

Понимание естественного текста в свете теории информации

Е.Г.Борисова

1. Моделирование понимания – часть динамического описания языка

Последние полвека лингвистика не ограничивается отражением системной организации языка. Все чаще описание языка предполагает моделирование речевой деятельности, что принято называть динамическим, поскольку оно отражает язык в динамике. Одним из первых (1) динамических описаний была порождающая модель Н.Хомского. К их числу можно отнести модель «Смысл – Текст», функциональные грамматики различных типов. Такой подход, развивавшийся преимущественно под влиянием практических потребностей прикладных направлений языкознания, позволил обнаружить недостаточность системного описания языка, т.е. такого, в котором устанавливалась зависимость между выражением и содержанием языкового знака. Выяснилось, что для прогнозирования правильного употребления знака нужна еще дополнительная информация. Такое описание получило название активных грамматик.(2)

Динамическая модель языка предполагает отражение таких видов деятельности, как говорение, письмо (активные, или продуктивные виды) и слушание и чтение (рецептивные виды). Хотя функциональные направления обращают внимание преимущественно на первые два вида деятельности, рецептивная деятельность тоже требует специального описания и моделирования (3). Это очевидно для систем анализа в компьютерной лингвистике, где распознавание сообщения представляет не менее (если не более) трудную задачу, чем порождение. В еще большей степени это верно в перлокутивном описании языка, т.е. при учете возможностей воздействия на слушателя. То же можно сказать и об описании языка с позиций редактирования. (4)

Наконец, важно иметь в виду, что достаточно подробное, отражающее важные свойства языковых единиц описание порождения речи связано с описанием ее восприятия. Порождение включает выбор говорящим наиболее оптимальных средств выражения своего замысла. А это невозможно без оценки возможного понимания нескольких вариантов: оптимально то, что наиболее легко и однозначно понимается слушающим.

Таким образом, необходима модель понимания сообщения, которая учитывала бы те особенности этого вида речевой деятельности, которые изучены в настоящее время.

Мы предлагаем здесь отдельные соображения, которые могли бы быть учтены при работе над такой моделью. Источником сведений, с нашей точки зрения, могут быть психолингвистические опыты с одной стороны, и наблюдения над процессом исправления текстов (редактированием) с другой. Во втором случае можно опираться на результат этих действий – опубликованные тексты.

2. Информационная основа процессов понимания

Понимание нередко трактуется как процесс обработки информации (5), на выходе которого – изменение картины мира адресата за счет добавления полученных сведений. Если принять такую точку зрения, затруднения в понимании следует интерпретировать как трудности в переработке информации. Отмечая различные типы затруднений, мы можем определить

основные параметры этого процесса и те типы информации, которые учитываются адресатом.

Перечислим те типы информации, содержащиеся в сообщении, которые уже в той или иной степени отмечались исследователями. Во-первых, выделяется лингвистическая и экстралингвистическая информация. Последнее – это сведения о мире, или энциклопедические знания, как это принято называть в компьютерной лингвистике, или культурный фон (бэкграунд), как это называется в лингвокультурологии. Сюда относится собственно то новое в сообщении, ради чего осуществляется коммуникация, а также общие знания участников ситуации, передача которых или навязана структурой языка и узусом (например, уменьшительность во фразе «*Пробейте билетик*», что связано с вежливостью), или необходима для «привязки» сообщения к нужному фрагменту картины мира (в нашем примере – информация о фрейме «поездка в автобусе»).

Лингвистическая информация представляет собой операционные сведения, необходимые для обработки сообщения. Сюда относится информация, содержащаяся в номинативном значении граммем (число, вид глагола), сведения, необходимые для референции участников ситуации и установлении отношений между ними, а также между отдельными событиями, упоминаемыми в сообщении (последовательность, обусловленность и т.п.).

Сюда же следует отнести и сведения по коммуникативной организации высказывания – операциональная информация о способах понимания сообщения. Как показывают исследования, необходимость в передаче такой информации возникает только при отступлении от определенных норм (дополнительное выделение темы или ремы *Уж он-то сделает*, снижение информационной насыщенности ремы *Туда он и пошел* и т.п.).

Наконец, следует выделить прагматическую информацию, включающую сведения о том, как расценивает говорящий взаимоотношения коммуникантов, эмоциональную окраску, возникающие ассоциации и т.п.

3. Процессы обработки информации

Таким образом, получение в ходе коммуникации сообщаемой информации означает переработку информации всех названных выше типов различными механизмами мозга. Перечислим соответствующие процессы и их представления, выработанные такими областями знания, как психология, психолингвистика, языкознание, лингвопрагматика, культурология.

Механизмы восприятия знака в психологии (когнитивной) делятся на рецепцию – возбуждение определенных рецепторов и семантизацию – сопоставление с паттернами, хранящимися в памяти, и опознание знака. В речи происходит одновременное распознавание сложных знаков – частей слов, слов целиком, иногда их сочетаний (6).

Лингвистические модели предусматривают целый ряд процедур, позволяющих извлечь из сообщения содержательную информацию, основываясь на операциональной, т.е. на сведениях об устройстве языка. Лингвистические модели описывают и различные приемы референции. Все это называется лингвистическим пониманием в противовес пониманию психологическому. Последнее предполагает обработку содержательной информации с целью изменения картины мира. Сюда входит и обращение к культурному фону, т.е. сведениям о мире.

Мы перечислили основные действия адресата при понимании, которые в обязательном порядке включают во все описания. Однако уже давно замечено, что деятельность адресата (слушающего) не сводится только к перечисленным процедурам. Адресат постоянно делает собственные выводы, позволяющие восполнить недосказанное, перейти от отдельных значений слов к смыслу сообщения, определить его истинность, выявить намерения говорящего и подготовить базу для ответа (7).

Частично это осознаваемые действия, которые могут быть вербализованы, т.е. адресат может проговорить вывод для себя: «Ага, он это сказал, значит, он показывает, что это больше не секрет от меня. Наверное, он хочет показать увеличение доверия ко мне. Небось, хочет таким образом что-то выведать» и т.п. Эта возможность относится преимущественно к тому виду понимания, которое принято называть психологическим, хотя и в этом случае по большей части этот процесс не осознается.

В части же лингвистического понимания осознание процесса характерно только при нарушении нормальных условий коммуникации, а это плохая слышимость, необычные слова, запутанно построенная фраза (последнее относится к письменной речи, так как в устной возврат к непонятой фразе практически невозможен).

Видимо, следует предусмотреть и такие действия слушающего, как прогноз содержания сообщения, основанный на восприятии ее части. Прогноз должен проверяться. Такие действия адресата позволяют воспринимать текст с ошибками, иногда даже не замечая их.

4. Нарушение понимания в свете теории информации

4.1. Из сказанного следует, что нарушение понимания может происходить, если какой-то знак не поддается расшифровке, например, не опознано слово или известные значения слов не удалось соединить осмысленным образом, например, *раскурить люльку* непонятно тем, кто не знает диалектного значения слова *люлька* – трубка. Однако и в этом случае нередко не наступает полного прекращения понимания. Во-первых, сообщение обычно бывает информационно избыточно, что позволяет восстановить нерасшифрованные места на основании других сообщений. Например, фраза со словом *ангрейд*, неизвестным большинству носителей русского языка, будет понятна в контексте: *Подвергните ваш компьютер ангрейду, и вы получите новые возможности для работы*. Во-вторых, при понимании всего сообщения иногда бывает возможно опустить какую-то непонятную часть.

В любом случае данное явление не вызывает вопросов относительно причины такого непонимания. Здесь интересны способы восстановления смысла, однако подробно мы это рассматривать не будем.

4.2. Второй случай непонимания происходит при формально сохраненном тексте, все части которого могут быть поняты, однако при работе в нормальном режиме – при чтении, слушании объяснения и т.п. он вызывает затруднения: для его понимания нужно вернуться к тексту еще раз, адресат не в состоянии перейти от лингвистического понимания к психологическому, понимание отстает от скорости восприятия информации. Все такие случаи так или иначе замечались редакторами, научными руководителями студентов, имиджмейкерами, оценивающими речь политика, и методистами, разбирающими урок педагога.

Мы предлагаем следующее объяснение нарушению понимания. Поскольку все типы информации при обработке хранятся в определенном порядке (иногда это устройство называется кратковременной, или оперативной памятью), затруднение понимания возникает тогда, когда это устройство переполняется. Понятно, что так происходит, если фраза оказалась слишком громоздкой, и до конца ее понимания надо держать в голове слишком много информации.

В таком случае можно вспомнить о «магическом числе 7 плюс-минус 2», которое, по мнению многих, характеризует объем оперативной памяти. Имеется в виду семь единиц хранения, которые в свою очередь могут быть сложными знаками, например, словами.

Если учитывать все виды информации, которые были перечислены выше и которые перерабатываются теми или иными механизмами при восприятии сообщения, окажется, что в целом приходится работать одновременно с большим объемом информации. И переполнение памяти может наступить за счет разных видов операций.

Рассмотрим с этой точки зрения ряд типичных случаев затруднения понимания, отмечаемых редакторами и специалистами по редактированию.

4.3. Невыраженность темы. Известный вид ошибок – отсутствие сведений для привязки новой информации к уже известной. Например, если в этом месте текста, который вы сейчас читаете, возникнет фраза *Говорящий использует в своей речи маркеры новизны*, адресат будет в затруднении относительно привязки этой новости к предыдущему контексту. Однако, скорее всего он догадается сразу или поймет смысл фразы из дальнейшего изложения.

И вообще, в большинстве случаев эта задача не является неразрешимой, и адресат правильно устанавливает связи сказанного с известным и осуществляет необходимую референцию. Сложность заключается в том, для этого требуются дополнительные усилия, которые препятствуют пониманию текста с нормальной скоростью. В этом смысле очень большую роль играют такие средства управления адресатом, как маркеры известности и новизны (частицы *же, ведь, а* и т.п.). Они позволяют заранее оценить необходимость тех или иных действий слушающего и сэкономить ему объем операционного устройства.

Видимо, с этими же задачами связана и такая ошибка как излишнее повторение одного и того же слова. Принято считать, что такое повторение делает текст «некрасивым». Иногда это действительно так. Однако в основном такие сложности возникают при составлении документов и других текстов, для которых эстетический аспект не столь значителен.

За этим запретом можно увидеть преодоление препятствия для отождествления имени с денотатом. Когда денотат попадает в тексте вторично, для его обозначения помимо называния имени можно использовать местоимения или другие описательные средства. Имя же существительное может использоваться и для нового денотата. Поэтому в соответствии с постулатами речевого общения его лучше не применять без необходимости (при вторичном назывании денотата имя создает больше неопределенности, чем местоимение). Ср. *На обочине мы заметили прохожего. Прохожий был высокий, крепкий, по виду – крестьянин, прохожий сразу выделялся среди толпы. Мы задали прохожему обычный вопрос...* Практически во всех случаях можно догадаться, что речь идет об одном и том же человеке. Однако поскольку для этой догадки требуются усилия, отрывок представляется неудачным с точки зрения выбора средств выражения.

Нечто аналогичное может происходить и с глаголом. Если вам нужно сообщить об одном и том же событии, то надо специальными средствами (частицами *же*, *вот* или местоимениями и наречиями *это*, *так*) показать, что слушающему уже должно быть известно об этом действии.

4.4. Длина предложения. Известно, что распространенной ошибкой является использование слишком длинных фраз. Здесь затруднение для понимания очевидно: в голове приходится держать все составные части предложения, что иногда переполняет объем оперативной памяти.

Однако нередко такой же ошибкой считают и слишком короткие фразы, рубленый текст. Так выглядит заметка школьника: «Мы позавтракали. Потом мы пошли на экскурсию. Мы пошли пешком. Шли два часа. Все немного устали. Когда пришли, нам рассказали о героях-панфиловцах. Мы пошли на поле. Там герои-панфиловцы остановили немецкие танки. На этом месте стоит памятник». При том, что текст понятен, он неудобен для чтения, требует каких-то нежелательных усилий.

Однако запрет на краткие фразы неабсолютен: в некоторых случаях столь короткими могут быть не только фразы, но и абзацы.

Не претендуя на окончательность выводов, мы предлагаем такую трактовку проблемы. Каждое законченное предложение требует определенной работы с этой единицей восприятия: необходимо обработать информацию о связях внутри предложения, о связях с предшествующим текстом. Кроме того, необходимо восстановить всю имплицитную информацию.

Дробное членение текста на предложения заставляет выполнять эту работу очень часто. Если автор имеет в виду значительный подтекст, малая длина фразы не вызывает возражений: работа и с небольшим предложением может оказаться достаточной. Однако нередко авторы таких "рубленых" не закладывают во фразы столько, и попытки извлечь имплицитную информацию, оказываются "холостыми проворотами". А поскольку они делаются, то все-таки загружают механизмы оперативной обработки сообщения и вызывают затрудненность понимания.

По ходу дела заметим, что характерная для детей и не очень грамотных людей рубленность текста соответствует скорости письма: более длинные фразы они не могли бы удерживать в голове из-за большого времени, которое им требуется на их написание. Если бы чтение шло так же медленно, то краткость фраз не вызывала бы возражений. Любому читателю достаточно вспомнить чтение на начальном этапе овладения иностранным языком, где из-за гораздо большего, чем при чтении на родном языке, количества операций возможный размер обрабатываемой фразы вообще невелик.

5. Из сказанного можно сделать ряд выводов для дальнейших усилий по построению модели понимания.

Во-первых, эта модель должна учитывать активную деятельность слушающего по переработке получаемого сообщения. В число совершаемых действий должны быть включены не только дешифровка знаков, но и самостоятельные выводы, прогнозирование действий говорящего и т.п.

Во-вторых, можно представить себе деятельность адресата как переработку информации. При этом надо учитывать, что информация может быть многих разных видов.

В-третьих, затруднения в понимании можно рассматривать как результат превышения некоторого объема памяти, что может быть вызвано

необходимостью большого количества операций по обработке информации. Единицей обработки можно считать в большинстве случаев предложение.

Из последнего тезиса следует, что рассмотрение возможных нарушений процесса понимания может служить основой для выводов относительно конкретных операций.

Примечания

1. Мы опираемся на терминологию и информацию из Алпатов 2000.
2. См. Борисова, А.Н.Латышева 2003.
3. В явном виде эта задача была сформулирована и до некоторой степени выполнена И.Г.Милославским в работе Милославский 2002.
4. Создание модели языка, соответствующей потребностям редактирования текстов, еще не стало задачей особого направления лингвистики. Однако имеющиеся разработки, например, Сенкевич 1968, показывают, что этот взгляд на язык может быть не менее плодотворен, чем в любом другом прикладном направлении языкознания, см. Борисова 2002.
5. Эта вполне распространенная точка зрения отражена, к примеру, в Солсо 1998.
6. См. Солсо цит.раб.
7. См. Имплицитность 2000.

Литература

- Алпатов В.М. История лингвистических учений. М. ЯРК. 2000.
- Борисова Е.Г. Формализуемые аспекты неудачи в понимании текста. – "Компьютерная лингвистика и интеллектуальные технологии". Труды Международного семинара Диалог – 2002. М. "Наука". 2002.
- Борисова Е.Г..Латышева А.Н Лингвистические основы РКИ. (Педагогическая грамматика русского языка). М. Флинта-Наука. 2003.
- Имплицитность в языке и речи. М. ЯРК. 2000.
- Милославский И.Г. культура речи и русская грамматика. М. "Ступени. Инфра-М" 2002.
- Сенкевич М.П., Феллер М.Г. Литературное редактирование. М. "Высшая школа". 1968.
- Солсо Р.Л. Когнитивная психология. М. "Тривола. 1998.