

ПСИХОЛИНГВИСТИКА

А.В. Венцов, В.Б. Касевич, Е.В. Ясунова (Санкт-Петербургский гос. ун-т)

ОРФОГРАФИЧЕСКИЙ ТЕКСТ И ТРАНСКРИПЦИЯ

С ТОЧКИ ЗРЕНИЯ ВОСПРИЯТИЯ РЕЧИ*

В настоящем сообщении излагаются материалы, связанные с очередным этапом разработки подходов к моделированию в Лаборатории применения вычислительных средств в гуманитарных исследованиях СПбГУ процессов восприятия речи. Напомним, что работа ведется на материале компьютерного корпуса русского языка, который на сегодняшний день составляет более миллиона словоформ (словоупотреблений), и составленного на базе данного корпуса частотного словаря словоформ русского языка (фактор частотности в данном сообщении не отражен).

В предыдущих сообщениях (Венцов, Касевич 1998; Kassevitch, Ventsov, Yagoulova 2000 и др.) описывались некоторые элементы общего подхода к проблеме, а также работа компьютерной программы автоматического расчленения на слова письменного русского текста, лишенного пробелов. Смысл последней задачи, как нетрудно видеть, заключается в том, чтобы имитировать важнейшее свойство звучащего текста – отсутствие границ между словами, выраженных сколько-нибудь систематически. Правда, в таком тексте «отсутствует» и интонация, которая в значительной степени способствует установлению границ между синтагмами и предложениями, а, следовательно, и частями словесных границ. В какой-то мере интонационные характеристики передаются знаками препинания, если их сохранять в беспробельном тексте. В нашем случае все знаки препинания были сведены к одному символу, который сигнализировал таким образом наличие некой не специфицированной границы.

Алгоритм сегментации был реально алгоритмом сегментации через идентификацию цепочки символов текста проверялись на их соответствие словарным единицам (здесь нужно напомнить, что наш словарь – это словарь словоформ, а не лексем). Использовалась простая версия модели когорт. Существенным отличием от известных версий было последовательное обращение к акцентным характеристикам слов: начальными сегментами слов, тождество которых (начальных сегментов) служило основанием для включения в когорту (класс слов-кандидатов), выступали не просто первые фонемы или кластеры, а предударно-ударные сегменты. Основной принцип работы программы заключался в том, что программа «отодвигала» правую границу потенциального слова до тех пор, пока и само это слово могло получить интерпретацию в терминах единиц словаря, и «остаток» цепочки, вошедшей в буфер (9 символов), тоже мог получить такую интерпретацию.

Не вдаваясь в детали работы программы, отметим лишь, что она оказалась чрезвычайно эффективной: примерно 98 % границ были простираны адекватно. Существенно отметить, что «прогон» программы осуществлялся на материале только лишь текстов из состава нашего корпуса, а потому исключались ситуации, когда программа работала бы со словами, не вошедшими в словарь. Это – серьезное ограничение, которое еще не позволяет смоделировать другую ситуацию: когда в текст входят новые слова (что особенно важно при усвоении языка в онтогенезе); пока с этим ограничением приходится мириться.

Из немногочисленных ошибок, допущенных программой, выделялись те, что совершенно явно не возникали бы, если бы мы имели дело не с графемами, а с фонемами, представленными соответствующими акустическими реализациями. Например, союз и достаточно регулярно присоединялся к глаголу в прош. времени ед.ч. муж.р., т.е. вместо тил и гулят мы получали тили гулят. Эта ошибка, разумеется, вызвана именно графическим (орфографическим) характером текста.

* Работа выполнена при частичной поддержке РФФИ (грант № 00-06-80247).

В любом случае последовательное приближение к работе со звучащим текстом предполагает, в качестве наиболее естественного развития исследовательской стратегии, обращение к транскрибированному тексту, как к некоторому приближению к звучащему тексту. Можно сказать, что по субстанции транскрипция есть графика, а по внутренним структурным соотношениям это уже фонетика.

Для преобразования графического представления текста и словаря в представление транскрикционное был использован один из первых вариантов транскриптора, разработанного на филологическом факультете Московского государственного университета им. М. В. Ломоносова в начале 90-х гг. Используемый тип транскрипции, в целях разумного приближения к облику звучащей речи, носит компромиссный характер: наряду с символами фонем в нем представлены символы основных аллофонов. Несколько подробнее это свойство транскриптора будет обсуждаться чуть ниже.

Прежде чем излагать результаты обращения к транскрибированной версии текста, кажется необходимым вкратце обсудить некоторые проблемы транскрипции и ее места в описании процессов речевой деятельности.

Прежде всего стоит напомнить, что широко распространенное представление о «фонетической транскрипции» в противопоставлении фонологической не вполне логично, если под фонетической транскрипцией понимать отражение реального звучания соответствующих единиц. Дело в том, что транскрипция как таковая не может не быть в той или иной степени фонологической – уже потому, что она по необходимости *дискретна*, в то время как речь принципиально *континуальна* (*непрерывна*). Следовательно, прибегая к дискретным символам любой традиционной транскрипции, мы уже вносим тем самым элемент фонологического представления.

Выше было упомянуто, что с данным транскриптором мы переводили в транскрикционное представление как тексты, так и словарные единицы. Это принципиально важное положение. Как уже было сказано, основным принципом работы программы было обращение к словарю, т.е. сличение цепочек символов текста с единицами словаря. Этот принцип с очевидностью предполагает, что текст и словарь используют *один и ту же систему единиц*, в противном случае сличение, наложение единиц, естественно, невозможно. Альтернативным решением могло бы быть приведение текстовых единиц к виду, сопоставимому с обликом словарных единиц. В сущности, именно на этом основывались многие теории восприятия речи, которые в качестве первого шага соответствующих процедур постулировали перекодирование «сырой» акустической информации текста (звукющей речи) в цепочки фонем, в терминах которых, предположительно, записаны и словарные единицы. В этом случае сопоставляются преобразованные в фонемный формат единицы текста с исходно фонемными экспонентами словарных единиц.

Такие представления, однако, входят в противоречие с низкой разрешающей способностью слухового анализатора человека, которая делает перекодирование звучащей речи в фонологическое представление как пофонемное ее (речи) кодирование до обращения к словарю крайне малозэффективной процедурой и, отсюда, процедурой маловероятной – по крайней мере, в качестве основной стратегии речевосприятия. Нельзя также не учитывать и обычно низкое качество речевого сигнала в наиболее частых ситуациях речевого общения.

Транскриптор, которым мы воспользовались с целью получения хотя бы предварительных материалов для сравнения двух версий работы программы сегментации – в графической и транскрикционной модальности, был разработан авторами (О. Ф. Кривновой и др.) для нужд автоматического синтеза речи. Насколько мы можем судить, использование или неиспользование тех или иных аллофонов оправдывалось прежде всего эмпирически: набор аллофонов вместе с набором отвечающих им акустических реализаций должен быть достаточным для получения качественного синтеза речи. Какова роль того же набора символов для анализа, восприятия речи?

Можно предположить, что каждый символ транскрипции отвечает одному из перцептивных эталонов, а некоторое множество эталонов, в свою очередь, отвечает фонеме (хотя возможны и более сложные схемы). Правда, данный транскриптор (как и все прочие,

известные нам) заведомо не покрывает всего потенциального множества перцептивных эталонов хотя бы потому, что в нем не представлены аллофоны для начально- и конечнослоговых согласных, а эти аллофоны, как известно, существенно различаются, и вполне вероятно, что им соответствуют самостоятельные перцептивные эталоны.

Чрезвычайно существенный вопрос заключается в том, что не вполне ясно, на каком этапе происходит обращение к фонемному коду, если мы осуществляем обращение к словарю и, следовательно, идентификацию словарных (лексических) единиц без использования фонемного кода.

Так же нeясна и роль аллофонов для реализации наиболее интересующей нас задачи – выявления межсловных границ. Транскриптор работает, естественно, с орографическим представлением текста, где слова разделены пробелами. Транскриптор различает аллофоны гласных прежде всего по степени и типу редукции. Редукция же, как известно, во многом определяется по отношению к ударению и границам слова. Отсюда, казалось бы, возникает возможность опираться при определении места словесной границы на тип аллофона. Это, однако, означало бы принятие положения Трубецкого о существовании «афонематических пограничных сигналов». Согласиться с этим положением трудно – фактически, невозможно – по соображениям принципиального характера. Информация о границах слова – это морфологическая информация, а аллофонические характеристики не могут «напрямую» связываться с морфологическими, это прерогатива характеристик фонологических. О том же говорят и эмпирические наблюдения. Если мы не различаем 1-е заударное *a* в *ракета* казалась и 2-е предударное в *ракета* оказалось (т.е. не различаем данные словосочетания) – при том, что «сильнее различаются все предударные от всех заударных, чем 1-й, 2-й и 3-й предударные между собой» (Бондарко 1979: 251), то это как раз и говорит о том, что при всей реальности различия аллофонов они не используются носителем языка как маркеры (афонематические) словесных границ.

Соответственно и в программу не вносились операции, которые использовали бы тип аллофона для определения межсловных границ (хотя сами по себе аллофонные различия использовались, поскольку они были отражены в словаре, хотя не слишком последовательно и не в полной мере).

Приведенный выше пример может одновременно служить иллюстрацией различий в результатах приложения программы к орографическому тексту, с одной стороны, и транскрибированному, с другой. Ясно, что для орографического представления проблема омонимии двух словосочетаний не возникает, в то время как для транскрипции она вполне реальна. Другого рода различия можно видеть на примере, который уже приводился: *тиль и гуляя*, который программа, работающая с орографическим текстом, представляет как *тиль гуляя*. В отличие от этого, приложение программы к транскрибированному тексту дает *тиль гуляя*. Иначе говоря, в обоих случаях программа делает (неустранимую – до привлечения информации более высоких уровней) ошибку, просто эти ошибки существенно разные (в первом случае ошибка дает меньшее семантическое искажение, но от этого мы, конечно, отвлекаемся).

Приведем один из примеров, где транскрибированный текст дает «лучшие» результаты по сравнению с орографическим (знак «плюс» указывает на ударность предыдущей гласной):

а мот⁺и мут⁺ж челове⁺к изве⁺стный лу⁺чшее го⁺ не найти⁺
а мот⁺и мут⁺ш ч'лав⁺э⁺к ызв⁺эн⁺й лу⁺тшъ й'иво⁺ н'и най⁺т'и⁺

Вполне очевидно, это различие объясняется тем, что в транскрипции вместо последовательности *его имеем ииво*, что и ведет к другой разбивке на слова.

При обращении к транскрипции – во всяком случае типа той, которую дает используемый транскриптор, – мы должны учитывать еще одну очень непростую проблему. Как уже упоминалось, на выбор аллофонов влияет позиция относительно позиции ударного слога и границ слова; но слово это – фонетическое. Не следует ли из этого, что и единица лексикона –

фонетическое слово? Разумеется, лексикон при этом разрастается, хотя, как показывают предварительные подсчеты, не более чем вдвое.

Мы не будем давать подробный сопоставительный анализ вариантов членения при переходе от орфографического представления текста к транскрипционному. Самый важный результат заключается в том, что для обоих вариантов текста эффективность программы примерно одинакова. В обоих случаях ошибки составляют приблизительно 2 %, хотя распределение ошибок, как мы видели, меняется.

Это – по крайней мере косвенно – говорит о том, что несмотря на существенное различие в субстанции речи, несмотря на несомненное частичное несовпадение стратегий обработки текста, общность стратегий, очевидно, перевешивает.

Вероятно, полученные результаты свидетельствуют также о том, что русская орфография вряд ли создает какие-то существенные трудности для восприятия орфографического текста по сравнению со звучащим (если считать, конечно, что транскрипция дает хорошее приближение к звучанию; кроме того, надо оговорить, что транскрибированию подвергается орфографический «учебный» текст, т.е. такой, в котором проставлены ударения и введена буква ё).

Подчеркнем в заключение еще раз некоторые вопросы, принципиально важные с теоретической точки зрения.

Мы уже отмечали, что в контексте наших задач возникает фундаментальный вопрос о месте фонологического представления и, тем самым, о соотношении кодов. Разные коды, вполне вероятно, реальны и внутри префонологического модуля, работающего с «сырой» акустической реальностью, поскольку характеристики сигнала сильно варьируют в зависимости от стиля речи, темпа и т.д. Отсюда следует, что в идеале нужен не один транскриптор, а некоторое их семейство.

Наконец, полная модель должна предусматривать и некие «линки», обеспечивающие переход от орфографического представления к фонологическому и наоборот. Как не раз отмечалось, у носителей языка, обладающих достаточно богатым опытом оперирования письменным текстом, фонологическое и орфографическое представления, соответствующие коды с неизбежностью интерферируют. (В частности, эмпирический опыт свидетельствует, что мы зачастую «окончательно осознаем» слово тогда, когда представляем себе его написание.) Все эти сложные соотношения тоже должны получить отражение в полной модели как восприятия, так и порождения речи.

Как всегда, вопросов здесь гораздо больше, чем ответов – да и не все вопросы получили эксплицитные формулировки.

ЛИТЕРАТУРА

- Бондарко Л.В. Редукция // Русский язык: Энциклопедия. М., 1979.
Венцов А.В., Касевич В.Б. Словарь для модели восприятия речи // Вестник СПбГУ. 1998. Вып. 3.
Kassevitch V.B., Ventsov A.V., Uzlosova E.V. The simulation of continuous text perceptual segmentation: A model for automatic segmentation of written text // Язык и речевая деятельность. 2000. Т. 3. Ч. II.