

О процедурах обработки фонетического слова при восприятии речи¹

В.Б.Касевич, Е.В.Ягунова
С.-Петербург

Авторами экспериментально показано, что в русском языке нет полной однозначности в противопоставлении фонетического слова и словосочетания за счет одно-/двуударности. Таким образом, фонологическое противопоставление слова словосочетанию в русском языке оказывается в некотором роде нечетким, допускающим варианты (подклассы знаменательных слов различаются по способности к клитисизации. Вероятно, представление о редукции должно быть расширено: общеизвестные положения о редукции сегментных единиц следует пополнить представлением о *просодической редукции* – в частности, *редукции ударения*. В представленном докладе планируется рассмотреть возможности нейтрализации противопоставления слова словосочетанию при *восприятии псевдотекста*, т.е. при восприятии текста в условиях отсутствия значения (внефонетической информации). Обращение к псевдотексту, получаемому путем систематической замены согласных исходного текста их парадигматическими «соседями», обеспечивает чистоту эксперимента – исключение обращения к внефонетической информации. Изучение проблемы фонетического слова ведется с использованием как инструментально-фонетических, так и аудитивно-перцептивных методик.

Заявленная тема предполагает рассмотрение как минимум двух нетривиальных вопросов. Первый касается понятия фонетического слова (ФС) с фонетической и фонологической точек зрения, второй относится к области процедурной: как и на каких этапах восприятия речевого сигнала человеком используется такая единица, как ФС.

Принятое в литературе универсальное определение ФС оказывается в сущности, тавтологическим: фонетическое слово есть отрезок текста, вычлняющийся по некоторым фонетическим (фонологическим) признакам; иначе говоря, «фонетическое слово есть слово, определяемое фонетически». Если изъять из определения термин *слово* (который сам нуждается в дефиниции), то под приведенное (традиционное) определение логически подпадают самые разные единицы: слог, морфема, ФС, синтагма и даже предложение.

В языках разного строя признаки ФС окажутся, естественно, различными. Для сингармонистических языков границы ФС совпадут со сферой действия правил сингармонизма, для тональных они, вероятно, будут связаны с закономерностями тонального сандхи и т.д. В акцентных языках ФС, как известно, выделяется по наличию ударения (сколько ударений – столько ФС).

Словесные ударения, акцентные контуры слов лишь создают условия для реализации просодической структуры текста. Просодическую структуру ФС можно считать своего рода фундаментом, над которым «надстраиваются» просодические структуры более высоких уровней, связанные с синтагмами и предложениями (Касевич и др. 1990). Просодические структуры более высоких уровней налагаются на слова, следовательно, их нельзя описать, не владея информацией о просодическом устройстве ФС. Таким образом, сам подход к описанию просодики текста – без чего невозможны представления о речевой деятельности – критическим образом зависит от представления о просодическом устройстве слова в языке.

¹ Работа выполнена при частичной поддержке РФФИ (грант № 03-06-80251)

Мы ограничим свое рассмотрение материалом русского языка. В русском языке ФС образуются знаменательными словами или же знаменательными словами с их клитиками – предлогами, частицами. При восприятии речи вычленение фонетических слов можно считать некоторым приближением к сегментации текста на слова как лексико-грамматические единицы: ФС не имеют самостоятельной функциональной значимости с точки зрения установления грамматических и семантических связей в тексте. Должны существовать механизмы перехода от ФС к лексико-грамматическому слову. Относительная сложность этих процедур зависит от «расстояния» между ФС и лексико-грамматическим словом. Но в первую очередь следует определить «степень надежности» вычленения самого ФС.

С целью выяснения указанного параметра был проведен ряд экспериментов. В частности, исследовалась возможность «перцептивной нейтрализации» противопоставления одного ФС двум. Наши данные свидетельствуют: в русском языке нет полной однозначности в противопоставлении ФС и словосочетания за счет одно-/двуударности. Имеет место своего рода нейтрализации противопоставления слова словосочетанию в русском языке (Касевич, Ягунова 2003). Об этом говорят результаты перцептивного эксперимента в ситуации выбора, в котором предъявлялись стимулы, соответствующие одному ФС или словосочетанию из двух знаменательных слов, совпадающему с первым членом пары по фонемному составу и месту ударения во втором слове, ср. *бархана – бар хана, вампир – вам пир*. Различение рассматриваемых пар для носителей языка представляет немалую сложность, для отдельных стимулов число ошибок доходит до 78% (изолированное предъявление членов пары) и 75% (предъявление в составе фразы с нейтральным – допускающий обе интерпретации – контекстом). Наше инструментальное исследование акцентной структуры слов и словосочетаний (долготных характеристик) также подтверждает гипотезу о возможности отсутствия просодического противопоставления слова и словосочетания. Например, на нашем материале можно видеть, что для членов пар типа *поварешка – пана Леика*, находящихся в конечнофразовой позиции, отсутствует противопоставление по соотношению длительностей ударного и безударного слога /ра/ (даже несмотря на то, что безударный слог стоит в позиции второго предударного и традиционно описывается как гласный второй степени редукции). Другой пример - постпозитивные местоимения, которые реализуются то как ударные, то как безударные (из них *вы* оказывается «наиболее безударным»). Вероятно, наряду с общеизвестными положениями о редукции сегментных единиц следует ввести представление о *просодической редукции* – в частности, *редукции ударения* (Касевич, Ягунова 2003).

Таким образом, уже в вычленении ФС носитель языка, скорее всего, пользуется вероятностными процедурами – доступная акустическая картина не дает ему абсолютно надежных ключей для однозначной сегментации текста на ФС.

В любом случае членение на ФС – это членение «второго порядка». Членение «первого порядка» – это членение на *перцептивные синтагмы*. Текст – сложным образом структурированный объект, который линейно развертывается во времени; в каждый момент времени слушающий может иметь дело лишь с некоторым сегментом развертывающегося текста, в то время как в конечном счете распознаванию и пониманию подлежит весь объект в целом (в наших исследованиях мы ограничиваемся уровнем «поверхностного восприятия», не рассматривая процессы понимания). Вполне очевидно, что звучащий текст не может быть воспринят слушающим целиком в режиме реального времени (в силу нейрофизиологических ограничений человеческой памяти и быстродействия). Членение на перцептивные синтагмы и выступает сегментацией «первого порядка». Оно осуществляется «принудительно» – именно потому, что, как сказано, слушающий должен превращать континуальный текст в последовательность «квантов», доступных для анализа.

Такие «кванты» должны удовлетворять следующим требованиям: во-первых, их объем должен быть приемлемым с точки зрения избранной слушающим стратегии восприятия и его нейропсихологических возможностей (прежде всего, объема оперативной памяти); во-вторых, вероятно, они должны быть интерпретируемы как сравнительно самостоятельные синтактико-семантические структуры. На данном этапе именно задача сегментации выдвигается на первый план: выделяемые сегменты-синтагмы очевидно не могут быть идентифицированы в целом через отождествление с единицами словаря, ибо подобных словарей нет и быть не может.

Напомним в этой связи основные результаты серии работ по перцептивной сегментации звучащего текста (Касевич и др. 1993). Исходная гипотеза работы как раз и заключалась в том, «что, воспринимая звучащую речь, человек ее каким-то образом сегментирует на линейные «кванты», которые в ситуации нормального естественного общения служат основанием для текущей перекодировки поверхностно-текстовых структур в структуры глубинно-семантические». Эксперимент проводился в трех разных модификациях, в качестве которых выступало восприятие интактного, беспазуального и семантически дезорганизованного текста (Венцов и др. 1993; Венцов, Касевич 2003; Слепокурова 1999). Показательно, что даже в условиях искажения (беспазуальная и семантически дезорганизованная модификации) около 87% перцептивных

синтагм оказались синтаксически структурированными единствами, т.е. эти перцептивные синтагмы завершаются в те моменты, когда могла бы произойти естественная смена синтаксических групп внутри предложения с точки зрения человека, текущим образом воспринимающего звучащий текст (Слепокурова 1999). Разумеется, весьма существенно, что синтаксические группы характеризуются соответствующим просодическим оформлением. Можно заключить на основании экспериментальных данных, что членение на перцептивные синтагмы обеспечено достаточно высоким уровнем надежности.

Обработка перцептивных синтагм, как следует из сказанного, принадлежит уровню восприятия, иерархически более высокому, чем уровень, на котором слушающий имеет дело с ФС. «Ниже» уровня обработки ФС естественным образом располагается, казалось бы, уровень обработки сегментных единиц языка – гласных и согласных. Однако в действительности это не совсем так. Текст (речевая цепь) принципиально неоднороден: лишь часть сегментов характеризуется полным типом произнесения и, по крайней мере, *может* интерпретироваться за счет самого по себе анализа фонетической данности, большая же часть принадлежит неполному типу и может быть распознана лишь в результате работы механизмов контекстной предсказуемости (т.е. *после* идентификации более крупной структурной единицы, чем фонема или слог – фактически слова). Иначе говоря, распознавание гласных/согласных и слов – это, в типичном случае, неразрывно связанные процедуры. На первых стадиях восприятия и распознавания речи человек определяет лишь самые общие, грубые, абстрактные фонетические и внефонетические характеристики выделяемого сегмента, которые в дальнейшем подлежат уточнению, развитию, а, возможно, и корректированию.

В этой связи возникает теоретически и практически очень важный вопрос о *соотношении* фонетической и внефонетической информации при восприятии речи. В данном сообщении описываются некоторые из предварительных результатов серии экспериментов по восприятию *псевдотекста*, лишенного внефонетической информации². Псевдотекст был получен путем замены в письменном варианте исходного текста согласных на их ближайшие аналоги (сонанты на другие сонанты и т.п.). Затем опытному диктору после длительной тренировки было предложено, прослушивая исходный текст, прочитать псевдотексты (две таблицы) с той же интонацией.

Изучение влияния просодической структуры на восприятие текста в нашем эксперименте было дополнено изучением влияния грамматической (или фонетико-грамматической структуры), опирающейся на флексии и предлоги как грамматические показатели. Два варианта псевдотекста отличались тем, что в одном (текст 2) операция замены не затрагивала флексии и предлоги, в другом (текст 1) исключений для грамматических показателей не делалось.

Общий объем текста – 3 предложения, 49 фонетических слов, 216 слогов. Испытуемыми (20 человек на каждую таблицу) выступали 40 студентов-филологов. Изучалась разборчивость текстов в терминах слогов, слов и т.д. Предметом данного сообщения является роль признаков ритмической (просодической) структуры.

В табл. 1 приведены данные по опознанию параметров ритмических структур (для ФС).

Таблица 1. Правильное опознание параметров ритмической структуры ФС (%)

	место ударения	число слогов
Текст 1	31	26
Текст 2	38	27

Признаки ритмической структуры ФС обладают сравнительно высокой распознаваемостью (особенно по сравнению со словесной разборчивостью). Показатели для текста 2 несколько выше, чем для текста 1 (см. табл. 1).

Параметр «место ударения» распознается лучше, чем число слогов. Это хорошо согласуется с известными данными по восприятию фильтрованной речи, речи в шуме и т.п. (Чистович, Кожевников 1965; Касевич и др. 1990). В то же время наши данные указывают не только на «перцептивные утраты» или перераспределения слогов, но и на появление дополнительных ударений, равно как и их утрат, ср. результаты по восприятию слов и словосочетаний, описанные выше. В таблице 2 приведены данные по ошибочному опознанию числа ФС (количества ударений) применительно ко всей совокупности ответов. Хотя слияние ФС представлено небольшим числом случаев, сам факт такой возможности кажется очень важным.

² В проведении эксперимента принимала активное участие студентка СПбГУ Людмила Глобус.

Таблица 2. Ошибочное опознание числа ФС

	случаи замены одного ФС на несколько	случаи замены нескольких ФС на одно
Текст 1	88 (9%)	12 (1%)
Текст 2	113 (12%)	5 (1%)

В таблице 3 приведены данные по числу тех ФС предъявляемого псевдотекста, которые допускают ошибочное опознание числа ударений в слове, т.е. вариации числа ФС, хотя бы для одного из испытуемых. Случаи возможной потери ударения, напр., *спОномы годобОна* → *с полногодовО*, ставят вопрос о роли соотношения «основное (фонологическое) vs. дополнительное ударение в процедурах восприятия речи».

Таблица 3. Число ФС предъявляемого текста, допускающих ошибочное опознание числа ФС

	число ФС текста, заменяемых на несколько ФС	число сочетаний ФС текста, заменяемых на одно ФС
Текст 1	25 (51%)	5 (10%)
Текст 2	40 (82%)	9 (18%)

Таким образом, наши данные по восприятию псевдотекста подтверждают выводы, сделанные на материале стимулов словесной и фразовой размерности: в русском языке нет полной однозначности в противопоставлении ФС и словосочетания за счет числа ударений. Имеет место своего рода нейтрализации противопоставления слова словосочетанию в русском языке.

Можно видеть также, что собственно фонетические параметры текста, сегментные и супraseгментные, в целом достаточны для получения информации, необходимой для адекватного распознавания текста – как минимум, для обеспечения фонетической базы восприятия. Необходимо еще установить, каково должно быть соотношение сегментов полного и неполного типа произнесения для обеспечения указанной задачи.

Уже минимальное структурирование псевдотекста за счет сохранения в нем флексий и предлогов создает предпосылки для более успешного распознавания текста. Это лишний раз свидетельствует о том, что в перцептивных процессах все аспекты текста вступают в сложное взаимодействие, природа которого, конечно, ясна далеко не полностью.

Литература

1. Бондарко Л.В., Вербицкая Л.А., Гордина М.В., Зиндер Л.Р., Касевич В.Б. Стили произношения и типы произнесения // ВЯ, 1974, №2, с.64-70.
2. Венцов А.В., Касевич В.Б. Проблемы восприятия речи. М., 2003.
3. Венцов А.В., Касевич В.Б., Слепокурова Н.А. Перцептивная сегментация звучащего текста // Проблемы фонетики I. М., 1993.
4. Зиндер Л.Р. Реальный поток речи и «реконструкция» фонемного состава слова // Теория языка: Методы его исследования и преподавания. Л., 1981.
5. Касевич В.Б. Семантика. Синтаксис. Морфология. М., 1988.
6. Касевич В.Б. Еще о понятии фонетического слова // Проблемы фонетики IV. М., 2001.
7. Касевич В.Б., Рыбин В.В., Шабельникова Е.М. О стратегиях сегментации текста (на материале китайского, японского и русского языка) // Востоковедение, 18. Л., 1993
8. Касевич В.Б., Рыбин В.В., Шабельникова Е.М. Ударение и тон в языке и речевой деятельности. Л., 1990.
9. Касевич В.Б., Ягунова Е.В. Ударение и фонетическое слово в русском языке // Проблемы социо- и психолингвистики. Вып.3. Пермь, 2003, с.19-25.
10. Слепокурова Н.А. Лингвистический анализ «перцептивных синтагм» // Материалы XXXI межвузовской научно-методической конференции преподавателей и аспирантов, СПб., 1999, вып.16. 1999, с.38-41.
11. Чистович Л.А., Кожевников В.А. и др. Речь. Артикуляция и восприятие. М.;Л., 1965.