

Некоторые свойства линейной структуры именных и предложных групп (Поверхностно-синтаксический анализ русского предложения)

Введение

Анализ именных и предложных групп (ИГ и ПГ) – одна из наиболее актуальных проблем автоматического анализа текста. Практически любая прикладная задача, связанная с автоматическим анализом естественного текста, предполагает построение ИГ и ПГ.

Ниже рассматриваются некоторые свойства линейной структуры ИГ и ПГ, в частности – проективность и рекурсивность, характерные для организации линейной структуры русского предложения. Способ использования этих свойств в значительной мере определяет особенности лингвистического базиса системы поверхностно-синтаксического анализа русского предложения (ПСА), разрабатываемой в настоящее время в Институте лингвистики РГГУ¹ [Кобзарева и др. 2001, Кобзарева 2002, Ножов 2003, Кобзарева 2004, Кобзарева 2005, Баталина и др. 2006, Кобзарева 2006, Кобзарева 2007].

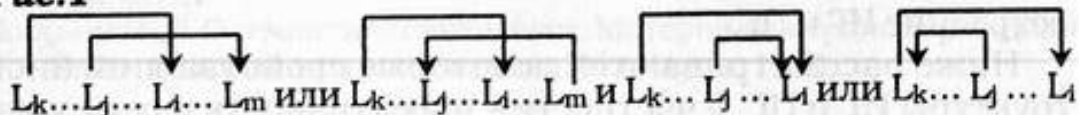
Задача ПСА, была поставлена давно, и существуют ставшие каноническими варианты ее решения для машинного перевода, описанные в [Мельчук 1964, Иорданская 1967, Кулагина 1979]. В них во всех большую роль играет порядок процедур анализа. Ниже обсуждаются возможности организации ПСА, отличной от описанных в [Мельчук 1964, Иорданская 1967, Кулагина 1979, Апресян 1989]. Система строится как много-ступенчатый комплекс модулей с выделенными в отдельный модуль алгоритмами сегментации [Кобзарева 2007]. Анализ на всех уровнях использует свойства сочинительной и подчинительной проективности и рекурсивности линейной структуры русского предложения (S) [Иорданская 1967, Падучева 1971, Кобзарева 2006, Кобзарева 2007]: фильтры для отсеивания ложных гипотез встроены непосредственно в процедуры анализа. Проективность рекурсивных структур подчинительных

и сочинительных [Кобзарева 2004, Кобзарева 2006, Кобзарева 2007] связей играет существенную роль и для построения динамических компонент анализа, и для упорядочения процедур анализа в целом.²

1. Понятие проективного фрагмента

Для организации анализа существенно, что проективность является свойством, производным от порядка слов, связанных отношениями подчинения или сочинения. Известно, что условие проективности графа по Теньеру (при изображении связей в виде дерева) соответствует в стрелочном изображении связей над линейной структурой предложения требованиям непересечения и необрамления [Иорданская 1967], т.е. не должно быть ситуаций типа (L_k – слово, k – его порядковый номер в предложении):

Рис. 1



Таким образом, если границами некоторого отрезка в предложении являются два слова L_k и L_i , связанные отношениями подчинения или сочинения ($L_k R L_i$ или $L_k C L_i$), то в этом отрезке (1) только слова-границы L_k и L_i могут иметь связи со словами, находящимися вне этого отрезка и (2) никакое слово внутри отрезка не может быть хозяином слова-границы. Отрезки, удовлетворяющие этим условиям, будем называть проективными фрагментами (ПФ).

Если внутри ПФ появляется слово, не являющееся непосредственным или опосредованным слугою слова-границы, то – в силу связности графа – оно будет иметь хозяина или слугу вне ПФ, что породит пересечение, или будет хозяином одного из слов-границ, т.е. возникнет обрамление. Таким образом, все слова внутри ПФ являются узлами ветки графа S с вершиною – одной из его границ.

2. Проективные фрагменты определительных именных групп и предложных групп

Объектом рассмотрения являются ПФ определительных именных групп (ИГ) и предложных групп (ПГ), а именно ПФ, границы которых задаются поверхностно-синтаксическими отношениями NRA , где N – существительное или его синтакси-

ческий аналог является хозяином А – полного прилагательного, причастия, местоименного прилагательного, порядкового числительного в необособленном согласованном определении в пре- или постпозиции или обособленном согласованном определении в препозиции и **PRN**, где Р – предлог. При рассмотрении особенностей ПФ этих связей нас будет занимать грамматически правильные для русского языка линейные конфигурации вне зависимости от вероятности их появления в текстах³.

Некоторые важные особенности линейной структуры определительных именных групп, а именно – место распространенного согласованного определения по отношению к определяемому слову, уже обсуждались в [Санников 1963], но это рассмотрение не обладает необходимой для анализа естественных текстов полнотой.

В рамках ПСА эти ПФ можно рассматривать с двух точек зрения.

1. Общий вопрос, определяющий порядок работы с текстом: на каком этапе анализа возможно и удобно с точки зрения общей стратегии ПСА строить эти связи?

2. Группа вопросов, относящихся к собственно процедуре построения связей, задающих эти ПФ: какие факторы в линейной структуре текста необходимо учитывать при построении ПФ этих групп? Что при построении определяющих их связей порождает потенциальную неоднозначность дешифровки? Какова наилучшая процедура анализа?

3. Построение проективных фрагментов ИГ и ПГ в общей архитектуре системы

Ответ на вопрос о том, на каком этапе лучше строить ПФ определительных ИГ и ПГ, связан с ролевыми особенностями ИГ и ПГ в структуре S.

В тексте мы имеем дело с предикативными словами (или группами связанных слов, выступающими в этой роли) и словами (или, опять-таки, группами связанных между собою слов), выступающими в роли имен актантов этих предикативных слов. При этом в группы, выступающие в качестве имен актантов, могут в свою очередь входить предикативные слова со своими актантами и т.д. Линейная структура манифестирует многоуровневую рекурсивную структуру с предикативными словами и актантами разных уровней связи в графе (Рис.2), где имена актантов выражены чаще всего ветками графа, представленными ПФ именных и предложных групп (ПГ рассматриваются как функциональный синтаксический аналог ИГ):

Рис.2



Пример 1: (см. Рис.2) проективный фрагмент имени субъекта сказуемого **Проплыла** задается синтагматической связью ветка R замороженная, а в него вложен ПФ имени обстоятельства места, заданный связью в R стакане.

Пример 2:

...раздутая трюизмами и арифметическими выкладками болтовня о гармонической личности **лезла** отовсюду, как сорная трава. (М)

ПФ субъекта предикативной вершины **лезла** с границами, определенными связью болтовня R **раздутая**, где предикативная вершина согласованного определения **раздутая** распространена двумя сочиненными слугами-ИГ.

Пример 3:

Постигаемые умом, **не данные** зрению линии (Флоренский, далее Ф)

Самая «длинная» связь линии R постигаемые задает ПФ с двумя соподчиненными предикативными вершинами, каждая со своим слугою-ИГ.

Пример.4:

Восвещенной лампой комнате, по пояс обнаженный, он **стоял**, первые портреты в газетах, растущих в богатой осадками, плачущей и кровоточащей стране, выходили расплывчатыми. (Набоков, далее Н).

В простом предложении с вершиной **стоял** внутри ПГ – ее слуги – появляется предикативная вершина **освещенной**, а в причастном обороте с вершиной **растущих** внутри ПГ – ее слуги – три соподчиненные предикативные вершины, у одной из которых – **богатой** есть ИГ-слуга осадками.

Структура «матрешек» ПФ в примерах 1-4 иллюстрируют возможность рекурсивных вложений, где А – слуги вершин опоясывающих ИГ\ПГ имеют собственных слуг, представленных вложениями проективных фрагментов ИГ\ПГ следующего уровня связей. При этом количество уровней вложений теоретически не ограничено ничем, кроме возможностей оперативной памяти.

При построении ПСА будем исходить из того, что в подавляющем большинстве случаев связи NRA и PRN задают границы ПФ определительных ИГ\ПГ, а слова внутри этих фрагментов могут быть только непосредственными или опосредованными слугами слов-границ, т.е. узлами веток, исходящих из одной из границ. Например, построив ПФ ИГ в предложении Это была могучая по силе интеллектуального характера семья (М)⁴ (Рис.3), мы при дальнейшем анализе связей внутри сегмента можем сначала рассматривать три компоненты (это, была, и вершина ПФ, ограниченного связью семья R могучая), а затем хозяев ПГ=по силе и ИГ=интеллектуального характера искать только внутри ПФ с вершиной семья.

Рис.3



Или, например, в предложении Он мог увлечь чванную, лишенную чувства юмора натуру царя (М) модуль предсегментации строит:

Рис.4



А в модуле внутрисегментного анализа (ВА) внутри ПФ натуру R лишенную достраиваются связи (Рис.5) лишенную R чувства юмора (программная версия ВА, представленная в [Баталина и др. 2006])

Рис.5



Или, например, в предложении На вопрос он издавал странный грудного тембра неопределенный звук, извлеченный из трубы неумелым музыкантом (М) такая последовательность действий облегчает, в частности, поиск хозяина N_{род} (грудного) тембра в препозиции к хозяину звук (программная версия [Баталина и др. 2006]):

Рис. 6



Для определения границ простых-главных, придаточных предложений и деепричастных и прочих обособленных оборотов, т.е. на этапе сегментации предложения, важно уметь определить функции знаков препинания (ЗП) и сочинительных союзов (СС) [Кобзарева 2004, Кобзарева 2005]. Как границы сегментов ЗП и СС разграничивают зоны влияния предикативных вершин сегментов и не являются границами, если сочиняют слова – не вершины сегментов [Кобзарева 2006]. Так как при этом они могут сочинять актанты предикативных слов – вершин, то для анализа сочинения удобно иметь уже построенные ПФ соответствующих ИГ и ПГ. Поэтому после снятия морфологических и морфосинтаксических неоднозначностей модуль предсегментации [Кобзарева 2007] строит связи, определяющие границы ПФ, вершины которых являются предикативными вершинами или именами актантов и которые можно рассматривать как единицы линейной структуры при сегментации: ПФ определительных ИГ и ПГ, конструкций с именами собственными, числами, сложных сказуемых. Их удается строить, используя только морфологические характеристики слов и минимальную грамматическую информацию о грамматическом управлении [Кобзарева и др. 2001, Кобзарева 2007].

Эти связи, в частности, обсуждаемые здесь NRA и PRN, задают границы ПФ, выступающих как единицы текста при сегментации. Внутри этих ПФ при сегментации нам не нужно заглядывать, что часто существенно упрощает линейную структуру.

«Связав» А (построив синтагму NRA в искомым ситуациях), мы исключаем эти А и из числа А, претендующих при дальнейшем анализе на роль вершины сегмента – обособленного согласованного определения – и на роли сказуемого и подлежащего. Из зоны обозрения элиминируются знаки препинания и сочинительные союзы, предикативные слова, номинативы внутри ПФ, которые могли бы быть претендентами на роль подлежащих или сказуемых, существенно разрешается падежная омонимия существительных и омонимия управления предлогов.

На этом этапе не ставится задача построить связный подграф слов внутри ПФ: эта задача решается уже после сегментации в рамках модуля построения всех связей внутри сегмента, причем построенные к этому этапу ПФ используются при определении «зон влияния» хозяев синтагм.

Обособленное согласованное определение с вершиной А (А-оборот) в препозиции к N-хозяину анализируется вместе с необособленными в отличие от обособленного согласованного определения в постпозиции. А-оборот в постпозиции строится, как другие сегменты, при сегментации, так как поиск его границ осложнен теми же проблемами, что и определение границ других сегментов: необходимо учитывать сочинение и \или ситуации разрыва А-оборота вложением других сегментов [Кобзарева 2004]. Например, в отрывке из Набокова А-оборот разорван вложением придаточного:

*...густые, очень темные брови, **составлявшие, когда он хмурился или внимательно слушал, одну сплошную черную черту**, но зато распахивающиеся, как легкие крылья, когда редкая улыбка обнажала на миг...*

Для ПФ ИГ в препозиции тоже необходимо, хотя это и встречается очень редко, учитывать возможность его разрыва вложением: *Летящий рысью **маленький, как мальчик, темный, как мулат, командир** алы-сириец, равняясь с Пилатом, что-то тонко крикнул и выхватил из ножен меч.* (Булгаков, далее Б).

Хотя, вообще говоря, фрагменты ИГ и ПГ могут быть и непроективны, случаи непроективности, встречающиеся, например, в поэтических текстах, такие как на Рис. 7 и 8, и чрезвычайно редкие в обычных прозаических текстах, в настоящей работе не рассматриваются.

Рис. 7



Ужален небольшою \ Крылатой я змеєю (Державин)

или

Рис. 8



Дай им цену, за которую любили, \ Чтоб за ту же и оплакивали цену (Бродский)

4. Согласование N и А

Необходимым условием построения связи NRA является согласование N и А [Мельчук 1993]. Известно, что помимо нюансов согласования в конструкциях с числительными (два больших \большие стола, большие \больших пять мальчиков) и др., для Вин. необходимо учитывать одушевленность N – потенциального хозяина А (в частности, и для ситуаций с сочи-

нением, когда возможен случай отсутствия согласования по числу): Вин п. есть у двух форм А – формы, омонимичной Род. п. и омонимичной Им.п., так как у Нодуш м.р. ед. и Нодуш мн. совпадают формы Вин. и Род. (*вижу маленького мальчика \ маленьких девочек \ маленького и большого мальчиков*), а у Ннеодуш м.р. в ед. ч. и Ннеодуш мн. совпадают Вин. и Им. (*вижу большой стол \ большие столы и стул \ большой и маленький столы \ маленькие чашки*).

Это существенно, например, при анализе цепочек $N_j A_k N_l$: если N_j и N_l – м.р., то в цепочке $N_j A_k N_l$ в ситуации *журнал_{им\вин} \ вин стар_{им\вин} \ вин мальчика_{род\вин} не занимал или девочек_{род\вин} \ вин маленьких_{род\вин} журналь_{им\вин} \ вин не занимали* для предупреждения построения ложных связей *мальчика_{род\вин} R стар_{им\вин} \ вин и журналь_{им\вин} \ вин R маленьких_{род\вин}*: А согласовано с Нодуш м.р. ед. и Нодуш мн. ч. тогда и только тогда, если в пересечение, кроме Вин, входит Род.п., а в пересечение падежей с Ннеодуш м.р. ед. и Ннеодуш мн. должен входить, кроме Вин, еще и Им.п.⁵

Ниже будут рассмотрены различные особенности линейной структуры, в силу которых согласование N и A является условием необходимым, но не достаточным для констатации связи N R A. В частности, для обычной структуры, когда потенциально распространенный A_k (один или цепочка соподчиненных) находится в препозиции к N_k – хозяину.

5. Рекурсивная структура ИГ и ПГ

Вершина ИГ (не «вырожденного» случая ИГ, когда у N нет А-слуг) может быть слугой предлога, при этом вся ИГ с вершиной – слугою предлога – оказывается внутри ПФ предложной группы. Внутри этой ИГ могут появиться ПФ как ПГ – слуг А, так и ИГ – слуг этого же А. Кроме того, появление любого слуги А провоцирует появление его слуг ИГ\ПГ. Появление каждой новой ИГ с А-слугою создает предпосылки появления новых вложений ИГ\ПГ – непосредственных или опосредованных слуг А. Так, у А может появиться слуга – инфинитив. Кроме того, появление каждого нового N_i внутри ПФ всегда означает возможность появления его слуг – N_j в слабоуправляемом Род.п. и, если N_i имеет валентности, в требуемых этими валентностями падежах.

Таким образом, при появлении А всегда вероятна матрица.

(1) *Близкое присутствие моря за окном томилло его, словно это огром-*

ное, липко-блестящее, лунной перепонкой стянутое пространство, которое он угадывал... (Н)

(2) Он вносит посторонние, чуждые этому духовному началу, черты (Ф)

(3) Много прохожих поглядывало с завистью на высокого гончара, который медленными шагами шел за своим товаром, заботливо окутывая глиняных своих щеголей ненавистным для них сеном (М)

Проблемы, инициируемые появлением «матрешки» – вложений ИГ и ПГ и ПГ в ИГ возникают, так как А – слуга ИГ в позиции к N – может иметь слуг (любых, так как появление любого слуги может инициировать появление слуг следующего уровня вложений).

Анализ рекурсивных структур ИГ и ПГ представляет собою некоторый комплекс проблем, которые удобно решать одновременно.

В каждом очередном вложении ИГ (без Р – хозяина или же слуги Р) с А – слугою потенциально содержится отпредикативная ветка графа, в линейной структуре которой могут появляться все новые и новые вложения ИГ и ПГ. В (4) Лежащий на ложе в грозовом полумраке прокуратор сам наливал себе вино в чашу... (Б) – два параллельных – на одном уровне вложения, а в предложении Набокова Измяв и отбросив последнюю газету, все высосав, все узнав, сжигаемый неотвязным зудом, желанием принять мне одному понятные меры, я сел за стол и начал писать способность А=сжигаемый управлять Тв.п. порождает вложение между ним и его хозяином N=я двух сочиненных ИГ, одно из которых – желание – может управлять инфинитивом, а появившийся инфинитив в свою очередь инициирует появление между ним и вершиной ПФ, ограниченного отношением я R сжигаемый двух вложенных ИГ – слуг этого инфинитива: мне одному и понятные меры. На этапе предсегментации будет построена матрешка ПФ:

Рис.9



6. Рекурсивная процедура анализа матрешек ИГ и ПГ

Если N является слугою предлога (Р), то возникает возможность вложений ИГ и ПГ между Р и N, его слугою.

Как обсуждалось в [Кобзарева 1004], сегментная структура русского предложения в письменном тексте, не являющемся записью устной речи или ее имитацией, проективна и рекурсивна. Как показано в [Кобзарева 2004, Кобзарева 2007], в ситуациях, когда мы анализируем «матрешку», где однотипные линейные структуры вложены друг в друга и отправной точкой построения каждой структуры является ее минимальная левая составляющая, анализ удобно вести справа налево, так как при этом мы начинаем анализ каждой матрешки с самого глубокого вложения.

При анализе ИГ и ПГ ситуация почти аналогична. Опорными элементами построения ИГ служат А, а исходными словами при построении ПГ – предлоги.

Необособленные согласованные А чаще всего находятся слева от N-хозяина, справа необособленными могут быть только нераспространенные А. Предлог обычно стоит слева от его слуги N⁶ (в постпозиции к существительному-слуге может появиться только предлог **ради (Бога ради)**). В силу сказанного, анализ ИГ и ПГ естественно начинать с поиска левой (по порядку следования в S) составляющей – А (А: полное прилагательное, причастие, местоименное прилагательное, порядковое числительное) и Р (предлог).

Структура ИГ и ПГ рекурсивна, и в принципе количество и глубина вложений заранее не определены. Оптимальной является процедура анализа, когда мы, двигаясь справа налево по тексту, анализируем последовательно справа налево все А и Р, находя для каждого очередного А и Р соответственно хозяина или слугу. При таком порядке анализа мы, если вложений несколько, анализируем каждую матрешку, начиная с самого глубокого вложения, и постепенно укрупняем строимый проективный фрагмент.

При итерации поиска для каждого очередного справа налево А или р соответственно хозяина или слуги используется один и тот же набор линейных конфигураций, разложенных для минимизации проверок на группы признаков ситуаций.

На Рис. 10 – условная схема фрагмента работы с предложением *Олицетворение пассивной и обреченной на всяческую пассивность мысли* (Ф) при построении ИГ и ПГ в ситуации, когда хозяин А находится справа.

Рис.10



Шаги анализа при поиске хозяина справа: 1. поиск хозяина $A=$ всяческую; 2. поиск слуги $P=$ на; 3. поиск хозяина $A=$ обреченной (на всяческую пассивность из рассмотрения уже исключается); 4. поиск хозяина $A=$ пассивной (обреченной на всяческую пассивность из рассмотрения при этом исключено). Каждый цикл анализа – итерация поиска для контекстной ситуации соответствующей базисной ситуации по соответствующим правилам.

В результате или строится соответствующая синтагма с учетом возможной неоднозначности, вытекающей или из морфологической неоднозначности, или из потенциальной неоднозначности линейной конфигурации, или констатируется, что у A есть или отсутствует N -хозяин справа, и при любом из этих результатов начинается поиск хозяина слева (если хозяин справа есть, возможна ситуация синтаксической омонимии).

Проективность фрагментов позволяет, как только ПФ построен, при анализе очередного следующего справа налево A или P исключить очередной построенный ПФ из рассмотрения, для ПГ – вместе с его границами, а для ИГ – все слова ПФ, кроме N -хозяина, тем самым уменьшив, и иногда существенно, зону анализа на очередном этапе.

7. Основные типы структур линейной конфигурации ПФ

Позиция в линейной структуре по отношению к потенциальному хозяину.

1. A необособленное в препозиции. Самая длинная связь с самым левым A потенциальной матрешки ограничивает фрагмент с несколькими соподчиненными и сочиненными между собою A – слугами одного и того же N : *Считалось, что это нелепая, истерическая, суеверная, сверхподозрительная, и чем-то привлекательная мать внушила сыну боязнь ада, которая терзала его всю жизнь* (N)

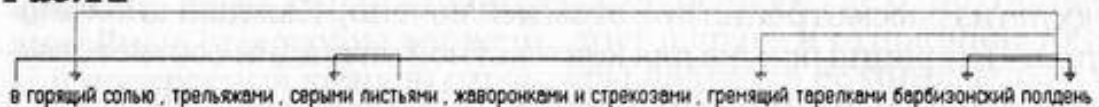
Рис. 11



При этом при каждой итерации процедуры анализа, т.е. при анализе каждого очередного A строится очередной ПФ, ограничивающий зону с собственными актантами, выраженными вложенными ПФ (Рис. 12).

Воздух, уксус и солнце уминались с зелеными тряпками в горячий солью, трельяжами, серыми листьями, жаворонками и стрекозами, гремящий тарелками барбизонский полдень (М).

Рис. 12



2. А в обособленном обороте в препозиции (Рис. 13, 14).

Неизбалованный такими желаниями, я отвечал умиленным согласием, и было решено, что он зайдет ко мне в гостиницу (Н)

Рис. 13



Маленький, коротконогий, в стариковской шубе до пят, в тяжелой шапке, он пытел, пока его не освободили от жарких бобров, и тогда сел на диван, протянув ножки, как ребенок (М)

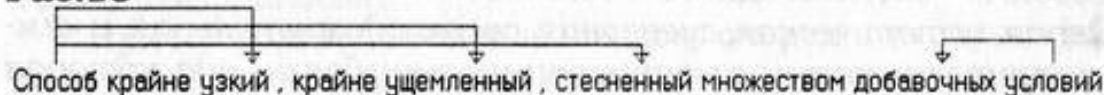
Рис. 14



3. Постпозиция без обособления: А стоит справа от Н (Рис. 15, 16).

Способ крайне узкий, крайне ущемленный, стесненный множеством добавочных условий, которыми определяется его возможность и границы применения (Ф)

Рис. 15



Он не слышал и, совершенно потерявшись, совершенно еще не сознавая происшедшего, а только стремясь отойти от чего-то стыдного, гнусного, нестерпимого, подвинулся к столику, где продавались билеты, вдруг повернул вспять и, очутившись у гардероба, протянул свой жетон (М)

Рис. 16



Процедура организована так, что могут быть построены ИГ и ПГ любых грамматически допустимых комбинаций исходных линейных конфигураций.

8. Некоторые проблемы поиска N-хозяина для A-согласованного определения в препозиции

Необходимое условие: справа от A есть N – потенциальный хозяин этого A.

Рассмотрим некоторые особенности анализа определительных конструкций с необособленным A в препозиции, когда справа от A найден N – потенциальный хозяин этого A: A ... N.

Один из распространенных вариантов такой ситуации (в силу регулярной омонимии Род. и Вин.): A и N – потенциальный хозяин – согласуются одновременно по Род. и Вин: **рисующего сына**. Между A и его потенциальным хозяином могут находиться N – слуги этого A, но, так как управление несильное, т.е. заполнение объектной валентности **рисующего** необязательно, то N может быть проинтерпретировано и как хозяин (вар. 1: **сына R рисующего**), и как слуга A (вар. 2: **рисующего R сына**).

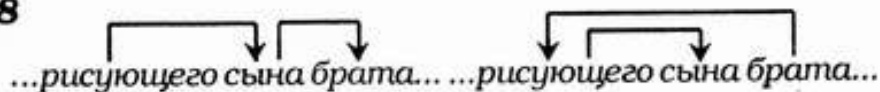
Если правее есть еще одно N подходящее на роль хозяина (в данном случае – в Вин. \ Род), для Вар. 1 возникает интерпретация (Рис. 17)

Рис. 17



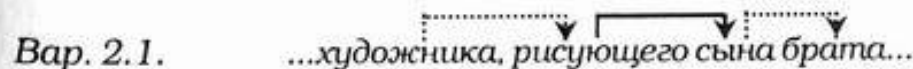
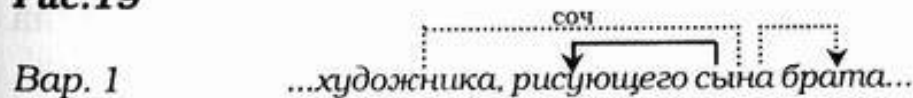
Вар. 2 позволяет построить два разных графа (Рис. 18):

Рис. 18



Если слева от A есть запятая, а левее этой запятой стоит еще одно N, согласующееся с A, то соответственно нужно зарезервировать возможность построения трех вариантов (Рис. 19)

Рис. 19



При этом для Вар. 1 и 3 анализ на данном этапе закончен, а для Вар. 2 нужно предусмотреть возможность появления линейной конфигурации, порождающей две интерпретации связей (Рис. 20):

Рис. 20

Вар. 2.2.

....художника, рисующего сына брата друга...

Вар. 2.3.

...художника, рисующего сына брата друга...

Вар. 2.2. предполагает очередной цикл поиска правее еще одного потенциального слуги \ хозяина А. Анализ будет закончен, как только окажется, что очередного такого N правее нет.

Мы видим, что при поиске хозяина справа от А неоднозначности может порождать (1) необязательность заполнения валентности А в совокупности с (2) возможностью совпадения падежей, по которым А и N согласуются, с падежами N, которыми А может управлять.

Неоднозначности такого типа могут возникать не только для Род \ Вин., но во всех случаях, когда падеж потенциального слуги А совпадает с падежом, общим у А и его потенциального хозяина. Во всех подобных случаях при анализе необходимо расширение контекста: ...любующимся сыном...: (...наблюдал за) отцом, любующимся сыном... \ (...наблюдал за) отцом, любующимся сыном братом...

Эту и еще одну важную проблему анализа ИГ иллюстрирует предложение *Перед ним был дуб, почти засохший этим летом, усыпанный **изъеденными гусеницами** молодыми побегами куст сирени, старая, завалившаяся кое-где и поломанная зимними ветрами изгородь, поросшая мхом* (Л.Н.Толстой). В ПФ *усыпанный **изъеденными гусеницами** молодыми побегами куст* (Рис.21) есть вложенная «матрешка» *изъеденными гусеницами* подходит и на роль слуги, и на роль хозяина причастия *изъеденными*, и только наличие второго претендента на роль хозяина – *побегами* – позволяет нам утверждать, что первый претендент на самом деле не хозяин, а слуга⁷.

Рис. 21

УСЫПА́ННЫЙ ИЗЪЕДЕ́ННЫМИ ГУСЕ́НИЦАМИ МОЛО́ДЫМИ ПОБЕГА́МИ КУСТ

Это предложение интересно еще и классическим случаем неоднозначности при интерпретации интерпозиции обособленного оборота *почти засохший этим летом*, причем стоящего не непосредственно перед потенциальным хозяином справа, а перед довольно длинным ПФ ИГ. Это распространенное определение можно интерпретировать и как обособленное

определение к *дуб* (R *засохший*) и как необособленное или обособленное (запятая после *летом* может быть проинтерпретирована и как правая граница обособленного определительного оборота, и как оператор сочинения *засохший* и *усыпанный*) пространственное определение в препозиции к *куст* (R *засохший*). В последнем случае запятая после *дуб* сочиняет *дуб* и *куст* и, если считать, что определение *почти засохший этим летом* обособлено, является еще и левой его границей.

Управление А практически всегда несильное, но возможность появления у А слуг нужно учитывать всегда, так как ни для какого А нельзя утверждать, что у него не может появиться слабоуправляемый слуга.

9. Ограничение зоны поиска N – хозяина А и N – слуги Р

8.1. Так как на этом этапе мы ставим задачу построить связи, определяющие границы проективной матрицы, границы поиска хозяина А – те части речи и знаки препинания, которые не могут находиться внутри такого ПФ, так как их появление там потенциально ведет к нарушению проективности.

К таким «нарушителям» проективности относятся любые вершины сегментов – сказуемые: глаголы в личной форме, краткие А, деепричастия, А – вершины обособленных оборотов, подчинительные союзы, свободные⁸ N (полнозначные и местоимения) в неомонимичном Им.п. и т.д.

8.2. Если при таких границах поиска мы не находим N, согласующегося с А, нужно учесть ситуацию возможного вложения в ПФ сегмента – в настоящее время алгоритм учитывает возможность вложения обособленных оборотов с **как** (*Летящий рысью маленький, как мальчик, темный, как мулат, командир алы-сириец, равняясь с Пилатом, что-то тонко крикнул и выхватил из ножен меч (Б.)*) и хотя (*Но какое оскорбление – северная, хотя и грамотная, речь раввина (М)*)

8.3. Не является границей поиска хозяина запятая непосредственно перед N- хозяином необособленного А. В-первых, из-за возможного обособления в препозиции перед местоименным существительным (Рис.22):

Изменяя и отбросив последнюю газету, все высосав, все узнав, сжигаемый неотвязным зудом, желанием принять мне одному понятные меры, я сел за стол и начал писать (Н)

Рис.22

СЖИГАЕМЫЙ НЕОТВЯЗНЫМ ЗУДОМ, ЖЕЛАНИЕМ ПРИНЯТЬ МНЕ ОДНОМУ ПОНЯТНЫЕ МЕРЫ, Я

Или же перед полнозначным существительным (Рис.23):
Грязная, на серой древесной бумаге, всегда похожая на корректуру, газетка...(М)

Рис.23



Но возможны и другие причины появления запятой непосредственно слева от N-вершины ИГ, например, у П.Флоренского:

Он вносит посторонние, чуждые этому духовному началу, черты.

Запятая непосредственно перед N может быть проинтерпретирована и как результат обособления уточняющего A=посторонние оборота.

10. Заключение

Рассмотрены линейные особенности проективных фрагментов ИГ и ПГ и механизмы их построения, работающие как фильтры проективности в процессе анализа, и некоторые свойства их линейной структуры, которые необходимо учитывать, чтобы строить ПФ ИГ и ПГ любых линейно-комбинаторных конфигураций.

Показано, что потенциально рекурсивные структуры ПФ ИГ и ПГ удобно строить, обрабатывая A и P в предложении справа налево, начиная с конца предложения. При этом анализ каждой матрешки проективных фрагментов начинается с самого глубокого вложения.

Рассмотренная модель использует сравнительно небольшой исходный набор возможных линейных конфигураций. Исключая при каждой очередной итерации процедуры анализа уже построенные ПФ, она дает возможность рассматривать самые разные их линейные комбинации, возникающие в тексте.

Такой способ организации анализа потенциально рекурсивных структур удобен в русском предложении при анализе синтаксических структур всех уровней (ср. [Кобзарева 2004])

Примечания

¹ Работа подготовлена при частичной поддержке гранта РФФИ №03-06-80434.

- 2 Система ориентирована на анализ любых неэллиптических русских предложений, не являющихся записью или имитацией устной речи.
- 3 В процессе отладки программных реализаций алгоритмов возник небольшой, но представительный корпус примеров из прозы Мандельштама, Набокова, Булгакова и др., подтверждающих необходимость рассмотрения излагаемых проблем для возможности работы с естественным текстом.
- 4 Приводятся в графическом виде примеры результатов анализа: с круглыми стрелками (кроме двух примеров со ссылкой на [Баталина и др. 2006]) – из частичной программной реализации системы [Ножов 2003], с квадратными ломаными стрелками связей – из дипломной работы студента ф-та Прикладной математики МАИ Пашковского П.В.
- 5 Можно считать, что у А есть Вин.одуш., совпадающий с Им., и Вин.неодуш., совпадающий с Род., но при любом представлении для Вин.п. нужна проверка одушевленности у потенциального хозяина.
- 6 Он может иметь слугу, общего с сочиненным с ним предлогом (до и после работы), или – при эллипсисе – не иметь слуги справа (*Вам с сиропом или без?*).
- 7 В ответ на вероятное возражение, что спасти в таких ситуациях может семантика, хочется отметить, что нельзя утверждать, что невозможен контекст, где *гусеницы м.б. изъеденными*.
- 8 Здесь важно, чтобы N в Им.п., которые не могут претендовать на роль подлежащего или сказуемого, были к этому моменту анализа уже обработаны, что в системе обеспечивают алгоритмы построения конструкций с именами собственными, числами и т.д.

Литература

- Апресян 1989 – Лингвистическое обеспечение системы Этап-2 / Апресян Ю.Д., Богуславский И.М., Иомдин Д.Л., Лазурский А.В., Перцов Н.В., Санников В.З., Цинман Л.Л. М.: Наука, 1989.
- Баталина А.М. и др. Опыт экспериментальной реализации алгоритмов поверхностно-синтаксического анализа / Баталина А.М., Епифанов М.Е., Кобзарева Т.Ю., Кушнарёва Е.В., Лахути Д.Г. // Компьютерная лингвистика и интеллектуальные технологии, Труды Международной конференции «Диалог 2006» (Бекасово, 31 мая – 4 июня 2006 г.) М.: Наука, 2006. С. 51-56.
- Иорданская Л.Н. Автоматический синтаксический анализ. 1961. Т.2. Межсегментный синтаксический анализ. Новосибирск: Наука, 1967.
- Кобзарева Т.Ю., Лахути Д.Г., Ножов И.М. Модель сегментации русского предложения // Труды Международной конференции «Диалог 2001». Аксаково, 2001. т.2. С.185-194.

- Кобзарева Т.Ю. Некоторые аспекты анализа сочинения при сегментации русского предложения // КИИ'2002. Труды восьмой национальной конференции по искусственному интеллекту с международным участием. М.: Физматлит, 2002. Т. 1. С. 192-198.
- Кобзарева Т.Ю. Принципы сегментационного анализа русского предложения // Московский лингвистический журнал М., 2004. Т. 8 №1. С. 31-80.
- Кобзарева Т.Ю. Омонимия и синонимия знаков препинания в русском тексте // Труды Международной конференции «Диалог'2005». М.: Наука, 2006. С. 233-237.
- Кобзарева Т.Ю. Рекурсивность и проективность сочинительных связей в русском тексте // Компьютерная лингвистика и интеллектуальные технологии Труды Международной конференции «Диалог 2006» (Бекасово, 31 мая – 4 июня 2006 г.). М.: Наука, 2006. С. 223-229.
- Кобзарева Т.Ю. Иерархия задач поверхностно-синтаксического анализа русского предложения // НТИ. Сер.2. 2007. № 1. С 23-35.
- Кулагина О.С. Исследования по машинному переводу. М.: Наука, 1979.
- Мельчук И.А. Автоматический синтаксический анализ. 1964. Том 1. Общие принципы. Внутрисегментный синтаксический анализ. Новосибирск. 1964.
- Мельчук И.А. Согласование, управление, конгруэнтность // ВЯ. 1993. №5. С. 17-57.
- Ножов И.М. Процессор синтаксической сегментации русского предложения. // НТИ. Сер. 2. 2003. № 11. С. 26-37.
- Падучева Е.В. О порядке слов в предложениях с сочинением: сочинительная проективность // НТИ. Сер. 2. 1971. №3. С. 14-18.
- Санников В.З. Место распространенного определения по отношению к определяемому слову в русской фразе // ВЯ. 1963. №1. С. 124-130.

Об использовании семантических отношений для описания изображений

Формальный аппарат, который используется в семантических исследованиях, часто основывается на отношениях вида R(A B). Такие отношения были предложены Н.Н. Леонтьевой для записи информации, содержащейся в тексте, например, **локализация (стена, картина) – картина находится (висит) на стене** [Леонтьева 1967], раньше, чем они стали использоваться в лингвистических моделях для представления семантики, в частности, основанных на глубинных падежах. Работа Филлмора, открывшая для западной лингвистики глубинные падежи [Fillmore 1968], появилась годом позже. Глубинные падежи Филлмора описывали падежную рамку глагола, например, **actor (John, break), theme (window, break) – John breaks the window**, т.е. семантическим действующим лицом глагола *break* является *John*, а объектом – *window*. Таким образом, семантические отношения (СО) могут представлять и отношения конкретных предметных областей, в которых им придается точный смысл, в частности, в области лингвистики, и разнообразные отношения, обозначенные словами, представляющими понятия, свойственные им в языке, при использовании для описания информации текста [Леонтьева 2006].

В данной статье рассматривается задача создания формализованных описаний содержания изображений с использованием СО, в частности, фотографий открытого пространства, например, пейзаж, городской вид, жанровая сцена на открытом воздухе, а также способы реализации этих отношений в тексте на русском языке в задаче автоматической генерации текстов. Статья продолжает тему, представленную в [Соколова, Болдасов 2006]. Описание изображения – ситуация наиболее тесной связи текста и предметного мира. Формализованное описание содержания изображения рассматривается как альтернатива текстовым описаниям, выполненным людьми, которыми могут сопровождаться изображения в коллекциях изображений.

1. Формализованное представление содержания изображения

Опыт показывает, что тексты описаний одного изображения, сделанные разными людьми, могут сильно различаться, даже если они представляют собой почти перечисление изображенных объектов, сравним, например, два описания одной фотографии, приведенные в [Добров и др. 2000]:

(1) *Ярко-синее небо, ветки с маленькими листочками, верхняя часть храма белого цвета с зеленым куполом.*

(2) *Башня с шатрового типа зеленым шпилем на фоне очень синего неба. Наверху что-то типа герба. Здание покрашено в розовый цвет с белой отделкой.*

Едва ли человек или тезаурус [Добров и др. 2000] или модель лингвистического анализа с перифразированием [Eteljanov et al. 2000] могут отождествить эти описания с одним и тем же изображением.

При интерпретации текстового описания картинки можно получить описание в терминах СО, например, для текста: *На фотографии изображен дом с тремя окнами. Рядом с домом – машина, может быть получено следующее описание:*

актант (фотография, изображать)

объект (дом, изображать)

количество (три, окно)

часть ((количество (три, окно)), дом))

рядом_с (машина, дом)

Достоинство этого представления состоит в том, что его легко получить и можно дальше интерпретировать, производя некоторые выводы [Леонтьева 2006]. Существенный недостаток состоит в том, что его можно получить только из текста, описывающего изображение, а не из самого изображения.

Можно было бы регламентировать процесс описания, но более интересной и радикальной является задача получения формализованного описания непосредственно из самого изображения. Такая возможность появляется, когда формализованное представление содержания изображения, пока только в идеале, получается как результат работы с изображением другой компьютерной системы, например, системы распознавания образов. В нашем случае это виртуальная, идеализированная система распознавания образов, обладающая знаниями, результат работы которой формулируется человеком. Назовем ее «распознаватель». Далее следует помнить, что мы забегаем вперед, и моделируя результат действий этого, пока

не существующего, распознавателя, пытаемся предвидеть, каким может быть формализованное представление содержания изображения, построенное им.

Распознаватель выделяет во внутрикомпьютерном представлении изображения – матрице пикселей, объекты и может фиксировать их взаимное расположение. Для распознавания объектов нужны знания. Это можно показать, в частности, на следующем примере. Каждый объект занимает некоторую поверхность фотографии, и те объекты, которые находятся ближе к наблюдателю, могут частично закрывать те, которые находятся дальше. Однако тот факт, что часть поверхности, занятая солнцем, закрывает на изображении участок неба, не означает, что солнце находится ближе к наблюдателю, чем небо. Таким образом, представление должно отражать не только особенности зрительного восприятия человека, но и его знания о мире. Далее рассматривается только зрительно воспринимаемая информация. Таковой не является, например, информация о том, что конкретного изображенного человека зовут Марк Шагал [Hollink et al. 2003], но эта информация может добавляться в представление из других источников знаний, например, из базы данных по художникам. В частности, в рассматриваемом материале – коллекции цветных фотографий, сделанных в начале прошлого века С.М. Прокудиным-Горским [Мир 1900-1917 в цвете], фотографиям часто сопутствует название населенного пункта или района, представленного на фотографии, род занятий изображенного человека и др. Для части этих фотографий студентами 4-го курса РГГУ были созданы текстовые описания и формализованные представления содержания изображений, которые послужили непосредственным материалом данного исследования.

В изображении распознаватель прежде всего распознает объекты и их свойства. Они описываются в виде XML представления, в котором все изображение представлено в виде дерева составляющих, а каждый объект представляется в виде узла этого дерева. От узла зависят его свойства и узлы, представляющие связанные с ним более мелкие объекты. Распознаватель устанавливает между объектами отношения двух типов: отношения **вложения** – вложение поверхности изображения, представляющей один объект (вложенный объект), в границы фрагмента изображения, представляющего другой объект (объемлющий объект), и **пространственные отношения** – ориентация объекта относительно границ изображения и других изображенных объектов. Распознаватель, как и наше зрение группирует признаки объектов, например, одинаковую направленность, цвет, и т.п., собирая их в единый фрагмент

изображения, соответствующий образу объекта. Кроме того, распознаватель группирует объекты так, как это сделал бы описывающий содержание фотографии человек. Формируются групповые объекты двух типов – <группа> однородных объектов, образующих общий фрагмент поверхности фотографии, например, семья, позирующая перед объективом, и <план>, объединяющий объекты, находящиеся приблизительно на равном расстоянии от наблюдателя, т.е. приблизительно на одной горизонтальной линии фона, хотя и не образующие общей поверхности, например: *На фотографии изображен ясный зимний день. [На первом плане мы видим детей, которые играют в снежки на большой лужайке.] (план 1) [На заднем плане – лес.] (план 2).*

Чтобы записать знания о постоянных свойствах и отношениях между объектами, а также избежать произвола в языковом выражении объектов и отношений, предполагается наличие специального ресурса – онтологии, содержащей, в частности, информацию об объектах и их назывании на русском и, возможно, других, языках. В онтологиях описываются совокупности понятий (концептуализации), представленные словами языка (значения) или понятиями конкретной области знания, между которыми устанавливаются отношения, свойственные данной совокупности понятий. Например, для онтологии лексических значений это отношения синонимии, гипо-гиперонимии, меронимии, и др., см., например, крупнейшую лексико-семантическую базу английского языка WordNet [Miller 1995].

Задача описания содержания фотографий открытого пространства не связана с узкой предметной областью, поэтому онтология должна содержать в основном понятия представленные значениями слов общеязыкового пласта лексики, но при этом она ограничена представлением конкретного аспекта восприятия мира – зрительного восприятия. Само понятие в онтологии может обозначаться словом и описываться как зрительный образ с указанием специальных сведений в терминах распознавания образов – характеристики контуров, текстуры, цвета и т.п. Кроме того, должны указываться зрительные признаки объекта, которые частично присутствуют в лингвистических описаниях, например, «вытянутый в длину пространственный объект – часть земли» (с текстурой земли, асфальта – *дорога*, с текстурой воды – *река, канал*). Зрительные признаки не обеспечивают решения задачи распознавания объекта, например, дорога на изображении может выглядеть как часть фона внизу фотографии, где ее протяженность не видна, но они определяют семантические типы объектов, которые формируются семантикой культуры и языка и отобража-

ются в языке. Мы не будем обсуждать возможности автоматического распознавания объектов на изображении в терминах этих двух видов признаков, считая, что для нашей задачи объекты опознаются человеком.

Композиция изображения описывается отношением вложения. Оно устанавливается между объектом <фото> и структурными элементами <план>, <группа>, а также между <объект>ами. Отношение вложения между объектами семантически неоднородно, поэтому в тексте оно выражается по-разному, в зависимости от семантики узлов объектов. В частности, и *окна* находятся на фрагменте поверхности фотографии, занятой *домом*, и *солнце* находится на поверхности, занятой *небом*. Но мы говорим *на небе находится солнце*, и не говорим **на доме находятся окна*. Семантическое различие интерпретации отношения вложения в данном случае состоит в том, что *солнце* является самостоятельным объектом по отношению к *небу*, хотя и связано с ним ассоциативно, а *окна* являются частью *дома* по определению. Таким образом, при отображении в текст отношения вложения должны быть интерпретированы с помощью онтологии.

На рис. 1 показан пример формализованного представления, в котором **окно вложено** в понятие **дом**, а понятия **небо** и **солнце** представлены как одноуровневые объекты. Между ними установлено отношение локализации, которое также является интерпретацией отношения вложения. Пространственные отношения представления помещены в специально созданную для этого отдельную составляющую <СО>.:

Рис. 1. Формализованное представление содержания изображения, которое может быть выражено, например, следующим текстом: «На первом плане мы видим дом. За ним – синее небо и солнце»

```
<фото id=«фото1»>
  <план id=«план1»>
    <объект понятие=«дом» id=«house1»>
      <объект понятие=«окно» id=«window1»/>
    </объект>
  </план>
  <план id=«план2»>
    <объект понятие=«небо» id=«sky1»>
      <свойство понятие=«синий» параметр=«цвет»/>
    </объект>
    <объект понятие=«солнце» id=«sun1»/>
  </план>
  <СО>
    <СО type=«локализация» domain=«sky1» range=«sun1»/>
  </СО>
</фото>
```

2. Выражение формализованного представления изображения средствами русского языка

Теперь рассмотрим возможности выражения формализованных описаний содержания изображений в виде текстов на русском языке. Прежде всего надо определить последовательность изложения, т.е. порядок упоминания объектов в тексте. Мы принимаем принцип планирования текста «от общего к частному», характерный для текстов типа отчет [Иорданская 1992]. Таким образом, структура текста определяется композицией изображения или структурой самого формализованного представления, просматриваемого в определенном порядке – от самого общего понятия <фото> до самых мелких деталей объектов. Стратегия «от общего к частному» рассматривается как нейтральная. В конкретных задачах могут использоваться другие стратегии. В частности, если пользователь заинтересован в конкретном объекте, могут описываться свойства и окружение только этого объекта. Дробность описаний в тексте зависит от конкретной задачи и характера изображения.

Относительно каждого объекта мы можем знать его онтологические свойства и отношения с другими объектами, а также зрительные свойства этого объекта на изображении – цвет, текстура, форма и др. Сопоставление объекту его языкового выражения основывается на образе объекта, представленном словом русского языка с учетом ограничений или признаков понятия. Ограничения неформально описываются в словарях, например, в словарных статьях «Нового объяснительного словаря синонимов русского языка» [Новый объяснительный словарь 1997], в котором понятие ДОМ 1 может быть выражено одним из следующих слов: *дом, здание, строение, постройка, корпус*, с учетом следующих признаков:

- размер, например, *здание* – большое;
- материал; например, *здание* – каменное, а *постройка* – деревянная;
- функциональность, например, *дом* – жилой, *постройка* – хозяйственная;
- изолированность, например, *постройка* – изолированная, *корпус* – один из ряда;
- время, например, *корпус* – современный.

Итак, каждый объект, представленный на изображении, может быть назван существительным русского языка или целой именной группой, включающей вложенные объекты. Но для построения предложений текста нужны не только существительные, но и глаголы – будущие сказуемые. В нашем слу-

чае они должны получаться в результате интерпретации отношений между объектами. Ниже мы подробно рассматриваем эти отношения.

Текст описания открытого пространства может начинаться с общей характеристики изображенного места, представленной следующими типами:

1. обобщения изображенных объектов до типа местности, например, *На картинке изображен горный пейзаж (городской вид)*; *Это фотография городской улицы*;

2. состояния погоды: *На картинке показан летний день после грозы*. Глагол *показан* употребляется только для представления состояния погоды.

3. указания точного названия местности, например, *На фотографии запечатлены Елисейские поля*;

4. указания места съемки: *Фотография сделана с возвышенного места, вероятно, крыши дома, (из окна высотного дома, с дома)*;

5. общей художественной характеристики изображения: *Фотография сделана в теплых желто-зеленых тонах*, эта характеристика может комбинироваться с другими: *Фотография в темных тонах, сделана с моста*.

Перечисленные возможности связаны с обобщением состава объектов или использованием дополнительной, не зрительной информации. Возможность формализованных построений таких обобщений – тема отдельного исследования.

Рассмотрим возможности прямого, без обобщений, выражения отношений из рассмотренного материала средствами русского языка.

2.1. Выражение отношения «вложение»

Рассмотрим сначала отношение вложения между понятиями с семантикой «носитель информации» и <объект>. Семантику «носитель информации» имеет само понятие <фото> и структурные элементы описания <план>. Текст описания содержания фотографии обычно начинается с отношения между объектом <фото> и представленными на ней объектами <объект>, прямо или косвенно вложенными в него. Поскольку фотография – это единственное, что дано человеку или системе генерации перед созданием текста, оно находится в фокусе внимания автора и является темой первого предложения. Тема может быть выражена подлежащим или обстоятельством места, помещенным в начальную позицию. Обстоятельство выражается предлож-

ной группой с предлогом «на». Выбор предлога определяется тем, что <фото> и <план> концептуализируются как плоскости. В русском языке роль объекта, концептуализированного как плоскость, и являющегося местоположением другого объекта, выражается предлогом «на», аналогичная роль объекта, не воспринимающегося как плоскость, выражается предлогом «в»¹. Этот признак не является чисто лексическим, а зависит от семантики употребления слова, в частности, *на небе солнце* (небо – плоское) vs. *в небе (летит) самолет* (небо не плоское).

Рассмотрим реализацию отношения **вложение (объект, фото)**. В нашем маленьком корпусе описаний не встретилось выражение понятия <фото> подлежащим, например, *Фотография изображает (объекты)*, а только тематизацией обстоятельства места: *На фотографии изображены (объекты)*. Отношение вложения между фотографией и объектом не выражается глаголами *запечатлен, показан*, которые используются только для характеристики содержания фотографии в целом (см. выше). Не встретилось также *представлен*, хотя интуитивно кажется возможным. Рассматриваемое отношение выражается в наших примерах тремя способами:

а) пассивной формой глагола изображения (*изображен*), подлежащим которого является изображенный объект, а обстоятельством места – <фото>. С учетом тематизации понятия <фото> используется предложение с инверсией подлежащего: *На фотографии изображены (объекты)*. Эта форма является наиболее употребительной;

б) глаголом зрительного восприятия с позиции наблюдателя с обстоятельством места в обобщенно-личном предложении: *На фотографии мы видим (объекты)*.

в) нулевой связкой, семантика которой соответствует семантике реализаций **а)** и **б)**. У этого способа есть ограничение: в качестве изображаемого называется объект, представляющий большую часть фона изображения, например, *На фотографии пересечение двух больших городских улиц*; *На картине река, уходящая вдаль*; или перечисленные объекты покрывают все пространство изображения, например, *На фотографии часть тротуара вдоль глухой стены*. В последнем случае часть тротуара и нижняя часть глухой стены составляют весь фон фотографии.

Отношение вложения между изображением и структурными элементами **вложение (план, фото)** в наших описаниях не реализуется, хотя, в принципе, это возможно, например, *На фотографии мы можем выделить два плана*.

Отношение **вложение (объект, план)** выражается более разнообразно, чем отношение **вложение (объект, фото)**. Понятие <план> при реализации также тематизируется и выражается обстоятельством места. Может использоваться способ **а)** и способы **б)** и **в)**:

а) На переднем плане изображен мостик;

б) На заднем плане мы видим забор.

в) На дальнем плане лес; На переднем плане – подземный переход. В отличие от отношения с <фото>, ограничений на тип объекта нет.

Кроме **а), б)** и **в)** используются и другие способы:

г) безличная форма с модальным словом *можно*, в которой роль модального слова состоит в обезличивании глагола: На переднем плане можно увидеть поле, на нем – дорогу.

д) краткое страдательное причастие глагола зрительного восприятия *виден*, *различим*: На заднем плане видны жилые постройки и небольшая полоска синего неба. Этот способ используется исключительно с «задним планом» или выражающими его наречиями или словосочетаниями, например, *вдали*, *на горизонте*, *вдалеке*. Изображенными объектами в связи с уменьшением масштаба в этом случае могут высокие, большие или капитальные, а также пространственные объекты – дом, лес, труба, строительный кран, горы, площадь, забор и др.

е) глагол *находится*: На переднем плане фотографии находится река. Глагол *находится* отражает другую общую перспективу: автор текста смотрит на фотографию не снаружи, а изнутри, как бы находясь внутри изображения. В этом случае он может ориентировать изображенные объекты относительно изображенного пространства, в частности, структурных объектов, хотя и мнимых. Заметим, что для реализации отношения с <фото> глагол *находится* не употребляется.

ж) связка, выраженная знаком тире (–) с <планом> в качестве подлежащего: *Передний план – речная отмель*. Здесь нулевая связка имеет другую семантику, чем в случае в), а именно, идентификация (есть). Выполняются следующие ограничения: весь план занят данным объектом; объект – часть земли.

Отношение вложения между двумя изображенными объектами **вложение (объект, объект)** представляется формой зависимого существительного в составе именной группы, например, *ступеньки крыльца, крыльцо с пятью ступеньками*. Напомним, что речь идет о двух предметных понятиях – объектах, а не об абстрактных понятиях, которые имеют другую семантику и источник, например:

- совокупность – букет цветов, комплекс зданий, толпа людей (количество объектов);
- форма – треугольник неба, полоска неба (форма объекта).

В контексте двух объектов генитивная конструкция, если рассматривать язык как средство коммуникации, представляет собой оператор, который придает объекту, представленному главным существительным именной группы, роль атрибута по отношению к объекту, представленному зависимым существительным в родительном падеже. В [Соколова, 2003] это отношение мы назвали «атрибут-объект». В наших примерах это: *крыша дома, крона дерева, шесть паромщика* и др. Иногда добавляется предлог, например, *пуговица от пальто*.

Обратная ситуация представлена тоже сочетанием двух предметных понятий, представляющих собой объемлющий и вложенный в него объекты. Она выражается именной группой, в которой объекту, представленному зависимым существительным, придается роль атрибута объекта, представленного главным существительным. Это отношение интерпретируется нами как отношение «объект-атрибут», например, *дом с тремя окнами, человек в кепке*. В отличие от языковой реализации отношения «атрибут-объект», языковая реализация отношения «объект-атрибут» разнообразна и чувствительна к семантике связываемых объектов [Соколова 2003]. Семантика объектов определяет и способ реализации этого отношения в виде предложной группы или с помощью глагола:

з) вложение (часть_объекта, объект) - в виде предложной группы с предлогом «с» или глаголом: *дом с тремя окнами, дом имеет три окна*. Таким же способом реализуется композитное отношение **вложение(вложение (объект 1, рука), человек)**, причем рука интерпретируется как место и не плоскость, например, *Человек с сумкой в руке*, где сумка – объект, или *Человек держит в руке сумку*.

и) вложение (одежда, человек) в виде предложной группы с предлогом «в»: *человек в кепке, на человеке надета кепка*. Если перечисленная одежда в основном закрывает тело одетого существа, то употребляется сочетание *быть одетым в*: *Человек одет в пальто и кепку, Человек одет в пальто*. **Человек одет в кепку* – некорректно.

Каков источник появления в тексте глаголов *иметь* и *быть одетым* в рассмотренных случаях? Возможной интерпретацией является такая: зрительное отношение вложения между двумя объектами интерпретируется на основе знаний о мире как функциональная локализация. Функциональная обознача-

ет дополнительную к местному отношению функциональную связь между двумя понятиями: одежда согревает, украшает, скрывает и т.д., часть объекта выполняет в нем определенную функцию. Эти отношения, с учетом семантики обоих связанных объектов, можно обобщить как функциональную локализацию (функ-лок):

функ-лок (X, Y), где X - часть объекта или ее разновидности – одежда, неотъемлемая часть, и др., а Y – целый объект.

Другой вид отношения между вложенным и объемлющим понятиями образуется между объектом и его свойством: **вложение (свойство, объект)**. Наиболее прямым выражением этого отношения является соединение имени объекта со значением свойства, например, *синее небо*, или сочетанием значения свойства с названием параметра, например, *небо синего цвета*. Однако эти отношения также могут реализоваться глаголами, например, *крыша, покрыта железом и окрашена в зеленый цвет*. Для выяснения, какой глагол нужно употребить, можно использовать следующий тест: Задать вопрос: *Как приобретено объектом это свойство?* И ответить на него. В ответе будет содержаться нужный глагол, например:

– Как крыша стала железной? – Ее покрыли железом. – Крыша покрыта железом.

– Как крыша стала зеленой? – Ее покрасили в зеленый цвет. – Крыша окрашена в зеленый цвет.

– Как небо стало синим? – Небо всегда имеет синий, голубой, белый, розовый, ..., черный цвет. *Небо имеет синий цвет / небо синее.*

2.2. Выражение пространственных отношений

Мы различаем два вида пространственных отношений – отношение локализации и отношения ориентации. Отношение локализации является интерпретацией отношения вложения: в тех случаях, когда связанные отношением вложения объекты не образуют атрибутно-объектных или функционально-локативных пар, они находятся в отношении локализации. Образы объектов локализованной пары частично пересекаются или один вложен в другой на изображении. Отношение локализации формулируется на основе наших знаний о свойствах объектов. Отношение **локализация (объект, объект)** реализуется в соответствии со сказанным выше предложной группой с предлогом «на» или «в» в зависимости от семантических признаков узлов. Поверхность земли всегда представляет плоскую поверхность, например, *На берегу большое количество церквей, старинных сооружений, деревьев.* Другие понятия,

содержащие плоскости, на которых могут находиться люди и другие объекты, также употребляются с предлогом «на» – *На балконе, на лавке, борту пароходика*, не плоские – с предлогом «в», например, *в лодке*.

Кроме нейтральных глаголов типа *находится, есть* и нулевой связки в наших описаниях два объекта, связанных отношением локализации, соединяются и другими глаголами местоположения, учитывающими дополнительные признаки семантики слов:

простирается (забор), *поднимается* (холм), *раскинулся* (городок), *проходит* (дорога), *расположен* (труба, часовня, баня, пристань), *лежат* (бревна), *покрыт/порос* (холм/берег/небо травой/кустами/лесом/облаками), *сидит* (мужчина крыльцо), *стоят* (люди, стога, ели и сосны, лавка, мужчина), *плывет* (пароходик), *отражается* (небо в реке), *ограждено* (здание забором).

Отношения ориентации уточняют композицию фотографии, описанную отношением вложения. При этом уточняется местоположение изображенного на фотографии объекта относительно частей фотографии или относительно других изображенных объектов.

Отношение ориентации связывает объект с частью изображения, например, **л_часть (объект, фото)** – объект находится на левой стороне фотографии. Части фотографии – это понятия, представляющие части поверхности фотографии: **п_часть, в_часть, н_часть, центр**, а также границы фотографии: **п_край, л_край, в_край и н_край**, а также углы: **лв_угол, лв_угол, пв_угол, пн_угол** – соответственно, правая часть, верхняя, и п.д., а также левый верхний угол, и т.д. В тексте они обычно выглядят как словосочетание, например, *В верхней части фотографии мы видим радугу*, или соответствующими наречиями – *вверху, внизу, справа, слева*. Интересно, что, хотя части изображения являются так же, как и сама фотография изображениями, т.е. поверхностями с представленными на них объектами, в тексте мы их выражаем отлично от объектов типа <фото> или <план>. С этими объектами всегда используется способ реализации с предлогом «в». Очевидно, что они концептуализируются языком не как плоскости. Они выражаются следующими способами:

в) нулевой связкой, например, *В верхней части изображения небо сероватого цвета*.

г) обобщенно-личным предложением с модальным словом *можем*: – *В левой части фотографии мы можем видеть людей, стоящих на набережной*.

Также отношение ориентации может содержать указание на ориентацию одного изображенного объекта относительно другого изображенного объекта с точки зрения наблюдателя. Они выражаются с помощью нулевой связки и предлога, выражающего ориентацию, например:

позади (лес, река) – *За рекой лес.*

впереди (открытое пространство, школа) – *Перед школой открытое пространство.*

слева (памятник, здание) – *Слева от здания (высокий белый) памятник.*

справа (стройматериалы, здание) – *Справа от здания стройматериалы.*

около (лодка, берег) – *У берега лодка. Рядом со школой лежат какие-то бревна.*

Несколько «одиночных» отношений ориентации могут быть заменены одним отношением **вокруг**, например, на фотографии изображены несколько (не менее трех) объектов примерно на одинаковом расстоянии от другого и между собой:

впереди (столб, дом), впереди (кусты, дом),

слева (клумба, дом) → вокруг ((столб, кусты, клумба), дом)

Это несмотря на то, что заднюю и правую части дома мы видеть не можем. Например:

Вокруг дома размещены клумба с цветами, фонарный столб, небольшие кусты.

Мы рассмотрели особенности формализованного описания содержания изображения и выделили типы понятий – объектов и отношений, которые могут быть выражены на естественном языке в связном текстовом описании содержания изображений, а также рассмотрели способы их выражения на русском языке. Эта ситуация является обратной по отношению к обычному толкованию и описанию значений лексических единиц в словарях. Построение работающей модели генератора текстов, описывающих изображения, позволит формализовать определенные аспекты лингвистических толкований значений слов, обозначающих воспринимаемые зрением объекты и некоторые условия употребления слов, связанных со зрительным восприятием.

Примечания

¹ В английском языке выбор делается между тремя возможностями – *at, on, in*.

Литература

- Добров Б.В. и др. Постановка эксперимента по индексированию корпуса изображений с помощью специализированного тезауруса / Добров Б.В., Сидоров А.В., Юдина Т.В., Лукашевич Н.В., Салий А.Д. // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Вторая всероссийская научная конференция, Протвино, 2000.
- Иорданская Л.В. Коммуникативная структура и ее использование в системе текстовой генерации // Международный форум по информатизации и документации. 1992. Т. 17, №2. С. 15-28.
- Леонтьева Н.Н. Об одном способе представления смысла текста // Труды 3-й Всесоюзной конференции по Информационно-поисковым системам и автоматизированной обработке научно-технической информации, Т. 2. М., 1967. С. 192-204.
- Леонтьева Н.Н. Автоматическое понимание текста: системы, модели, ресурсы. М.: Академия, 2006.
- Мир 1900-1917 в цвете. Все фотографии С.М. Прокудина-Горского, восстановлены в цвете www.prokudin-gorsky.ru.
- Новый объяснительный словарь синонимов русского языка / Под общ. ред. Ю.Д. Апресяна. Вып. 1. М.: Языки рус. культуры, 1997.
- Соколова Е.Г. Семантические свойства имен для генерации запросов к базе данных // Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции «Диалог'2003». Протвино, 2003. С. 640-647.
- Соколова Е.Г. Болдасов М.В. Принципы построения семантических аннотаций содержания изображений // Компьютерная лингвистика и интеллектуальные технологии: Труды Международной конференции «Диалог'2006» (Бекасово, 31 мая – 4 июня 2006). М.: Изд-во РГГУ, 2006. С. 474-480.
- Fillmore, Charles J. The Case for Case. // *Universals in Linguistic Theory / Bach and Harms (Ed.)*. New York: Holt, Rinehart, and Winston. 1968. P. 1-88.
- Emeljanov G. V. Tree grammars in the problems of searching for images by their verbal descriptions / Emeljanov G. V., Krechetova T. V., Kurashova E. P. // *Pattern recognition and image analysis*. 2000. Vol. 10, №4. P. 520-526.
- Hollink L., Shreiber G., Wielemaker J., Wielinga B. Semantic annotation of image collections // *Knowledge Capture 2003. Proceedings Knowledge Markup and Semantic Annotation Workshop, Florida, USA, October 2003 / Handschuh S., Koivunen M., Dieng R., Staab S., eds*. P. 41-48. Также: <http://www.cs.vu.nl/~guus/papers/Hollink03b.pdf>
- Miller G.A. WordNet: a lexical database for English // *Communications of the ACM*. 1995. 38(11). P. 39-41. Также: <http://www.acm.org/pubs/articles/journals/cacm/1995-38-11/p39-miller/p39-miller.pdf>. WordNet: <http://wordnet.princeton.edu/>.

О построении курса «Морфология» для специальности «Теоретическая и прикладная лингвистика»

Программа и формы контроля знаний

Предисловие

Образовательное значение и задачи курса

Курс «Морфология», или «Общая морфология» традиционно входит в учебные планы теоретической подготовки специалистов в области лингвистики.¹ В совокупности курсы фонетики, морфологии, синтаксиса и семантики, предметом которых является собственно языковая система, составляют ядро теоретического цикла, поскольку они формируют понятийный аппарат, используемый всеми другими лингвистическими дисциплинами: лексикографией, историей языка, социолингвистикой, психолингвистикой, лингвостилистикой и др. Будучи на первый взгляд узкоспециальной областью лингвистики, морфология занимает особое место в контексте всего гуманитарного образования. Именно в процессе изучения морфологии студенты вплотную сталкиваются с феноменом языкового знака, знакомятся с его структурой, типами, свойствами, функциями. Осознание знаковой природы языка способствует пониманию глубоких содержательных связей лингвистики с теми науками, которые в своих целях и под своим углом зрения изучают знаковые системы – с философией, семиотикой, культурологией, литературоведением. Еще одна важная задача курса – продемонстрировать многообразие конкретных реализаций в языках мира тех общих принципов, которые характеризуют язык как естественную семиотическую систему, присущую человеку.

Методические особенности курса

Лекционная часть. Названные задачи в значительной мере определили методику и композицию курса, который читается автором с 2001 года в Институте лингвистики РГГУ. В качестве ключевого для данного курса выбрано понятие язы-

ковой знак, через которое раскрываются основные понятия морфологии: *морфема, словоформа, грамматическая категория, лексема, парадигма, словоизменение, словообразование*. В качестве иллюстрации базовых понятий морфологии мы стремились использовать материал самых разных, с генетической, типологической и географической точек зрения, языков мира. Курс рассчитан на 28 часов лекционных и 28 часов практических занятий.

При изучении теоретической части курса в качестве обязательного чтения студентам предлагаются прежде всего главы книги «Курс общей морфологии» И. А. Мельчука. Эта пятитомная монография имеет, разумеется, совершенно иной формат, чем привычное учебное пособие. По существу это «энциклопедия» современной морфологии. В ней обсуждается практически весь спектр теоретических проблем морфологии, содержится подробное описание и критический анализ различных подходов к их решению, приводятся многочисленные, разнообразные и тщательно откомментированные примеры как из хорошо изученных, так и из «экзотических» языков, дается подробная библиография, включающая работы общетеоретического характера и работы, посвященные частным проблемам. В обязательном порядке студенты знакомятся с теми разделами книги, которые непосредственно касаются вопросов, включенных в программу учебного курса. Вместе с тем организация и способ подачи материала в монографии Мельчука дают им возможность при желании самостоятельно расширить свое представление о морфологии не только в плане общей теории, но и в плане объяснения морфологических особенностей конкретных языков.

Практическая часть. Основная цель лабораторных занятий – обучить студентов технике морфологического анализа, поэтому значительное место в практикуме отводится решению «лингвистических задач». В этих задачах студентам предлагается описать в терминах изучаемой теории специально подобранные примеры из какого-то конкретного (в том числе и неизвестного им) языка. Дидактический материал такого рода вообще довольно давно используется в процессе обучения лингвистов. В частности, одним из первых учебных пособий по морфологии, в котором представлен большой корпус лингвистических задач, является вышедший впервые в 1946 году и с тех пор неоднократно переиздававшийся классический учебник Ю. Найды (E. Nida, *Morphology*). К сожалению, из-за отсутствия необходимых сборников задач и методических

разработок в отечественных центрах подготовки лингвистов к этому материалу обращаются крайне редко. Исключением в этом отношении является Отделение теоретической и прикладной лингвистики Филологического факультета МГУ и Институт лингвистики РГГУ, где регулярно готовится специальный материал для школьных Олимпиад по лингвистике и математике, которые с 1965 года проводились в МГУ, а с 1988 года проводятся совместно МГУ и РГГУ. В качестве источника при подготовке практических заданий для данного курса использовались в основном задачи, предлагавшиеся школьникам на этих Олимпиадах. Однако, поскольку олимпиадные задачи в принципе рассчитаны на другую аудиторию и другие цели, для студентов готовятся определенным образом переработанные их варианты. Особенно существенно изменяется часть «Задание»).

Формы контроля

Основной формой промежуточного контроля является письменная работа, в которую наряду с теоретическими вопросами включаются лингвистические задачи по изучаемой теме. В конце курса студентам предлагается открытый тест, состоящий из 5-10 теоретических вопросов по разным темам программы. Образцы лингвистических задач и вопросы, предлагаемые в итоговых тестах, приводятся во второй части предлагаемой публикации.

Часть I. Программа курса «Морфология»

Тема 1. Предмет морфологии. Место морфологического компонента в модели естественного языка.

Соотношение понятий *слово, словосочетание, предложение* в современной лингвистике. Слово как центральный объект изучения морфологии. Неоднозначность понимания термина *слово* и попытки ее разрешения: содержание терминов *словоформа, лексема, вокабула*.

Понятие языковой *знак* – устойчивое соответствие между элементами звуковой и ментальной субстанций языка. Морфология – наука о структуре и свойствах языковых знаков.

Словосочетание и предложение – объекты, не имеющие знаковой природы; основания для непризнания словосочетания и предложения языковыми знаками: открытость класса, отсутствие жестких ограничений на сочетаемость. Представление словосочетаний и предложений в языковой модели.

Вопрос об универсальности морфологического компонента в моделях конкретных языков – существуют ли языки без морфологии?

Тема 2. Языковые знаки. Общие положения

2.1. Противопоставление *односторонние vs. двусторонние* объекты языковой модели.

Типы односторонних объектов:

- *фон, просодема* – языковые объекты, интерпретируемые в звуковую субстанцию языка;
- *сема* – языковой объект, интерпретируемый в ментальную субстанцию языка;
- *языковая операция* – операция, применяемая к языковым объектам; основные типы языковых операций, актуальных для морфологического компонента – *чередование, повтор, конверсия*.

Языковой знак (ЯЗ) – один из типов двусторонних объектов языковой модели. Формализм для представления языковых знаков (по И. А. Мельчуку): *языковой знак* – объект вида $\langle 'Y', X, \Sigma \rangle$, где *'Y* – *означаемое*, *X* – *означающее*, Σ – *синтактика* (сочетаемость) пары $\langle 'Y', X \rangle$.

Означаемое ЯЗ; соотношение понятий *означающее* и *смысл*; *означаемые* в разных языках.

Означающее ЯЗ: типы означающих: *означающие-сущности* (сегментные – цепочки фонов; *супрасегментные* – комплексы просодем), *означающие-операции*.

Синтактика ЯЗ: примеры синтактик: *позиционный класс морфа, часть речи, грамматический род, модель управления, тип склонения/спряжения, информация о чередованиях в означающем языковых знаков*.

2.2. Противопоставление *элементарные vs. неэлементарные* языковые знаки. Эвристическая процедура, позволяющая установить элементарность/неэлементарность ЯЗ: «квадрат Гринберга». *Уникальные языковые знаки. Связанные языковые знаки*.

Понятие *представимость* одних языковых знаков через другие. *Метаоперация языкового объединения* в языковой модели: *объединение означаемых ЯЗ + объединение означающих ЯЗ* (объединение сегментных означающих, сегментного и супрасегментного означающего, сегментного означающего и применяемой к нему операции: *чередования, повтора или конверсии*) + *объединение синтактик ЯЗ*. Понятие *квази-представимость* языковых знаков: *квазипредставимость по*

означаемому – понятие *мегаморф* (= слово – вешалка); квази-представимость по означаемому – понятие *фразема*; *морфоиды* – компоненты фраземы.

Элементарный языковой знак – языковой знак, не представимый через другие языковые знаки. Классификация элементарных языковых знаков по характеру означающего:

- сегментные – *морф* (означающее – цепочка фонов);
- несегментные, в том числе:
 - супрасегментные – *супрафикс* (означающее – просодема);
 - операционные – *апофония* (означающее – операция чередования);
 - редупликация* (означающее – операция повтора)
 - конверсия* (означающее – операция конверсии).

Нулевой знак – знак, означающее которого не имеет ни сущностного, ни операционного соответствия: <Y, Λ, Σ>.

Пустой знак – знак, означаемое которого не имеет соответствия в семантическом представлении (=в представлении ментальной субстанции языка): <Λ', X, Σ>.

Морфема как множество находящихся в отношении дополнительного распределения морфов (*алломорфы* одной морфемы). Возможные соотношения между означающими алломорфов. Понятия *представляющий морф*, *супплетивные морфы*.

2.3. Словоформа как один из типов языкового знака. Противопоставления *словоформа vs. морф* и *словоформа vs. словосочетание*.

Эвристические критерии выделения словоформ:

- *автономность сильная* – способность образовывать полное высказывание (высказывание, которое может быть выделено двумя абсолютными паузами);
- *автономность слабая*: а) *отделимость* (другим, сильно автономным, языковым знаком); б) *непереместимость* (компонентов словоформы); в) *широкая сочетаемость* словоформы с другими словоформами в предложениях конкретного языка;
- *цельнооформленность*: наличие в словоформе единственного из возможных альтернативных наборов словоизменительных аффиксов (окончаний);
- *фонетическая выделенность*: акцентуационная – наличие одного ударения; сингармоническая; наличие особых чередований на границах словоформы (*внешние сандхи*), отличных от чередований на границах морфов внутри словоформы (*внутренние сандхи*) и под.; понятие *пограничный сигнал*;
- словоформа – «*синтаксический остров*» (утрата некоторыми компонентами словоформы присущей им в независимом ста-

тусе способности присоединять другие словоформы);

- словоформа – «анафорический остров» (невозможность замены части словоформы анафорическим местоимением);
- внутренняя согласованность системы словоформ (наличие парадигматического сходства между словоформами данного языка): фонетическая, морфологическая и синтаксическая согласованность системы словоформ.

Неабсолютный характер перечисленных критериев. Трудные случаи выделения словоформ.

2.4. Клитики – просодически несамостоятельные словоформы. Позиционные типы клитик:

- проклитики – клитики, присоединяющиеся к непосредственно следующим за ними просодически самостоятельным словоформам;
- энклитики – клитики, присоединяющиеся к непосредственно предшествующим им просодически самостоятельным словоформам.

Клитика как обязательный или необязательный вариант просодически самостоятельной словоформы. Типы языковых знаков, имеющие тенденцию к реализации в виде клитик: личные местоимения, артикли, вспомогательные глаголы, предлоги, союзы, частицы. Особенности порядка расположения клитик: особый порядок, жесткий порядок.

Тема 3. Элементарные языковые знаки в составе словоформы

3.1. Морфы в составе словоформы. Противопоставление *корни vs. аффиксы*. Эвристические соображения, лежащие в основе этого противопоставления:

- со стороны означаемого: корни как обозначения более конкретных, автономных понятий;
- со стороны означающего: корни как более длинные морфы;
- со стороны синтактики: корни как потенциально самостоятельные (необходимые и достаточные для образования словоформы) единицы; единицы, ответственные за большинство внешних синтактик словоформы;
- по объему класса: множество корней по объему значительно превосходит множество аффиксов;
- типологическая универсальность класса корней (не существует языков без корней, но существуют языки без аффиксов).

Относительный характер перечисленных критериев (примеры).

3.2. Позиционные типы аффиксов. Классификация аффиксов (по И. А. Мельчуку) с точки зрения их способности а) разрывать (линейно) корень, б) быть разорванными корнем:

- *конфиксы* – аффиксы, которые не разрывают корень и сами не разрываются корнем;
- *циркумфиксы* – аффиксы, которые не разрывают корень, но разрываются корнем;
- *инфиксы* – аффиксы, которые разрывают корень, но сами не разрываются корнем;
- *трансфиксы* – аффиксы, которые разрывают корень и сами разрываются корнем.

Иерархия этих четырех типов аффиксов с точки зрения их распространенности в естественных языках (конфиксы > циркумфиксы > инфиксы > трансфиксы).

Классификация конфиксов:

- *интерфиксы* – конфиксы, за которыми и перед которыми в словоформе обязательно присутствует корень;
- *суффиксы* – конфиксы, которые не являются интерфиксами и перед которыми в словоформе обязательно присутствует корень;
- *префиксы* – конфиксы, которые не являются интерфиксами и за которыми в словоформе обязательно следует корень.

Лингвистические универсалии Дж. Гринберга, связанные с суффиксами и префиксами:

- если в языке есть префиксы, то в нем, как правило, есть и суффиксы;
- если в языке есть и префиксы, и суффиксы, то суффиксов в нем, как правило, больше.

Порядок расположения суффиксов и префиксов в словоформе. «Жесткие» ограничения на место префикса/суффикса в цепочке аналогичных конфиксов. Принятый в лингвистических описаниях способ указания обычного места конфикса в словоформе (... -3 -2 -1 0 +1 +2 +3 ...). Возможная длина цепочек префиксов и суффиксов в естественных языках.

Некоторые общие закономерности, связывающие расположения конфиксов в словоформе и их значения:

- словообразовательные конфиксы располагаются, как правило, ближе к корню, чем словоизменительные;
- семантические словоизменительные конфиксы располагаются, как правило, ближе к корню, чем синтаксические.

Прочие типы аффиксов (циркумфиксы, инфиксы, трансфиксы) – теория, история, типология.

3.3. Композиты (сложные слова) – соединение двух или более корней в пределах одной словоформы. Композиты с прозрачной, с синхронической точки зрения, семантической связью корней *vs.* композиты с непрозрачной для современного носителя языка семантической структурой. Модифицируемый и модифицирующий корни в композитах.

Структурные особенности композитов с прозрачной семантической связью корней:

- части речи компонентов сложения: в языках представлены практически все теоретически мыслимые сочетания;
- форма корней в композитах: корень сохраняет форму, представленную в независимых образованиях, *vs.* корень выступает в измененной форме, чаще всего в усеченной;
- порядок следования корней: допустимы оба теоретически возможных порядка: модифицирующий корень → модифицируемый корень; модифицируемый корень → модифицирующий корень;
- наличие / отсутствие интерфикса.

Семантическая классификация композитов. Особый вклад древнеиндийской грамматической традиции в изучение композитов. Семантические отношения между корнями, входящими в состав композита:

- координативное, или сочинительное, (др.-инд. название – *двандва*);
- актантное, или дополнительное (др.-инд. – *татпуруша*);
- атрибутивное, или определительное, (др.-инд. – *кармадхарайя*).

Особенности синтактики композитов:

- эндоцентрические конструкции: синтактика композита = синтактике модифицируемого компонента);
- экзоцентрические конструкции: синтактика композита не совпадает с синтактикой составляющих (др.-инд. *бахуврихи*).

3.4. Понятие инкорпорация. Работа Э. Сепира «Проблема именной инкорпорации в языках американских индейцев» (1911 г.), положившая начало теоретическому осмыслению этого явления. Инкорпорация как особый способ образования композитов: включение в состав глагольной словоформы одного или нескольких корней (или основ), представляющих подчиненные ей в предложении словоформы таким образом, что полученная в результате словоформа-композит (*инкорпо-*

ративный комплекс) также является глагольной словоформой. *Инкорпорирующие языки* – языки, для которых особенно характерно наличие инкорпоративных комплексов (наряду, как правило, с другими типами композитов): палеоазиатские языки (чукотский, корякский, аляutorский и др.), некоторые языки американских индейцев (в частности языки ирокезской группы), некоторые языки австралийских аборигенов.

Структурные особенности инкорпоративных комплексов. Синтаксические роли инкорпорированных корней (основ): прямое дополнение (наиболее распространенный случай), подлежащее, косвенное дополнение, обстоятельство. Количество модифицирующих корней в инкорпоративном комплексе: один (наиболее распространенный случай); возможность инкорпорации нескольких корней.

Параллельное существование в инкорпорирующих языках инкорпоративных комплексов и близких к ним по смыслу словосочетаний (ср. усл. *вод(а)-пьет* и *пьет воду*). Факторы, определяющие выбор способа оформления одного и того же синтаксического отношения посредством инкорпоративного комплекса или посредством словосочетания: семантические (референциальный статус зависимой словоформы; постоянность описываемой экстралингвистической ситуации), синтаксические (синтаксическая роль зависимой словоформы – подлежащее, дополнение и под.).

3.5. Несколько замечаний о знаках-операциях

Апофония. Происхождение апофоний: некогда контекстно обусловленные чередования, утратившие в ходе исторического развития языка такого рода обусловленность.

Редупликация. Понятие *иконичность* языкового знака: *иконический* знак – языковой знак, структура означающего которого в той или иной степени ассоциируется с его означаемым. Редупликации как знаки, обладающие большой степенью иконичности. Типы значений, чаще всего выражаемые с помощью редупликации: 'множественность', 'повторяемость действия', 'длительность действия', 'интенсивность действия или качества' и под.

Конверсия. Классификация конверсий (по И. А. Мельчуку):

- *категориальная* (изменяющая часть речи);
- *рекция* (изменяющая грамматический род или модель управления);
- *парадигматическая* (изменяющая тип склонения или спряжения);
- *смешанная* (например, рекционная + парадигматическая).

Тема 4. Языковые значения, которые могут выражаться средствами морфологии.

4.1. Лексические vs. грамматические значения. Базовые различия этих двух типов значений:

- сходство инвентарей лексических значений разных языков / своеобразие инвентарей грамматических значений;
- открытость и изменчивость инвентаря лексических значений / закрытость и консервативность инвентаря грамматических значений;
- более непосредственная связь с внеязыковой действительностью лексических значений / большая абстрактность грамматических;
- слабая структурированность лексических значений / четкая структурированность грамматических.

4.2. Семантические (непосредственно отражаемые в представлении смысла высказывания) vs. **синтаксические** (отражаемые в представлении смысла опосредованно, через синтаксические отношения) языковые значения.

4.3. Словоизменительные vs словообразовательные грамматические значения.

Словоизменительная грамматическая категория (СГК) данного языка как содержащий не менее двух элементов ряд взаимоисключающих друг друга в одной словоформе грамматических значений, обладающий свойствами:

- 1) обязательности присутствия одного из этих грамматических значений в означаемых словоформ (из определенного класса);
- 2) регулярности внешнего выражения элементов данного ряда.

Словоизменительное грамматическое значение (=граммема) – элемент определенной словоизменительной грамматической категории.

Словообразовательное грамматическое значение (=дериватема) – грамматическое значение, не являющееся граммемой данного языка.

Наиболее общие различия в свойствах граммем и дериватем:

- большая абстрактность граммем / большая конкретность дериватем;
- широкая сочетаемость граммем / ограниченная сочетаемость дериватем;
- стандартность способов выражения граммем / нестандартность способов выражения дериватем;

- более близкое расположение к корню аффиксов, выражающих дериватемы, / более удаленное – аффиксов, выражающих граммемы;
- существенность граммема / несущественность дериватем для синтаксических правил.
- неспособность аффиксов, выражающих граммемы, в отличие от аффиксов, выражающих дериватемы, изменять часть речи исходной основы.

Кумулятивное выражение граммема – вхождение более одной граммемы в означаемое одного аффикса (например, окончания русских существительных содержат граммема СГК «падеж» и граммема категории «число»).

Аналитическая форма как особый объект морфологии: словосочетание, в котором одна словоформа представляет лексическое значение, а остальные только уточняющие его словоизменительные грамматические значения. Критерии, позволяющие различать аналитические формы и свободные словосочетания.

4.5. Основные типы словоизменительных грамматических категорий.

4.5.1. Семантические СГК. Общие положения. Две точки зрения на высказывание:

- действие, состоящее в передаче информации от одного участника коммуникации (*Говорящего*) к другому ее участнику, или участникам (*Адресату, или Слушающему*): высказывание – *речевой акт* (=языковой факт); *Говорящий, Адресат* – участники речевого акта;
- описание определенного фрагмента внеязыковой действительности: высказывание – *сообщаемый факт* со своими, не обязательно совпадающими с *Говорящим* и *Адресатом*, участниками.

Шифтеры – языковые знаки, значение которых содержит отсылку к участникам, времени, месту определенного речевого акта (например, Я, ЗДЕСЬ, СЕЙЧАС). *Шифтерные* (содержащие в своих граммемах отсылку к параметрам речевого акта) и *нешифтерные* (не содержащие такой отсылки) СГК.

Примеры семантических СГК:

- категории существительных, выражающие *пространственные значения*: *дейксис, локализация, ориентация I, ориентация II*;
- категории глагола, выражающие *временные значения*: *абсолютное время, относительное время*;

- категории, выражающие количественные (в широком смысле) значения: число (существительных), степень сравнения (прилагательных), прогрессив (глагола);
- категории, выражающие качественные значения: вежливость, определенность, модальность, посессивность, принадлежность.

4.5.2. Синтаксические СГК. Общие положения: понятие синтаксическая зависимость; понятие синтаксическая структура предложения.

Типы синтаксической информации, выражаемой синтаксическими категориями: роль синтаксической вершины предложения, роль синтаксического хозяина, роль синтаксически зависимого элемента.

Примеры синтаксических СГК:

- категории, маркирующие роль синтаксической вершины: глагольная категория *финитность*, именная категория *предикативность*;
- категории, маркирующие роль синтаксического хозяина: именная категория *состояние*, понятие *изафет*;
- категории, маркирующие роль синтаксически зависимого элемента: *управляемый падеж существительного*; процедура «выделения падежей» Колмогорова-Успенского-Зализняка; примерный универсальный инвентарь падежей.

4.5.3. Лингвистическая теория маркированности языковых элементов в морфологии.

Разработка теории маркированности для фонологического компонента языковой модели в рамках Пражской лингвистической школы в 30-х годах XX-го века (в частности в работах Н. С. Трубецкого применительно к фонологии и Р. О. Якобсона – для морфологии).

Основные различия в свойствах немаркированных/маркированных элементов.

1. Различия в синхронических свойствах:

- поведение в позиции нейтрализации (представлен, как правило, немаркированный элемент морфологической оппозиции);
- возможность нулевого формального выражения (обычно у немаркированного элемента);
- регулярность выражения (меньшая степень регулярности выражения для немаркированного элемента);
- частотность в текстах (выше для немаркированного элемента);
- дистрибуция (шире для немаркированного элемента);

- воспринимаемость (более легкая для немаркированного элемента);

2. Различия в диахронических свойствах:

- усвоение при изучении языка детьми или иностранцами (более быстрое для немаркированного элемента);
- утрата при различных речевых расстройствах (более поздняя для немаркированного элемента);
- время появления в языке (более раннее для немаркированного элемента);
- время исчезновения (более позднее для немаркированного элемента);
- возможность быть заимствованным в другой язык (большая для немаркированного элемента);
- склонность к пиджинизации и креолизации (большая у немаркированных элементов).

Примеры граммем, наиболее часто выступающих в языках как немаркированные: 'именительный падеж существительного', 'единственное число существительного', 'настоящее время глагола', 'финитная форма глагола', 'положительная степень прилагательного'.

Тема 5. Лексема, парадигма, вокабула

5.1. Лексема. Вспомогательные понятия. *Лекса* – словоформа или аналитическая форма. *Лексемное значение* – компонент означаемого лексы, включающий лексические значения и словообразовательные грамматические значения (дериватемы) и не включающий словоизменительных грамматических значений (граммем).

Лексема – множество лекс, содержащее все лексы с одинаковым лексемным значением.

5.2. *Парадигма (парадигматическая схема)* – полный набор всех возможных комбинаций граммем, характеризующих лексемы определенного класса (например, парадигма русских существительных).

Конкретная парадигма – соответствие между парадигматической схемой определенного класса лексем и множеством лекс конкретной лексемы из данного класса (парадигма существительного «город»).

Стандартная парадигма – конкретная парадигма, в которой каждой комбинации граммем, предусмотренной для лексем данного класса, соответствует ровно одна лекса.

Нестандартные парадигмы – конкретные парадигмы, в которых: либо 1) каким-то из возможных комбинаций граммем

соответствует более одной лексы – *вариативные парадигмы*; либо 2) каким-то из возможных комбинаций граммем не соответствует ни одна лекса – *дефектные парадигмы*, либо 3) одновременно имеют место обе описанные ситуации – *нестандартность смешанного типа*. Типы дефектности парадигмы: *семантически мотивированная* (например, частое отсутствие форм мн. числа у «неисчисляемых» существительных, ср. русск. «молоко»); *случайная* (например, отсутствие формы Род. мн. у русского существительного «мечта» – *мечт).

5.3. Вокабула (словарная статья) – множество близких по значению лексем, лексы которых имеют тождественные означающие.

Полисемия как семантическое отношение между лексемами, входящими в одну вокабулу.

Омонимия – отсутствие, по крайней мере с синхронической точки зрения, семантического сходства между лексемами при наличии тождества означающих, образующих их лекс. Омонимичные лексемы как представители разных вокабул. Расплывчатость границы между полисемией и омонимией.

5.4. Основа, окончание.

Основа словоформы – языковой знак, компонент словоформы, который не включает ее словоизменительных аффиксов.

Окончание словоформы – аффикс (совокупность аффиксов), который выражает словоизменительные грамматические значения (граммемы) данной словоформы.

Стандартное соотношение между основами словоформ одной лексемы: тождество их основ. Нестандартное соотношение основ в пределах лексемы: в некоторых случаях в одну лексему могут быть включены словоформы с разными основами (русск. человек ~ люди) – *супплетивизм основ*. Источники возникновения супплетивизма в языке: фонетические изменения; лексическая конвергенция.

5.5. Морфологическая структура основы. Типы основ:

- *простая основа* – основа, единственным компонентом которой является корневой морф;
- *производная основа* – основа, образованная от некоторой другой (*исходной, производящей*) основы путем присоединения того или иного языкового знака.

Словообразование (=основообразование) – описание возможных способов образования производных основ (= возможных словообразовательных моделей).

Тема 6. Классификация лексем, базирующаяся на синтактиках их основ

6.1. Противопоставление *внутренние vs внешние* синтактики основы.

Внутренние синтактики основы – синтактики, обеспечивающие построение правильных словоформ с данной основой.

Внешние синтактики основы – синтактики обеспечивающие построение правильных словосочетаний со словоформами, содержащими данную основу.

6.2. Примеры классов лексем, учитывающих внутренние синтактики основ.

Тип словоизменения – множество лексем, характеризующихся одинаковым набором окончаний.

Акцентный тип – множество лексем, словоформы которых (соответствующие друг другу по комбинации граммем) имеют одинаковую просодическую характеристику. *Схема ударения* – способ распределения ударения между основой и окончанием в словоформах одной лексемы (частный случай акцентного типа).

6.3. Примеры классов лексем, учитывающих внешние синтактики основ.

6.3.1. Противопоставление *частей речи* как универсальное, присущее в том или ином варианте всем языкам, противопоставление лексем. Разные подходы к выделению частей речи в лингвистической теории:

- *семантический* – по сходству называемых внеязыковых объектов (предмет, действие, качество/свойство, число и под.);
- *синтаксический* – по сходству возможных синтаксических ролей (подлежащее, дополнение, сказуемое, определение и под.);
- *морфологический* – по сходству парадигматических схем; использование для обозначения такого рода классов термина *грамматический разряд* (вместо термина *часть речи*).

Основные части речи из универсального инвентаря: существительное (N), глагол (V), прилагательное (Adj), местоимение (Pron), числительное (Num), наречие (Adv), предлог (Prep), союз (Conj), артикль (Art).

6.3.2. Понятие *согласовательный класс* – множество лексем существительных, словоформы которых, имеющие одинаковый состав граммем, требуют одного и того же набора синтаксически связанных с ними словоформ прилагательных, местоимений, глаголов.

Грамматический род и именной класс как основные разновидности согласовательного класса.

Основные критерии для признания согласовательного класса в конкретном языке *грамматическим родом* существительного:

- небольшое число классов (2-4);
- связь с биологическим полом;
- немотивированное распределение по классам у объектов, для которых признак пола нерелевантен;
- отсутствие показателя класса в самом существительном;
- нетривиальное соотношение классов с типами склонения;
- неиспользование изменения класса для выражения словоизменительных и словообразовательных значений.

Основные критерии для признания согласовательного класса *именным классом*:

- большое число классов (5-20, иногда более);
- отсутствие прозрачной связи с биологическим полом;
- наличие очевидного семантического распределения по классам;
- наличие, как правило, специального некумулятивного показателя класса в самом существительном;
- отсутствие корреляции со склонением существительных;
- регулярное использование изменения класса (конверсии) для выражения словоизменительного значения (числа) и словообразовательных значений;
- наличие также некумулятивных классовых показателей в зависимых словоформах.

Ареалы, языковые семьи, для которых характерно наличие родов или именных классов: противопоставление по родам характерно в частности для многих индоевропейских языков (русский, немецкий, французский, испанский и др.), для семитских языков (арабский, иврит); именные классы представлены, например, в языках семьи банту (Африка), в нахско-дагестанских языках (Кавказ), в папуасских языках (Новая Гвинея).

Часть II. Формы контроля знаний по курсу «Морфология»

Образцы задач для практических занятий и контрольных работ

Задача 1.

(Составлена на основе задачи А. Н. Журина: *Журинский А.М. Лингвистика в задачах. М.: Индрик, 1995. С. 48*)

Условие

Даны словоформы туркменского языка (в упрощенной транскрипции МФА) и их переводы, расположенные в другом порядке:

baglarıjızda, baglarımız, bagıjnda, bagım, bagımnda, bagda, bagımız, baglarıj, bagıjız, baglar

сады, в саду, наш сад, наши сады, в твоём саду, твои сады, в ваших садах, ваш сад, мой сад, в моём саду

Задание

1. Установите, какой перевод соответствует каждой из словоформ:

2. Предложите элементарные знаки для представления туркменских словоформ; опишите их означаемые, означающие и синтактики;

3. Опишите структуру туркменской словоформы.

Задача 2.

(Составлена на основе задачи Е. Н. Савиной: Алпатов В.М., Вентцель Р.Д. и др. Лингвистические задачи. М.: Просвещение, 1983. С. 19)

Условие

Даны венгерские словоформы (в транскрипции МФА) с переводами на русский язык.

ha:boru:	война	køru:t	кольцевая дорога
ha:boru:ro:l	о войне	kørgø:l	о кольце
byn	вина	ba:tor	смелый
byntelen	безвинный	ba:torʃa:gro:l	о смелости
bynro:l	о вине	ba:tortalan	несмелый
u:t	дорога	føld	земля, поле
u:tro:l	о дороге	føldtelen	безземельный
kesery	горький	burgona	картофель
keseryʃe:g	горечь	burgonaføld	картофельное поле
keseryʃo:	горькая соль	ʃøte:t	темный
køŕ	кольцо	ʃøte:tʃe:g	темнота

Задание

1. Предложите элементарные знаки для представления венгерских словоформ; опишите их означаемые, означающие и синтактики.

2. Укажите, какие из приведенных словоформ можно рассматривать как (в определенном отношении) «неправильные», объясните, чем они отличаются от других словоформ, представленных в задаче.

3. Какими фонетическими средствами маркируются (как правило) границы венгерской словоформы.

4. Какие из перечисленных словоформ являются словоформами-композиатами? Охарактеризуйте исчерпывающим образом структурный тип каждой из этих словоформ.

Задача 3.

(Составлена на основе задачи А. Н. Журинского: *Журинский А.М. Лингвистика в задачах*. М.: Индрик, 1995. С. 53)

Условие

Даны фразы на языке курия (банту, Кения) в упрощенной транскрипции МФА с переводами на русский язык:

- | | | |
|-----|---------------------------------|----------------------------|
| 1. | agatotʃo kagokegetʃa egete | - Зайчик режет палку |
| 2. | agatʃutʃu gakogotegerera | - Птенчик слушает тебя |
| 3. | akaboe gakugutiga | - Шакальчик оставляет тебя |
| 4. | ekeboe gikugitiga egekone | - Шакал оставляет банан |
| 5. | igitʃutʃu kigukibukia egekondo | - Птенец будит обезьяну |
| 6. | egekondo kegokegonga ekegogo | - Обезьяна хвалит гитару |
| 7. | egetotʃo gikuguturia | - Заяц спасает тебя |
| 8. | agakondo gakogetegerera ekegogo | ? |
| 9. | ekeboe kegokogonga | ? |
| 10. | ? | - Шакальчик спасает гитару |
| 11. | ? | - Палочка будит обезьяну |
| 12. | ? | - Птенец оставляет тебя |

Задание

1. Предложите элементарные знаки для описания словоформ языка курия; опишите их означаемые, означающие и синтактики.

2. Какие части речи можно выделить в курия, ориентируясь на материал задачи?

3. Что вы можете сказать о словоизменительных грамматических категориях и синтактиках в курия?

4. Опишите структуру словоформ разных частей речи.

5. Какими фонетическими средствами маркируются границы словоформы в языке курия?

6. Заполните пропуски.

Вопросы для итогового теста

1. Перечислите *односторонние* объекты языковой модели.
2. Перечислите не менее 6 типов морфологических синтактик.
3. Опишите формализм для представления языкового знака, используемый в модели «Смысл-Текст».
4. Перечислите 5 основных классов элементарных языковых знаков (различаемых по характеру означающего). Приведите примеры.

5. Что понимается под *представимостью* языкового знака через другие языковые знаки. Поясните на примерах.
6. Что понимается под *квазипредставимостью* языкового знака? Опишите основные типы квазипредставимости знака. Приведите примеры.
7. Перечислите основные критерии выделения словоформы.
8. Дайте определение понятия *клитика*. Назовите два основных позиционных типа клитик. Приведите примеры.
9. Перечислите не менее 5 типов языковых знаков, имеющих тенденцию реализовываться в виде клитик. Приведите примеры.
10. Приведите классификацию аффиксов, учитывающую их способность разрывать корень или быть разорванным корнем. Укажите относительную частотность выделенных классов. Приведите примеры.
11. Назовите два основных типа конфиксов. Сформулируйте общие закономерности расположения конфиксов по отношению к корню.
12. Перечислите 5 базовых различий между *корнями* и *аффиксами*.
13. Дайте определение *компози́та*. Объясните, что такое композиты в *сильном* и *слабом* смысле. В каких формах может выступать корень в составе композита? Укажите возможные порядки следования корней в композитах. Приведите примеры.
14. Перечислите 3 основных типа семантических отношений между корнями в композитах. Укажите традиционные (санскритские) названия этих типов. Приведите примеры.
15. Что характеризует *инкорпорацию* как особый тип композитов. Какую синтаксическую роль чаще всего выполняет модифицирующий корень по отношению к модифицируемому в инкорпоративном комплексе.
16. Что понимается под *иконичностью* редупликации. Охарактеризуйте семантический инвариант редупликации. Назовите (не менее 6) типов значений, для выражения которых в языках часто используется редупликация.
17. Дайте определение *конверсии* как особого типа элементарного языкового знака. Перечислите основные типы конверсий. Приведите примеры.
18. Назовите 4 базовых различия между *лексическими* и *грамматическими* значениями.
19. Дайте определение понятия *словоизменительная грамматическая категория*. Приведите примеры словоизменительных грамматических категорий из известных вам языков.
20. Дайте определение понятия *дериватема*. Приведите примеры дериватем.
21. Перечислите 4 базовых различия между *граммемами* и *дериватемами*.
22. Дайте определение *нулевого знака*. Перечислите требования, которыми предлагается обусловить введение в описание нулевых знаков.
23. Дайте определение *пустого знака*. Приведите примеры.

24. Объясните понятие *речевой акт (языковой факт)*, назовите основных его участников.
25. Объясните, что такое *сообщаемый факт* и его участники.
26. Дайте определение грамматической категории *Дейксис*. Приведите возможные в конкретных языках наборы ее граммем.
27. Дайте определение грамматической категории *Локализация*, укажите универсальный инвентарь ее граммем.
28. Дайте определение грамматической категории *Ориентация I*, приведите универсальный инвентарь ее граммем.
29. Дайте определение грамматической категории *Абсолютное время*. Приведите универсальный инвентарь ее граммем. Укажите возможные в конкретных языках наборы граммем этой категории.
30. Дайте определение грамматической категории *Относительное время*, приведите универсальный инвентарь ее граммем.
31. Назовите не менее 3 *синтаксических словоизменительных грамматических категорий*. Укажите для каждой из них возможные в конкретных языках наборы граммем.
32. Перечислите не менее 5 *семантических словоизменительных грамматических категорий*. Укажите для каждой из них возможные в конкретных языках наборы граммем.
33. Дайте определение грамматической категории *Финитность*. Перечислите основные *нефинитные* формы.
34. Дайте определение грамматической категории *Предикативность*, перечислите ее граммемы.
35. Перечислите не менее 5 различий в синхронических свойствах между *маркированными* и *немаркированными* языковыми элементами.
36. Перечислите не менее 4 различий в диахронических свойствах *маркированных* и *немаркированных* языковых элементов.
37. Объясните соотношение понятий *лекса*, *словоформа*, *аналитическая форма*.
38. Дайте определение понятий *лексема* и *вокабула*. Приведите примеры.
39. Дайте определение понятия *парадигма*. Приведите примеры.
40. Дайте определение понятия *дефектная парадигма*. Перечислите основные типы дефектных парадигм. Приведите примеры.
41. Перечислите 3 основных подхода к выделению *частей речи*. Укажите (кратко), в чем суть каждого из них.
42. Дайте определение *согласовательного класса*. Перечислите признаки, которые дают основания квалифицировать согласовательный класс как *грамматический род*. Укажите, в каких ареалах, в каких языковых семьях распространены грамматические роды.
43. Дайте определение *согласовательного класса*. Перечислите признаки, которые дают основания квалифицировать согласовательный класс как *именной класс*. Назовите, в каких ареалах, в каких языковых семьях распространены именные классы.

Примеры итоговых тестов

Тест 1

1. Перечислите *односторонние* объекты языковой модели.
2. Приведите классификацию аффиксов, учитывающую их способность разрывать корень или быть разорванным корнем. Укажите относительную частотность выделенных классов. Приведите примеры.
3. Определите понятие *дефектная парадигма*. Перечислите основные типы дефектных парадигм. Приведите примеры.
4. Объясните понятие *речевой акт (языковой факт)*, назовите основных его участников.
5. Перечислите признаки, которые дают основания квалифицировать согласовательный класс как *грамматический род*. Укажите, в каких ареалах, в каких языковых семьях распространены грамматические роды.

Тест 2

1. Опишите формализм для представления языкового знака, используемый в модели «Смысл-Текст».
2. Перечислите основные критерии выделения словоформы.
3. Дайте определение *компози́та*. Объясните, что такое композиты в *сильном* и *слабом* смысле. В каких формах может выступать корень в составе композита? Укажите возможные порядки следования корней в композитах. Приведите примеры.
4. Дайте определение *пустого знака* и *нулевого знака*. Приведите примеры.
5. Дайте определение грамматической категории *Ориентация II*, перечислите универсальный инвентарь ее граммем.

Примечания

- ¹ См., к примеру, варианты программы этого курса, в разные годы вводившиеся на филологическом факультете МГУ им. М.В.Ломоносова: Плунгян В.А. Морфология // Программы и учебный план Отделения теоретической и прикладной лингвистики. М., 1996. С. 16-22; Он же. Общая морфология // Программы и учебный план Отделения теоретической и прикладной лингвистики. Изд. 2-е, испр. и доп. М.: МГУ, 2002. С. 14-24

Основная литература

- Зализняк А. А. Русское именное словоизменение: с приложением избранных работ по современному русскому языку и общему языкознанию. Ч. 1. М.: Языки рус. культуры, 2002. С. 19-100.
- Мельчук И. А. Курс общей морфологии. Т. I. М.: Языки рус. культуры, 1997. Гл. II-IV. С. 107-239 (к теме 2); гл. V. С. 240-319 (к теме 4); гл. VI, С. 320-349 (к теме 5).
- Мельчук И. А. Курс общей морфологии. Т. II. М.: Языки рус. культуры, 1998. Гл. I-III, С. 7-371 (к теме 4).
- Мельчук И. А. Курс общей морфологии. Т. III. М.: Языки рус. культуры, 2000. Ч. IV. Гл. I-II, С. 159-296 (к теме 6).
- Мельчук И. А. Курс общей морфологии. Т. IV. М.: Языки рус. культуры, 2001. Гл. I. С. 9-36 (к теме 2); гл. II-IV, С. 38-367 (к теме 3).

Дополнительная литература

- Задачи по лингвистике / Алексеев М. Е., Беликов В. И., Евграфова С. М., Журинский А. Н., Муравенко Е. В. Ч. 1. М.: РГГУ, 1991.
- Гринберг Дж. Некоторые грамматические универсалии, преимущественно касающиеся порядка значимых элементов // Новое в лингвистике. Вып. V, М.: Прогресс, 1970. С. 114-162.
- 200 задач по языковедению и математике / Алпатов В. М., Вентцель Р. Д. и др. М.: МГУ, 1972.
- Журинский А. Н. Лингвистика в задачах. М.: Индрик, 1995.
- Зализняк А. А. О понимании термина «падеж» в лингвистических описаниях. I // Зализняк А. А. Русское именное словоизменение: с приложением избранных работ по современному русскому языку и общему языкознанию. М.: Языки рус. культуры, 2002. С. 613-647.
- Лингвистические задачи / Алпатов В. М., Вентцель Р. Д. и др. М.: Просвещение, 1983.
- Лингвистический энциклопедический словарь. М.: Сов. энциклопедия, 1990 (и последующие переиздания).
- Плунгян В. А. Общая морфология: введение в проблематику. М.: Эдиториал УРСС, 2000.
- Реформатский А. А. Введение в языковедение. М.: Аспект Пресс, 1996.
- Сепир Э. Проблема именной инкорпорации в языках американских индейцев // Сепир Э. Избранные труды по языкознанию и культурологии. М.: Прогресс-Универс, 1993. С. 355-385.
- Смирницкий А. И. Морфология английского языка. М.: Изд-во лит-ры на иностранных языках, 1959.
- Якобсон Р. О. О структуре русского глагола // Якобсон Р. Избранные работы. М.: Прогресс, 1985. С. 210-221.
- Якобсон Р. О. Нулевой знак // Якобсон Р. Избранные работы. М.: Прогресс, 1985. С. 222-230.
- Matthews P. H. Morphology. Cambridge: Cambridge University Press, 1991.

Nida E. A. Morphology: The descriptive analysis of words. Ann Arbor: the University of Michigan Press. 1946 (и последующие переиздания).
Spencer A. Morphological Theory. An Introduction to Word Structure in Generative Grammar. London: Basil Blackwell, 1991.

Моделирование языка при автоматическом переводе (в рамках сущностного подхода к лингвистическому описанию)

Программа спецкурса для студентов-лингвистов

Предисловие

Данный курс может рассматриваться как продолжение и углубление курса «Проблемы и методы автоматического перевода», программа которого уже была опубликована (см. Московский лингвистический журнал, 2003, Том 7, №1. С. 125-154). Если в том курсе ставится задача возможно более широко охватить различные лингвистические подходы и формализмы, применяемые при создании систем автоматического перевода, рассматривая эти подходы и формализмы в их исторической перспективе и в сопоставлении друг с другом, то здесь более подробно и содержательно изучаются методы решения основных проблем автоматического перевода в рамках одного из разработанных в этих целях подходов – так наз. **сущностного подхода** к лингвистическому описанию. Данный подход целесообразно выделить в качестве предмета самостоятельного изучения, поскольку он составляет лингвофилософскую основу таких важных понятий и инструментов современной лингвистики, как система зависимостей, аппарат валентностей, компонентный принцип описания семантики и т.п. В то же время он позволяет непосредственно интерпретировать и эффективно использовать в лингвистическом описании отношения наследования, механизмы умолчаний и другие средства, выработанные в рамках объектно-ориентированного подхода в искусственном интеллекте и программировании и может рассматриваться как его принципиальный аналог применительно к лингвистике.

Подчеркнем, что содержание данного курса и курса «Проблемы и методы автоматического перевода» практически не пересекается, так что в принципе каждый из них может читаться независимо от другого. От слушателей данного курса ожидается только, что они уже знакомы с основными лингвистическими понятиями (в частности, владеют навыками построения представлений синтаксической структуры предложения

в терминах составляющих и зависимостей), а также знакомы с общим строением классической теории порождающих грамматик Н.Хомского и модели «Смысл \leftrightarrow Текст» И.А.Мельчука и А.К.Жолковского.

Курс рассчитан на 66 часов учебного времени и включает четыре основных части:

1) Общая модель языковой действительности и разграничение реляционного и сущностного подходов к описанию языка в целях АП;

2) Моделирование морфологических и морфолого-синтаксических аспектов языка при сущностном подходе к его описанию в целях АП;

3) Моделирование синтаксических и семантико-синтаксических аспектов языка при сущностном подходе его описанию в целях АП;

4) Моделирование семантики при сущностном подходе к описанию языка в целях АП.

Соответственно, он может читаться либо как единый годичный курс с зачетом после изучения первых двух его частей и с экзаменом в конце года, либо как два взаимосвязанных семестровых курса:

(1) Сущностный подход к языку и моделирование морфологических и морфолого-синтаксических аспектов языковой системы при АП;

(2) Моделирование семантико-синтаксических и семантических аспектов языковой системы при АП в рамках сущностного подхода к языку, каждый из которых завершается экзаменом. В таком варианте он читался, в частности, для студентов III-го курса ФТиПЛ РГГУ в 1994-1995 и 1997-1998 учебных годах).

В результате усвоения данного курса студенты должны, с одной стороны, получить понятие о лингвофилософских основаниях выделения сущностного подхода и о месте, которое он занимает среди различных направлений традиционного и современного языкознания с учетом общих представлений о языке и языковой деятельности, и, с другой стороны, усвоить конкретный понятийный аппарат данного подхода, изучить содержательные лингвистические критерии, выработанные в его рамках для определения общих и частных лингвистических оппозиций и категорий, в терминах которых может задаваться при этом подходе конкретно-лингвистическая информация различных уровней, и приобрести навыки организации и записи такой информации с помощью ряда конкретных формализмов, разработанных для этих целей в отечественных исследованиях по автоматическому переводу.

Соответственно, лекционные занятия в составе данного курса должны систематически сопровождаться семинарскими занятиями, на которых студенты тренируются в практическом применении сведений, полученных ими из лекций. Поэтому курс должен читаться в относительно немногочисленных группах – желательно не более 10 слушателей.

При оценке знаний студентов по данному курсу также необходимо учитывать как его теоретические, так и практические аспекты, включая в каждый экзаменационный билет один вопрос по теории рассмотренных лингвистических моделей и один вопрос по практическому описанию тех или иных языковых явлений в терминах изученных формализмов.

Программа

Часть 1. Общая модель языковой действительности и разграничение реляционного и сущностного подходов к описанию языка в целях АП

Тема 1. Стратификационные модели естественного языка в теории автоматического перевода

Стратификационная лингвистическая теория С.М. Лэма. Четыре уровня описания и три типа единиц, выделяемых на каждом уровне. Внутри- и межуровневые связи единиц разных типов. Порядок обработки текстов входного и выходного языков при моделировании перевода на основе данной модели.

Модель «Смысл \Leftrightarrow Текст». Понятие уровня представления текста и его отличие от аналогичного понятия стратификационной лингвистики С.М. Лэма. Основные уровни модели. Принципы дифференциации основных уровней на поверхностный и глубокий подуровни. Особенности реализации этих принципов применительно к морфологическому, синтаксическому и семантическому уровням.

Глубинно-синтаксический уровень представления текста в модели «Смысл \Leftrightarrow Текст». Особенности его лексики и синтаксиса. Соотношение глубоких и базовых синтаксических структур. Уровень базовых синтаксических структур как уровень языка-посредника при автоматическом переводе; достоинства и недостатки такого определения языка-посредника.

Уровень семантического представления текста в модели «Смысл \Leftrightarrow Текст». Его отличия от уровня семемных структур в стратификационной модели С.М. Лэма.

Компоненты описания языка в модели «Смысл \Leftrightarrow Текст». Словарь как полноправный и даже центральный компонент такого описания. Два возможных подхода к реализации этого принципа. Связь тезиса о центральной роли словаря в лингвистическом описании с рядом современных направлений в программировании и в исследованиях по искусственному интеллекту.

Лингвистические уровни, выделяемые в хомскианской теории языка. Их отличие от уровней стратификационных моделей.

Определение характера отношений между языковыми уровнями в работах Пражской лингвистической школы.

Тема 2. Основные функции языкового знака и общая модель лингвистической компетенции.

Принципиальное противопоставление языковой компетенции и языковой деятельности. Стабильность и вариативность той и другой. Проблема независимых оснований для модели языковой компетенции.

Определение модели языковой компетенции как многомерного лингвистического пространства, оси которого формируются логико-семиотическими функциями языкового знака. Основные из этих функций: обозначение, членение, обобщение. Их принципиальная несводимость друг к другу. Распределение лингвистических уровней по трем осям пространства языковой компетенции, соответствующим указанным функциям языкового знака.

Уровни оси обозначения. Соссюрианское понятие значимости и план собственно знаковой структуры языка. Соотношение плана собственно знаковой структуры с планами выражения и содержания. Интерпретация с точки зрения этого соотношения свойства асимметричного дуализма языкового знака и связанных с ним противопоставлений (вокабула vs. лексема, омонимия vs. смысловая неопределенность, синонимия vs. вариантность и др.). Степень нелинейности языковых структур в разных планах оси обозначения.

Основные уровни (ярусы) оси членения, выделяемые в рамках каждого из трех планов оси обозначения. Соотношение между собой ярусов, относящихся к разным языковым планам. Ярусы, выделяемые с учетом межплановых соотношений включаемых в них единиц, Уровни классических стратификационных моделей, соотносящиеся с осью членения.

Уровни оси обобщения в их отношении к планам оси обозначения. Теоретико-множественное объединение и пересече-

чение как два основных источника лингвистического обобщения. Глубинные и поверхностные подуровни модели «Смысл \Leftrightarrow Текст», с точки зрения оси обобщения.

Проблемы метрики языкового пространства. «Непараллельность» его основных осей и «несимметричность» переходов по каждой из них.

Тема 3. Моделирование языковой деятельности человека на базе трехмерного пространства языковой компетенции.

Языковая деятельность (языковое функционирование) как движение в пространстве языковой компетенции. Принципиальная множественность видов и моделей языкового функционирования.

Понятие языкового «центра». Лингвистическая избыточность, обусловленная частичным дублированием информации, относящейся к «центру» языка, в разных планах оси обозначения. Возможности варьирования видов и моделей языкового функционирования, связанные с этой избыточностью.

Три базовых модели анализа/синтеза, определяемые выбором основного для них плана оси обозначения. Специфика синтеза текста на основе плана выражения. Синтез на основе плана содержания и взаимодействие языка и мышления. Достоинства и недостатки анализа на основе плана содержания; опыт применения таких моделей анализа в прикладных системах обработки текстов. План собственно знаковой структуры как основа и анализа и синтеза при функционировании языка в режиме «текст-текст».

Варьирование направления движения по осям обобщения и членения как еще один источник множественности видов и моделей языкового функционирования. Обобщение/конкретизация и агрегация/сегментация при создании формализованных лингвистических описаний (моделирующем процессы первичного овладения языком), с одной стороны, и при обработке текстов на основе уже созданных лингвистических описаний (моделирующей понимание/порождение текстов на уже усвоенном языке), с другой.

«Анализ через синтез» и «синтез через анализ», с точки зрения осей обозначения, членения и обобщения.

Моделирование языковой компетенции в ориентации на конкретную задачу как «сужение» универсального лингвистического пространства применительно к этой задаче.

Тема 4. Общая формальная модель переводческой деятельности человека и проблемы автоматического перевода.

Перевод как языковая деятельность типа «текст–текст». Три вида собственно переводческих операций: поиск переводных эквивалентов для элементарных единиц перевода, формирование переводных эквивалентов для неэлементарных единиц перевода, межъязыковые преобразования.

Уровень осуществления собственно переводческих операций, с точки зрения оси обозначения. Виды языкового функционирования при переводе, связанные с движением относительно этой оси, и их моделирование в системах автоматического перевода. Возможности варьирования при переводе требуемой «глубины» анализа переводимого текста.

Обобщение / конкретизация в процессе перевода. Неполнота смысловых соответствий между оригиналом и переводом, ее возможные причины и основные пути ее преодоления. Три вида перефразирования при переводе: до-переводческое, после-переводческое и межъязыковое. Синтез и анализ при построении текста перевода. Редактирование текста перевода как необходимый компонент адекватной модели переводческой деятельности.

Уровни осуществления собственно переводческих операций, с точки зрения оси членения (проблема единиц перевода). Минимальные единицы автоматического перевода. Переводческая идиоматика и методы ее описания. Максимальные единицы автоматического перевода. Возможности и пределы их «минимизации» с учетом характера маркировки в языке структурных связей соответствующего уровня (яруса).

Основные типы знаний, необходимые в формализованном описании языка, моделирующем лингвистическую компетенцию человека, для поддержки автоматического перевода как модели его переводческой деятельности.

Тема 5. Реляционный и сущностный подход к организации лингвистических знаний при их формализованном описании.

Оппозиция «часть–целое» в языке и противопоставление реляционных и сущностных подходов к организации лингвистического описания. Соотношение обобщения и членения языковых единиц при описании синтагматических аспектов различных языковых уровней в рамках подходов того и другого типа. Преимущества и недостатки реляционного и сущностного способов лингвистического обобщения.

Понимание синтагматических отношений как свойств связываемых ими лингвистических сущностей. Роль эксплицитных и имплицитных повторов в организации синтагматической системы языка. Отображение языковой синтагматики в терминах валентностей лингвистических сущностей как основной механизм ее моделирования при сущностном подходе. Статус понятий валентности и валентностной модели управления применительно к различным планам оси обозначения. Валентности собственно знакового плана языка на различных уровнях (ярусах) членения.

Словарь и грамматика в реляционных и сущностных описаниях. Возможности и проблемы их согласования друг с другом и оптимизации их соотношения при расширении и корректировке лингвистического описания в целом.

Механизмы наследования и их роль в организации лингвистических описаний сущностного типа. Основные виды наследования, действующие на оси обобщения и на оси членения. Языковой «центр» и проблемы наследования «по оси обозначения».

Относительная трудоемкость реляционных и сущностных описаний в зависимости от их ориентации на максимально полный учет информации о каждой из охватываемых описанием словарных единиц (при возможной неполноте их инвентаря) или на максимально полное перечисление самих имеющих в языке словарных единиц (при возможной неполноте информации об отдельных единицах).

Общие принципы обработки текстов на основе формализованных лингвистических описаний, организованных в рамках сущностного подхода. Сущностный подход как естественная основа для интерпретации задач обработки текста как «задач удовлетворения ограничений».

Часть 2. Средства и методы моделирования морфологических и морфолого-синтаксических аспектов языковой системы при сущностном подходе к ее описанию

Тема 6. Разграничение морфологии и синтаксиса при их моделировании в целях автоматического перевода.

Ярусы морфологии и синтаксиса в рамках общей модели языковой компетенции и языковой деятельности человека. «Содержательные» и «формальные» критерии их разграничения. Место аналитического формообразования, с точки зрения этих критериев.

Проблема разграничения уровней оси членения (ярусов) в собственно знаковом плане естественного языка. Минимальные, максимальные и промежуточные единицы лингвистического яруса. Принцип гомогенности каждого лингвистического яруса относительно допустимых в его пределах лингвистических сущностей, отношений и процессов.

Рекурсивный критерий установления принадлежности языковых единиц к определенному ярусу лингвистического описания на основании принципа его гомогенности. Понятие безусловных сущностей яруса и специфических для него отношений и процессов. Общая схема установления принадлежности конкретной лингвистической сущности к сфере распространения того или иного яруса. Принципиальная возможность «пересечения» сфер распространения смежных ярусов при определении состава входящих в них лингвистических сущностей на основе данного критерия.

Конкретизация общей схемы рекурсивного критерия разграничения лингвистических ярусов применительно к ярусам морфологии и синтаксиса. Слово как единица, пограничная между данными двумя ярусами. Постулат о принципиальной конечности безусловно морфологических сущностей и потенциальной бесконечности синтаксических. Возможности его интерпретации в синтагматическом и в парадигматическом смыслах. Специфически морфологические и специфически синтаксические отношения и процессы, с точки зрения этого постулата.

Проблема разграничения слова и фразеологического сочетания слов. Некоторые возможности моделирования фразеологических словосочетаний, не допускающих внутренних перестановок, одновременно как морфологических и синтаксических сущностей.

Тема 7. Моделирование морфологических знаний и процедур при сущностном подходе к лингвистическому описанию.

Возможности применения понятия валентности при моделировании морфологических и морфонологических (морфографических) аспектов языковой системы в рамках сущностного подхода. Проблема активных и пассивных валентностей этих ярусов.

Морфонологические (морфографические) чередования и их описание в валентностных терминах. Словарь и грамматика графико-морфологических лингвистических сущностей.

Интерпретация общих словоизменительных категорий в терминах морфологических валентностей лексических основ.

Морфологические валентности аффиксов. Стандартная и нестандартная сочетаемость морфологических сущностей в валентностном описании морфологии.

Проблема лингвистической классификации словоизменительных категорий в свете требований компактности их представления в машинной форме. Соотношение этих требований с требованиями лингвистической адекватности описания.

Синтагматическая информация, релевантная для моделирования словоизменения. Организация этой информации в словаре и грамматике формализованного лингвистического описания с использованием механизмов наследования. Суммирующее наследование и наследование по умолчанию как основные из используемых здесь механизмов.

Морфологический анализ при сущностном подходе к описанию морфологии. Основные типы процедур такого анализа: сегментация анализируемой цепочки единиц выражения на ее минимальные подцепочки, допускающие собственно знаковую интерпретацию; переход к ее представлению на уровне морфологического яруса собственно знакового плана языка; верификация формируемых вариантов этого представления с учетом сочетаемости включаемых в них сущностей в пределах данного яруса. Неоднозначность сегментации и интерпретации графико-морфологических объектов и возможности ее устранения на морфологическом уровне.

Морфологический синтез при сущностном подходе к описанию морфологии. Словарь А.А.Зализняка как основа такого описания для русского языка. Лексико-морфологическая избыточность русского языка и определяемые ею отношения наследования морфологической информации от графико-морфологических сущностей, выделяемых в составе русских лексем, к самим лексемам. Графико-морфологическая грамматика русского языка, опирающаяся на инвентари конечных и начальных компонентов русских лексем. Использование этой грамматики для обработки лексем, отсутствующих в словаре, и для оптимизации общего объема словаря.

Тема 8. Моделирование морфолого-синтаксических аспектов языка на различных этапах автоматического перевода при сущностном подходе к лингвистическому описанию.

Морфолого-синтаксический ярус лингвистического описания как уровень, охватывающий область интерференции морфологии и синтаксиса. Основные виды морфолого-синтаксических сущностей.

Распространение аппарата морфологических валентностей на описание морфолого-синтаксической синтагматики как средство повышения устойчивости и надежности морфологических моделей. Устранение неоднозначности сегментации и интерпретации морфологических и графико-морфологических сущностей средствами данного аппарата. Механизм вытесняющего наследования и его использование при формировании морфологических представлений для морфолого-синтаксических аналогов словоформ.

Языки формальных лингвистических операторов/предикатов для описания морфолого-синтаксических явлений в процедурных терминах. Организация процедурных знаний, представленных в терминах таких языков, при сущностном подходе к лингвистическому описанию. Использование процедурных знаний морфолого-синтаксического уровня в целях анализа. Основные типы операторов (предикатов), служащих для снятия неоднозначности анализируемых единиц по их по линейному контексту. Примеры процедур данного типа.

Возможности и ограничения моделирования на морфолого-синтаксическом уровне процедур межъязыкового перехода. Линейная структура текста как объект перевода. Основные типы операторов (предикатов), требуемые для записи межъязыковых процедур морфолого-синтаксического уровня, и примеры самих процедур.

Языковая избыточность, связанная с линейной структурой текста, как основа для морфолого-синтаксического синтеза по неполной и/или неточной входной информации. Возможности такого синтеза для русского языка.

**Часть 3. Средства и методы моделирования
синтаксических и семантико-синтаксических аспектов
языковой системы при сущностном подходе
к лингвистическому описанию**

Тема 9. Структуры зависимостей как основной вид представления языковой синтагматики при сущностном подходе к ее описанию. Лингвистические основания расстановки стрелок зависимостей.

Отношение <синтаксис–семантика> и его отличия от отношения <морфология–синтаксис> в трехмерной модели языковой компетенции. Двусторонний – «семантико-синтаксический» – характер конкретных синтагматических отношений синтаксического яруса в плане собственно знаковой структуры

языка. Синтаксические и семантические отношения как обобщение семантико-синтаксических отношений данного яруса.

Представление семантико-синтаксических отношений в терминах зависимостей и составляющих как проявление различия между сущностным и реляционным подходами к языковой синтагматике. Эндоцентрические и экзоцентрические конструкции при том и другом подходе. Требование древовидности структуры зависимостей для эндоцентрических конструкций.

Основные критерии расстановки и ориентации семантико-синтаксических зависимостей в представлении эндоцентрических конструкций. Критерии свертывания/развертывания словосочетаний, связанных стрелками зависимостей. Критерий семантической неотделимости. Критерий морфологического согласования. Критерий морфолого-синтаксического управления. Критерий семантического управления. Критерий структурной аналогии. Критерий автономной семантической интерпретируемости. Критерий трансформационной неотделимости. Критерий трансформационного подобию. Критерий максимальной проективности.

Применение имеющихся критериев расстановки и ориентации связей семантико-синтаксических зависимостей к материалу конкретных языков.

Представление в терминах зависимостей основных видов экзоцентрических конструкций. Сочинительные, спецификативные, анафорические зависимости. Структуры зависимостей для эллиптических конструкций.

Объединение различных типов зависимостей в составе единого графа и проблема его древовидности.

Тема 10. Зависимости и валентности.

Семантико-синтаксические валентности языковых сущностей как обобщение собственно знаковых зависимостей между текстовыми инстанциями этих сущностей путем разбиения множества таких зависимостей на классы эквивалентности. Основные принципы классификации текстовых зависимостей в этих целях. Требование однократной реализации валентностей в тексте. Условия допустимых нарушений этого требования. Трансформационный критерий классификации зависимостей. Критерий классификации зависимостей на базе компонентного анализа семантики связываемых ими сущностей.

Проблема актантов (комплементов) и сирконстантов (адьюнктов). Связь этой проблемы с ориентированностью от-

ношений зависимости в синтаксической структуре и с необходимостью выбора в качестве носителя актантной валентности на такое отношение только одной из двух связываемых им сущностей. Роль противопоставления актантов и сирконстантов в коммуникативной организации предложения и текста.

Понятие контр-актантной валентности. Противопоставление актантных и контр-актантных валентностей в его соотношении с противопоставлением актантов и сирконстантов. Синтаксические и семантические аспекты этого соотношения. Принципиальная конечность инвентаря и актантных и контр-актантных валентностей, сопоставляемых одному носителю.

Прямая и косвенная реализация валентностей. Понятие контекстного валентностного наследования.

Энциклопедическая информация и возможности ее представления в валентностных терминах. Понятие ассоциативных валентностей.

Внутрифразовые и межфразовые валентности.

Вероятностные и статистические сведения о языковой синтагматике при ее описании в терминах семантико-синтаксических валентностей.

Тема 11. Принципы организации семантико-синтаксических знаний о языковой синтагматике при сущностном подходе к ее описанию.

Основные виды лингвистических знаний, представляющие семантико-синтаксические аспекты языковой синтагматике при сущностном подходе к лингвистическому описанию.

Знания об отдельной валентности: семантическая интерпретация валентности; семантические требования к лингвистическим сущностям, способным заполнять эту валентность в тексте; варианты ее текстовой реализации.

Знания об отдельном варианте реализации валентности: направление текстовой зависимости, реализующей эту валентность, при данном варианте ее реализации; требования к аналитическому оформителю текстовой сущности, заполняющей эту валентность; требования к самой этой сущности; позиционные сведения. Три вида требований к заполнению валентностей: морфологосинтаксические, семантические, лексические. Виды позиционных сведений о варианте реализации валентности: позиционное соотношение сущности, заполняющей валентность, и ее аналитического оформителя друг с другом и с сущностью – носителем валентности; требо-

вания к текстovým сущностям, допустимых между ними; допустимость какой-либо из этих сущностей в качестве источника разрешенной непроективности.

Знания о соотношении между собой единиц, заполняющих отдельные валентности данной лингвистической сущности и/или сущностей, связанных с ней в структуре, в тех или иных вариантах реализации этих валентностей.

Формальные виды требований к описываемым лингвистическим сущностям: разрешения, запреты, исключения из запретов. Оценочная информация в составе таких требований.

Организация грамматических знаний семантико-синтаксического типа в виде иерархии статей грамматических лингвистических сущностей, связанных друг с другом и с лексическими сущностями отношениями наследования, Грамматические сущности, выделяемые в рамках оси обобщения (= классифицирующие грамматические признаки). Суммирующее наследование и наследование по умолчанию как основные механизмы наследования, действующие на этой оси. Грамматические сущности, выделяемые с учетом отношений членения в ярусах морфологии и синтаксиса (resp., морфологические и валентностные категории). Механизмы взаимодействия между наследником и наследодателем при их соотношении по оси членения.

Знания о семантико-синтаксических аспектах языковой синтагматики в статьях лексических сущностей. Собственно лексемы и идиоматические сочетания лексем: нестандартные механизмы наследования.

Тема 12. Семантико-синтаксические описания сущностного типа и их использование на различных этапах обработки текстов в целях автоматического перевода.

Проблема отбора знаний о языковой синтагматике, включаемых в конкретное лингвистическое описание, и увязки формального аппарата их представления в этом описании с характером осуществляемых на его основе процедур обработки текстов.

Формальные средства для записи синтагматической информации семантико-синтаксического типа в лингвистических описаниях, ориентированных на задачи анализа. Представление синтагматических отношений, устанавливаемых при анализе между лингвистическими сущностями анализируемого текста, как свойств этих сущностей – реализаций

их валентностей. Основные реляционные признаки (функции), используемые в этих целях. Представление синтагматической структуры как списка лингвистических сущностей с их реляционными признаками.

Возможности моделирования процедур выявления и обработки синтагматических структур семантико-синтаксического уровня операторными и предикатными средствами. Основные типы требуемых операторов (предикатов) и примеры процедурных описаний, записанных в их терминах. Универсальные алгоритмы анализа структур зависимостей на основе валентностного представления семантико-синтаксической синтагматики. Проблема интеграции универсальных процедур анализа с локальными операторными (предикатными) процедурами.

Структуры зависимостей и структуры составляющих на этапе межъязыкового перехода. Проблема «предпереводческого анализа» и «предпереводческого перефразирования». Основные типы межъязыковых знаний семантико-синтаксического уровня и их организация при сущностном подходе к лингвистическому описанию. Семантико-синтаксические условия выбора и формирования переводных эквивалентов для выделенных в переводимом тексте сущностей и отношений (структурных и позиционных).

Процедурные средства представления межъязыковых знаний семантико-синтаксического типа. Их соотношение с аналогичными средствами морфолого-синтаксического уровня. Возможности моделирования межъязыкового перехода «по умолчанию» на основе неполных сведений о семантико-синтаксической организации переводимого текста. Примеры записи соответствующих процедур.

Представление входной информации семантико-синтаксического уровня для целей синтеза при сущностном подходе к лингвистическому описанию. Этапы и проблематика синтеза при таком подходе.

Организация семантико-синтаксических знаний и процедур их использования в модели русского синтеза, реализующей данный подход. Возможности учета в этих целях словообразовательной структуры лексем. Формальный аппарат для их записи в этой модели. Примеры словарных и грамматических описаний, представленных в его терминах. Языковая избыточность и механизмы умолчаний синтаксического и семантико-синтаксического уровней. Механизм «семантико-синтаксического согласования» как средство снятия смысловой вариативности, связанной с не эксплицированными во вход-

ном представлении текста аспектами его содержания, по их «следам» (возможно, множественным) в его составе. Механизм «лексико-семантического согласования» и его использование при синтезе для моделирования «конкретизирующего перевода» в случаях семантической неопределенности (нерасчлененности) единиц оригинала. Соотношение этого механизма с понятиями лексических и энциклопедических функций

Часть 4. Средства и методы моделирования семантики при сущностном подходе к лингвистическому описанию

Тема 13. Общие требования к семантическому описанию при компонентном подходе к семантике

Компонентный подход к семантике и основные требования к аппарату семантического описания лингвистических сущностей. Постулат о «непрерывности семантической записи» и принцип членимости значений лингвистических сущностей. Принципы парадигматической и синтагматической системности семантических свойств лингвистической сущности. Семантические признаки, семантические падежи и семантические компоненты. Смысловая парадигматика и синтагматика лингвистической сущности в их соотношении друг с другом.

Основные типы семантической информации о лингвистической сущности: толкование, правила его контекстуальных модификаций, набор семантических классов, в которые входит данная лингвистическая сущность, семантическая интерпретация ее валентностей, семантические требования, предъявляемые ею к их заполнению. Взаимосвязь этих типов информации между собой.

Понятие идеальной системы элементарных семантических единиц. Требования к системе таких единиц: конечность инвентаря элементарных единиц, различных по своему сигнификативному содержанию, при бесконечности их комбинаторных возможностей; содержательность и полнота этого инвентаря; его минимальность; соотносимость элементарных семантических единиц с соответствующими им единицами естественного языка.

Принципы выделения элементарных семантических сущностей (семантических элементов). Индивидуализирующие и классифицирующие семантические элементы. Категориальные и энциклопедические семантические элементы.

Принципы выделения элементарных синтагматических семантических отношений. Ориентированные и неориентированные семантические отношения.

Соотношение семантических сущностей и синтагматических семантических отношений в идеальной системе элементарных семантических единиц.

Тема 14. Язык формальных толкований лингвистических сущностей.

Язык формальных толкований (ЯФТ) как средство представления парадигматических и синтагматических аспектов лингвистической семантики при сущностном подходе к лингвистическому описанию.

Словарь языка формальных толкований. Средства записи референциальных индексов и энциклопедических семантических сущностей. Категориальные семантические сущности и их валентности. Определение синтагматических семантических отношений как производных семантических единиц – «дериватов» категориальных семантических сущностей по соответствующим их валентностям. Парадигматические отношения включения и выводимости, определенные на множестве категориальных семантических сущностей и их дериватов. Основные категориальные семантические сущности языка формальных толкований. Сравнение их инвентаря с аналогичными инвентарями других семантических языков.

Коммуникативная организация элементарных синтагматических сочетаний семантических сущностей и отношений в формальном толковании. Понятие семантической коммуникативной организации такого сочетания. Три базовых коммуникативных режима: нексус, юнкция, реляция. Противопоставление нексуса и юнкции в естественно-языковом синтаксисе. «Синтаксические» лексические функции как естественно-языковой аналог режима реляции. Возможность рекурсивного комбинирования базовых коммуникативных режимов. Средства представления коммуникативных режимов в языке формальных толкований.

Понятие правильной семантической формулы и ее формальное определение. Механизмы наследования валентностей семантических формул, основанные на отношениях членения («от части к целому»). Парадигматические отношения между семантическими формулами: равнозначность, включение, выводимость, юнктивная деривация. Наследование свойств семантических формул на основе этих отношений.

Тема 15. Возможности использования языка формальных толкований для описания семантики в связи с задачами автоматического перевода.

Словарные и контекстуальные толкования языковых сущностей. Проблема омонимии и полисемии в формализованном описании лексики.

Толкования и семантическое представление текста.

Толкования и языковая парадигматика. Отношения синонимии, конверсности, семантической близости, логического следования. Понятия родо-видовых и валентностных лексико-семантических классов. Правила выведения признаков тех и других классов из толкований языковых сущностей.

Толкования и «семантическая грамматика». Наследование парадигматических и валентностных свойств языковых сущностей по семантическим основаниям. Основные типы грамматических сведений в словарных статьях категориальных семантических сущностей. Семантические требования к заполнению валентностей как один из видов этих сведений. Примеры толкований и семантических требований к валентностям конкретных лексических единиц.

Использование языка формальных толкований в сложных случаях структурного анализа, восстановления эллипсиса, перефразирования. Проблема контекстного валентностного наследования и связанного с ним семантического согласования «разноуровневых» актантов в предикатно-актантных конструкциях. Примеры описаний соответствующих явлений средствами рассматриваемого языка.

Семантические и прагматические явления, представляющие трудности для описания в терминах языка формальных толкований. Возможности дополнения этого языка процедурными средствами.

Литература

- Апресян Ю.Д. и др. Лингвистическое обеспечение системы ЭТАП-2. М.: Наука, 1989 (к теме 1); с. 43-45 (к теме 4); с. 40-59, 224-235 (к теме 8); с. 266-275 (к теме 11); с. 29-33 (к теме 12).
- Апресян Ю.Д. Толкование лексических значений как проблема теоретической семантики // Изв. АН СССР. Сер. лит-ры и языка, 1969, Т. XXVIII, Вып. 1. С. 11-24 (к теме 14).
- Апресян Ю.Д. Экспериментальное исследование семантики русского глагола. М.: Наука, 1967 (к теме 14).

- Арсентьева Н.Г. От семантических свойств слова – к поверхностно-синтаксическим: постановка задачи и результат машинного эксперимента // Научно-техническая информация. Сер.2. 1976. №9. С. 25-32 (к теме 15).
- Богуславский И. Исследования по синтаксической семантике. М.: Наука, 1985 (ко всему курсу).
- Вардуль И.Ф. Основы описательной лингвистики. М., 1977 (к теме 2).
- Вежбицка А. Семантические универсалии и «примитивное мышление». Толкование эмоциональных концептов // Вежбицка А. Язык. Культура. Познание. М.: Русские словари, 1997. С.291-319, 3288-341 (к теме 13).
- Есперсен О. Философия грамматики. М., 1957 (к теме 14).
- Железняков М.И., Невлева Т.Н., Новицкая И.М., Смирнова Л.Н., Цейтин Г.С.. Опыт построения модели типа «Текст -> действительность» с использованием ассоциативных сетей // Машинный фонд русского языка: предпроектные исследования. М., 1988. С.140-167 (к теме 9).
- Жолковский А.К., Мельчук И.А., Шаляпина З.М. Экспериментальный фрагмент англо-русского толково-комбинаторного словаря в процессе перевода. М., 1971 (к теме 4).
- Жолковский А.К., Мельчук И.А. К построению действующей модели языка «Смысл \leftrightarrow Текст» // Машинный перевод и прикладная лингвистика. Вып. 11. М., 1969. С.5-35 (к теме 5); с. 5-35 (к теме 13).
- Жолковский А.К., Мельчук И.А. О семантическом синтезе // Проблемы кибернетики. Вып. 19. М., 1967. С.177-238 (к теме 12); с.177-238 (к теме 14).
- Засорина Л.Н., Берков В.П. Понятие валентности в языке // Вестник Ленинградского университета. Сер. истории, языка и литературы. Л., 1961. Вып. 2. № 8. С. 133-139 (к теме 5).
- Канович М.И., Шаляпина З.М. Аппарат R-отсылок как универсальное средство синтаксического синтеза (на опыте разработки системы русского синтеза RussLan) // Диалог 2006: Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции. М.: Изд-во РГГУ, 2006. С. 207-213 (к теме 12).
- Канович М.И., Шаляпина З.М. Лингвистическая избыточность как фактор организации лексико-морфологической базы в модели русского словоизменительного синтеза // ДИАЛОГ'95: Труды Международного семинара по компьютерной лингвистике и ее приложениям. Казань, 1995. С. 133-139 (к теме 7).
- Канович М.И., Шаляпина З.М. Русский морфологический и морфолого-синтаксический синтез по неполной и/или неточной входной информации // Русский язык как иностранный: лингвистические проблемы. М.: ИКАР, 1997. С. 137-147 (к теме 8).
- Карцевский С. Об асимметричном дуализме лингвистического знака // Звегинцев В.А. История языкознания XIX-XX веков в очерках и извлечениях. Часть 2. М., 1965. С. 85-90 (к теме 2).
- Каунельсон С.Д. О грамматической категории // Вестник Ленинградского университета. Сер. истории, языка и литературы. Л., 1948, № 2 (к теме 5).

- Кибрик А.Е. Критерии определения направления синтаксической связи в синтагме // Структурно-математические методы моделирования языка. Часть 1. Киев, 1970. С. 60-61 (к теме 9).
- Кибрик А.Е. О соотношении понятий синтаксического подчинения с понятиями согласования, управления и примыкания // Проблемы теоретической и экспериментальной лингвистики. М.: МГУ, 1977. С. 161-179 (к теме 9).
- Козинский И.Ш. Некоторые грамматические универсалии в подсистемах выражения субъектно-объектных отношений. Дисс. ... канд. филол. наук. М., 1979 (к теме 3).
- Костыркин А.К. Применение технологий ограничений в синтаксическом анализе предложения // ДИАЛОГ 2000: Труды международного семинара по компьютерной лингвистике и ее приложениям. Т. 2. Прикладные проблемы. Протвино, 2000. С. 199-208 (к теме 5).
- Кулагина О.С. Исследования по машинному переводу. М.: Наука, 1979 (к теме 4).
- Лемб С.М. Стратификационная лингвистика как основа машинного перевода // Научно-техническая информация. 1964. № 10. С. 33-42 (к теме 1).
- Леонтьева Н.Н., Волковыская Е.В., Копылова О.В., Молчанова Т.В., Штернова О.А. Словарь энциклопедических функций и его роль в автоматическом индексировании // Научно-техническая информация. Сер. 2. 1978. № 7. С. 23-29 (к теме 12).
- Леонтьева Н.Н. Автоматическое понимание текстов: системы, модели, ресурсы: Учеб. пособие для студ. лингв. фак. вузов. М.: Академия, 2006. 304 с (ко всему курсу).
- Мартемьянов Ю.С. К описанию текста: язык валентно-юнктивных эмфазных отношений // Машинный перевод и прикладная лингвистика. Вып. 13. М.: МГПИИЯ им. М.Тореза, 1970 (к теме 14).
- Марчук Ю.Н. Проблемы машинного перевода. М.: Наука, 1983. С. 106-109 (к теме 8).
- Мельников Г.П. Системология и языковые аспекты кибернетики. М., 1978 (к теме 3).
- Мельчук И.А. Опыт теории лингвистических моделей «Смысл \leftrightarrow Текст». М.: Наука, 1974. 314 с (ко всему курсу).
- Мельчук И.А. Уровни представления языковых высказываний и общее строение модели «Смысл \leftrightarrow Текст» // Актуальные проблемы лексикологии и словообразования. Новосибирск: НГУ, 1972 (к теме 1).
- Мишулич Л.И., Червоненкис А.Я. Специализированная диалоговая система // Вопросы разработки прикладных систем. Новосибирск, 1979. С. 111-129 (к теме 3).
- Модина Л.С., Шаляпина З.М. Графико-морфологическая грамматика японского языка в действующей модели анализа японских письменных текстов // Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки: Материалы III Межд. конф. М.: МГУ, ИСАА, СПб.: СПбГУ, 1995. С. 120-123 (к теме 7).
- Модина Л.С., Шаляпина З.М. Лексико-грамматическая классификация японских лингвистических сущностей и ее функции в моде-

- ли японско-русского автоматического перевода ЯРАП // Языки Дальнего Востока, Юго-Восточной Азии и Западной Африки: Материалы IV Межд. конф. М.: МГУ, ИСАА, СПб.: СПбГУ, 1997 (к теме 12).
- Модина Л.С., Шалыпина З.М. Межъязыковые операции в модели японско-русского автоматического перевода ЯРАП // Электронные библиотеки и базы данных по истории Евразии в средние века. Вып. 7. М.: ИВ РАН, 1999. С. 172-231 (к теме 8).
- Модина Л.С., Шалыпина З.М. Принципы анализа японских текстов при моделировании японской лексико-морфологической системы // Диалог 1996: Труды международного семинара по компьютерной лингвистике и ее приложениям. М., 1996. С. 169-174 (к теме 7).
- Модина Л.С., Шалыпина З.М. Принципы организации лингвистических знаний в объектно-ориентированной модели лексико-морфологической системы японского языка // Диалог 1995: Труды Международного семинара по компьютерной лингвистике и ее приложениям. Казань, 1995. С. 198-205 (к теме 5); с. 198-205 (к теме 7).
- Нагао М., Цудзиси Дз., Накамура Дз. Машинный перевод с японского языка на английский // ТИИЭР. Труды Института инженеров по электротехнике и радиоэлектронике. 1986. Т. 74, №7. С. 114-118 (к теме 15).
- Нариньяни А.С. Формальная модель: общая схема и выбор адекватных средств // Препринт №107 ВЦ СО АН СССР. Сер. «Экспериментальная разработка формальной модели русского языка». Вып. 1. Новосибирск, 1978. 29 с (к теме 3).
- Падучева Е.В. Некоторые проблемы моделирования соответствия между текстом и смыслом в языке // Известия Академии наук. Серия литературы и языка. 1975. №6. С. 548-559 (к теме 1).
- Падучева Е.В. О семантике синтаксиса. М., 1974 (к теме 9).
- Попов Э.В. Общение с ЭВМ на естественном языке. М., 1982 (к теме 3).
- Пумпянский А.Л. Информационная роль порядка слов в научной и технической литературе. М.: Наука, 1974 (к теме 4).
- Ревзин И.И. Современная структурная лингвистика. Проблемы и методы. М.: Наука, 1977 (к теме 9).
- Севбо И.П. Графическое представление синтаксических структур и стилистическая диагностика. Киев: Наукова думка, 1981 (к теме 9).
- Соссюр Ф. де. Курс общей лингвистики // Соссюр Ф. де. Труды по языкознанию. М.: Прогресс, 1977. С. 31-273 (к теме 2).
- Танака Х., Мидзогути Ф. Система машинного перевода, учитывающая семантику // Дзёхо сери. 1985. Т.26, № 1. С. 1191-1196 (к теме 15).
- Шалыпина З.М. Три функции языкового знака и проблемы их отражения в лингвистическом описании // Проблемы функциональной грамматики. М.: Наука, 1985. С. 103-120 (к теме 2).
- Шалыпина З.М. и др. Об англо-русском многоаспектном словаре с грамматическим обеспечением // Вопросы информационной теории и практики. Сб. № 27. М: Советское радио, 1975. С. 64-87 (к теме 11).

З. М. Шаляпина

- Шаляпина З.М., Канович М.И., Костыркин А.В., Модина Л.С.* Лексико-синтаксическая информация в словаре для автоматического перевода (В рамках сущностного подхода к лингвистическому описанию) // Словарь в современном мире. Материалы Третьей международной школы-семинара. Иваново, 2000. С. 117-121 (к теме 11).
- Шаляпина З.М., Канович М.И., Модина Л.С.* Экспериментальный комплекс ЯРАП как инструмент сопоставительных лингвистических исследований // Общее и восточное языкознание. Сб. науч. тр. М.: Современный писатель, 1999. С. 64-79 (к теме 8).
- Шаляпина З.М.* Автоматический перевод как моделирование переводческой деятельности человека // Международный форум по информации и документации. 1980. Т. 5, № 26. С. 16-20 (к теме 4).
- Шаляпина З.М.* Англо-русский многоаспектный автоматический словарь (АРМАС) // Машинный перевод и прикладная лингвистика. Вып. 17. М.: МГПИИЯ, 1974. С. 7-67 (к теме 11).
- Шаляпина З.М.* Грамматика и ее соотношение со словарем при словоцентрическом подходе к языку (на опыте формализованного лингвистического описания) // Вопросы языкознания, 1991. №5. С. 42-54 (к теме 5).
- Шаляпина З.М.* К проблеме коммуникативной организации текста и ее отражения в семантической записи // Вычислительная лингвистика. М.: Наука, 1976. С. 53-80 (к теме 14).
- Шаляпина З.М.* К проблеме моделирования языковой компетенции и языковой деятельности человека в рамках общей модели языковой действительности // Референция и проблемы текстообразования. М., 1988. С. 214-232 (ко всему курсу).
- Шаляпина З.М.* Межфразовые связи как способ реализации лексических валентностей // Аспекты изучения текста. М., 1981. С. 80-87 (к теме 10).
- Шаляпина З.М.* Многофункциональность аппарата лексических функций и некоторые возможности его применения и развития // Восток – Запад: Вторая международная конференция по модели «Смысл ↔ Текст». М.: Языки слав. культуры, 2005. С. 447-457 (к теме 12).
- Шаляпина З.М.* О двух типах кореферентности валентностей в предикатно-актантных конструкциях // Типология конструкций с предикатными актантами. Л.: Наука, 1985. С. 46-50 (к теме 15).
- Шаляпина З.М.* О морфолого-синтаксическом статусе некоторых японских грамматических единиц (в связи с общими проблемами разграничения морфологии и синтаксиса в естественном языке // Актуальные проблемы японского языкознания. М.: Наука, 1986. С. 69-103 (к теме 6).
- Шаляпина З.М.* Об одном возможном подходе к задаче анализа японских текстов при их автоматической обработке // Проблемы компьютерной лингвистики. Минск., 1997. С. 110-144 (к теме 8); с. 110-144 (к теме 12).
- Шаляпина З.М.* Об одном подходе к выделению фразеологических

- единиц в связи с задачами автоматической обработки текстов // Фразаография в машинном фонде русского языка. М.: Наука, 1990. С. 188-195 (к теме 6).
- Шаляпина З.М. Об одном способе формализации понятия лингвистической нормативности в целях автоматической обработки текстов // Формальное представление лингвистической информации. Новосибирск, 1982. С. 20-37 (к теме 10).
- Шаляпина З.М. Об одном формализме для записи толкований слов и словосочетаний // Обработка текста и когнитивные технологии. Вып. 1. М., Пуцдино, 1997. С. 73-110 (к теме 14).
- Шаляпина З.М. Оппозиция «часть-целое» и сущностный подход к моделированию языковой компетенции // Роман Якобсон: тексты, документы, ис следования. М.: РГГУ, 1999. С. 541-551 (ко всему курсу).
- Шаляпина З.М. Семантико-синтаксический анализ в системе англо-русского автоматического перевода (АРАП). М., 1974. 58 с. (Публ. проблемной группы по экспериментальной и прикладной лингвистике, вып. 47) (к теме 12).
- Шаляпина З.М. Семантические элементы как основа лексикографического описания: общие принципы и формальный аппарат // Машинный перевод и прикладная лингвистика. Вып. 15. М., 1972. С. 54-98 (к теме 13).
- Шаляпина З.М. Синтаксис и семантика в системе англо-русского автоматического перевода АРАП // Актуальные вопросы практической реализации систем автоматического перевода. Часть 1. М.: МГУ, 1982. С. 184-204 (к теме 9).
- Шаляпина З.М. Словарь как единая форма организации формализованного лингвистического описания // Proceedings of the Fourteenth International Congress of Linguists. Berlin: Akademie-Verlag, 1991. P. 1244-1247 (к теме 5).
- Шаляпина З.М. Текст как объект автоматического перевода // Текст и перевод. М.: Наука, 1988. С. 113-129 (к теме 4).
- Шаляпина З.М. Формальный язык для записи толкований слов и словосочетаний // Проблемы кибернетики. Вып. 36. М.: 1979. С. 247-278 (к теме 14).
- Allerton D.J. Valency and the English Verb. L.: Academic Press, 1982 (к теме 10).
- Bever T.G., Katz J.J., Langedoen D.T. An integrated theory of linguistic ability. Hassocs (Sx.), 1977 (к теме 1).
- Chomsky N. Aspects of the theory of syntax. N.Y., 1965 (к теме 2).
- Danes F. On Linguistic Strata (Levels) // Travaux linguistiques de Prague. 4. Prague, 1971. P. 127-143 (к теме 1).
- Helbig G. Zum Status der Valenz und der semantischen Kasus // Deutsch als Fremdsprache, Leipzig, 1979. H.2. S.65-78 (к теме 10).
- Helwig P. Dependency unification grammar // Proceedings of COLING 1986. Bonn, 1986. P. 195-198 (к теме 10).
- Chauché J. Les systèmes ATEF et GETA // T.A. Informations. 1975. №2. P.21-39 (к теме 8).
- Jäppinen H., Lehtola A., Valkonen K. Functional Structures for Parsing

- Dependency Constraints // Proceedings of COLING 1986. Bonn, 1986. P. 461-463 (к теме 9).
- Ružicka R. Three aspects of valence // Valence, semantic case and grammatical relations. Amsterdam: John Benjamins, 1978. P. 47-53 (к теме 10).
- Sakamoto Y., Ishikawa T., Satoh M. Concept and Structure of Semantic Markers for Machine Translation in Mu-Project // Proceedings of COLING 1986. P. 13-19 (к теме 14).
- Schmidt P. Valency Theory in a Stratificational MT-System // Proceedings of COLING 1986. Bonn, 1986. P. 307-312 (к теме 10).
- Sgall P. Zur Frage der Ebenen im Sprachsystem // Travaux linguistiques de Prague. 1. Prague, 1966. P. 95-106 (к теме 1).
- Sgall P. Cases in a stratificational description // Prague Studies in Mathematical Linguistics. 3. Praha, 1972. P. 201-211 (к теме 13).
- Somers H.L. On the validity of the complement-adjunct distinction in valency grammar // Linguistics. 1984. Vol. 22. №4. P.507-530 (к теме 10).
- Somers H.L. The need for MT-oriented versions of Case and Valency in MT // Proceedings of COLING 1986. Bonn, 1986. P. 118-123 (к теме 10).
- Starosta S., Nomura H. Lexicase Parsing: A Lexicon-driven Approach to Syntactic Analysis // Proceedings of COLING 1986. Bonn, 1986. P. 127-132 (к теме 5).
- Trnka B. On the linguistic sign and the multilevel organization of language // Travaux linguistiques de Prague. 1. L'École de Prague d'aujourd'hui. Prague: Academia, 1966. P.33-40 (к теме 1).
- Vater H. On the possibility of distinguishing between complements and adjuncts // Valence, semantic case and grammatical relations. Amsterdam: John Benjamins, 1978. P. 21-45 (к теме 10).
- Wierzbicka A. Semantic primitives // Linguistische Forschungen, Band 22. Frankfurt, 1972. 235 p (к теме 13).
- Yokoyama Sh., Hanakata K. Conceptual Lexicon Using an Object-Oriented Language // Proceedings of COLING 1986. Bonn, 1986. P. 226-228 (к теме 11).
- Zernik U., Dyer M. Disambiguation and Language Acquisition through the Phrasal Lexicon // Proceedings of COLING 1986. Bonn, 1986. P. 247-252 (к теме 5).

Пиктский памятник «30 бруде» в свете общих проблем ранней эпиграфики

Почти двухвековое изучение истории и языка народа, известного под условным обозначением «пикты», как это ни странно, дало поразительно мало общепризнанных интерпретаций дошедших до нас более, чем скудных сведений, в том, что касается их языка и этнической принадлежности. Относится это и к более общим проблемам (были пикты индоевропейцами или нет), и к частностям (прочтения отдельных текстов, записанных огамическим или латинским письмом). Как мы уже отмечали [Михайлова 2000], принципиальная невозможность надежной реконструкции языка и культуры пиктов осознавалась практически всеми, кто в той или иной степени затрагивал данную проблематику. Более того, сам Р.А.С.Макалистер, который еще в 1940 г. решился дать положительное прочтение пиктских надписей [Macalister 1940], сам же позднее называл проблему пиктов «кошмаром британской этнологии» [Macalister 1996, p. 47].

Ввиду недостаточности археологических и языковых данных о пиктах, и при почти полном (к нашему удивлению) отсутствии надежных антропологических описаний имеющихся погребений, этно-лингвистическая атрибуция пиктов действительно во многом определяется либо изначальной убежденностью исследователя, либо тем конкретным материалом, с которым он предпочитает работать в первую очередь. Так, анализ топонимики дает скорее бриттские данные, имена королей и знаменитые загадочные «пиктские символы» – обычно ставят всех в тупик, а их звериный орнамент напоминает скифо-сарматский. Невольная предубежденность, как нам кажется, в данном случае вполне объяснима, однако мы полагаем, что ее влияние не должно распространяться за рамки собственно интерпретации имеющихся данных. Понятным также является и невольное нежелание «замечать» материал, в концепцию исследователя не укладывающийся. Однако, как мы с удивлением обнаружили, при работе с «пиктским мате-

риалом» нередко случаи его достраивания, пренебрежения существенными деталями или даже их искажения, – как мы полагаем также невольные, – но при этом все же странные и досадные. Приведем несколько примеров.

Так, в книге Энтони Джексона, посвященной описанию, анализу и систематизации пиктских символов, содержится раздел «Огам и язык пиктов» [Jackson 1984, p. 174-200]. Основная идея автора состоит в том, что огамические знаки пиктов, в отличие от гойдельского огама, должны были передавать не звуки, организованные в слова, а числа календаря, причем – календаря лунного. Естественно, поэтому исходным числом у него получается число 28, а количество штрихов в каждой надписи должно быть кратно семи. Расположение штрихов, таким образом, соответствует уже не традиционному огамическому буквенному коду (для которого, отметим, маркированным является число 5), но некоей особой нумерологической символике. Джексон тщательно анализирует и детально расписывает 36 надписей, естественно, принимая во внимание, что многие из них значительно повреждены, иными словами – у многих оказывается отбитым верх, и ответить на вопрос, сколько именно знаков было на камне изначально, часто просто невозможно, ввиду темноты самого текста сохранившейся надписи. Но именно это дает Джексону право и основание с легкостью достроить надпись до нужного ему числа штрихов-символов. Не менее свободно обращается он и со словоделением надписи, хотя именно в пиктском огаме, в отличие от традиционного ирландского, кроме часто условных и не всегда надежных пробелов между «словами», обычно уже ставились точки-маркеры.

Концепции Э.Джексона не принято принимать всерьез, о чем свидетельствует и скандально-полемический тон, которым написана его книга. Однако, такое же пренебрежение к практическому материалу мы встречали и в работах гораздо более серьезных исследователей языка пиктских надписей. Так, в свое время В.П.Калыгиным нам был предоставлен во временное пользование отпечаток статьи К.Форсайт, лично подаренный ею. Работа посвящена анализу пиктской огамической надписи, нанесенной на круглый плоский камень-кольцо, обнаруженный во время раскопок 1970 г. в погребении на Оркнейских островах (см. рис. 1а). Огам написан вдоль кольцевой оси и, по предположению К.Форсайт, должен читаться против часовой стрелки. При таком чтении «расшифровывается» древнеир-



рис. 1а

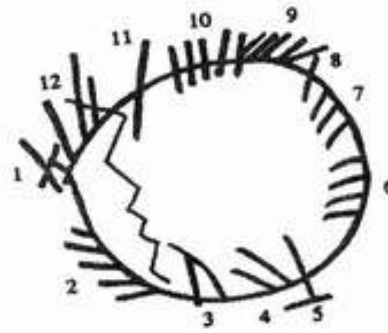


рис. 1б

ландский текст BENDDACT ANIM L «благословение имени эЛЬ». Все исследование выполнено очень основательно, даны параллели из близких по содержанию надписей, характерных для древнеирландской монастырской традиции, фоном к работе выступает не менее тщательно описанный археологический материал. Лежащая в основе концепции Форсайт идея, что надпись должна читаться против часовой стрелки, аргументируется ею особо, что вполне естественно, так как хорошо известно, что «счастливым» направлением в древне-кельтской традиции (да и не только в ней) было направление по часовой стрелке. Как она пишет, «в ирландском языке *dessel* означает направление "по солнцу" или "направо", а также имеет вторичное значение "удача, благоприятное положение", причем аналогичное значение наблюдается и в других кельтских языках» [Forsyth 1995, p.683]. Возможное чтение в противоположную сторону объясняется ею довольно сложным образом: по ее мнению, каменное кольцо-амулет не было положено поверх трупа, но покойный носил его еще при жизни в качестве оберега надписью, видимо, наружу. При этом текст «для него» виртуально прочитывался по часовой стрелке, как если бы он был прочерчен, например, на стекле или прозрачной пластмассе. В принципе, такое странное объяснение кажется все же возможным, однако взгляд на саму надпись ясно показывает, что это не так: на камне очень отчетливо видна прочерченная резчиком круговая ось (см. рис. 1б), характер штриха которой, более глубокий в начале и мелкий к концу, ясно показывает, что движение совершалось именно по часовой стрелке. Более того, то же можно сказать и о самих огамических знаках. Характер их расположения на оси ясно показывает, что резчик разместил их не очень точно, и к концу у него осталось свободное место (по интерпретации К. Форсайт это должен был бы быть

To Victor from Katherine
Actually, I'm rather doubtful myself.
Any ideas?

28 November 1999

Glasgow

рис. 2

неоправданный пробел между 1-м и 2-м знаками, т.е. между В и Е).

Однако, будучи честны по отношению к Форсайт, отметим, что на подаренном ею В.П.Калыгину оттиске помещена очень интересная посвятельная надпись (см. рис. 2). Надпись выполнена латинским курсивом, язык надписи – предположительно английский. Восстанавливается следующий текст: «To Victor from Kath [нрзб., предположительно – Katherine]. Actually, I am rather doubtful myself. Any ideas?». Что для нас особенно важно, надпись датирована 28 ноября 1999 г., Glasgow, причем аутентичность надписи и датировки подтверждаются нашими сведениями о том, что в ноябре 1999 г. В.П.Калыгин действительно находился в Шотландии. Сомнения, которые испытывала К.Форсайт в конце 1999 г., предстают особенно значимыми на фоне того, что в 1997 г. в своей книге «Язык в стране пиктов», снабженной задиристым подзаголовком The case against 'non-Indo-European' Pictish, она достаточно уверенно пишет об этой же надписи – «этот текст следует читать задом наперед» [Forsyth 1997, p. 36].

Аналогичное пренебрежение к чисто палеографической работе с материалом, как нам кажется, встречается и при анализе известной латинской надписи на так называемом «Дростеновом камне» (Drosten Stone) (см. рис.3). Надпись выполнена латинским полу-курсивом, разделена на четыре строки, датируется IX в. и находится в хорошем состоянии. Общепринятым считается следующее ее прочтение: (1)DROST-EN (2)IPEUORET (3)ETTFOR (4)CUS. На эту надпись в свое время опирался еще Р.Макалистер в своих попытках реконструкции пиктского «словаря», и слово ipe в данном случае переводимое им предположительно как «племянник, сын сестры» было для него очень важным. Он предлагал следующее членение тек-

ста – DROSTEN IPEUOR ETETTFOR CUS, и в целом надпись трактуется им следующим образом: «Дроста – его племяннику – ему Эттету – ему крест» (надпись действительно выполнена на основании креста). Данное членение текста на слова кажется нам все же странным, поскольку при этом возникают «перебивки» через строку, совершенно неоправданные и вызывающие особое удивление на фоне того, что на камне совершенно отчетливо видны разделители-маркеры в виде групп из трех равноудаленных точек. Причем эти маркеры присутствуют и в прориси, сделанной самим же Макалистером (см. [Macalister 1940, p. 195])!

«Прочтения» Макалистера и его «финская теория» происхождения языка пиктов в настоящее время почти не имеют последователей, однако предположительная апеллятивная природа слова IPE отмечалась и в классической работе К. Джексона, который видел в нем слово не-кельтского и, более того, не-индоевропейского происхождения [Jackson 1955, p. 140]. Тем не менее, совершенно новая и смелая интерпретация этой надписи была предложена современным шотландским исследователем Т. Кланси, который предположил, что это слово следует расчленить и прочесть P как R, что в целом приводит его к следующему прочтению: DROSTEN I RE UORET ETT FORCUS – «Дростен во время (ирл.!) Ворета и Форкуса» [Clancy 1993]. Более того, такое прочтение дает ему основание датировать надпись временем правления пиктского короля Вурада сына Баргота, который, согласно «Пиктской хронике», правил с 839 до 842 г. Не говоря уже о том, что такое прочтение, как отчасти пони-



рис. 3

мает и сам автор, с большим трудом согласуется с реконструируемыми им же самим именными парадигмами пиктского языка, оно не представляется разумным именно по причине наличия в надписи все тех же разделительных маркеров, которые Т.Кланси почему-то просто предпочитает не замечать! Иными словами, вся его серьезная аргументация и все его интересные рассуждения лингвистического и исторического характера рассыпаются в прах, стоит лишь взглянуть на саму надпись!

Мы смеем предположить, что аналогичное нежелание вчитываться в сам текст, смотреть прежде всего на оригинал, руководило и рядом исследователей, анализировавших знаменитый список «30 бруде» из так называемой «Пиктской хроники».

Этот памятник предположительно был составлен в конце VI – начале VII в., как полагал Э.МакНил – в монастыре в Абернети (Abernethy [MacNeill 1940, p. 15]). Оригинал, естественно не сохранился, и в настоящее время мы располагаем несколькими списками, различающимися по объему (некоторые – крайне фрагментарны). Реконструируемый текст «Пиктской хроники» был написан в основном на латыни и содержал список королей, начиная с мифического предка Круитни (Cruithne, кстати – ирландское обозначение пиктов), и кончая вполне исторической фигурой – королем Кеннетом сыном Альпина, который в 871 г. отказался от титула *rex pictorum*, положив тем самым начало рождению уже собственно Шотландии. Отдельные фрагменты списков пиктских королей встречаются также и в ирландской обработке (в осн. в «Желтой книге Лекана» и в «Книге Баллимота XIV-XV вв).

Собственно перечни пиктских королей, которые и принято называть «Пиктской хроникой», дошедшие в нескольких списках (основных – 7), принято делить на три раздела (или – секции), не равноценные с точки зрения исторической (см. об этом подробнее – [Wainwright 1955, p. 16-18]). Раздел 1 открывается упоминанием о правлении эпонимического предка Круитни, следом за которым правили его семь сыновей, на смену которым пришли еще 6 или 7 также мифических правителей. Раздел 2, который содержится только в трех рукописях и который как раз и является объектом нашего интереса, представляет собой список из «30 бруде», которые правили в общей сложности 150 лет. За этим списком в данных рукописях следует уже собственно исторический перечень пиктских королей, находящий верифицирующие параллели и в ирландских хрониках (в основном – «Ульстерских анналах»). В рукописях, в которых нет списка 30 бруде, «исторический» перечень следует обычно непосредственно сразу за мифическими предками и загадочных «30 бруде» в них просто как бы вообще не существует. С точки зрения хронологической, говоря очень условно, время правления 30 бруде попадает примерно на период с сер. VII в. до конца V в. до н.э., что в итоге помещает время правления мифического Круитне уже в начало 2 тыс. до н.э. Но, естественно, не следует это понимать буквально.

Список «бруде» интересен и как историографический фе-

номен, и как материал для лингвистического анализа, т.к. содержащиеся в нем пиктские имена поразительно необъяснимы и имеют подозрительную «неиндоевропейскую» окраску. Можно также вспомнить, что Бруде – было именем пиктского короля, который по свидетельству Бэды Достопочтенного в конце VI в. принял от св. Колума Килле христианство, говоря с ним при этом через переводчика (об этом пишет также и Адамнан и практически нет ни одной работы о языке пиктов, в которой не упоминался бы этот знаменательный и показательный, навязший в зубах эпизод!). Однако, сейчас мы хотели бы остановиться лишь на самом принципе организации этого списка и, описав существующие его интерпретации, предложить нашу, строящуюся на неожиданной ирландской параллели. Собственно говоря, всего лишь ради двух слов, которые мы назовем в конце, мы и пускаемся в столь пространные рассуждения.

Итак, список «30 бруде» организован довольно странным образом: он состоит из 15-ти пар имен «бруде» ('вождей, королей?'), причем каждое второе имя снабжено префиксом *ur-*, например: *brude Pant, brude Urpant, brude Leo, brude Urleo, Brude Gant, Brude Urgant, Brude Cal, Brude Urcal...* Надо отметить, что обозначенный в рукописи список 30 бруде, как и принято его называть, на самом деле состоит из 28 (29?) имен и, соответственно, 14, а не 15 пар (одна «пара» считается выпущенной при переписывании). Время правления этих королей – 150 лет, т.е. на каждого приходится в среднем 5 лет правления.

Такое странное расположение имен, естественно, всегда привлекало к себе внимание, и уже в 1940 г. Э.МакНил предположил, что «эта дубликация возникла, видимо, в результате удвоения прежнего списка из 15 королей, поскольку список завершается словами: *regnauerunt cl. annis ut diximus*. Список мог стать результатом деятельности наивного историка, для которого срок в 150 лет казался слишком длинным и он предположил, что их было вдвое больше и каждый правил в среднем пять лет, а не десять» [Macneill 1940, p. 12]. Значение префикса *ur-* он еще никак не объясняет, однако его главная идея – королей было не 30, а 15 – считается в настоящее время общепринятой.

Именно так рассматривается список бруде и в вышедшей недавно книге П.Дунбавина «Пикты и древние бритты», однако более чем настоятельное желание автора непременно увидеть в реликтах пиктского языка финно-угорский субстрат, приводит его к следующему объяснению: «Префикс *ur-* заставляет

вспомнить финское *uros* 'муж', а также *urho* 'герой' и *urhea* 'храбрый'; список в целом должен читаться следующим образом: *Бруде Гест, Бруде Герой-Гест, Бруде Понт, Бруде Герой-Понт* и т.д.» [Dunbavin 1998, p. 34].

Данная интерпретация, особенно на фоне системы доказательств в книге Дунбавина, наверное, не должна рассматриваться как серьезная. В настоящее время общепринятой является интерпретация, предложенная в 1983 г. Дж.Куком [Koch 1983] и принятая, что для нас важно, А.А.Королевым [Калыгин, Королев 1989, p. 200]. Согласно этой теории префикс *ur-* восходит к **wer-* 'после, на, через' (ср. ирл. *for*), и список «бруде» в целом должен быть перекомпанован следующим образом: *ur gest brude, pont brude, ur pont brude, leo brude* etc. – «после Геста короля Понт король, после Понта короля Лео король...». Имена собственные при этом, однако, предстают как неизменные, в чем Кук видел проявление ранней апокопы и упрощения падежной парадигмы.

ḡoṣ xij. ḡo. xv. fivaid xxij.
 ḡoṣolguitch. lxxv. denleam. c.
 olfriceta. lx. fuidid gaed hiecha
 ch. i. fivgurnach. xl. Wurgvistr
 xxx. **S**incedont. uq. xxx. hu
 de regnaimto hūlūma r allnā
 p centū. l. ānoꝝ ꝥaāū. xl. vna
 nis regū. i. hūte pnto. b. urpār
 b. leab. leo. b. uleo. b. gant. b.
 urgant. b. gūech. b. urguith. b.
 fecir. b. urfecir. b. cal. b. urcal. b.
 cuit. b. urcut. b. fec. b. urfec. b.
 ru. b. eru. b. gant. a. urgart. b. and
 b. uenid. b. iup. b. x. urup. b. grid
 b. irgrid. b. mud. b. urmud.

Glgrid. c. l. ān reg tharam
 c. moꝝleo. xv. deoalmid
 xl. amiood ānus arwis. vii. deo
 oꝝd. l. flicidhmad. v. deoꝝtēc. fi
 tm. xl. v. fambuel. xxx. amioꝝtr
 xl. deo arūmōs. xx. dūti. l. ru. c
 fācūnūthioꝝ a quo gartuāt. m
 regūa. vere. i. y. ā. reg. b. rech. āi.
 hūchū. vii. v. iup. g. uanec. xxx
 āi reg. samūadachama. iij. āi.

рис. 4

Данная интерпретация выглядит очень убедительной, однако сам текст перечня королей, если взглянуть в него пристально, заставляет все же в ней усомниться. В так наз. рукописи А, наиболее полной версии «Пиктской хроники» он дословно, если «вчитаться» в факсимильное воспроизведение соответствующей страницы (см рис. 4), выглядит следующим образом:

i. brude pant. b. urpant / b. leab.
 leo (?). b. uleo. b. gant b / urgant. b. gui-
 th. b. urguith. b / fecir. b. urfecir. b. cal.
 b. urcal. b / cuit. b. urcut. b. fec. b. ur-
 fec. b / ru. b. eru. b. gart. brugart. b.
 cinid / burcinid. b iup. bruriup. b grid /
 b irgrid. b. mud. b. urmud [Wainwright
 1955, p. 96].

Т.е. сам характер словоделения, точки-маркеры, а также в целом расположение имен в списке и сам факт, что слово *brude* открывает его, а не замыкает, показывает, что концепция Дж. Кука не соответствует, по крайней мере – позиции фиксатора

данного текста (отчасти, отметим это понимал и он сам, полагая, что изначальный текст списка мог быть устным «распевом»). Но уместно ли говорить о том, что «было на самом деле», если речь идет о мифических правителях?

Означает ли это, что мы склонны вернуться к идее, что этих правителей-бруде было все-таки 30 (28)? Наверное, нет. Идея увидеть в нем нечто совершенное новое была подсказана нам совершенно случайно, когда в тексте русского перевода саги «Видение Фингена» мы вдруг наткнулись на сходную пару имен: «...пал [...] от руки Голла и Ирголла»¹. В примечании к этому месту С.В.Шкунаевым сказано скупно: «Голл и Ирголл – вожди фоморов» [Шкунаев 1991: 266].

Мы не склонны, конечно, прямо отождествлять пиктов с фоморами, легендарным населением Ирландии, обитавшем, согласно преданию, на острове до его заселения другими племенами, но не содержится ли в списке их правителей аналогичный намек на достаточно распространенный архаический институт *парного правления*?

Примечания

¹ В оригинале: *dorochair [...] for Goll 7 Irgull* [AF, p. 330].

Литература

- Калыгин В.П., Королев А.А. Введение в кельтскую филологию. М., 1989.
- Михайлова Т.А. К проблеме реконструкции языка пиктов (предварительные замечания) // Проблемы изучения дальнего родства языков на рубеже третьего тысячелетия. Доклады и тезисы международной конференции. М., РГГУ, 2000.
- Шкунаев 1991 – Предания и мифы средневековой Ирландии / Пер. и прим. С.В.Шкунаева. М., 1991.
- AF – *Airne Fingein* / Ed. J.Vendryes. Dublin, 1953.
- Clancy T. The Drosten Stone: a new reading // *Proceedings of the Society of Antiquaries of Scotland*. Vol. 123, 1993.
- Dunbavin P. *Picts and Ancient Britons. An Exploration of Pictish Origins*. Oxford, 1998.
- Forsyth K. The Ogam-inscribed spindle-whorl from Buckoy: evidence for the Irish language in pre-Viking Orkney? // *Proceedings of the Society of Antiquaries of Scotland*. Vol. 125, 1995.
- Forsyth K. *Language in Pictland. The case against 'non-Indo-European Pictish'*. Utrecht, 1997.

T. A. Михайлова

Jackson K. The Pictish Language // The Problem of the Picts. New York, 1955.

Jackson A. The Symbol stones of Scotland. Orkney Press, 1984.

Koch J. The loss of final syllables in Brittonic // BBCS, vol. 30, 1983.

Macalister R.A.S. The Inscriptions and Language of the Picts // Essays and Studies presented to Professor Eoin MacNeill. Dublin, 1940.

Macalister R.A.S. The Archaeology of Ireland. London, 1996 (1949).

MacNeill E. The language of the Picts // Yorkshire Celtic Studies. Vol. II. 1940.

Wainwright F.T. The Picts and the problem // The Problem of the Picts. New York, 1955.