

Ингве В. Гипотеза глубины // Новое в лингвистике. Вып. 4. М.: Прогресс, 1965. С. 126-138.

Кузнецова А.И. Морфемная глубина слова в русском языке / слова малой и большой глубины как периферийное явление русского языка // Проблемы структурной лингвистики. 1984. М.: Наука, 1988а. С. 71-93.

Кузнецова А.И. Параметрическое исследование периферийных явлений в области морфемки: Автореф. дис. ... д-ра филол. наук. М., 1988б. 38 с.

Курилович Е. Очерки по лингвистике. М.: Изд-во иностр. лит. 1962. 202 с.

Миракян Э.Б. Испытание гипотезы Ингве на материале армянского языка // Лингвистические исследования по общей и славянской типологии. М.: Наука, 1966. С. 206-213.

Москович В.А. Глубина и длина слова в естественных языках // Вopr. языкознания. 1967. № 6. С. 17-32.

Муравичкая М.П. Простое и сложное предложение в структуре реферативного текста // Лингвистические проблемы автоматизации редакционно-издательских процессов. Киев: Наук. думка, 1986. С. 66-97.

Оливерус З.Ф. Морфемы русского языка. Fr.: Universita Karlova, 1976. 198 в.

Слипченко Л.Д. Фонемная структура лексики английского языка: Автореф. дис. ... канд. филол. наук. Киев, 1974. 21 с.

Клименко Н.Ф. Про закономірності сполучуваності афіксів у сучасній українській мові // Мовознавство. 1968. № 6. С. 32-39.

Клименко Н.Ф. Глубина слова современной украинской мовы // Структурні рівні мови і методи їх дослідження. К.: Наук. думка, 1972. С. 17-21.

Клименко Н.Ф. Система афіксального словотворення сучасної української мови. К.: Наук. думка, 1973. 193 с.

Клименко Н.Ф., Карпіловська С.А. и др. Морфемно-словотвірний фонд української мови як дослідницька та інформаційно-довідкова система // Мовознавство. 1990а. № 5. С. 41-50.

Клименко Н.Ф., Карпіловська С.А. Морфемні структури слів у сучасній українській мові // Мовознавство. 1990б. № 6. С. 3-15.

Перебийніс В.С. Кількісні та якісні характеристики системи фонем сучасної української літературної мови. Киев: Наук. думка, 1977. 270 с.

Словник української мови. Т. I-XI. Киев: Наук. думка, 1970-1980.

Частотний словник сучасної української художньої прози. Т. I-II. Киев: Наук. думка, 1981.

Ліценко І.Т. Морфемний аналіз. Словник-довідник. Т. I-2. Киев: Наук. думка, 1980-1981.

Gruber J. Functions of the lexicon in formal descriptive grammars. Bloomington, 1972. 172 p.

Martinet A. Éléments de linguistique générale. P.: Librairie Armand Colin, 1961. 224 p.

А.И. Кузнецова

ПРОБЛЕМА ДЕКОДИРОВАНИЯ СЛОВА НА ОСНОВЕ АФИКСАЛЬНЫХ БИМОРФОВ

В последние годы проблемы автоматического распознавания, обработки и синтезирования текста активно привлекают к себе внимание лингвистов. При решении этих задач обсуждаются также вопросы кодирования и декодирования слова. Существующие способы кодирования (лемматизация, получение сверток из одних согласных с сохранением гласных первого слога - так называемый метод буквенного сжатия слова, усечение правой части слова до тех пор, пока имеется возможность для работающего с кодом в любой нужный момент однозначно восстановить слово по остающейся левой части, и др.) основаны на том, что зв'язки словаря распознавания берутся слова. Естественно, что если исходными компонентами декодируемых слов служат начальные сегменты слова, являющиеся своего рода идентификаторами слов и выполняющие одну-единственную функцию - быть однозначным средством спознания слова, то словарь получается весьма громоздким. Так, используя при кодировании слов методику усечения правой части слова, И.С. Караулов построил "Частотный словарь семантических множителей русского языка" (1960) с помощью ЭВМ, преследуя, впрочем, иную цель, чем простое декодирование слова. По словам А.С. Караулова, "правая граница кода никак не связана с морфемным членением слова, она может про-

ходить и внутри корня (поздр^{авить}_{авление}), и по его конечной бук-
ве (лад^{но}_{ить}), справа от нее (представи^{тель(ство)}_{тель(ный)})"

(Караулов, 1980, с.11). В работе осуществляется проверка "на омонимию квазиоснов (исходных сегментов слова, подлежащих развертыванию в слово - А.К.), преследующая цель избежать возможности неоднозначного декодирования какой-то из них" (основ) (Караулов, 1980, с.13). Так, предположительно выделенный сегмент рассм- допускает три продолжения: рассм ^{атривать}_{отрение} и, ^{ваться}

следовательно, должен быть заменен на рассмо-, рассме- (Караулов, 1980, с.13). Отказываясь от частного анализа грамматических элементов - морфем или частей речи, - автор указанной работы сосредоточил свое внимание на статистическом изучении словарного запаса. Однако при таком подходе не удается "выявить ясных отношений между наблюдаемыми статистическими закономерностями и вероятностной структурой языка в целом" (Караулов, 1980, с.18). Кроме того, при таком подходе для различных слов нарушается единство метода декодирования, так как исходным сегментом слова, подлежащим развертыванию в слово, могут быть и отдельные морфемы, и их сочетания, и часть основы, вообще не совпадающая ни с одним из морфов. Чаще всего исходным сегментом, подвергаемым декодированию, являются незначимые отрезки слов.

Метод декодирования, опознавания слова по начальным элементам можно применять на морфном уровне. Если за основу анализа текста брать морфный уровень, то текст представляется как последовательность морфов, появление которых может быть предсказано с большим или меньшим трудом. Распознавание морфов, следующих за исходным, становится все более легким по мере того, как увеличивается количество правильно угаданных морфов.

Проблема прогнозирования слова на основе начальных биморфов - это в начальной степени вопрос определения наиболее существенного признака, облегчающего процесс декодирования, распознавания слова. В качестве исходного сегмента слова на морфном уровне могут быть префикс + префикс (иначе - префиксальные биморфы), префикс + корень, корень + суффикс.

В настоящей статье анализу подвергаются только префиксальные биморфы. Когда центр тяжести декодирования слова переносится с уровня слов на уровень морфем, единицами словаря распознавания становятся не слова, а морфемы, что дает экономию памяти как минимум в 10 раз по сравнению со словарем слов и в 100 раз по сравнению со словарем словоформ (Асиновский, 1986).

После одних префиксальных биморфов декодировка слова затруднена, так как может быть получено много продолжений, после других - однозначна. Однозначно предсказуемые продолжения, естественно, сильно облегчают распознавание текста в целом.

Чтобы лучше понять специфику процесса развертывания префиксальных биморфов в слова, можно сопоставить этот процесс с декодированием слов, имеющих в своем начале другие морфемы: корень + суффикс, префикс + корень и префикс + префикс + префикс.

Наиболее продуктивными моделями русского языка (см. Кузнецова, Ефремова, 1986) являются 20 беспрефиксальных и 55 однопредфиксальных моделей; с помощью этих 75 наиболее продуктивных моделей образуется около 17 тыс. слов русского языка, в то время как приблизительно 35 тыс. слов образуется с помощью нескольких тысяч моделей. Модели, по каждой из которых образуется от 500 до 1000 слов, возможны только беспрефиксальные, но обязательно суффиксальные (суперактивной моделью является модель корень + окончание, которая включает около 2750 корней, функционирующих в качестве слова).

Корневые окружения (насчитывается от 100 до 300 корней) могут быть без префикса или с одним префиксом и, как правило, с одним суффиксом, а если с двумя суффиксами, то обязательно без префиксов; таких моделей среди продуктивных только 6. Все в этом случае самосогласованно: по закону средней глубины русского слова (Кузнецова, 1988а) приблизительно 87% русских слов состоит из трех - пяти морфем с г. ком на четырех морфемах. Отсюда видно, что предсказание продолжений беспрефиксальных начальных биморфов, представляющих собой сочетание корень + суффикс, производиться относительно легко, особенно для корней с малой словообразовательной парадигмой.

Например, биморф ход-ов- имеет единственное продолжение ходовой, а у биморфа ход-ул- может быть развертывание в три слова: ходила, ходильность, ходильный.

Развертывание слова по первому префиксу и корню труднее для

тех корней, гнезда которых достаточно велики, например, после биморфа префикс + корень проход - могут быть продолжения (помимо самого слова проход): проходить, проходимы, проходимец, проходимка, проходимость, проходка, проходной, проходческий, проходчица, проходящий и некоторые другие, не зафиксированные в "Словаре морфем русского языка" (Кузнецова, Ефремова, 1986).

Распознавание слов может вестись также по начальным префиксальным биморфам и полиморфам, т.е. по трем префиксам (более трех префиксов в русском слове не встречается, за исключением нескольких слов, начинающихся с отрицания не-, типа не-при-с-по-соб-л-енн-ость). Слово, имеющее в своем составе три префикса, восстанавливается практически однозначно. Полиморфов такого рода всего около 30 в русском языке (приводимые в скобках продолжения, позволяющие расшифровать слово, не включают однокорневых слов):

вознена(видеть)	невозна(градимый)
воспроиз(водить)	неиспо(ведимый)
довоо(ружение)	неос(поримый)
набезоб(разничать)	непрево(иденный)
назна(нку)	непрас(боримый)
невзна(чай)	невро(ворот)
невдо(мек, гад)	слоза(ранку)
обезо(ружить, пасить)	пересна(ржать)
обезза(разить)	перепроиз(водить)
обезна(дежить)	перераспре(деление)
обезоб(раживать)	предопре(делить)
обесс(мыслить, мертвить)	предраспо(лагать)
перевос(ручать)	приспо(собить)
перес(видетельствовать, мыслить)	присово(купить)
позапро(шлый)	порасс(казать, просить)

Интересно, что полиморфные префиксы, представляющие собой последовательность из трех префиксов, в подавляющем большинстве случаев имеют в качестве продолжения лишь один корень (иногда - в виде разных алломорфов, например -вед-/-зв-/-вол-: воспроизведение, воспроизвести, воспроизводить; -лаг-/-лож-: предрасполагать, предрасположить; -ряд-/-ляж-: переснарядить, перелаяжить и др.) и только в нескольких единичных случаях - по

два различных корня, как в следующих парах: невдомек, невдогад; обезопасить, обезоружить и т.п.

Тернарные цепочки префиксов встречаются в русском языке менее чем с 30 различия корневыми морфемами. Иногда цепочку из трех префиксов можно представить как единый морфемный блок, подобный суффиксальным морфемным блокам (Кузнецова, 1987), или как префиксальную цепочку, сросшуюся с корнем, в результате чего в слове произошло опрощение, что наиболее очевидно происходит со связанными корнями (названку, позапрошлый, обезабыздить, обезображенный и некоторые другие).

Небольшой эксперимент, направленный на восстановление слова по трем начальным префиксам, показал, что из 30 тернарных последовательностей среди 10 испытуемых ни один человек не смог опознать только пять слов, начинающихся с префиксов не-о-с-, о-бес-с- и по-рас-с-*. Эти трехэлементные префиксальные блоки, как и некоторые другие, могут со временем превратиться в двухморфемные: не-о-с-(поримый) - не-о-(споримый), о-бес-с-(мыслить) - о-бес-(смыслить), о-бес-с-(мертвить) - о-бес-(смертвить), по-рас-с-(казать) - по-рас-(сказать), по-рас-с-(просить) - по-рас-(впросить).

Механизм декодирования слова по начальным морфемам можно подробнее рассмотреть на примере префиксальных биморфов.

В статье, в частности, выясняется вопрос о том, в силу каких причин облегчается распознавание слов по начальным префиксальным биморфам.

Предсказание значимой последовательности морфем исходя из начальных двух или трех префиксов относится к сфере установления правил аранжировки морфем, выявления основных ограничений на порядок следования морфем, обнаружения корреляций элементов морфемного уровня.

* Во время проведения эксперимента при декодировании было много случаев ложной омонимии, когда за префиксы принимались части корней: попона (вместо попо-мнить, попо-льзоваться), при-нести (вместо при-не-водить), распад (вместо распа-лубка) и др.

Нередко появлялись не запрещаемые языком, но не засвидетельствованные словарями образования различных способов глагольного действия (совершенностей), как например: довзбить, довзодушевить, довытягивать, довыничивать, довыкрасить, допропасть, допросеивать вместо фиксируемых словарями довыискать, довозбудить, до-выполнить, допродать и т.п.

В области морфематики до сих пор или просто констатировались ограничения и приводились списки сочетаний (Ефремова, 1966, 1970; Кузнецова, Ефремова, 1986; Кузнецова, 1986б), или анализировались морфематические причины ограничений на порядок следования отдельных морфов.

Занимаясь декодированием слова по исходному полиморфу (чаще — биморфу), можно попытаться выявить причины, затрудняющие определение продолжений префиксальных биморфов в одних случаях и облегчающие однозначность решения в других.

В качестве материала исследования были отобраны по "Словарю морфем русского языка" А.И.Кузнецовой и Т.Ф.Ефремовой (1986) все двух- и трехэлементные сочетания префиксов (при этом префиксальные полиморфы из трех морфов составили всего 1% от биморфов). 330 отобранных биморфов и полиморфов были сопоставлены с префиксальными биморфами, имеющимися в "Частотном словаре русского языка" под ред. Л.Н.Засориной (1977). Дальнейшему анализу были подвергнуты только те единицы, которые оказались зафиксированными в обоих словарях: это 243 исходных префиксальных биморфа с продуктивными и частотными компонентами (или обоими, или каким-либо одним из двух, что дополнительно проверялось по работе (Кузнецова, Лавренова, 1975). Эти биморфы обладают ненулевой вероятностью появления в тексте, судя по "Частотному словарю русского языка" (1977), и состоят из 46 различных префиксов. Возможны еще 87 префиксальных биморфов, имеющих нулевую вероятность появления в тексте (по тому же словарю).

При автоматическом прочтении письменного текста окрестность морфемного состава слова очень важно не только знать ограничения в ранжировке морфов фонетического и структурно-позиционного характера, но и учитывать частоту и продуктивность каждого морфа. Знание частот префиксальных биморфов существенно сокращает и упорядочивает перебор гипотез в процессе автоматического распознавания).

Сочетание префикс + префикс весьма информативно с точки зрения ограничений на возможные продолжения биморфов, поскольку наличие двух и более префиксальных морфов в слове — явление довольно редкое, что, согласно теории информатики, делает такого рода биморфы высокоинформативными.

Наиболее существенный параметр, характеризующий префикс, вступающий в сочетание с другим префиксом, — частота. С понижением

частоты префиксальных морфов теряется их способность выступать в качестве компонентов соответствующих биморфов (Гольдштейн, 1988). С.Р.Гольдштейн доказал, что "между частотой префиксальных морфов и количеством биморфов, компонентами которых они могут являться" (1980, с.37), существует определенная связь, а именно: чем больше частота префиксов (см.Кузнецова, Лавренова, 1975), тем больше и количество биморфов, в которых они участвуют. С убыванием частоты префикса ослабляется способность префиксальных морфов выступать в роли одного из компонентов биморфа (как первого, так и второго), хотя этот процесс идет неравномерно, с отдельными значительными выбросами.

С точки зрения частоты компонентов можно выделить группы 1) префиксальных биморфов со вторым компонентом более частотным, чем первый, 2) биморфов с обратной структурой. Биморфы первого типа, по данным С.Р.Гольдштейна (1988), составляют 52% от общего числа различных биморфов и покрывают 60% появлений префиксальных биморфов в тексте; на долю биморфов второго типа, в которых более частотным является первый компонент, приходится 48% от общего числа различных биморфов и 40% от числа появлений их в тексте.

Между частотой и продуктивностью компонентов в пределах высокочастотных и высокопродуктивных префиксов существует корреляция. Связь числа биморфов с продуктивностью компонентов более слабая, чем с частотой. Число префиксальных биморфов заметно убывает, если компонентами являются малопродуктивные префиксы.

Префиксальные биморфы, у которых второй компонент более продуктивен, чем первый, составляют 58%, а префиксальные биморфы с более продуктивным первым компонентом — 42% от общего числа различных биморфов. Частота в тексте префиксальных биморфов со вторым более продуктивным компонентом составляет 60% случаев появления префиксальных биморфов в рассматриваемой выборке (Частотный словарь..., 1977), а на долю биморфов с первым более продуктивным компонентом приходится около 40% появлений всех биморфов в тексте. Это абсолютно такая же пропорция, какая отмечалась для префиксальных биморфов со вторым компонентом более частотным, чем первый, в противоположность биморфам, имеющим в качестве более частотного первый компонент.

Учитывая соотношения частот компонентов и их продуктивности, все множество префиксальных биморфов целесообразно разделить на

четыре группы, приписав каждому элементу множества одну из следующих помст:

II - биморфы*, у которых первый компонент менее продуктивен и менее частотен, чем второй;

IO - биморфы, у которых первый компонент менее продуктивен, но более частотен, чем второй;

OI - биморфы, у которых первый компонент более продуктивен, но менее частотен, чем второй;

OO - биморфы, у которых первый компонент более продуктивен и более частотен, чем второй.

Данные результаты показывают, что больше всего биморфов таких, у которых первый компонент одновременно менее частотный и менее продуктивный по сравнению со вторым; они подтверждают на материале префиксальных биморфов существование корреляции между продуктивностью и частотой, что наблюдается в 75,7% случаев (тип II и OO).

Независимо от частоты первый компонент увеличивает информативность префиксального биморфа как целого.

С точки зрения теории информации малочастотные элементы являются наиболее информативными и, наоборот, высокочастотные - малоинформативными. Отсюда можно предположить, что префиксы с малой частотой, несущие больше информации, плохо сочетаются между собой (или совсем не сочетаются) и потому могут быть легко продолжены, а само слово легко декодировано; префиксы же с большой частотой и малоинформативные сочетаются хорошо, как и префиксы с различными частотами (большими и малыми), в связи с чем таких биморфов много и, естественно, слова, с них начинающиеся, угадываются с трудом.

Так, биморф ИС-ПО-, судя по "Словарю морфем..." (Кузнецова, Ефремова, 1986), имеет 19 продолжений, использующих 5 разных корней, но из них ненулевых по вероятности появления всего II (с корнями -вед- и -льз-), при этом два продолжения (-льзова- и -льзовани-) покрывают 80% случаев появления в тексте словоформ, содержащих префиксалы и биморф ИС-ПО-.

* Биморфы каждого типа составляют следующие проценты от общего числа различных префиксальных биморфов (Гольштейн, 1988): II - 103 (42,4%); IO - 35 (14,4%); OI - 24 (9,9%); OO - 81 (33,3%).

Биморф НА-ИГД- имеет три продолжения (например, напринимать, наприлевоочный), но число ненулевых по вероятности появления в тексте продолжений может быть только одно - слово например, покрывающее 100% текста (Частотный словарь..., 1977).

Префиксальный биморф В-НА- имеет 10 продолжений (внаброс, внаем, внаймы, внакидку, внакладе, внакладку, внакрой, внатяжку, внахлестку, вначале), из них только четыре - ненулевых по вероятности появления в тексте, из которых, в свою очередь, лишь одно (наречие вначале) дает 91% покрытия текста, и т.д.

В общей сложности около 40 биморфов с частотой от II80 до 45 имеют в сумме около 640 продолжений, из них с ненулевой вероятностью появления - приблизительно 350 продолжений, из которых, в свою очередь, только 130 продолжений дают покрытие текста 80% и более (в приводимой далее таблице опущены префиксальные биморфы, не начинающиеся с не-).

Абсолютная частота биморфов	Префиксальные биморфы по убыванию их частоты	Продолжения биморфов		Наиболее вероятные (ед)	Покрытие текста наиболее частотными продолжениями, %
		Рас-смотренные	Ненулевые по вероятности		
II80	Про-из-	27	18	-водство, -ведение, -водственный, -водить, -вести, -водитель, -водительность (7)	80
-I60	Рас-с-	22	15	-казать, -казывать, -каз (3)	81
	С-по-	28	21	-способность, -способ, -способный, -способствовать (4)	83
	О-пре-	18	8	-деленный, -делиться, -делять, -делиться (4)	83
	Вс-по-	11	7	-мнить, -мнить (2)	85
	По-з-	7	4	-волять, -волять (2)	98
	В-н-	13	10	-утренний, -утри-утрь, -утрь (4)	83
	Ис-по-	19	11	-льзовать, -льзовать (2)	80
	На-при-	3	1	-мать (1)	100
	Рас-по-	26	15	-ложить, -ложить -лагать, -ложить, -лагать-ся, -ложиться (6)	80

Продолжение таблицы

Абсолютная частота биморфов	Префиксальные биморфы по убыванию их частоты	Продолжения биморфов			Покрытие текста наиболее частотными продолжениями, %
		Рассмотренные	Ненулевые по вероятности	Наиболее вероятные (ед)	
	Про-но-	7	5	-ходить, -хождение (2)	82
	Пред-при-	9	9	-ятие, -нять (2)	88
	Из-об-	25	14	-ражение, -разьть, -речение, -разить, -рести, -ретать, -ретатель (7)	84
	О-бес-	71	22	-печить, -печивать, -печение (3)	81
	Бес-по-	19	10	-лезный, -цадный, -рядочный, -рядок, -мощный (5)	86
	Во-о-	16	12	-ружение, -руженный, -ружить, -ружать (4)	80
	Пред-по-	20	12	-лагать, -ложение, -локоть, -читать, -лагаться, -честь (6)	80
	В-за	13	5	-имный, -мен (2)	83
	На-по-	11	7	-минать, -мнить (2)	84
	Со-от-	13	9	-ветствовать, -ветствующий, -ветствие (3)	81
	Раз-о-	25	13	-ружение, -чтение (2)	80
	Рас-про-	21	10	-стражение, -странять-ся, -стренить, -страненный, -странять (5)	88
	Пред-у-	28	13	-предить, -преждать, -смагивать, -смотреть, -смагивать (5)	82
110 -30	При-об-	8	7	-реть, -рести, -ретье-ние (3)	95
	Сс об-	13	8	-ражение, -разить, -разьть (3)	82
	Пред-о-	24	13	-ставить, -ставлять, -ставление, -ставлять-ся, -стеречь, -судительный (6)	81

Окончание таблицы

Абсолютная частота биморфов	Префиксальные биморфы по убыванию их частоты	Продолжения биморфов			Покрытие текста наиболее частотными продолжениями, %
		Рассмотренные	Ненулевые по вероятности	Наиболее вероятные (ед)	
	Ви-с-	24	11	-казать, -казывать, -казываться, -казаться, -казывание (5)	89
	Со-в-	22	11	-местный, -падать, -падение (3)	80
	Пре-о-	8	5	-долеть, -доление, -долсвать (3)	82
	Рас-пре-	8	7	-деление, -делиться, -делять (3)	85
	Под-с-	15	7	-читать, -казать, -чет, -казывать (4)	90
55-45	Бес-пре-	8	5	-дельный, -пытство-ный, -рывный (3)	85
	В-по-	6	4	-следствия (1)	85
	По-на-	21	11	-добиться, -чалу (2)	86
	Вос-при-	10	6	-нимать, -ятие, -нять (3)	80
	Без-на-	8	6	-казный, -казанный (2)	87
	В-на-	8	4	-чале (1)	91
	По-за-	17	8	-быть, -вчера, -видо-вать, -нимать (4)	30

Как видно из таблицы, появление в тексте того или иного префиксального биморфа предопределяет следующие за ним морфы, что зависит от частоты и продуктивности сочетаний префиксов и позволяет восстанавливать слова, хотя к с разной степенью точности. Учет всех затронутых в статье вопросов имеет большое значение для задач автоматического распознавания текста, проведенного на морфемном уровне, а в ряде случаев позволяет однозначно предсказать возможное продолжение, декодировать основу слова.

Л и т е р а т у р а

- Асиновский А.С. Выступление в дискуссии // Машинный фонд русского языка: идеи и суждения. М.: Наука, 1986. С.157-158.
- Гольдштейн С.Р. Количественное исследование префиксальных биморфов русского языка. М.: Изд-во МГУ, 1988. 74 с.
- Ефремова Т.Ф. Из наблюдений над структурой современного русского языка на уровне морфов // Семантические и фонологические проблемы прикладной лингвистики. М.: Изд-во МГУ, 1968. С.45-55.
- Ефремова Т.Ф. Опыт описания современного русского языка на уровне морфов: Автореф. дис. ... канд. филол. наук. М., 1970. 24 с.
- Караулов Ю.С. Частотный словарь семантических множителей русского языка. М.: Наука, 1980. 207 с.
- Кузнецова А.И. Морфемные блоки и их роль в морфемной и словообразовательной структуре русского языка // Морфемика. Принципы и методы системного описания. Л.: Изд-во ЛГУ, 1987. С.97-106.
- Кузнецова А.И. Морфемная глубина слова в русском языке (слова малой и большой глубины как периферийное явление русского языка) // Проблемы структурной лингвистики. 1984. М.: Наука, 1986а. С.52-66.
- Кузнецова А.И. Параметрическое исследование периферийных явлений в области морфематики: Автореф. дис. ... д-ра филол. наук. М., 1988. 38 с.
- Кузнецова А.И., Ефремова Т.Ф. Словарь морфем русского языка. М.: Рус.яз., 1986. 1136 с.
- Кузнецова А.И., Лавренова О.А. О существовании корреляции между продуктивностью и употребительностью аффиксов в русском языке // Исследования по структурной и прикладной лингвистике. М.: МГУ, 1975. С.83-99.
- Частотный словарь русского языка / Под ред. Л.Н.Засориной. М.: Рус.яз., 1977. 935 с.

А.И.Моисеев

СТРУКТУРНЫЕ МОДЕЛИ ЯЗЫКОВЫХ ЕДИНИЦ (слог и слово)

Языковые единицы и явления поддаются моделированию. Моделировать можно, в частности, слог и слово.

Модель — это обобщенное и в той или иной мере абстрактное представление моделируемого явления обычно в сокращенном, свернутом виде и в другом материальном выражении; примером могут служить математические модели физических процессов.

Языковые модели можно представить в форме условных записей, например, слог в виде сочетаний символов гласных (Г) и согласных (с) звуков: стол — ссГс; слово — в виде сочетания знаков ударного (┌) и неударного (┐) слогов, например, стена: ┐┌. Морфемный состав слов можно записать в условных символах морфем — корня (К), приставки (П), суффикса (С), окончания (О), постфикса (П); приставки и постфиксы при этом будут дифференцированы позицией знака (П): у-мы-ть-ся (П-К-С-П). Нулевую морфему (морфологически значимый материальный нуль) принято обозначать через (∅).

Модели слога и слова могут быть различные. Возможны следующие модели (типы моделей): принципиальная (абстрактная и конкретизированная), минимальная, полная, максимальная (сolutный максимум), оптимальная (относительный или оптимальный максимум).

1. Принципиальная модель. Она должна отразить общие черты строения моделируемой единицы, без конкретизации и без доведения до исчерпанности (в абстрактной разновидности). Модель фиксирует обязательные, постоянные компоненты строения единицы (структурная константа) и указывает на возможность факультативных компонентов (структурные переменные), исчерпывая (конкретизированная разновидность принципиальной модели) или не исчерпывая (абстрактная разновидность) их. Графически это можно показать с помощью скобочной записи: обязательный элемент вне скобок, факультативные — в скобках.

Обязательным компонентом слога является гласный звук, факультативными — согласные звуки в препозиции или/и в постпозиции к гласному. Обязательной фонетической частью слова является ударный слог (┌), факультативными — безударные слоги (┐) в препозиции или/и постпозиции к ударному. Обязательную часть мор-