

ГЛАГОЛЬНО-ИМЕННЫЕ ТРАНСФОРМАЦИИ ПРИ АНГЛО-РУССКОМ МАШИННОМ ПЕРЕВОДЕ

VERBAL AND NOMINAL TRANSFORMATIONS IN THE ENGLISH-RUSSIAN MACHINE TRANSLATION

Козеренко Е.Б. (kozerenko@mail.ru), Институт проблем информатики РАН

В данной работе рассматривается проблема создания формализованных лингвистических представлений глагольно-именных трансформаций для англо-русского и русско-английского машинного перевода. Исследуются соотношения признаков «глагольности» и «номинативности» в языковых структурах английского и русского научного дискурса. Основой системы формализованных представлений языковых объектов в лингвистических процессах является когнитивная трансферная грамматика (КТГ), разработанная для систем машинного перевода.

1. Введение

Данная работа посвящена важному аспекту разработки систем машинного перевода, связанному с необходимостью моделирования синтаксических и семантических преобразований, неизбежно происходящих в процессе перевода с английского языка на русский и с русского языка на английский. Глагольно-именные трансформации – центральный вопрос формирования структуры переводного высказывания. Новое содержание проблеме языковых трансформаций придают современные реалии: необходимость проектировать и развивать обучающие компоненты систем машинного перевода и обработки текстовых знаний на основе уже существующих и вновь создающихся корпусов параллельных текстов. Значительные вычислительные ресурсы современных систем позволяют накапливать и использовать ранее переведенные текстовые фрагменты, эффективно поддерживать компоненту «переводческой памяти», обеспечивать машинный перевод, основанный на прецедентах [1,2].

Наши исследования в значительной степени базируются на концепциях функционального подхода, который был применен нами для многоязычной ситуации. При разработке лингвистического процессора, обеспечивающего англо-русский и русско-английский трансфер, нами было предложено понятие полей функционального переноса (ПФП), явившихся базисом сегментации языковых структур для решения задач машинного перевода. Основная идея такого подхода состоит в принятии гипотезы о том, что в основе грамматических структур лежат структуры когнитивные (ментальные фреймы). Поле функционально-семантического переноса отражает взаимодействие элементов разных языковых уровней в многоязычной ситуации [3]. В результате формализации описания синтаксических и семантических свойств полей функционального переноса и установления межъязыковых соответствий (вначале для языковой пары русский-английский) была разработана грамматика когнитивного переноса, или когнитивная трансферная грамматика (КТГ)[4].

Ориентация нашей работы на функционально-семантические категории, соотнесенные с системой грамматических единиц и их окружением, продолжает традицию рассмотрения языковых объектов грамматическими теориями, использующими содержательную категоризацию. В отечественной лингвистике это концепции А.А. Потебни, А.А. Шахматова, И.И. Мещанинова, Л.В. Щербы, В.В. Виноградова, работы А.В. Бондарко, В.С. Храковского, М.А. Шелякина, Г.А. Золотовой. Близкие и по сей день актуальные идеи были выдвинуты в работах Е. Куриловича, Г. Гийома, Э. Бенвениста. Грамматическая теория А.А. Потебни идет от мыслительного содержания к его языковой категоризации, связанной с формой как способом представления «внеязычного содержания». Функциональная направленность концепции А.А. Потебни определяла понимание и грамматической формы [5].

Грамматические категории в трактовке А.А. Шахматова во многом походили на категории семантические. Так, например, включая в сферу грамматических категорий средства, имеющие разную структуру, он относил к категории повелительного наклонения не только соответствующие глагольные формы, но и восклицания типа *молчать!*, *стоп!*, *пошел вон!*, *цыц!* [6]. Разработка понятия функции, являющегося центральным в функциональной грамматике, связана с широкой проблематикой функций языка [7,8,9,10,11,12,13,14,15]. Функции связаны со значениями языковых единиц, но они не тождественны им. Исследование функции некоторой языковой формы

включает анализ ее значения (или ряда значений в случае многозначности).

На современном этапе лингвистических исследований и разработок необходимо синергетическое сочетание функционального и уровневого подходов. Функциональный подход интегрирует языковые средства (синтаксические, лексические, словообразовательные и словоизменительные), принадлежащие разным уровням языка, на основе их функционально-семантических характеристик.

2. Функция и категория

Отношение между понятиями *категория* и *функция* можно иначе выразить как *исходное структурное значение* и *реализуемое структурное значение*. Подобная концепция была высказана Леонардом Блумфилдом в его работе «Язык».

Так, с точки зрения Л. Блумфилда, каждая лексическая форма связана с грамматическими формами в двух направлениях. С одной стороны, лексическая форма, даже когда она взята сама по себе, абстрагировано, обнаруживает значимую грамматическую *структуру*. С другой стороны, лексическая форма в любом конкретном высказывании, являясь особой языковой формой, всегда сопровождается той или иной грамматической формой. Она выступает в определенной функции, и случаи в которых преимущественно данная лексическая форма встречается, составляют в совокупности ее грамматическую функцию. Лексические формы, выполняющие какие-либо общие функции, принадлежат к одному формальному классу. На основе различных функций могут возникать частично совпадающие формальные классы. Так, выполнение функции действующего лица характерно для субстантивных выражений и для типично инфинитивных словосочетаний (*to scold the boys would be foolish* «бранить мальчиков было бы глупо» [16]).

Здесь Блумфилд называет функцией действующего лица субъектную функцию, тем самым подчеркивая ее «именной» характер. То есть функциональная потенция «быть субъектом предложения» исходно заложена в категории имени существительного. Глагольная же единица (в данном случае инфинитив), исполняя функцию, свойственную имени, как бы «надевает» на себя новую – именную форму – поверх своей исходной, глагольной. Таким образом, получается *композиционная категория* Глагол-Имя, или VerbNoun в нотации КТГ.

В работах А.В. Бондарко функциональная грамматика трактуется как грамматика, направленная на изучение и описание функций единиц строя языка и закономерностей функционирования этих единиц во взаимодействии с разноуровневыми элементами окружающей среды. Грамматика данного типа рассматривает в единой системе средства, относящиеся к разным языковым уровням, но объединенные на основе их семантических функций; при описании языкового материала используется подход «от семантики к ее формальному выражению» («от функций к средствам») как основной, определяющий построение грамматики, в сочетании с подходом «от формы к семантике» («от средств к функции»). Под единицами строя языка подразумеваются, прежде всего, грамматические формы слова и синтаксические конструкции, а также единицы «строевой лексики» (по Л.В. Щербе)[17]: модальные и фразовые глаголы, слова типа *вчера, обычно, часто, прежде, долго* и т.п.

Для описания функционально-семантической динамики языка Ю.С. Степанов [9] вводит понятие *функтора* как языкового средства транспозиции одного множества языковых единиц в другое множество языковых единиц того же языка. Функция есть свойство или значение функтора. Приводится пример функции: если принять за исходное множество единиц русские глаголы типа *сообщать, выразить, исполнять* и т.п., а за производное множество единиц *сообщение, выражение, исполнение* и нечто *сообщается, исполняется*, то отношение первого множества ко второму будет функцией, а языковые формы *–ение, нечто – (ает)ся* будут языковыми средствами этой функции, функторами.

Преобразования лексем с полной заменой означающего, например, *врач - лечить* (вместо *врачевать*). - это перифразы, по мнению Ю.С. Степанова. Под трансформациями понимаются, прежде всего, преобразования предикаторов в имена и имен в предикаторы: *бежать – бег, учитель – учительствовать*, при этом сохраняется частичное тождество формы – корень или основа слова и определенное тождество семантики. Трансформации постоянно выступают как одно из двух главных средств – наряду с перифразами – создания высказываний. В фокусе нашего внимания в данной работе находятся именно те языковые явления, которые Ю.С. Степанов называет трансформациями – мы рассматриваем имена и предикаты, сохраняющие частичное тождество формы. Это личные глагольные формы, инфинитивы, отглагольные существительные и герундии¹.

Понятие функции является одним из центральных в коммуникативной грамматике Г.А. Золотовой. Функция – это предназначенность элемента к определенному способу существования в системе, к определенно-

¹ Термин герундий мы употребляем в смысле английской грамматики: герундий это фактически отглагольное существительное, сохраняющее, однако, больше признаков «глагольности», чем собственно отглагольное существительное, прежде всего, это свойство иметь прямое дополнение, например *driving a car*.

му служению этой системе [18]. Если за целое принимаем предложение в его коммуникативном назначении, то функции его элементов, его составных частей определяются как их строительные, комбинаторные потенции, реализуемые в построении предложения.

Определение предикативной конструкции в коммуникативной грамматике: предикативной считают синтаксическую конструкцию, соотносящую имя признака (в широком смысле) с именем его носителя в категориях времени, модальности и лица. В таком понимании, восходящем к В.В. Виноградову, предикативность рассматривается как общее для любого предложения грамматико-семантическое значение данной единицы синтаксического уровня [19,20].

На предикатные потенции исходно номинативных языковых единиц указывал Ю.Д. Апресян, отмечая, что предметные существительные в ряде случаев могут быть эквивалентны по значению пропозиции. [21]: *После нас хоть потоп., Мы остались дома из-за дождя.* При переводе подобных конструкций с русского языка на английских в некоторых случаях требуется восстановление эксплицитно выраженного глагольного предиката (особенно это важно при создании систем перевода на основе языка-посредника и в системах искусственного интеллекта, использующих глубинно-семантические представления). Однако при переводе методом трансфера можно получить понятный, хотя и не всегда грамматичный перевод, поскольку механизм «свертки предикатов» часто используется и в других языках, включая английский.

Отсутствие полного совпадения между английскими и русскими языковыми конструкциями в научно-технических текстах можно обнаружить при изучении сравнительной частоты употребления в них отдельных частей речи, что важно для построения систем перевода, включающих машинное обучение.

Для рассмотрения семантики способов конфигурирования языковых структур мы пользуемся понятием структурного знака [22], предложенным в семиотической лингвистике С.К. Шаумяном. При этом слово также рассматривается нами не с точки зрения его лексического значения, а как функционально-категориальная единица, минимальный структурный знак. Такой подход принимается нами как определенный этап исследований структурных знаков, продиктованный необходимостью максимально полного извлечения семантической информации из возможных способов конфигурирования языковых объектов и изучения когнитивных механизмов линеаризации языковых структур. На когнитивный характер способов линеаризации структур, используемых естественным языком, было указано в работах Н.И. Лауфер [23].

Семиотическая лингвистика вводит понятие *суперпозиции функций*, полагая, что каждый языковой объект обладает исходной *первичной* функцией, а происходящие в действующем языке сдвиги значений – это наложение вторичной и других функций на исходную. Таким образом, использование инструмента суперпозиции [20] категорий дает возможность выразить функциональные свойства языковых объектов. В наших формально-грамматических представлениях категориально-функциональных свойств языковых объектов (элементов и групп) мы широко используем суперпозицию.

3. Изофункциональные трансформации при переводе имен и предикатов

Функционально-семантический подход, исследующий отношения «функциональной синонимии» разнородных и разноуровневых единиц языка, чрезвычайно актуален в настоящий момент, когда проводятся эксперименты по выявлению изофункциональных и изосемичных языковых структур из параллельных текстовых корпусов. Именно этот подход позволяет найти соответствия в текстах на разных языках. В самом деле, заранее нельзя с полной достоверностью определить, каким именно образом была переведена та или иная языковая структура в текстовом корпусе. Поэтому необходимо строить и исследовать различные гипотезы при проектировании лингвистического процессора. Функции реализуются при взаимодействии языковых объектов и их контекстов.

В данной статье мы рассматриваем трансформации имен и предикатов и исследуем средства языковой синонимии в двуязычной ситуации для языковой пары русский – английский.

3.1. Трансформации глагол → имя

Для научного изложения в целом характерен признак номинативности, т.е. более широкое использование существительных, чем в других функциональных стилях. При этом сопоставительный анализ переводов показывает, что в русском языке эта тенденция выражена более четко, и при переводе английские глаголы нередко заменяются существительными. Проведенные нами статистические исследования параллельных текстов позволяют сделать вывод о том, что русский текст приблизительно на 35% более номинативен, чем английский. Рассмотрим следующие примеры глагольно-именных трансформаций при англо-русском переводе.

(1) *The engine is the source of power that makes the wheels go round and the car move.*

Двигатель служит источником энергии для вращения колес и движения автомобиля.

go round → вращения

(2) *The fuel system is designed to store liquid gasoline and to deliver it to the engine cylinders in the form of vapor mixed with air.*

Система питания предназначается для заправки жидким топливом и подачи его в цилиндры в виде смеси паров бензина с воздухом.

to store and to deliver → для заправки и подачи

(3) *A similar approach has marked the EU's efforts to expand the current club of 15 countries to embrace former communist countries further east.*

Точно таким же подходом характеризуются усилия ЕС по расширению нынешнего клуба 15 стран дальше на восток путем присоединения к нему бывших коммунистических стран.

to embrace → по расширению

Нами были проведены исследования на материале имеющихся в нашем распоряжении параллельных переводов научных статей и отдельно взятых примеров высказываний с исследуемыми конструкциями, а также мы обращались к опросу экспертов-переводчиков. Наиболее продуктивные типы глагольно-именных трансформаций при англо-русском переводе коррелируют со следующими функциональными значениями.

- Обстоятельства цели и следствия, выраженные инфинитивом (58% в письменных переводах и 71% при опросе респондентов – профессиональных переводчиков).

(4) *In order to understand the phenomenon one should consider the laws of motion.*

Для понимания этого явления надо рассмотреть законы движения.

In order to understand → Для понимания

(5) *To explain the pressure of gases and their diffusion, it is generally supposed that all substances are made of very minute particles called molecules.*

Для объяснения давления газов и их диффузии обычно предполагают, что все эти вещества состоят из очень мелких частиц, называемых молекулами.

To explain → Для объяснения

- Составное сказуемое с инфинитивом (be + инфинитив) (51% в письменных переводах 59% при опросе респондентов – профессиональных переводчиков).

(6) *The method in use is to smash up the simplest nuclei.*

Используется метод расщепления простейших ядер.

to smash up расщепления

(7) *The difficulty will be to obtain the substance in question.*

Трудность будет состоять в получении рассматриваемого вещества.

to obtain → в получении

- Инфинитив после относительных местоимений *which* и *whom* с предшествующим предлогом часто переводится отглагольным существительным с предлогом *для*; в этом случае относительное местоимение с предлогом не переводится (48% в письменных переводах и 52% при опросе респондентов – профессиональных переводчиков).

(8) *In vacuum molecules have large space in which to move.*

В вакууме молекулы имеют большое пространство для движения.

in which to move → для движения

- Адъективные трансформации инфинитива (существительное + инфинитив в определительной функции) (практически 100% в обоих случаях).

(9) *The amount of polonium to be obtained from a uranium mineral can be simply calculated.*

Количество полония, которое должно быть получено из урана, можно довольно просто подсчитать.

to be obtained → которое должно быть получено из урана

(10) *Polonium was the first of the radioactive elements to be isolated by the chemists.*

Полоний был первым из радиоактивных элементов, выделенных химиками.

to be isolated → выделенных

- Инфинитив в функции второго дополнения (глаголы *cause, get, lead, make* + инфинитив) (42% в письменных переводах и 58% при опросе респондентов – профессиональных переводчиков).

(11) *Pressure causes ice to melt.*

Давление приводит к таянию снега.

to melt → к таянию

(12) *The force caused the body to move.*

Эта сила вызвала движение тела.

to move → движение

- Инфинитив, стоящий в начале предложения и выполняющий функцию подлежащего (31% в письменных переводах, 44% при опросе респондентов – профессиональных переводчиков).

(13) *To define exactly what is meant by the total heat in a body is at present still not possible.*

Точное определение того, что имеется в виду под общим нагревом тела, в настоящий момент все еще невозможно.

To define exactly → Точное определение

- Оборот “for + существительное + инфинитив” выполняет функции различных членов предложения (в научной литературе чаще всего функции обстоятельства цели или следствия) – 32% в письменных переводах и 43% при опросе респондентов – профессиональных переводчиков.

(14) *The temperature was too low for the substance to decompose.*

Температура была слишком низкой для разложения этого вещества.

to decompose → для разложения

3.2. Трансформации имя → глагол

Перевод английского герундия (the Gerund) на русский язык вызывает затруднения, связанные с его двойственной природой - это неличная форма глагола, выполняющая в предложении функции, исконно присущие имени существительному: подлежащего, дополнения; а также функции определения и обстоятельства, свойственные, соответственно, имени прилагательному и наречию. Морфологически герундий совпадает с действительным причастием английского языка (the Participle I, i.e. the Participle Present), которое также может играть роль определения и обстоятельства, но не может быть ни подлежащим, ни дополнением в предложении. Эта ситуация является неиссякаемым источником ошибок при переводе даже для человека-переводчика (это одна из центральных тем в курсе теории и практики перевода), а в существующих системах машинного перевода различие форм причастия и герундия вообще не происходит, отдельные виды конструкций реализованы лишь фрагментарно. Рассмотрим примеры герундия в возможных функциях.

- Герундий в функции подлежащего

(15) *Reading books is useful.*

Читать книги полезно.

- Герундий в функции прямого дополнения

(16) *I like reading.*

Я люблю читать.

- Герундий в функции предложного дополнения

(17) *He insisted on taking part in the conference.*

Он настаивал на участии (на том, чтобы принять участие} в конференции.

- Герундий в функции определения

(18) *The reading room was empty.*

Комната для чтения была пуста.

Читальный зал был пуст.

- Герундий в функции обстоятельства

(19) *On arriving in Moscow he called them immediately.*

По прибытии в Москву (прибыв в Москву) он незамедлительно им позвонил.

Как только он прибыл в Москву, он тут же им позвонил.

Действие или состояние, выраженное герундием, может иметь свой субъект. В этом случае перед герун-

дием всегда стоит: существительное в притяжательном падеже или притяжательное местоимение (*his arriving, John's arriving*), или существительное в общем падеже, которые и являются субъектом действия либо состояния, выраженного герундием. Такие герундиальные обороты переводятся придаточным предложением с союзом *что* (*чтобы*) с предшествующим ему местоимением *to*. Герундий передается личной формой глагола в функции сказуемого этого придаточного предложения, а стоящее перед ним существительное или местоимение — существительным (местоимением) в именительном падеже в функции подлежащего. Переводить герундий отглагольным существительным далеко не всегда возможно, поэтому при переводе на русский язык самым частотным вариантом будет личная форма глагола в придаточном предложении (69% в письменных переводах, 64% при опросе респондентов – профессиональных переводчиков).

(20) *He insisted on my being examined by a physician.*

Он настаивал на том, чтобы меня осмотрел врач.

(21) *These substances are alike in having high melting points.*

Эти вещества сходны тем, что они имеют высокие точки плавления.

Система правил трансфера для англо-русского машинного перевода вначале была построена по принципу одновариантных правил, когда переводное соответствие подбиралось как наиболее широкий способ перевода некоторой конструкции, пусть не всегда совершенно грамматичный, однако же обеспечивающий «понятность» перевода в наибольшем числе случаев [4]. При этом подходе предпочтение отдавалось всегда тому варианту, который был по форме ближе всего к исходной английской конструкции: для того, чтобы избежать трансформаций при переводе, которые всегда приводят к появлению «шумов» и резкому увеличению вычислительных затрат и, соответственно, программистских усилий. Однако в настоящее время нами разработана система многовариантных правил трансфера фразовых структур, при этом функциональные значения языковых единиц отражены в расширенной системе категориально-функциональных атрибутов. Языковые структуры представлены в виде иерархии правил когнитивной трансферной грамматики, которая является разновидностью унификационно-порождающей грамматики. Структуры атрибутов и значений и их преобразования задаются в виде контекстно-свободных и мягко контекстно-зависимых продукционных правил. Отношения зависимости реализуются через механизм головных вершин фразовых структур, а сами фразовые структуры задают линейные последовательности языковых объектов [25].

В новом варианте многовариантной когнитивной трансферной грамматики (МКТГ) отражено явление многозначности синтаксических структур и предусмотрены некоторые механизмы разрешения неоднозначности посредством включения в систему правил разбора и перевода статистической информации о возможных контекстах языковых структур. Продолжают формироваться обучающие наборы данных для расширения и модификации правил. Для сокращения числа избыточных правил (которые будут неизбежно возникать на этапе обучения), в систему закладываются не только шаблоны языковых структур, по аналогии с которыми будут порождаться правила, но и правила-фильтры на основе полей функционального переноса.

4. Многовариантный перевод и композиционные категории

Инструментом задания категориальных и функциональных значений языковых структур служат композиционные категории. Метод построения композиционных категорий позволяет реализовать суперпозицию категориальных значений языковых объектов для осуществления многовариантного перевода. Так *герундий* мы относим к категории «VerbNounIng», инфинитив в субъектной функции – к «toPlusInfinitiveSubj», инфинитив в объектной функции – к «toPlusInfinitiveObj», русское деепричастие и английское Participle I в адвербиальной функции – к категории «ParticipleAdv», личные формы глагола – к «VerbFinit», используются также и другие категории.

[Category: VerbNounIng]

asking questions

[Category: toPlusInfinitiveSubj]

She is known to be a skilled typist

[Category: toPlusInfinitiveObj]

We feel them to be sensitive readers

Разработаны многовариантные правила функционально-семантического переноса, посредством которых задаются наиболее вероятные способы перевода рассматриваемых языковых структур. Например, инфинитивные конструкции в функции обстоятельства цели/следствия могут переводиться следующим образом:

(20) *Hydrogen and oxygen unite to form water.*

Водород и кислород соединяются, образуя воду.

Водород и кислород соединяются и образуют воду.
Водород и кислород соединяются с образованием воды.

Схема многовариантного англо-русского трансфера инфинитивной конструкции *to form water* будет выглядеть следующим образом.

[Category: VerbInf] → {*to form water* }
OR {[Category: ParticipleAdv]; {*образуя воду* }
[Category: VerbFinit]; {*образуют воду* }
[Category: VerbNounIng]} {*с образованием воды* }

В следующем примере инфинитивный оборот соотнесен с ранее стоящим наречием *too* и выполняет функцию обстоятельства следствия.

(21) *The temperature was too low for the substance to decompose.*
Температура была слишком низкой для разложения этого вещества.
Температура была слишком низка, для того чтобы вещество разложилось.
Температура была слишком низка, для того чтобы вещество могло разложиться
Температура была слишком низка, и разложения вещества не произошло.

Схему трансфера инфинитивной конструкции можно представить в виде следующих правил.

[Category: VerbInf] →
OR {[Category: VerbNounIng];
[Category: VerbFinit];
[Category: VerbModalPhrase];
[Category: SentenceNeg]}.

При осуществлении анализа предложений и последующей свертки сегментов конструкций в узлы для трансфера, этим узлам будут присваиваться соответствующие метки (VerbNounIng, SentenceNeg, и т.д.), которые запускают необходимые процедуры трансформаций при переводе.

5. Исследование транскатегориальных переносов в существующих системах машинного перевода

Мы рассмотрели примеры перевода предложений с герундиальными и инфинитивными оборотами с английского языка на русский существующими машинными переводчиками. В большинстве систем доминирует подход выбирать эквивалент, наиболее близкий по форме к исходной языковой конструкции. Однако отчетливо прослеживается тенденция к включению трансформаций, даже если это не всегда пока еще улучшает качество перевода. Следует отметить, что имеет место большое сходство переводов в системах, основанных на совершенно различных принципах, что указывает на ситуацию интенсивной диффузии систем МП и различных подходов к лингвистическому моделированию. Рассмотрим примеры перевода предложений с герундиями в системах, которые в настоящее время лидируют в области англо-русского и русско-английского МП.

(22) *He could not help joining the discussion.*
Он не мог не принять участия в обсуждении.
Он не мог не выступить в прениях.

Переводы предложения (22) существующими системами машинного перевода:

Он не смог помочь соединить обсуждение. Babelfish
Он не смог помочь соединению дискуссии. ЭТАП
Он не мог сдержать присоединение к обсуждению ПроМт
Он не мог помочь присоединяясь к обсуждению. CognitiveTroll
Он не мог сдержать присоединение к обсуждению. SDL

(23) *Carrying out experiments is a must with every scientist.*
Проведение опытов необходимо для каждого ученого.
Проводить эксперименты необходимо каждому ученому.

Переводы предложения (23) существующими системами машинного перевода:

Проведение опытов - необходимость с каждым ученым. ЭТАП
Выполнение экспериментов - необходимость с каждым ученым. SDL
Вынося эксперименты является настоящей необходимостью с каждым ученым. Cognitive Troll
Носить из экспериментов будет сусло с каждым научным работником. Babelfish.

Далее мы поставили эксперимент, который заключался в том, чтобы исследовать поведение систем МП в ситуациях совпадения лексических значений конstituентов предложения при различии их категориально-функциональных значений. Для этой цели была использована фраза известного математика Н.Н. Константинова, которая постоянно приводится в брошюрах Московских математических олимпиад:

(24) Награждение наградами награжденных, не награжденных наградами на награждении, будет происходить по средам с 16.00 до 18.00 в комнате 202 МЦНМО.

Ниже приводятся переводы этой фразы различными системами МП:

Rewarding by awards награжденных, not награжденных awards on rewarding, will occur on Wednesdays from 16.00 till in a room 202 МЦНМО. (SDL).

Rewarding by awards награжденных, not награжденных awards on rewarding, will occur on Wednesdays from 16.00 till in a room 202 МЦНМО. (ПроМТ).

Rewarding awards of rewarded, not rewarded with awards on rewarding, will occur on wednesday with 16.00 to 18.00 in room 202 МЦНМО. (Cognitive Troll).

Rewarding with awards rewarded, not rewarded with awards on rewarding will happen in room 202 over the media since 16.00 before 18.00 MTSNMO. (ЭТАП).

Rewarding by the rewards of those rewarded, not rewarded with rewards on the rewarding, will occur on the media from 16.00 to 18.00 in the room 202 MTSNMO. (Babelfish).

6. Заключение

Актуальность проблемы моделирования трансформаций глагольных и именных конструкций для систем машинного перевода и извлечения знаний из текстов обусловлена тем, что до сих пор эти явления мало исследованы с точки зрения возможностей их компьютерных реализаций и, соответственно, недостаточно учтены в действующих системах МП. Настоятельная потребность в создании функционально-семантических представлений глагольно-именных трансформаций также вызвана тем, что дальнейшее развитие систем МП ведется с использованием машинного обучения на параллельных корпусах и правила, задающие функциональную синонимию языковых конструкций позволяют извлечь необходимую информацию и избежать формирования избыточных правил и «шумов». Для достижения этой исследовательской цели мы производим дальнейшее уточнение и расширений полей функционально-семантического переноса (ПФП) и развитие формализма когнитивной трансферной грамматики (КТГ) в сторону ее многовариантности (МКТГ) для учета многозначности синтаксических конструкций. Для построения категориально-функциональных представлений мы вводим композиционные категории, например VerbNounIng, VerbInf, ParticipleAdv и им подобных. При формировании процедур снятия неоднозначности мы используем статистические данные, полученные при изучении параллельных текстов научных и патентных документов нашего экспериментального корпуса. Так, при ранжировании предпочтительности вариантов трансфера нами используется статистика соотношения «глагольности-номинативности» русского и английского научного дискурса.

Дальнейшие наши исследования связаны с расширением числа типов трансформаций в англо-русском и русско-английском переводе и построением лингвистических представлений для многоязычной ситуации.

Список литературы

1. Козеренко Е.Б. Логико-статистические методы представления языковых структур в машинном переводе. // Труды Международной конференции Диалог'2005 "Компьютерная лингвистика и интеллектуальные технологии", М.: "Наука", 2005.
2. Knight, K. Automating knowledge acquisition for machine translation. AI Magazine 18, pp.81-96, 1997.
3. Козеренко Е.Б. Моделирование переноса функциональных значений для англо-русского машинного перевода. // Труды Международной конференции Диалог'2004 "Компьютерная лингвистика и интеллектуальные технологии", М.: "Наука", 2004.
4. Kozerenko, E.B. Cognitive Approach to Language Structure Segmentation for Machine Translation Algorithms // Proceedings of the International Conference on Machine Learning, Models, Technologies and Applications, June, 23-26, 2003, Las Vegas, USA.// CSREA Press, pp. 49-55, 2003.
5. Потебня А.А. Из записок по русской грамматике. М., 1958. Т. I-II. 536 с.
6. Шахматов А.А. Синтаксис русского языка. 3-е изд. М.:Эдиториал УРСС, 2001. 624 с.
7. Якобсон Р.О. Разработка целевой модели языка в европейской лингвистике в период между двумя войнами // Новое в лингвистике. М., 1965. Вып. 4. С. 372-377.

8. Якобсон Р.О. Шифтеры, глагольные категории и русский глагол. // Принципы типологического анализа языков различного строя. М., 1972.
9. Степанов Ю.С. Имена, предикаты, предложения (семиологическая грамматика). М.: Едиториал УРСС, 2004.
10. Halliday M.A.K. An Introduction to Functional Grammar. / Ed. by Edward Arnold, London, 1985.
11. Halliday M.A.K. System and Function in Language // Halliday M.A.K. Selected Papers, London, 1976. 250 p.
12. Слюсарева Н.А. Проблемы функционального синтаксиса современного английского языка. М., 1981, 206с.
13. Шведова Н.Ю. Один из возможных путей построения функциональной грамматики русского языка // Проблемы функциональной грамматики. М., 1985. С. 30-37.
14. Звегинцев В.А. Функция и цель в лингвистической теории // Проблемы теоретической и экспериментальной лингвистики. М., 1977. С. 120-146.
15. Гак В.Г. К типологии функциональных подходов к изучению языка // Проблемы функциональной грамматики. М., 1985. С. 5-15.
16. Блумфилд Л. Язык. Изд. 2-е стереотипное. – М.: Едиториал УРСС, 2002. 608 с.
17. Щерба Л.В. Языковая система и речевая деятельность. Л., 1974. 428 с.
18. Золотова Г.А., Ониненко Н.К., Сидорова М.Ю. Коммуникативная грамматика русского языка. М., 2004.
19. Виноградов В.В. Избранные труды: Исследования по русской грамматике. М., 1975.
20. Виноградов В.В. Русский язык: (Грамматическое учение о слове). 2-е изд. М., 1972. 614 с.
21. Ю.Д. Апресян. Синтаксис и семантика в синтаксическом описании. «Единицы разных уровней грамматического строя языка и их взаимодействие». М., 1969, стр. 304-306.
22. Шаумян С.К. Семиотическая Лингвистика как Объяснительная Наука. // Труды Международной конференции Диалог'2005 «Компьютерная лингвистика и интеллектуальные технологии», М.: “Наука”, 2005, с. 507-513.
23. Лауфер Н.И. Линеаризация компонентов сочинительных конструкций. В сб. «Моделирование языковой деятельности в интеллектуальных системах». Ред. Кибрик А.Е., Нариньяни А.С. М: «Наука», с. 167-176, 1987.
24. Shaumyan, S. A Semiotic Theory of Language. Indiana University Press, 1987.
25. Козеренко Е.Б. Лингвистические аспекты информатики // Системы и средства информатики. Специальный выпуск Научно-методологические проблемы информатики. Москва, ИПИРАН, 2006, 88-111.