации человеком через устную речь или скоропись не превышает 10 символов (букв, фонем)/сек.

ед. Объемы памяти этого порядка имеют только машины большой мощности, например, советская БЭСМ-6, американские IBM-370 и IBM-360 [3].

Во-вторых, между «электронным мозгом» и высшей нервной деятельностью человека имеются качественные различия, которые гораздо в большей степени, чем количественные расхождения, ограничивают лингвистические возможности ЭВМ. Структурная единица машинной памяти — ячейка (машинное слово) может одновременно принимать или передавать информацию только одной такой же ячейке. В связи с этим современная машина решает лингвистические задачи по принципу строгой очередности, последовательности, одноканальности передачи информации [4] и ее поэлементной запоминаемости. Напротив, структурная единица мозга человека — нейрон — может, с одной стороны, передавать через концевые ответвления своего аксона нервный импульс одновременно тысячам других нейронов, а, с другой, этот нейрон может получать импульс от тысяч других нейронов, с которыми он имеет прямую связь через свои дендриты. В итоге, нейроны через свои связи с другими нейронами как бы замыкаются сами на себя, образуя нейронные кольца, которые в свою очередь объединяются в системы нейронных колец [5]. В результате каждый нейрон входит сразу в несколько нейронных колец, причем возбуждение одного нейрона в одном кольце может активизировать другие кольца и системы. Такое предположение о построении и механизмах человеческой памяти не только проливает свет на способность мозга к ассоциативной записи и выборке информации, но и хорошо согласуется с современными представлениями о парадигматической структуре языка — структуре, образующейся путем взаимопересечения разных парадигм. При записи этой структуры в мозгу человека каждая лингвистическая единица входит одновременно в целый пучок парадигм или, как их называл Ф. де Соссюр, ассоциативных групп [6]. Это взаимоналожение парадигм, являющееся предпосылкой для метафорического употребления лингвистических единиц в речи и развития у них многозначности, определяет существование языка как открытой, вечно изменяющейся системы. Вместе с тем, взаимоналожение парадигм позволяет человеку осуществлять целенаправленный лингвистический поиск, не прибегая к сканированию (последовательному перебору) материала, как это делает лишенная ассоциативной памяти ЭВМ [7].

---

[3] Более подробные численные оценки возможностей современных ЭВМ с точки зрения решения различных лингвистических задач см. в работах: Р. Г. Пиотровский, Экстралингвистические и внутриязыковые вопросы при переработке текста в системе «человек — машина — человек», сб. «Вопросы социальной лингвистики», Л., 1969, стр. 42—44; его же, Машинный перевод (некоторые итоги и перспективы), сб. «Проблемы структурной лингвистики», М., 1972; ср.: А. П. Крайзмер, С. А. Матюхин, С. Г. Майоркин, Память кибернетических систем (основы мнемологии), М., 1971, гл. III—XI.

[4] Появление мультипрограммных машин (Минск-32, БЭСМ-6, ЕС-1020) практически ничего не меняет в общих принципах машинного решения лингвистических задач, поскольку возможности обмена информацией между параллельно работающими в ЭВМ программами весьма ограничены.

[5] О кольцевых связях в нервной системе см.: А. Н. Радченко, Моделирование основных механизмов мозга, М., 1968, стр. 6 и 103—148.

[6] Ф. де Соссюр, Курс общей лингвистики, М., 1933, стр. 123—124.

[7] Разработки приемов ассоциативного программирования и ассоциативных запоминающих устройств ведутся уже давно; ср.: Г. Аронов, В. Щеглов, Кодирование информации и ее ассоциативный поиск, сб. «Электронно-вычислительная техника и программирование», 4, М., 1971, стр. 73—79; F. Chew Woo, Plated wire content-addressable memories with bit-steering technique, «IEEE transactions of electronics and computers», XVI, 5, 1967; N. Y. Findler, On a computer language which simulates associative memory and parallel processing, «Cybernetica», X, 4, 1968. Однако ЭВМ, имитирующих ассоциативную память человека, пока не существует.

Чем тоньше и сложнее лингвистическая задача, тем большее число парадигм и дистрибуций, а также реализующих их нейронных колец (или систем колец) вовлекается в одновременный ассоциативный поиск. Если такая сложная задача решается на машине, то весь пучок расчленяется и превращается в последовательность отдельных парадигм и дистрибуций, которые записываются одна за другой в памяти ЭВМ. Одновременный ассоциативный целенаправленный поиск заменяется последовательным перебором со слабой целенаправленностью [8]. При этом быстрота действия ЭВМ, являющаяся основным преимуществом машины перед человеком, растрачивается на сканирование материала. В этих случаях медленно работающий человеческий мозг благодаря использованию целенаправленного ассоциативного поиска избегает непроизводительных затрат и решает сложную семантико-синтаксическую задачу скорее, чем это делает машина.

Напротив, когда речь идет о «рутинных» лингвистических задачах, т. е. о таких задачах, которые предусматривают поиск и упорядочение лингвистических объектов по одной, максимум двум-трем парадигмам, машинное алгоритмическое их решение дает более точный и быстрый результат, чем ассоциативный поиск, осуществляемый человеком.

Наибольший эффект дает использование ЭВМ при выборе из текста и сортировке большого числа лингвистических объектов, при условии, что этот выбор и сортировка осуществляются на основе небольшого числа формальных критериев [9]. Без этих первичных рутинных переработок текста практически невозможно любое серьезное лексикографическое, статистико-грамматическое, фонетико- и графемно-статистическое, лексико-статистическое, а теперь и методическое исследование.

Не касаясь автоматических обработок текста, проводившихся в интересах прикладной лингвистики (машинный перевод, информационный поиск, оптимизация преподавания иностранных языков) [10], остановимся подробнее на некоторых результатах применения ЭВМ при решении чисто лингвистических задач.

Одним из основных направлений в исследованиях Института языкознания АН Казахской ССР является нормализация современного литературного языка, лингвистическое освоение произведений современных писателей и литературного наследия классиков казахской литературы, а также лексикографическое изучение древнетюркских памятников.

В ходе планирования этой работы стало ясно, что получить в обозримый период (7—10 лет) исчерпывающее описание норм современного казахского литературного языка и исследовать историю его лексики можно при условии, если к этой работе будет привлечено несколько сотен языковедов. При этом основная затрата труда и времени ложится на просмотр

---

[8] О современных методах целенаправленного поиска, сортировки и упорядочения лингвистической информации см.: К. И. Курбаков, Кодирование и поиск информации в автоматическом словаре, М., 1968; В. С. Крисевич, Серийный автоматический поиск при обработке информации с помощью ЭВМ, сб. «Статистика текста», II — Автоматическая переработка текста, Минск, 1970.

[9] Делались попытки использовать для этих целей разного вида электромеханические устройства, в частности счетно-аналоговые машины (САМ); ср.: B. Quémada, La mécanisation dans les recherches lexicologiques, «Cahiers de lexicologie», I, Bésançon, 1959; Й. Келемен, Машинный сбор и обработка данных на службе лексикографии, «Лексикология и лексикография», М., 1972, стр. 10—12. Однако эти приемы лингвистической механизации значительно уступают по точности и скорости обработке языкового материала на ЭВМ и поэтому серьезных перспектив развития не имеют.

[10] См.: А. В. Зубов, Переработка текста естественного языка в системе «человек — машина», «Статистика речи и автоматический анализ текста», Л., 1971, стр. 307 и сл.; Л. А. Попова, М. А. Щукина, Создаем частотные словари-минимумы, «Вестник высшей школы», 1972, 4, стр. 32.

миллионов страниц текста, выбор из них отдельных словоформ и словосочетаний, а затем на их сортировку и упорядочение, т. е. на такую нетворческую работу, которая легко поддается автоматизации. Собственно научный лексикографический и историко-лингвистический анализ этого материала занимает не более 10% времени, затрачиваемого на разработку лексикографической проблемы.

Не располагая необходимыми штатными возможностями и одновременно считая нецелесообразным растрачивать труд высококвалифицированных специалистов на рутинную работу, казахские лексикографы переложили работу по выборке, упорядочению и пересчету лингвистических единиц на ЭВМ «Минск-22». Машине были поручены следующие операции по первичной лексикографической и морфологической обработке публицистических, научно-технических текстов и произведений казахских классиков: а) составление алфавитно-частотных словников; б) построение обратных словарей; в) составление алфавитно-частотных списков словосочетаний.

В течение 1969—1972 гг. с помощью ЭВМ получены следующие результаты: составлен на машине и издан «Обратный словарь казахского языка»; готовится выпуск «Алфавитно-частотного словаря романа М. Ауэзова „Путь Абая“», полученного на основе текстов объемом около полумиллиона словоупотреблений; составлены частотные, алфавитно-частотные, обратно-частотные словари по пьесам М. Ауэзова, публицистическим, художественным, научно-техническим текстам по казахскому эпосу, а также по древнетюркским текстам объемом более одного миллиона словоупотреблений; на обширном материале проведен статистический анализ морфологической структуры существительных, глаголов и прилагательных; определены информационные характеристики (меры) казахского языка; исследованы статистика и дистрибуция графем казахского языка.

Алфавитно-частотные словники и списки словосочетаний используются при построении толкового и фразеологического словаря современного казахского языка. Обратный частотный словарь является фундаментом для полного описания морфологии современного казахского языка, для отбора материала в нормативную грамматику, а также при осуществлении информационных измерений грамматики и лексики в современном тексте.

Об эффективности работы машины по первичной обработке текста можно судить по следующему примеру. Словник для романа М. Ауэзова «Путь Абая» получен двумя научными работниками с помощью ЭВМ за 8 месяцев. Для выполнения этой задачи вручную двум человекам понадобилось бы 10 лет изнурительной механической работы.

Кроме того, ЭВМ используется в Институте языкознания АН Казахской ССР при исследовании древнетюркских памятников, а также графемной, фонемной и лексической комбинаторики киргизских, узбекских, каракалпакских и других тюркских текстов [11].

При выполнении описанных выше лингвистических заданий машина производит вместо человека «однопарадигматические» мыслительные операции, сравнения лингвистических объектов с точки зрения их графического тождества, осуществляет операцию суммирования, а также подменяет человека, сортирующего вручную словарные карточки. Выполняя эти рутинные операции, ЭВМ заметно облегчает труд языковеда, экономя его время и предоставляя возможность концентрировать свои усилия на лингвистическом анализе собранного и упорядоченного машиной материала.

---

[11] С. К. Кенесбаев, Р. Г. Пиотровский, К. Б. Бектаев, Инженерная лингвистика и тюркология, «Советская тюркология», 1970, 6, стр. 5—7.

Наряду с выполнением рутинных лингвистических задач современные ЭВМ обладают и некоторыми «гуманистическими» возможностями. В результате формализации и алгоритмизации эвристических лингвистических задач машина способна извлекать из текста смысловую информацию, переводя ее на информационный или любой естественный язык. На этом пути реально осуществляется машинный перевод, автоматическое индексирование, аннотирование и реферирование.

Используемые здесь алгоритмы и программы опираются на сравнение формальных признаков текста с предварительно заложенными в ЭВМ формализованными семантическими эталонами. При этом используется два подхода — иконический и алгоритмический.

Идея иконического подхода состоит в прямом отождествлении текстового сегмента, состоящего из одного или нескольких контактирующих словоупотреблений, с аналогичным сегментом в памяти ЭВМ,— сегментом, значение которого заранее формализовано и выражено средствами машинного, информационного или какого-либо естественного языка. На иконическом подходе строятся, например, автоматические двуязычные словари, в том числе словари машинных оборотов. Этот подход особенно удобен в тех случаях, когда распознается смысл текстов, порожденных подъязыками, представляющими собой закрытую и строго фиксированную совокупность лингвистических единиц (словоупотреблений, словосочетаний, предложений). Примером могут служить английский и русский подъязыки переговоров «земля — воздух» («аэродром — пилот»), каждый из которых состоит из фиксированного числа фраз-штампов (в английском подъязыке их 755, в русском — 820) типа англ. *may I take off?* — русск. *разрешите взлет*. Аналогичным образом строятся подъязыки коммуникативных систем «земля — вода», «вода — вода».

Построенная на иконическом принципе система большого англо-русского автоматического словаря (АС) группы «Статистика речи» [12] уже сейчас может быть использована для распознавания смысла и перевода английских и русских текстов, порожденных подъязыками указанного типа.

Иконический подход является примером машинного решения лингвистической задачи с позиции «грубой силы». В память ЭВМ закладывается большая входная и выходная информация (например, упомянутый выше англо-русский АС, включая около 15 тыс. входных словоформ и оборотов, содержит до 300 тыс. выходных лексических единиц). Зато алгоритмы отождествления словоформ текста с единицами словаря и выдачи выходных результатов сравнительно просты, поскольку ЭВМ рассматривает каждую словоформу как неизменяемую самостоятельную лексему. При этом как входной, так и выходной языки (в нашем случае английский и русский) превращаются в лишенные морфологии «изолирующие» языки.

При алгоритмическом подходе закладываемая в ЭВМ информация сравнительно невелика. Например, составленный в группе «Статистика речи» экспериментальный немецко-русский АС для перевода текстов по электронике содержит около 3 тыс. немецких и такого же количества русских машинных основ [13]. Зато алгоритмы морфологического анализа и синтеза здесь не только сложны, но требуют часто искусственного изменения типологии входного и выходного языков. Например, при построении алго-

[12] В. А. Вертель, Е. В. Вертель, В. С. Крисевич, Р. Г. Пиотровский, Л. И. Трибис, Автоматические словари в системе бинарного вероятностного МП, сб. «Инженерная лингвистика» («Уч. зап. ЛГПИ им. А. И. Герцена», 458), 1971.

[13] М. Г. Зореф, Машинные основы и машинная морфология в немецко-русском автоматическом словаре. Автореф. канд. диссерт., Кишинев, 1972, стр. 18.

ритма немецко-русского (или французско-русского) морфологического перевода лингвист должен однозначно указывать границы между машинными основами и окончаниями, часто не совпадающими с традиционными основами и флексиями. При этом фузионная структура индоевропейского слова заменяется «квазиагглютинативным» членением.

Действующие промышленные системы автоматической переработки текста представляют собой обычно комбинацию алгоритмического и иконического подходов. Примером может служить система лингвистического обслуживания группы «Статистика речи», которая объединяет построенную на алгоритмическом принципе программу русского аннотирования и индексирования английского научно-технического текста [14], и иконическую программу его подстрочного перевода (ср. выше).

Дальнейшее развитие автоматической переработки смысловой информации идет, во-первых, по линии устранения многозначности лексических и грамматических единиц текста, опирающегося на анализ их контекстного окружения, во-вторых, по пути разработки синтаксического анализа и синтеза предложения и, наконец, по линии тезаурусного реферирования текста[15]. Первые экспериментальные алгоритмы этих типов уже реализованы на ЭВМ «Минск-22» и БЭСМ-4.

Программы извлечения и формализации смысловой информации имеют первостепенное значение с точки зрения развития информационного дела и автоматизации управления. С помощью этих программ может быть преодолен барьер между потоками научно-технических и деловых документов, написанных на естественных языках, с одной стороны, и призванными принимать и обрабатывать эти потоки информации автоматическими системами управления (АСУ) и автоматическими информационными системами (АИС), с другой. Существо этого барьера заключается в том, что, будучи созданными не алгоритмической памятью ЭВМ, но интуитивно-ассоциативным мыслительным аппаратом человека, научно-технические и деловые тексты имеют неформализованный и поэтому непонятный для машины вид. Распознавание смысла документа и его формального представления в виде индекса, аннотации или реферата на понятном машине информационном языке осуществляется в настоящее время в АСУ и АИС вручную. Работа эта требует больших затрат времени и средств, а результаты ее во многом зависят от квалификации, опыта, внимания специалиста-индексатора. Поэтому, если на компановку, сортировку и анализ всего множества машинных индексов документов в АСУ и АИС уходят секунды и минуты, то на ручное индексирование и аннотирование документов уходят часы и даже сутки. В итоге основное преимущество АСУ и АИС, заключающееся в больших скоростях переработки текстовой информации, по существу, сводится на нет.

Преодолеть этот парадокс можно только путем автоматического р а с п о з н а в а н и я с м ы с л а документа и формализованного представления этого смысла на языке машины (или, если это необходимо, на естественном языке). Эту задачу и призвана выполнять описанная выше программа индексирования, аннотирования и реферирования текста.

Если обратиться к теоретическим проблемам инженерной лингвистики, то основным ее методологическим вопросом является отношение машинных моделей к языковой действительности. С одной стороны, реше-

[14] См. :. Н. П. Р а х у б о, Е. С. Т а р а с о в а, А. Н. П о п е с к у, К вопросу об автоматическом индексировании и аннотировании научно-технических текстов, «Уч. зап. Калининск. гос. пед. ин-та им. М. И. Калинина», 81, 1971.

[15] См.: А. Н. П о п е с к у, М. С. Х а ж и н с к а я, Тезаурусный метод составления алгоритма для автоматического реферирования французского научно-технического текста, «Автоматическая переработка текста», Кишинев, 1972.

ние с помощью ЭВМ лингвистических задач служит надежным средством логической экспликации и моделирования различных лингвистических объектов и их функций. Действительно, если языковое явление смоделировано в форме алгоритма, и этот алгоритм, будучи перенесенным с ассоциативного субстрата человеческого мышления на последовательно-логический механизм ЭВМ, устойчиво выдает правильный лингвистический результат, можно не сомневаться, что некоторая существенная сторона интересующего нас языкового явления понята правильно [16].

С другой стороны, встает вопрос о том, насколько адекватно могут описать машинные модели лингвистическую действительность.

Несовместимость ассоциативной эвристики человеческого мозга и последовательно-алгоритмического функционирования ЭВМ, о которой мы говорили выше, вызывает к жизни две лингвистические антиномии. Первую назовем условно парадоксом человека и робота, вторую обозначим как парадокс Ахиллеса и черепахи.

Лингвистический парадокс человека и робота состоит в следующем. Согласно второй теореме Гёделя, непротиворечивость данной формальной системы можно доказать только с помощью другой, более мощной системы. Но непротиворечивость этой второй системы может быть показана только методами третьей, еще более сильной системы и т. д. Стремясь к стопроцентной формализации языка, мы должны создать на основании нашего неформализованного эвристического знания языка и его описаний предельно мощную формализацию $L$. Однако эта формализация $L$ обязательно будет содержать выражение $\Phi$, которое окажется неразрешимым в системе $L$. Вместе с тем более мощное описание языка $L'$, в котором можно было бы разрешить выражение $\Phi$, мы построить уже не в силах, поскольку возможности нашего неформализованного эвристического знания исчерпаны.

Не имея возможности создавать все более и более мощные формальные модели языка, асимптотически приближающие нас к его стопроцентной формализации, мы лишены возможности строить машинные модели языка, практически близкие к его стопроцентной формализации. Неполнота машинной формализации особенно отчетливо проявляется при машинном семиосисе. Образуемый в ЭВМ искусственный знак, состоящий из означающего естественного языка и означаемого, включающего формализованные значения, всегда беднее знака естественного языка.

Лингвистический парадокс Ахиллеса и черепахи, отражающий соссюровскую антиномию синхронии и диахронии, состоит в том, что формальное закрытое описание языка, ориентированное на такой синхронный срез, который совпадает с началом разработки этой формализации, оказывается в результате действия диахронических процессов, действующих в открытой системе естественного языка, несколько устаревшим к моменту реализации этого описания на ЭВМ. Этот парадокс служит еще одним препятствием при построении стопроцентного машинного описания языка.

Однако неполноту машинной формализации языка не следует рассматривать как свидетельство бесперспективности инженерно-лингвистических исследований. Инженерная лингвистика сосредоточила свое внимание на построении таких машинных моделей, которые перерабатывают, в основном, легко формализуемые и алгоритмизуемые стороны языка. Что касается тонких, неформализуемых или трудноформализуемых лингвистических задач, то их решение по-прежнему должно осуществляться человеком.

[16] Без реализации на ЭВМ никогда нельзя быть уверенным в логической корректности алгоритма, даже в том случае, если он прошел «ручную» проверку, т. е. проверку, осуществляемую «квазилогическим» автоматом — человеческим мозгом. Именно поэтому мы обсуждаем только прошедшие машинную реализацию алгоритмы.