4
П-71
47

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР

Проблемная группа по экспериментальной
и прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 47

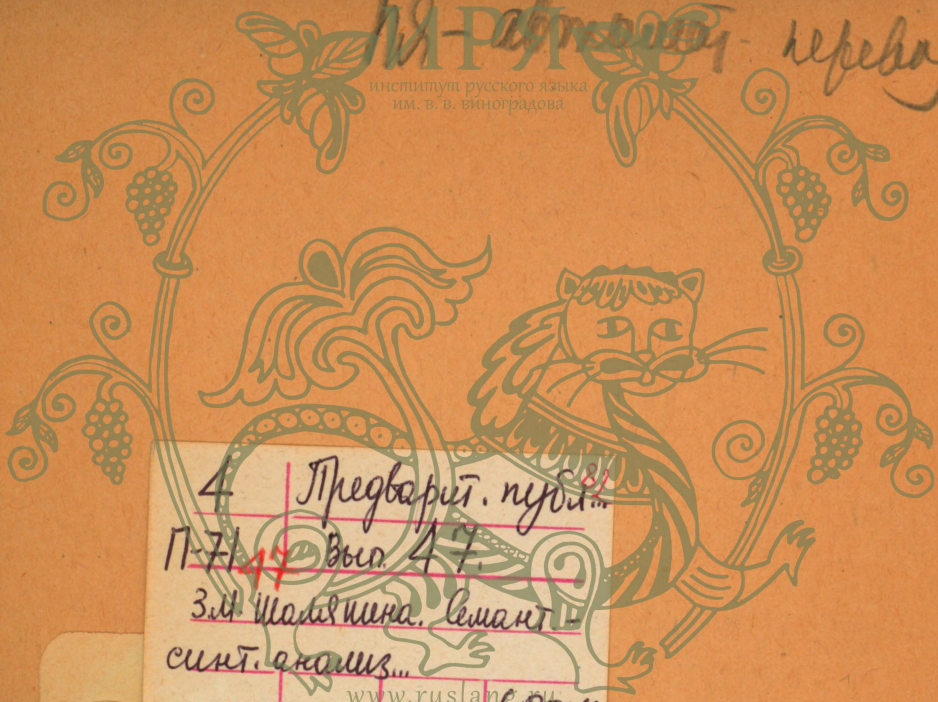
З. М. ШАЛЯПИНА

СЕМАНТИКО-СИНТАКСИЧЕСКИЙ АНАЛИЗ В СИСТЕМЕ АНГЛО-РУССКОГО АВТОМАТИЧЕСКОГО ПЕРЕВОДА (АРАП)

1. Общие принципы анализа и описание ручного эксперимента.
2. Материалы анализа фразы:
   This is termed amplitude-modulated c-w Jamming.

Москва — 1974

ИРЯЗ
институт русского языка
им. в. в. виноградова



4 Предварит. публ.
П-71 Вып. 47.
З.М. Шалягина. Семант.-
синт. анализ...
68007
12 л. Рисунок

www.ruslang.ru

68007

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР

Проблемная группа по экспериментальной
и прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 47

З.М.Шаляпина

СЕМАНТИКО-СИНТАКСИЧЕСКИЙ АНАЛИЗ В СИСТЕМЕ АНГЛО-РУССКОГО АВТОМАТИЧЕСКОГО ПЕРЕВОДА (АРАП)

1. Общие принципы анализа и описание ручного эксперимента.
2. Материалы анализа фразы: This is termed amplitude-modulated c-w jamming.

Москва - 1974

4
П-71
17
БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР
68004

Работа над созданием системы англо-русского автоматического перевода (АРАП) начата сотрудниками Лаборатории машинного перевода МГПИИЯ им. М.Тореза в 1969 г. на основе программы разработки этой системы, предложенной О.С.Кулагиной и И.А.Мельчуком [7, 16]. До 1973 г. эта работа концентрировалась в первую очередь на составлении англо-русского многоаспектного автоматического словаря (АРМАС), являющегося, по замыслу авторов системы АРАП, одним из важнейших ее компонентов (см. об этом словаре [15]). В настоящее время одной из центральных задач становится описание английского синтаксиса и согласование этого описания с имеющимся фрагментом словаря, с тем чтобы обеспечить возможность использования готовых компонентов системы АРАП для автоматического анализа английских текстов. Работа над синтаксическим описанием английского языка ведется по двум основным направлениям. С одной стороны, близится к завершению работа над моделью английского поверхностного синтаксиса, составляемой И.А.Мельчуком и Н.В.Перцовым [9, 10]; с другой стороны, разрабатывается формальное синтактико-семантическое описание комбинированного типа, имеющее своей целью объединение поверхностно-синтаксического и глубинно-синтаксического анализа текста в один этап и существенно опирающееся на особенности структуры словарной статьи в АРМАС. Одним из этапов разработки этого описания и проверки готовых его фрагментов является ручной эксперимент по анализу отдельных английских фраз. Данная работа представляет собой предварительное сообщение об общем строении данного описания, а также о ходе эксперимента, проводимого на его базе.

Основная особенность разрабатываемой комбинированной семантико-синтаксической модели английского языка состоит в том, что она строится в форме набора отдельных описаний, каждое из которых имеет в принципе ту же структуру, что и статьи слов и фразеологических сочетаний в АРМАС (см. описание структуры словарной статьи в [15], а также перечень ее разделов и подразделов в Приложении к данной работе). Далее мы будем называть эти описания г р а м м а т и ч е с к и м и с л о в а р н ы м и с т а т ь я м и, а статьи слов и фразем, в отличие от них, - л е к с и ч е с к и м и,

В заглавии грамматических статей стоят не имена лексем, как в лексических статьях, а символы словообразовательных значений (*деривационные символы* – ДС) или грамматические признаки (см. ниже).

В грамматических статьях, озаглавленных деривационными символами (эти статьи естественно назвать *деривационными*, или *словообразовательными*), записываются все те сведения, которые необходимы для получения словарной статьи производного слова, образованного с применением данного словообразовательного значения, если имеется словарная статья того слова, от основы которого образовано данное производное. Так, в статье, озаглавленной ING1 (где ING1 – морфема английского отглагольного существительного), будет указано, что производная единица X, образованная с помощью данной морфемы, является, в отличие от исходной единицы Y, не глаголом, а существительным; что с точки зрения системы лексических функций, она представляет собой $S_0(Y)$; что ее субъектная валентность оформляется существительным с предлогом of или by и т.д.

Грамматические статьи, имеющие в заглавии грамматический признак (или конъюнкцию таких признаков) подразделяются, в зависимости от того, что за признаки стоят в их заглавии, на два подтипа: *обобщенно-синтаксические* и *морфологические* статьи.

В заглавии обобщенно-синтаксических статей указываются так называемые синтаксические признаки – символы частей речи и других синтаксических классов (V – глагол, S – существительное, A – прилагательное, Adv – наречие, Prep – предлог, Pron – местоимение, Aux – вспомогательный глагол и т.д.), используемые в данной синтаксической модели. Эти признаки[1] приписываются конкретным словам в р. ЗИ сопоставленных им словарных статей (лексических и грамматических), образуя так называемую *синтакти-*

[1] Используется, в основном, тот же список синтаксических признаков, который разработан И.А.Мельчуком и Н.В.Перцовым в рамках модели поверхностного синтаксиса английского языка.

к у этих слов. Обобщенно-синтаксическая статья, озаглавленная признаком (или конъюнкцией признаков) X, содержит все те сведения и правила, которые применимы при обработке любого слова C, если оно входит в синтаксический класс X (т.е. имеет этот признак в р. ЗИГ своей статьи). Например, в статье, озаглавленной A (прилагательное), указывается, что любое слово C данного класса может подчиняться существительному по связи слабого управления; что если управляющее существительное стоит справа от C, то между ними не могут стоять предлоги, подчинительные союзы или личные глаголы; что если C подчинено глаголу-связке, то слово, заполняющее субъектную валентность этого глагола, должно удовлетворять семантическим требованиям к первому месту модели управления C, записанным в его лексической словарной статье, и т.д.

Наконец, морфологические словарные статьи имеют в заглавии морфологические признаки, которые, согласно используемой модели морфологии[1], могут входить в характеристику словоформы в ее глубинно-морфологическом представлении (см. ниже). Информация, записываемая в таких статьях, характеризует уже не лексическую единицу как таковую, а ту форму, в которой она употреблена в тексте. Так, в статье, озаглавленной fin (где fin - признак личной формы глагола), отмечается, что субъектная валентность любого глагола C, стоящего в личной форме, может быть заполнена существительным; что наличие в тексте слова, заполняющего эту валентность, является обязательным, если только непосредственно перед C не стоит as, than или какое-либо другое слово из ограниченного списка (имеются в виду случаи типа ... as is well-known); что C может подчиняться существительному по связи слабого управления (определительные придаточные типа the man I saw); что такое существительное всегда стоит слева от C и т.д. Могут даваться и более сложные правила, требующие как-либо модифицировать информацию, имеющуюся в лексической статье слова с данным морфологическим признаком. Например, в статье $part_{pass}$ 1 (где $part_{pass}$ 1 - признак причастия, используемого при образовании английского пассива) указывается, что при обработке слова C, име-

[1] Эта модель целиком изложена в [13]; краткое ее описание см. в [2].

щего этот признак, следует игнорировать сведения о способах заполнения первых двух валентностей С, записанные в его лексической словарной статье.

Описанный способ организации грамматической информации позволяет предложить сравнительно простой (с точки зрения общей последовательности действий) алгоритм автоматического анализа английской фразы. Идея алгоритма состоит в том, что для каждой словоформы фразы рассматриваются все варианты ее глубинно-морфологического представления (с учетом лексической и лексико-грамматической омонимии) и для каждого из этих вариантов устанавливаются все синтаксические связи, которые предусмотрены моделью управления хотя бы в одной из словарных статей, описывающих эту единицу (т.е. в статье соответствующей лексемы - простой или производной - или в статьях ее синтаксических и морфологических признаков), и которые при этом не противоречат никаким семантико-синтаксическим ограничениям, зафиксированным в этих статьях. По окончании обработки всех словоформ фразы установленные связи объединяются в единое дерево зависимостей (или, в случае омонимии, в несколько разных деревьев), с учетом требований к правильной синтаксической структуре (см. [4]). Снятие всех видов омонимии, включая синтаксическую, не составляет особого этапа, а осуществляется в ходе выполнения других задач анализа. Никакой сегментации обрабатываемой фразы в процессе анализа не предусматривается (так что данный алгоритм относится к числу алгоритмов интегрального типа).

Прежде, чем более подробно рассматривать последовательность действий при анализе, определим некоторые символы, которые могут при этом потребоваться.

C_0 - рабочая словоформа в рассматриваемом ее значении, точнее - глубинно-морфологическое представление (ГМП) этой словоформы в данном значении.

ГМП словоформы в общем случае имеет вид[1]:

$$[\ldots[[Д(\Sigma_Д) + ДС_1(\Sigma_{ДС_1})] + ДС_2(\Sigma_{ДС_2})] + \ldots + ДС_к(\Sigma_{ДС_к})](\Sigma; Х),$$

[1] Данная формула построена на базе определения ГМП словоформы в [8] и полностью характеризованной лексемы в [3].

где Л – заглавие лексической словарной статьи, сопоставленной данной словоформе; $ДС_1$ – деривационный символ (символ одного из словообразовательных значений, с помощью которых было образовано рассматриваемое слово); Σ_X – синтактика единицы X (набор синтаксических признаков из р. III словарной статьи этой единицы); Σ – синтактика всей словоформы; χ – морфологическая характеристика словоформы (набор морфологических признаков, сопоставленных ей при анализе). В ГМП конкретной словоформы любой из компонентов $ДС_1$, $\Sigma_{ДС_1}$, Σ, χ может быть в принципе пустым. Заметим, что в связи с явлением словосложения в определении ГМП словоформы следовало бы предусмотреть возможность появления более чем одного символа лексической словарной статьи. В разрабатываемой системе АП, однако, это явление пока не учитывается и все сложные слова целиком задаются в словаре, что и позволяет нам ограничиться приведенным выше упрощенным определением.

Легко видеть, что ГМП словоформы целиком определяет множество сопоставленных ей лексических и грамматических словарных статей. Таким образом, символ $С_0$ может употребляться также как обозначение множества всех статей, сопоставленных рабочей словоформе.

$ДС_0$ – рассматриваемая деривационная словарная статья.

$Л_0$ – рассматриваемая лексическая словарная статья.

$Мф_0$ – рассматриваемая морфологическая словарная статья.

$П_0$ – рассматриваемая грамматическая словарная статья (статья морфологического или синтаксического признака).

$Синт_0$ – рассматриваемая обобщенно-синтаксическая словарная статья.

$ДС(С_0)$ – словарная статья основы словоформы $С_0$ (если эта основа является производной).

$Л(С_0)$ – множество лексических словарных статей, сопоставленных словоформе $С_0$. В нашей системе, где, как уже сказано, словосложение не рассматривается, множество $Л(С_0)$ совпадает с $Л_0$.

$Мф(С_0)$ – множество морфологических словарных статей, сопоставленных словоформе $С_0$ (= множество признаков, входящих в характеристику χ этой словоформы).

$П(С_0)$ - множество грамматических словарных статей, сопоставленных словоформе $С_0$ (= множество синтаксических и морфологических признаков $С_0$).

$Синт(С_0)$ - множество обобщенно-синтаксических словарных статей, сопоставленных $С_0$ (= множество признаков, входящих в синтактику $\sum$ этой словоформы).

$Синт(Л_1)$, где $Л_1$ - заглавие конкретной лексической статьи, - множество обобщенно-синтаксических статей, заглавия которых даны в р. ЗП статьи $Л_1$.

$Синт(Мф_1)$, где $Мф_1$ - заглавие конкретной морфологической статьи, - множество обобщенно-синтаксических статей, заглавия которых даны в р. ЗП статьи $Мф_1$.

$Синт(ДС_1)$, где $ДС_1$ - конкретный деривационный символ, - множество обобщенно-синтаксических статей, заглавия которых даны в р. ЗП статьи $ДС_1$.

$[X_1 + ДС_1]$ - словарная статья производной единицы, образованной от единицы X_1 с помощью словообразовательного значения $ДС_1$. Единица X_1 здесь может быть простой (представлять собой заглавие некоторой лексической словарной статьи $Л_1$) или производной (т.е. иметь вид $[X_j + ДС_j]$).

Перейдем теперь к описанию общей схемы процесса автоматического анализа в системе АРАП. Она может быть представлена в виде следующей последовательности шагов.

I. Построение сокращенного глубинно-морфологического представления анализируемой фразы.

Проводится морфологический анализ всех словоформ анализируемой фразы, и каждая словоформа заменяется получаемым для нее в результате этого анализа **сокращенным** ГМП или (в случае омонимии) дизъюнкцией всех возможных для нее сокращенных ГМП. Сокращенное ГМП отличается от полного тем, что в его составе отсутствует синтактика. Таким образом, общий вид такого ГМП можно представить формулой:

$$[\ldots [[Л + ДС_1] + ДС_2] \ldots + ДС_к] (Х),$$

где Л – исходная лексема, $ДС_i$ – деривационный символ, χ – набор морфологических признаков. Для словоформ, образованных от непроизводных (относительно данного словаря) основ, сокращенное ГМП имеет вид $Л(\chi)$.

Заметим, что уже на этапе построения сокращенного ГМП фразы могут быть разрешены некоторые виды омонимии. Например, при морфологическом анализе словоформы termed отбрасываются все те значения лексемы term, в которых она выступает как существительное.

2. Построение рабочего глубинно-морфологического представления фразы.

Рабочее ГМП фразы, во-первых, должно содержать эксплицитные указания на порядок расположения словоформ данной фразы в ее линейном представлении (поскольку обработка отдельных ГМП словоформ может вестись не в порядке следования этих словоформ в тексте) и, во-вторых, должно быть полным – т.е. содержать синтактику. Первое требование обеспечивается тем, что перед каждым из имеющихся ГМП словоформ ставится римская цифра, означающая порядковый номер соответствующей словоформы в данной фразе. Что касается синтактики, которую следует включать в состав того или иного ГМП словоформы, то ее выявление представляет собой более сложный процесс. Он включает следующие операции.

Пусть сокращенное ГМП рабочей словоформы $С_0$ состоит из имени лексемы $Л_0$, деривационных символов $ДС_1$, ..., $ДС_к$ и морфологических признаков $МФ_1$, ..., $МФ_n$. Словарные статьи всех этих единиц выбираются из словаря и образуют некоторое множество $Р_0$, которое мы будем называть рабочим корпусом словарных статей. Рассматривается раздел 311 этих статей и на основании записанных в нем данных строится синтактика рабочей словоформы с одновременным пополнением множества $Р_0$. При этом может быть несколько случаев, в зависимости от того, имеются ли в этом разделе хотя бы в одной из привлекаемых статей правила вида:

Если X то ВОЗМОЖНО Y ;
Если X то СНЯТЬ Усл (z) = Sbl (Y);
СНЯТЬ Усл (z) = Sbl (Y),

где X - некоторое высказывание, записанное на применяемом символическом языке ЯЗОП (см. [1], а также [15]), Y - синтаксический признак или их конъюнкция, Z - один из символов: $Л_0$, ДС(C_0), Мф(C_0), Синт(C_0), П(C_0), C_0.

Если таких правил нет, то все признаки, записанные в р.ЗП статей из рабочего корпуса, включаются в синтактику рассматриваемого ГМП словоформы, образуя в нем конъюнкцию. При этом полезно до окончания процесса построения синтактики C_0 сохранять сведения о том, из какой статьи взят тот или иной набор признаков, имеющийся в составе этой синтактики.

Если в р. ЗП некоторой статьи из P_0 есть правило вида Если X то ВОЗМОЖНО Y , то признаки из Y включаются в синтактику C_0 только при условии, что в нашем случае выполнено требование X .[1]

Если в какой-либо статье из P_0 есть правило вида Если X то СНЯТЬ Усл (z) = Sbl (Y) , то, как и в предыдущем случае, проверяется, выполнено ли требование X. Если оно не выполнено, то правило не учитывается; если же X выполнено, то те статьи из P_0, которые задаются выражением Z , обрабатываются не по общему методу (изложенному в двух предыдущих абзацах), а так же, как при обнаружении правила вида СНЯТЬ Усл(z) = (Y), - см. ниже.

Если в одной из привлекаемых статей найдено правило СНЯТЬ Усл (z) = Sbl (Y) , это значит, что статьи из P_0, задаваемые выражением Z , должны обрабатываться особым образом. А именно, из р. ЗП этих статей переносятся в синтактику C_0 только те признаки, которые не указаны в составе Y . В некоторых случаях ни один из признаков, записанных в р. ЗП статьи Z , не переносится

---

[1] Это требование, в зависимости от его конкретного вида, может предъявляться либо к составу анализируемой фразы, либо к составу словарных статей, входящих в рабочий корпус. В последнем случае оно содержит символы Усл (от русск. условие) и Sbl (от англ. symbol).

в синтактику рабочей словоформы. Это имеет место, если в какой-либо статье из P_0 найдено правило: СНЯТЬ Усл (z) = Sbl ($П_1$), где $П_i$ - переменная, означающая произвольный признак, - или же сокращенный вариант того же правила: СНЯТЬ Синт (z) . Если статья, к которой отсылает выражение z в составе правила СНЯТЬ Усл (z) = Sbl (Y) , уже обработана (т.е. записанные в ней синтаксические признаки полностью или частично перенесены в синтактику C_0), происходит обращение не к самой этой статье, а к тем признакам в составе имеющегося фрагмента синтактики, которые были получены из статьи z (выше отмечалось, что информация об источнике тех или иных признаков в синтактике C_0 сохраняется вплоть до окончания ее построения). При обнаружении среди них признаков, задаваемых выражением Y , эти признаки из синтактики C_0 убираются.

Каждый раз при занесении в синтактику рабочей словоформы некоторого синтаксического признака, обобщенно-синтаксическая статья, имеющая этот признак в своем заглавии, выбирается из словаря и включается в рабочий корпус P_0, после чего ее р. ЗП обрабатывается по изложенному выше способу.

Построение синтактики некоторого ГМП словоформы заканчивается тогда, когда ни в одной из имеющихся в рабочем корпусе статей в р. ЗП нет признаков, чьи статьи не были бы уже включены в P_0.

После того, как выполнены все операции, предусмотренные на данном этапе, каждое из ГМП словоформ анализируемой фразы приобретает вид:

$$N\,[\,\ldots\,[[Л + ДС_1] + ДС_2]\,\ldots + ДС_к]\ (\Sigma;\chi),$$

где N - порядковый номер словоформы во фразе, Σ - построенная описанным образом синтактика.

3. Построение рабочих словарных статей для всех ГМП словоформ анализируемой фразы.

Для каждого из ГМП словоформ рассматриваемой фразы на данном этапе строится рабочая словарная статья, имеющая точно такую же структуру, что и лексическая или грамматическая статья АРМАС. Единственная разница состоит в том, что поскольку рабочая статья

строится уже после того, как проведен морфологический анализ, в нее нет смысла включать морфологическую информацию, так что такая статья не содержит морфологической зоны (зона I).

Заглавием рабочей словарной статьи некоторого ГМП словоформы является само это ГМП; раздел 3I1 рабочей статьи представляет собой синтактику данного ГМП словоформы, полученную на предыдущем этапе. Что касается всех прочих разделов зон 2, 3 и 4, то их построение осуществляется вполне аналогично построению синтактики (= раздела 3I1). А именно, рассматривается тот же рабочий корпус P_0 словарных статей, который использовался для данного ГМП словоформы на предыдущем этапе (точнее, окончательный вариант P_0, полученный по завершению построения синтактики данной словоформы). Для каждого номера раздела N берутся все разделы с данным номером, имеющиеся в статьях рабочего корпуса, и записанные в них сведения переносятся в раздел N рабочей статьи по правилам, описанным в пункте 2 применительно к обработке р. 3I1 (с той разницей, что вместо конъюнкций признаков теперь рассматриваются конъюнкции высказываний, допустимых в рамках используемого символического языка ЯЗОП).

В некоторых уточнениях нуждается только способ применения правил вида СНЯТЬ Усл (z) = Sbl(Y) , где Y – любое высказывание, допустимое в ЯЗОП. При наличии такого правила в разделе N(≠ 3I1) некоторой статьи P_0 в одноименный раздел рабочей статьи не переносятся (из соответствующих разделов статей, заданных выражением z) не только те высказывания, которые имеют вид Y , но и те, которые имеют вид Если U то Y [1]. При этом, если выражение Y в составе правила СНЯТЬ Усл (z) = Sbl(Y) содержит переменные, то считается, что высказывание T имеет вид Y всегда, когда T совпадает с Y с точностью до этих переменных (вместо переменных в T могут стоять любые единицы из области их значений).

При необходимости исключить из рабочей статьи все высказывания определенного вида в некоторых разделах словарных статей могут применяться сокращенные варианты общего правила СНЯТЬ Усл(z)= = Sbl(Y) . В настоящее время такие сокращения встречаются только в разделах 32I и 322.

[1] Для раздела 3I1 это не оговаривалось в связи с тем, что в нем не встречаются высказывания такого вида.

В р. 321, содержащем модель управления рассматриваемой единицы, применяются сокращения:

СНЯТЬ W(z),

где W – одна из валентностей единицы z; запись соответствует развернутому правилу: СНЯТЬ Усл (z) = Sbl (W = X) – и означает требование не включать в рабочую статью никакие сведения о способах заполнения валентности W, имеющиеся в статье z;

СНЯТЬ МУ(z)

(соответствует конъюнкции СНЯТЬ Усл(z) = Sbl (W_1 = X) И СНЯТЬ Усл(z) = Sbl (W_2 = X) И ... И СНЯТЬ Усл(z) = Sbl(W_n = X), где n – число всех валентностей единицы z; означает требование вообще не включать в рабочую статью сведений о модели управления – МУ – единицы z).

В р. 322, содержащем семантические требования (СТ) к единицам, способным заполнять те или иные валентности модели управления, применяется сокращение:

СНЯТЬ СТ (W(z)).

Сокращение соответствует полному правилу СНЯТЬ Усл(z) = Sbl(W=X), где X – набор семантических признаков (конъюнкция или дизъюнкция).

4. Определение порядка обработки отдельных ПМ словоформ данной фразы при семантико-синтаксическом анализе.

Все имеющиеся рабочие ПМ словоформ располагаются в некотором порядке, соответствующем последовательности их обработки при выявлении синтаксической структуры анализируемой фразы. В проводимом в настоящее время ручном эксперименте ПМ словоформ рассматриваются в порядке следования этих словоформ во фразе, однако, в отличие от алгоритмов предсказуемостного типа, при нашем подходе допустима и любая другая последовательность (что создает возможность оптимизации в будущем предлагаемой процедуры обработки текста с привлечением статистических и других критериев).

5. Формулировка гипотез о синтаксических связях, реализующих в данной фразе валентности рабочей словоформы C_0.

Берется первое из ГМП словоформ, подлежащих обработке на данном этапе (C_0), и рассматривается раздел 321 его рабочей словарной статьи (раздел содержит перечень всех возможных для C_0 синтаксических валентностей, или модель управления C_0).

При наличии в разделе 321 правил вида $W = X$ или $W_j = X$ (где W – одна из валентностей C_0: M_1, G_m или G_c [1], – а X – конъюнкция синтаксических и/или морфологических признаков) во фразе ищутся такие ГМП словоформ C_h, в синтактике и характеристике которых имеются все признаки из X, без учета того, как расположены эти словоформы относительно C_0 и других словоформ фразы. При их обнаружении между ними и C_0 устанавливаются гипотетические синтаксические связи вида:

(1) $C_0 \xrightarrow{1} C_h$
(2) $C_h \xrightarrow{m} C_0$
(3) $C_h \xrightarrow{c} C_0$

в зависимости от характера валентности W. Далее мы будем говорить о таких связях, что они реализуют валентность M_1, G_m или G_c единицы C_0, соответственно. Заметим, что количество связей с одним и тем же индексом, которые можно на данном этапе установить с одной и той же словоформой, заранее ничем не ограничивается и зависит, следовательно, только от того, сколько найдется в данной фразе словоформ, имеющих все требуемые морфологические и синтаксические признаки.

Если некоторая гипотетическая связь устанавливается по правилу, в котором какие-то признаки стоят в угловых скобках, при записи этой связи соответствующие признаки в ГМП словоформы C_h также заключаются в угловые скобки (впоследствии заключенные в скобки единицы снимаются при переходе к глубинно-синтаксической структуре фразы).

---

[1] 1 – индекс связи сильного управления (одна из цифр 1, 2, 3,...), m – индекс связи слабого управления (от англ. modifying), c – индекс координативной связи (от англ. coordinate).

Если в р. 321 рабочей словарной статьи есть правило вида:

$$M1 = \langle \text{word} \xrightarrow{\alpha} \rangle X \quad \text{или}$$

$$M1 = X \langle \xrightarrow{\alpha} \text{word} \rangle ,$$

где word – некоторое конкретное слово (заглавие лексической словарной статьи), выступающее в данном случае в роли показателя сильного управления рабочей словоформы C_0, а α – индекс типа связи, то вначале проверяется, есть ли во фразе единица word и только при ее наличии ищутся единицы с признаками X; связи, реализующие в этом случае валентность M1, будут иметь вид, соответственно,

$$C_0 \xrightarrow{1} \langle \text{word} \xrightarrow{\alpha} \rangle C_h \quad \text{или}$$

$$C_0 \xrightarrow{1} C_h \langle \xrightarrow{\alpha} \text{word} \rangle .$$

По правилам вида $M1 = \langle \text{word} \xrightarrow{\alpha} \rangle X$ предусматривается также устанавливать "усеченную" связь:

$$C_0 \xrightarrow{1} \langle \text{word} \xrightarrow{\alpha} \rangle$$

(независимо от того, есть ли в данной фразе единицы с признаками X). Такие связи оказываются необходимыми при анализе некоторых случаев "разрешенной ненасыщенности" (см. [4], стр. 30-37), напр., определительных оборотов типа the book I spoke of, где предлог of (являющийся показателем сильного управления глагола speak) не имеет подчиненного существительного.

6. Разрешение омонимии на основе сведений об обязательности реализации валентностей C_0.

Рассматривается раздел 323 рабочей статьи C_0 и в нем ищутся правила вида:

Есть W;

Если Y то есть W.

Для каждого из таких правил проверяется, есть ли среди установленных согласно п. 5 гипотетических связей такие, которые реализуют указанную в этом правиле валентность. Если таких связей нет, а условие Y (если оно имеется в правиле) выполнено, обрабатываемое ПМП словоформы C_0 отбрасывается (как нереализуемое в контексте данной фразы) и все ранее установленные с C_0 гипотетические

связи снимаются[1].

Далее проверяется, нет ли в данном разделе правила вида:

Нет W ;

Если Y то нет W

и при их наличии выясняется, были ли установлены гипотетические связи, реализующие валентность W ; если такие связи есть, а условие Y в данном контексте выполнено, эти связи отбрасываются.

Следует заметить, что правила последних двух видов (запрещающие реализацию некоторой валентности) вполне аналогичны по своему действию правилам вида: СНЯТЬ Усл(C_o) = Sbl(W = X) или СНЯТЬ W(C_o) , - записанным в р. 321. Какой из данных двух способов записи удобнее использовать в окончательном варианте разрабатываемой системы, предполагается уточнить в ходе дальнейшей работы.

При практической работе по анализу конкретных текстов раздел 323 целесообразно рассматривать до установления гипотетических синтаксических связей. В таком случае валентности, реализация которых в этом разделе запрещается, могут попросту не учитываться, а при наличии у слова хотя бы одной обязательной валентности, которая в данном контексте не реализуется, соответствующее ЛМП можно исключать из дальнейшего анализа без рассмотрения всех прочих его валентностей. Именно так и делается в осуществляемом в настоящее время ручном эксперименте (см. ниже). При изложении общей схемы анализа, однако, нами принят несколько иной порядок перечисления отдельных операций, что вызвано стремлением максимально четко продемонстрировать ту роль, которую играют сведения различных разделов словарной статьи в процессе анализа.

7. Проверка синтаксической адекватности гипотетических связей между C_o и C_h , согласно свойствам C_o.

Рассматривается раздел 312 рабочей статьи C_o, который содержит

---

[1] Эта последняя операция предполагает, что всякий раз, когда между какими-либо двумя ЛМП словоформ устанавливается некоторая гипотетическая связь, каждое из них снабжается информацией об участии его в данной связи. Очевидно, что это не представляет технических трудностей.

правила вида:

Если $C_h = W(C_o)$ и Y то Z, –

где Y и Z – требования, характеризующие C_o и связанную с C_o единицу C_h с точки зрения наличия у них тех или иных признаков, их взаимного расположения во фразе, а также, возможно, наличия между ними или слева/справа от какой-либо из них слов определенных синтактико-морфологических классов. Эти правила играют роль фильтров, через которые "просеиваются" все установленные гипотетические связи, реализующие какие-либо валентности C_o. А именно: для каждой связи, реализующей валентность W единицы C_o, и каждого правила указанного вида проверяется, выполнены ли в данном контексте требования Y и Z. Если оказывается, что требование Y (оно, в частности, может быть пустым) выполняется, а требование Z – нет, рассматриваемая гипотетическая связь отвергается. Если проверка требования Y на данном этапе не дает окончательного результата в связи с явлением омонимии (напр., если в правиле требуется, чтобы между C_o и C_h не было глагола, а между ними имеется слово, которое в одном из значений является глаголом, а в другом – существительным), результатом этой проверки является некоторое условие, характеризующее совместимость данной синтаксической связи с тем или иным ГМП словоформы, которое само непосредственно в этой связи не участвует. Условия такого рода используются позднее на этапе объединения отдельных гипотетических связей между словами анализируемой фразы в единое дерево зависимостей.

Если проверяется связь вида $C_o \xrightarrow{i} \langle word \xrightarrow{\alpha} \rangle C_h$, кроме р. 312 рабочей статьи C_o рассматривается также р. 312 рабочей статьи единицы word, правила которого позволяют выяснить приемлемость связи $word \xrightarrow{\alpha} C_h$.

Если в результате проверки имеющихся гипотетических связей по правилам р. 312 некоторые из этих связей отбрасываются, то прежде, чем переходить к выполнению каких-либо из последующих проверок, полезно снова обратиться к разделу 323 рабочей статьи C_o и выяснить, не привела ли "чистка" гипотетических связей к то-

6800 7

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР

му, что какие-то из обязательных валентностей C_0 не имеют в нашем контексте реализующих их связей. Если это так, рассматриваемое ІМП словоформы и все установленные с ним связи отбрасываются и в дальнейшем анализе не участвуют (см. п. 6 - разрешение омонимии).

8. Проверка синтаксической адекватности гипотетических связей между C_0 и C_h, согласно свойствам C_h.

Рассматриваются гипотетические связи, реализующие валентности C_0. Для всех связей вида $C_h \xrightarrow{m} C_0$ и $C_h \xrightarrow{o} C_0$ проверяется раздел 331 рабочей словарной статьи C_h и выясняется, нет ли там правил, запрещающих эти связи. Для связей вида $C_0 \xrightarrow{i} C_h$ таким же образом обрабатывается раздел 332 рабочей статьи C_h. Правила этих разделов по содержанию и форме записи аналогичны правилам разделов 323 и 312 и применяются сходным образом (разница состоит только в том, что в правилах рр. 331 и 332 описываются связи, которые не являются реализацией валентностей единицы, стоящей в заглавии данной рабочей статьи, а в которых сама эта единица заполняет валентности других единиц). Если в результате какие-то из гипотетических связей, установленных для C_0, отбрасываются, то прежде, чем продолжать анализ, вновь проверяется по разделу 323 статьи C_0, не привело ли это к нарушению требований обязательной реализации тех или иных валентностей C_0 и не следует ли в связи с этим отказаться в данном контексте от рассмотрения C_0 и перейти к другому ІМП данной омонимичной словоформы (см. п. 6).

9. Проверка семантической адекватности гипотетических связей между C_0 и C_h.

Семантическая адекватность связей проверяется по разделу 322 словарной статьи той единицы, которая прогнозирует эти связи (имеет соответствующие валентности) - т.е., согласно описываемому пути анализа, единицы C_0. При этом для каждой связи, реализу-

ющей некоторую валентность W словоформы C_0 выясняется, нет ли в разделе 322 статьи C_0 правил вида:

$$W = X,$$

где X - список символов семантических классов, в которые могут входить лексические единицы, если они связаны с C_0 по связи, реализующей данную валентность W (= семантические требования к заполнению валентности W). Если такое правило имеется, происходит обращение к разделу 213 (формальное толкование) единицы C_h , связанной с C_0 по рассматриваемой гипотетической связи, и выясняется, находится ли это толкование в отношении включения (см. [14]) с выражением X. Если это не так, рассматриваемая связь отвергается.

Как и в пп. 7 и 8, если те или иные связи в результате проводимых проверок отбрасываются, выясняется, не оказались ли при этом нереализованными какие-либо из обязательных валентностей C_0, - что приводит к исключению C_0 из дальнейшего анализа.

10. Определение следующего ГМП словоформы, подлежащего обработке.

Согласно требованиям к порядку рассмотрения ГМП словоформ данной фразы, выработанным в п. 4, определяется то ГМП словоформы, которое должно обрабатываться после C_0. Это ГМП становится рабочим, и для него проделываются все операции, предусмотренные в пп. 5-9.

11. Определение сочетаемости друг с другом различных гипотетических связей между словоформами анализируемой фразы.

После окончания обработки всех ГМП словоформ данной фразы рассматриваются разделы 324 и 325 (условия на синтаксическую и семантическую сочетаемость связей) рабочих статей тех ГМП, которые не были отвергнуты в ходе анализа. На основе записанных в них правил вырабатывается информация о несовместимости или неот-

делимости тех или иных из имеющихся гипотетических связей в составе одного и того же дерева зависимостей.

12. Построение окончательной синтаксической структуры анализируемой фразы.

В соответствии с требованиями к виду правильной синтаксической структуры и сведениями о сочетаемости различных гипотетических связей друг с другом или с теми или иными ГМП словоформ данной фразы (эти сведения вырабатываются на этапе II, а также частично на этапах 7 и 8) для анализируемой фразы строятся все возможные для нее деревья зависимостей. Для каждого из них далее проверяется, выполняются ли в нем условия, записанные в разделе ЗИЗ рабочих словарных статей тех ГМП словоформ, которые использованы в этом дереве (раздел содержит сведения о синтактико-комбинаторных особенностях этих ГМП словоформ с точки зрения общих свойств синтаксической структуры фразы, т.е. требования типа проективности, которые не могут быть проверены на более ранних этапах анализа). Деревья, для которых выполнены все соответствующие условия считаются окончательными. Если таких деревьев несколько, выбор между ними должен производиться уже на основании более широкого контекста, выходящего за пределы одной фразы.

x

x x

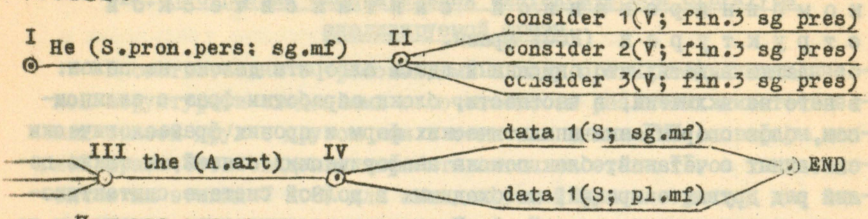
Дерево зависимостей, которое может быть получено по окончании такой обработки, не является в полной мере ни поверхностным, ни глубинным. С одной стороны, каждая словоформа анализируемой фразы представлена в этом дереве некоторым ГМП, так что в этом смысле его можно считать поверхностным. С другой стороны, ветви получаемого дерева обозначены индексами уже не поверхностных, а глубинных синтаксических отношений, а при его построении применяется, кроме поверхностно-синтаксической информации (синтактико-морфологическое оформление связей, порядок слов и т.д.), так-

же семантическая – что дает основание говорить о нем как о глубинном, тем более что практически все поверхностные единицы, которые не должны присутствовать в глубинно-синтаксической структуре фразы, в нашем дереве получают особую помету (заключаются в угловые скобки). Далее мы будем называть дерево такого вида комбинированной синтаксической структурой (КСС) фразы.

Легко видеть, что описанный здесь алгоритм далеко не полон: в него не включены, в частности, блоки обработки фраз с эллипсисом, блок свертывания аналитических форм и прочих фразеологически связанных сочетаний, блок поиска анафорических связей, а также целый ряд других операций, необходимых в полной системе синтактико-семантического анализа текста. Предложенное описание, однако, и не претендует на полноту; его следует рассматривать скорее как иллюстрацию тех общих принципов, в соответствии с которыми предполагается разрабатывать алгоритм анализа и используемую в нем комбинированную синтактико-семантическую модель английского языка, а также как пояснение к приводимым ниже материалам эксперимента по анализу английского текста.

Осуществляемый в настоящее время ручной эксперимент включает только синтактико-семантическую обработку фраз; морфологический анализ, а также составление рабочих словарных статей для каждого ГМП словоформ данной фразы предполагается проведенным. В связи с этим в качестве исходной информации выступает не реальный английский текст, а некоторое его вспомогательное линейное представление, в котором каждая словоформа заменена множеством всех ГМП, которые могут быть получены для нее при морфологическом анализе (таким образом, в этом представлении в явном виде указываются все случаи лексической и лексико-грамматической омонимии). При этом фразу удобно изображать в виде некоторого графа, где каждой словоформе сопоставлен один узел, а от этого узла вправо отходят столько ребер, сколько разных ГМП имеет данная словоформа. Узел обозначается порядковым номером словоформы в анализируемой фразе, над ребрами пишутся те ГМП, которые этим ребрам соответствуют. Пространственное расположение узлов определяется порядком обработки

ГМП отдельных словоформ при анализе и в описываемом варианте эксперимента изоморфно порядку слов во фразе. Заканчивается фраза узлом, обозначенным символом END (конец). Приведем в качестве иллюстрации вспомогательное представление фразы He considers the data:



Процесс синтактико-семантического анализа заключается в составлении по вспомогательному представлению фразы трех таблиц и матрицы гипотетических связей, на основании которых строится затем комбинированная синтаксическая структура фразы. Ниже дается неформальное словесное описание осуществляемых при этом операций, которое следует рассматривать как расширенный комментарий к публикуемым в этом и следующем выпусках предварительных публикаций ПГЭПЛ материалам анализа конкретных английских фраз. В связи с этим примеры в дальнейшем изложении почти не приводятся: читателю предлагается обращаться непосредственно к материалам анализа.

Каждая из таблиц, составляемых при анализе фразы, состоит из шести столбцов. В первом из них пишется порядковый номер обрабатываемой словоформы и имя вокабулы соответствующей лексической единицы[1], во втором – номер рассматриваемого значения этой единицы, в третьем и четвертом – ее синтактика и характеристика соответственно. Пятый столбец содержит применяемые при анализе правила. Для наглядности они выписываются целиком с указанием, в каком разделе и в какой статье словаря АРМАС (из тех, что использовались при составлении рабочей статьи данного ГМП словоформы) записаны эти правила. Способ записи: если правило взято из лексической словарной статьи, перед ним пишется только номер раздела и порядковый номер, под которым в этом разделе рассматриваемой лексической статьи стоит данное правило; если правило взято из грамматической статьи, после но-

[1] В вокабулу объединяются омонимичные единицы.

мера раздела в круглых скобках пишется признак (или набор признаков) или же – в случае словообразовательных статей – символ словообразовательного значения, указанный в заглавии этой статьи. Порядок применения конкретных правил в описываемом эксперименте целиком определяется порядком их записи в соответствующих словарных статьях. В таблицы выписываются только те правила, которые применимы в данном контексте. В шестом столбце указываются результаты применения каждого из выписанных правил. Таким образом, каждая из трех таблиц имеет вид:

1	2	3	4	5	6
Порядковый номер словоформы и имя вокабулы	Номер рассматриваемого значения	Синтактика рассматриваемого ГМП словоформы	Морфологическая характеристика	Применяемое правило	Результаты применения правила

При составлении первой таблицы выполняются операции, предусмотренные в 5-ом, 6-ом, 7-ом и 8-ом пунктах приведенной выше общей схемы анализа, в следующем порядке: вначале в рабочей статье ГМП словоформы, выписанном в столбцах 1-4 таблицы 1, просматривается раздел 323 и выясняется, есть ли у данной единицы валентности, которые должны быть обязательно заполнены в тексте. Если такая валентность обнаружена (найдено правило вида Есть W или Если U то есть W , где U – условие, которое в данном контексте выполняется), соответствующее правило выписывается. Затем в разделе 321 рабочей статьи находятся все правила, задающие возможности заполнения этой валентности, и во фразе ищутся кандидаты на ее заполнение (см. п. 5 общей схемы). Для каждого из найденных кандидатов проверяется, проходит ли он через фильтры, записанные в рр. 312, 331 и 332 рабочей статьи (см. пп. 7 и 8 общей схемы). Если все правила этих разделов для рассматриваемого кандидата выполняются, в столбце 6, напротив соответствующего правила из раздела 321, записывается устанавливаемая гипотетическая

связь[1]. Если какое-то правило раздела 312, 331 или 332 запрещает использование данного кандидата, оно выписывается в столбец 5, а в столбце 6 напротив него пишется в квадратных скобках, какая связь запрещается этим правилом: [Нет $c_i \xrightarrow{\alpha} c_j$], где c_i и c_j - соответствующие ГМП словоформ. В тех случаях, когда некоторое правило П pp. 312, 331 или 332 не позволяет сделать безусловный вывод о приемлемости/неприемлемости той или иной связи в рассматриваемом контексте, но дает информацию о несовместимости этой связи с определенными ГМП словоформ анализируемой фразы, данная связь записывается в столбце 6 таблицы 1, а правило П и вытекающее из него условие относительно сочетаемости этой связи с имеющимися во фразе ГМП словоформ помещаются, соответственно, в столбцы 5 и 6 таблицы 3 (см. ниже).

Если для рассматриваемой обязательной валентности рабочего ГМП словоформы не устанавливается ни одной реализующей ее связи, в столбце 6, напротив правила из р. 323, утверждающего обязательность этой валентности, пишется: Нет X, где X - рабочее ГМП. Далее это ГМП не обрабатывается.

Если множество связей, реализующих данную обязательную валентность C_0, не пусто, то после обнаружения всех возможных кандидатов на ее заполнение вновь просматривается р. 323 рабочей статьи. Если C_0 имеет еще какие-то обязательные валентности, они обрабатываются тем же способом; если же оказывается, что все обязательные валентности C_0 уже рассмотрены, начинается поиск гипотетических связей для всех прочих валентностей данного ГМП словоформы, за исключением тех из них, для которых в р. 323 найдено правило Нет W или Если U то нет W (где U в нашем контексте выполнено). На этом этапе отсутствие кандидатов на заполнение какой-либо из валентностей уже не является основанием для того, чтобы отказаться от дальнейшего рассмотрения данного ГМП.

---

[1] Для большей компактности записи вместо полных ГМП словоформ в столбце 5 таблицы 1 указывается, как правило, только порядковый номер словоформы во фразе, номер значения лексемы, выступающей в рассматриваемом ГМП этой словоформы, и ее морфологическая характеристика. В большинстве случаев этого достаточно для идентификации соответствующего ГМП. В случаях, когда словоформа может быть проанализирована с использованием нескольких разных вокабул, в записи связи, устанавливаемой с одним из ГМП этой словоформы, указыва-

Все связи, устанавливаемые в таблице 1, нумеруются подряд в порядке их установления; номера записываются в круглых скобках перед записью самой связи в столбце 6.

После того, как все ГМП словоформ анализируемой фразы обработаны указанным способом, составляется вторая таблица, соответствующая пункту 9 общей схемы анализа. В этой таблице для каждого ГМП словоформы, не отброшенного на предыдущих этапах анализа, в пятом столбце пишутся семантические требования, указанные в разделе 322 рабочей статьи данного ГМП для тех его валентностей, которые в анализируемой фразе реализуются хотя бы одной гипотетической связью (зафиксированной в таблице 1). Здесь же пишутся семантические классы, к которым принадлежат, согласно их толкованиям (р. 213), слова, вступающие с данным ГМП словоформы в соответствующие гипотетические связи. В столбце 6 напротив выписанных правил указывается, какие из этих связей должны быть, по имеющимся семантическим требованиям, отвергнуты (пишется Нет (i) , где i - номер отвергаемой связи по таблице 1). Если то или иное правило раздела 322 не дает возможности отбросить никакую из установленных связей, оно не выписывается. После проверки всех таких правил выясняется, не были ли в результате их применения отвергнуты все гипотетические связи, реализующие в анализируемой фразе одну из обязательных валентностей обрабатываемого ГМП. Если это имеет место, в столбец 5 выписывается правило из раздела 323, задающее обязательность этой валентности, а в столбце 6 пишется: Нет X, где X - данное ГМП.

Построение последней таблицы осуществляется тем же путем, что и первых двух, но с использованием на этот раз правил разделов 324 и 325 рабочей словарной статьи. Эти правила рассматриваются в применении к тем гипотетическим связям анализируемой фразы, которые остались после семантического анализа. В столбце 6 записываются получаемые условия на сочетаемость этих связей. Если при этом какие-то связи отвергаются безусловно, то, как и в таблице 2, проверяется, не являлось ли их наличие обязательным для ГМП той единицы, валентность которой они реализуют. При положительном результате проверки отвер-

---

ется также имя соответствующей лексемы (см. словоформу УI в приводимых материалах анализа).

гается само это ГМП. Как уже отмечалось, в таблице 3 записываются также те условия на сочетаемость связей, которые могут быть получены на первых этапах анализа при рассмотрении правил разделов 312, 331 и 332.

Чтобы показать, насколько эффективным явился анализ с точки зрения снятия омонимии, после построения третьей таблицы дается новое вспомогательное линейное представление анализируемой фразы, уже не включающее всех тех ГМП словоформ, которые были отвергнуты при построении таблиц.

Во всех случаях, когда на некотором этапе анализа отвергается ГМП словоформы, с которым ранее были установлены какие-то связи, номера этих связей в таблице 1 зачеркиваются (данная операция, впрочем, диктуется только соображениями наглядности и с формальной точки зрения необходимой не является).

Для дальнейшего анализа удобно воспользоваться матричной формой представления структур зависимостей[1]. Матрица в нашем эксперименте строится следующим образом; каждому из оставшихся после составления таблиц ГМП словоформы сопоставляется строка и столбец; альтернативные ГМП одной и той же словоформы записываются рядом, и соответствующие им столбцы (строки) объединяются под одной римской цифрой - порядковым номером данной словоформы во фразе. Если в таблице 1 были установлены (и впоследствии не отвергнуты) связи вида $C_0 \xrightarrow{i} \langle \text{word} \xrightarrow{\alpha} C_h \rangle$, $C_0 \xrightarrow{i} \langle \text{word} \xrightarrow{\alpha} \rangle$ или $C_0 \rightarrow C_h \langle \xrightarrow{\alpha} \text{word} \rangle$, единице word сопоставляется, кроме тех, которые включены в матрицу по общему правилу, еще один столбец и одна строка, входом в которые служит запись $\langle \text{word} \xrightarrow{\alpha} \rangle$ (или $\xrightarrow{\alpha} \text{word} \rangle$, в зависимо-

---

[1] Данная форма используется в целом ряде работ по автоматическому переводу. В частности, она подробно рассматривается Г.С.Слуцкером и О.С.Кулагиной [5, 6, 12] в связи с системой французско-русского машинного перевода. В нашем эксперименте, однако, в матрицу включается большее количество сведений, чем в указанной системе, где информация о наличии омонимов, индекс типа связей, сведения об их обязательности и т.д. записываются отдельно и в матрицу как таковую не включаются. Еще одним отличием является то, что в упомянутых работах клетка на пересечении строки C_f и столбца C_k заполняется при наличии связи $C_k \rightarrow C_f$, а не $C_f \rightarrow C_k$, как у нас. Принципиального значения все эти отличия не имеют.

сти от вида имеющейся в таблице 1 связи). На пересечении строки c_f и столбца c_k указываются в круглых скобках номера всех гипотетических связей вида $c_f \xrightarrow{\alpha} c_k$, имеющихся в таблице 1; после номера связи через двоеточие записывается ее индекс α. Номера связей, имеющих вид: $c_o \xrightarrow{i} \langle word \xrightarrow{\alpha} \rangle c_h$, в матрице повторяются дважды: в строке, озаглавленной $\langle word \xrightarrow{\alpha} \rangle$ (где номер связи сопровождается индексом α), и в одноименном столбце (где после номера связи указывается индекс i сильного места, реализуемого этой связью). Номера всех других связей появляются в матрице ровно один раз. Если некоторая связь реализует обязательную валентность, ее номер в матрице подчеркивается. После того, как получена такая матрица, можно переходить к построению деревьев зависимостей для данной фразы, что сводится к "чистке" матрицы в соответствии с общими требованиями к виду правильной синтаксической структуры, а также теми условиями на сочетаемость имеющихся в анализируемой фразе гипотетических связей, которые зафиксированы в таблице 3.

Требования к виду правильной синтаксической структуры применительно к матрице данного вида выражаются следующим образом.

Однозначность каждой словоформы в контексте означает, что в окончательной матрице каждой римской цифре (порядковому номеру словоформы) должны быть сопоставлены ровно один столбец и одна строка (при наличии неразрешенной лексико-грамматической омонимии окончательных матриц должно быть несколько).

Единственность вершины соответствует требованию, чтобы в матрице был ровно один пустой столбец (единица, являющаяся входом в него, ничему не подчинена).

Связность дерева требует, чтобы в каждом из других столбцов имелся хотя бы один символ (при этом, по условию единственности "хозяина" - слова, управляющего данным, - не более, чем один).

Единственность заполнения валентностей "сильного" управления означает, что цифровые индексы типа связей в каждой строке не должны повторяться.

Наконец, обязательность заполнения тех или иных валентностей сильного управления (обозначаемая в матрице подчеркиванием связей,

реализующих эти валентности) может быть выражена требованием, чтобы при обработке матрицы в строках, где имеются подчеркнутые номера, сохранялся один из каждой группы таких номеров (в группу объединяются номера связей с совпадающими индексами). Снятие всех подчеркнутых в некоторой строке номеров с одним и тем же индексом типа связи приводит к вычеркиванию из матрицы всей этой строки и сопряженного с ней (= соответствующего тому же ГМП словоформы) столбца.

В описываемом ручном эксперименте обработка матрицы производится в следующей последовательности.

1. Поиск потенциальной вершины синтаксического дерева фразы. Рассматриваются все ГМП словоформ, которым соответствуют пустые столбцы. Если они имеют признаки, запрещающие их употребление в роли вершины фразы (прилагательное A, наречие Adv , союз Conj, абсолютная форма глагола abs и т.д.), соответствующие им строки и столбцы из матрицы вычеркиваются (в публикуемых материалах - заштриховываются). Если таких признаков нет, ГМП словоформ, имеющие пустые столбцы считаются безусловными претендентами на роль вершины данной фразы. Если безусловных претендентов на эту роль не оказывается (в частности, если все они отвергаются в ходе построения дерева), необходимо для каждого ГМП, не имеющего противоречащих признаков, проверить гипотезу о том, что он является вершиной анализируемой фразы.

2. Поиск безусловных синтаксических связей.

Рассматриваются однозначные словоформы - т.е. словоформы, которым в данной матрице сопоставлен ровно один столбец. Связи, являющиеся в таких столбцах единственными, считаются безусловными для всех вариантов анализа, в которых в качестве вершины рассматривается ГМП любой другой словоформы (если в роли вершины выступает ГМП данной словоформы, связи, имеющиеся в ее столбце, естественно, не учитываются). Если в строках, соответствующих однозначным словоформам есть обязательные (подчеркнутые) номера связей, причем индексы этих обязательных связей в пределах строки не повторяются, все такие связи также считаются безусловными (в любом варианте анализа).

Для каждой из установленных безусловных связей проверяется, нет ли в таблице 3 сведений о ее сочетаемости с другими связями или ГМП словоформ. Если какие-либо связи или ГМП противоречат данной безусловной связи, они из матрицы вычеркиваются, если же какие-то связи неотделимы от данной, они также получают статус безусловных и для них производятся те же проверки. Каждый раз при вычеркивании из матрицы каких-либо связей или ГМП словоформ проверяется, не приводит ли это к появлению новых безусловных связей.

3. Построение связного синтаксического дерева для каждого претендента на роль вершины.

После выявления всех безусловных связей для каждого данного претендента на роль вершины фразы проверяется, не может ли с помощью этих связей быть построено связное дерево всей фразы. Если да, это дерево считается окончательным, и проверяются другие варианты анализа (другие претенденты на роль вершины).

Если при данном вершинном ГМП безусловные связи не обеспечивают связного дерева, проверяется, есть ли у этого ГМП "слуги" (т.е. заполнена ли хотя бы одна клетка в соответствующей ему строке). Если их нет, а фраза содержит более одной словоформы, рассматриваемый вариант анализа отвергается: данное ГМП не может быть вершинным. Если при этом соответствующий ему столбец пуст, это ГМП вообще вычеркивается из матрицы.

Если у данного вершинного ГМП имеются слуги, то проверяется, есть ли среди них такие, которые вводятся не безусловными связями. Для каждой из таких связей выясняется, не позволяет ли ее присоединение к безусловным получить связное дерево (с учетом имеющихся в таблице 3 условий на сочетаемость этой связи с другими). Если получаемая при этом структура продолжает быть несвязной, к ней присоединяются новые связи (из числа имеющихся в матрице и притом совместимых с уже включенными в эту матрицу связями). Если таких связей не оказывается, данный вариант анализа отвергается и проверяется следующий. Основной принцип, применяемый при каждом "ветвлении" хода анализа (обнаружении нескольких альтернативных способов его про-

должения) состоит в том, что в первую очередь проверяется наличие, говоря в терминах О.С.Кулагиной, "уникальных гипотез" для словоформ, чьи ПМП еще не включены в построенный к данному моменту фрагмент дерева фразы.

В публикуемых материалах анализа обработка матрицы гипотетических связей рассматриваемой фразы подробно комментируется.

К настоящему времени по изложенному методу обработано более 50 английских фраз длиной от 5 до 25 словоформ, взятые из одного и того же текста. В обработке отдельных фраз принимали участие сотрудники ЛМП Л.А.Афонасьева, Л.Л.Ельницкий, Т.В.Заворыгина, В.А.Матвеенко, Н.Н.Павлова, Т.В.Пивоварова, Т.Г.Потемкина, А.С.Чехов, З.М.Шаляпина и М.Г.Шаталова, а также сотрудник информационного центра главного управления пропаганды тематики и методики ВДНХ СССР В.А.Альперин.

*

* *

Рукопись данной статьи была прочитана О.С.Кулагиной и И.А.Мельчуком, которым автор искренне благодарен за сделанные замечания. Автор выражает также свою признательность всем сотрудникам, осуществлявшим анализ конкретных текстов, прежде всего – Л.Л.Ельницкому, В.А.Матвеенко и А.С.Чехову, в обсуждении с которыми уточнялись многие детали описываемого здесь алгоритма.

П Р И Л О Ж Е Н И Е

Перечень разделов словарной статьи АРМАС.

Зона 0. Общий статус описываемой единицы C_0.

01. Заглавие словарной статьи.
011. Поверхностное представление C_0.
012. Отсылка к главной словарной статье.
013. Отсылка к статье основного графического варианта.
014. Предметная область (в лексических статьях терминов).

Зона 1. Морфологические сведения.

11. Часть речи.
121. Тип парадигмы.
122. Сведения о неполноте парадигмы.
123. Сведения о наличии добавочных форм.

Зона 2. Семантические сведения.

211. Лексическая производность C_0.
212. Переводные эквиваленты C_0.
213. Формальное толкование C_0.
214. Индивидуальные правила включения единицы C_0 в семантическую запись текста.
22. Примеры употребления C_0 в тексте.
23. Грамматическое значение C_0.

Зона 3. Синтаксические сведения.

311. Синтаксические признаки.
312. Сведения о пределах поиска в тексте единиц, синтаксически связанных с C_0.

313. Общие требования к синтаксической структуре сочетаний, содержащих C_0.
321. Модель управления C_0.
322. Семантические требования к словам, заполняющим валентности C_0.
323. Сведения об обязательности или недопустимости реализации в тексте тех или иных валентностей C_0.
324. Сведения о синтаксической сочетаемости реализаций различных валентностей C_0, а также валентностей C_0 и валентностей слов, синтаксически связанных с C_0.
325. Сведения о семантической сочетаемости различных заполнений валентностей C_0 и единиц, связанных с C_0.
331. Сведения о слабых и координативных слугах C_0.
332. Сведения о сильных хозяевах C_0.

Зона 4. Лексико-синтаксические сведения трансформационного типа.

41. Стандартные лексические функции, аргументом которых является C_0.
42. Индивидуальные лексические функции, аргументом которых является C_0.
43. Фразеологические сочетания, прогнозируемые единицей C_0.
44. Правила преобразования глубинно-синтаксической структуры текста при замене C_0 переводным эквивалентом.
45. Правила синонимических преобразований синтаксической структуры текста, связанные с особенностями C_0.
46. Правила восстановления связных синтаксических структур при наличии в тексте эллипсиса, нарушающего их связность.
47. Правила восстановления полных синтаксических структур при наличии в тексте других видов эллипсиса.

Л И Т Е Р А Т У Р А

1. Н.Г.Арсентьева, З.М.Шаляпина. Язык для записи лингвистической информации в автоматическом словаре. – Машинный перевод и прикладная лингвистика. Вып. 16. М., 1972.

2. Л.А.Афонасьева, Л.Л.Ельницкий, Т.В.Заворыгина, Н.Н.Павлова, Н.В.Перцов, Т.В.Пивоварова, Т.Г.Потемкина, И.И.Убин, А.С.Чехов, З.М.Шаляпина, М.Г.Шаталова, К.О.Эрастов. Об англо-русском автоматическом словаре с грамматическим обеспечением (предварительные результаты. – В кн.: Автоматическая обработка текстов (тематический сборник) [в печати].

3. Н.А.Еськова, И.А.Мельчук, В.З.Санников. Формальная модель русской морфологии. I. Формообразование существительных и прилагательных. М., 1971, 71 стр. [ИРЯ АН СССР. ПГЭПЛ. Предв. публ.].

4. Л.Н.Иорданская. Автоматический синтаксический анализ. М., 1967.

5. О.С.Кулагина. Использование машин в исследованиях по машинному переводу. – Проблемы кибернетики. Вып. 10. М., 1963, стр. 205-215.

6. О.С.Кулагина. Алгоритм синтаксического анализа в системе французско-русского машинного перевода. Препринт № 13 за 1970 г. Москва, 47 стр. [ИПМ АН СССР].

7. О.С.Кулагина, И.А.Мельчук, К.О.Эрастов. Об одной возможной системе машинного перевода. М., 1971, 72 стр. [ИРЯ АН СССР. ПГЭПЛ. Предв. публ. Вып. 21].

8. И.А.Мельчук. Уровни представления языковых высказываний и общее строение модели "Смысл $\Leftrightarrow$ Текст". – В кн.: Актуальные проблемы лексикологии и словообразования. Вып. I. Новосибирск, 1972, стр. 5-35.

9. И.А.Мельчук, Н.В.Перцов. Фрагмент модели английского поверхностного синтаксиса (предварительное сообщение). М.,1973. [ИРЯ АН СССР. ПГЭПЛ. Предв. публ. Вып. 35, стр. 3-18].

10. И.А.Мельчук, Н.В.Перцов. Поверхностно-синтаксические отношения (ПСО) в английском языке. [ИРЯ АН СССР. ПГЭПЛ. Предв. публ. В печати].

11. Образцы словарных статей англо-русского автоматического семантического словаря. М., 1970, 58 стр. [ИРЯ АН СССР. ПГЭПЛ. Предв. публ. Вып. 9].

12. Г.С.Слуцкер. Получение всех допустимых вариантов синтаксического анализа текста при помощи машины. - Проблемы кибернетики. Вып. 10. М., 1963, стр. 215-225.

13. Формальная грамматика английского словоизменения (третий этап работы по теме "Создание формальной грамматики английского словоизменения (формальное описание английской морфологии)" по хоздоговору между ВИНИТИ Госкомитета по науке и технике при СМ СССР и АН СССР и МГПИИЯ им. М.Тореза). М., 1972, 147 стр. [МГПИИЯ им. М.Тореза. ЛМП. Отчет 3 (1972)].

14. З.М.Шаляпина. Семантические элементы как основа лексикографического описания: общие принципы и формальный аппарат. - Машинный перевод и прикладная лингвистика. Вып. 15. М., 1972, стр. 54-98.

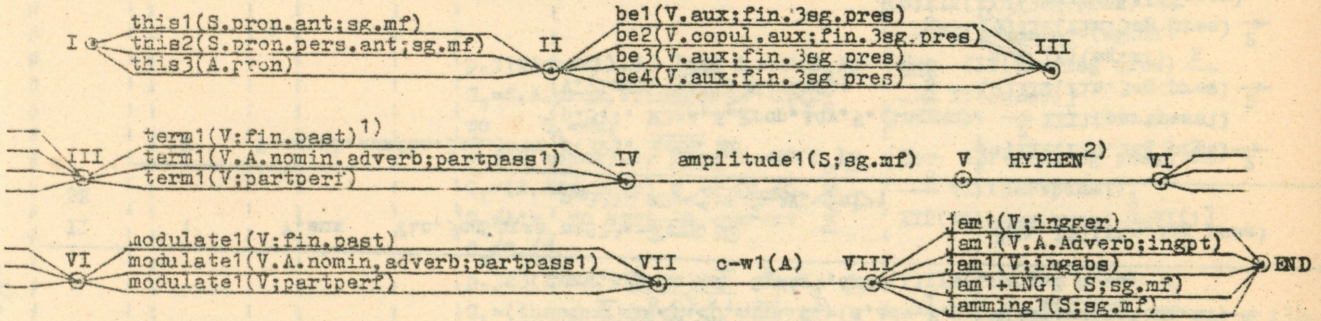
15. З.М.Шаляпина. Англо-русский многоаспектный автоматический словарь (АРМАС). - Машинный перевод и прикладная лингвистика. Вып. 17. М., 1974 [в печати].

16. O.S.Kulagina, I.A.Mel'čuk. Automatic Translation: some theoretical aspects and the design of a translation system. - In: Machine Translation. Amsterdam, 1967, pp. 137-171.

МАТЕРИАЛЫ АНАЛИЗА ФРАЗЫ:

This is termed amplitude-modulated c-w jamming

Вспомогательное линейное представление 1



1) Слово term имеет в словаре 5 значений, в первом из которых оно является глаголом, а в остальных - существительным. В процессе морфологического анализа значения 2-5 отбрасываются в силу несовместимости основы существительного с суффиксом -ed.

2) Дефис, как и все знаки препинания, считается особым словом (исключая случаи употребления его в составе сложных слов, целиком заданных в словаре - ср. словоформу c-w) и в качестве такового имеет собственную словарную статью.

Таблица I.

1	2	3	4	5	6
I THIS	1	S.pron.ant	sg.mf	p.323. Нет Gm;	———
	2	S.pron. pers.ant	sg.mf	- " -	———
	3	A.pron		p.321(A).1. Gm=S и не it3,wh; p.312(A).1. Если C_1=G(C_0) и C_1= =S и $C_0 < C_2 < C_1$ то C_2=(A,Adv,S, crd,hyphen) и $C_2 \neq$ Art,def;	[Нет IV1(sg.mf) $\xrightarrow{m}$ I, VIII[jam1+ING](sg.mf) $\xrightarrow{m}$ I, VIIIjamming1(sg.mf) $\xrightarrow{m}$ I]
II BE	1	V.aux	fin.3sg pres	p.323. Есть M2 p.321. M2=<to1 $\xrightarrow{2}$ V.<inf>;	Нет III1(fin.3sg pres)
	2	V.copul.aux	fin.3sg pres	p.323. Есть M2 p.321. M2=A,S,Prep,Adv,V.<ingger>, V.<inf>, <to1 $\xrightarrow{2}$ V.<inf>;	(1)III2(fin.3sg pres) $\xrightarrow{2}$ $\xrightarrow{2}$ III1(partpass1) (2)III2(fin.3sg pres) $\xrightarrow{2}$ $\xrightarrow{2}$ IV1(sg.mf) (3)III2(fin.3sg pres) $\xrightarrow{2}$ $\xrightarrow{2}$ VIIIjam1(ingger) (4)III2(fin.3sg pres) $\xrightarrow{2}$ $\xrightarrow{2}$ VIIIjam1(ingpt) 1)

1) Связь (4) устанавливается по правилу M2 = A.

1	2	3	4	5	6
II	2	V.copul.aux	fin.3sg.pres	p.312(copul).5. Если C_1=M2(C_0) и C_1=S,A.pron,V.ingger и $C_1 \prec C_0$ то C_1=wh; p.312(copul).1. Если C_1=M2(C_0) и C_1=(A,Adv,Prep) и $C_0 \prec C_2 \prec C_1$ и $C_2 \neq$Adv то есть C_3=parent и $C_0 \prec C_3 \prec C_1$; p.323(fin). Если нет $C_1 \cap C_0$ и C_1=(than,as,whatever,whoever) и нет $C_2 \prec C_0$ и C_2=(ord,what,whatever,whichever) то есть M1(C_0); p.321. M1=S,V.⟨ingger⟩, ⟨to1 $\xrightarrow{2}$ V.⟨inf⟩;	(5)II2(fin.3sg.pres) $\xrightarrow{2}$ VIII[jam1+ING1](sg.mf) (6)II2(fin.3sg.pres) $\xrightarrow{2}$ VIIIjamming1(sg.mf) [Нет II2(fin.3sg.pres) $\xrightarrow{2}$ I1(sg.mf)] [Нет II2(fin.3sg.pres) $\xrightarrow{2}$ VI1(partpass1), II2(fin.3sg.pres) $\xrightarrow{2}$ VII1] (7)II2(fin.3sg.pres) $\xrightarrow{1}$ I1(sg.mf) (8)II2(fin.3sg.pres) $\xrightarrow{1}$ I2(sg.mf) (9)II2(fin.3sg.pres) $\xrightarrow{1}$ IV1(sg.mf) (10)II2(fin.3sg.pres) $\xrightarrow{1}$ VIIIjam1(ingger) (11)II2(fin.3sg.pres) $\xrightarrow{1}$ VIII[jam1+ING1] (sg.mf)

1	2	3	4	5	6
II	2	V.copul.aux	fin.3sg.pres	p.321(fin). Gm=S; p.312(fin).2. Если C_1=Gm(C_0) и C_1≠(what,whatever,whoever)и нет C_2<C_1 и C_2=(whatever,whichever) то C_1<C_3<C_0 и C_3=S;	(12)II2(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIIIjamming1(sg.mf) [Нет I1(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres), I2(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres)]
	3	V.aux	fin.3sg.pres	p.323(fin). Если нет C_1∩C_0 и C_1=(than,as,whatever,whoever) и нет C_2<C_0 и C_2=(ord,what,whatever,whichever) то есть M1(C_0); p.321. M1=S;	(13)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ I1(sg.mf) (14)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ I2(sg.mf) (15)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ IV1(sg.mf) (16)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIII[jam1+ING1](sg.mf) (17)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIIIjamming1(sg.mf)
				p.321(fin). Gm=S; p.312(fin).2. Если C_1=Gm(C_0) и C_1≠(what,whatever,whoever) и нет C_2=(whatever,whichever) то C_1<C_3<C_0 и C_3=S;	[Нет I1(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres), I2(sg.mf)$\xrightarrow{m}$ II2(fin.3sg.pres)]

1	2	3	4	5	6
II	4	V.aux	fin.3sg.pres	p.323. Есть M2; p.321. M2=Prep,Adv;	Нет II4(fin.3sg.pres)
III TERM	1	V	fin.past	p.312(fin).10. Если $C_1 \cap C_0$ и C_1= =fin и нет C_2(parent)$\langle C_0$ то есть $C_3 \langle C_4 \langle C_1$ и C_3=S и C_4=S;	Нет III1(fin.past)
		V.A.nomin. adverb	partpass1	p.323(partpass1).1. СНЯТЬ Усл(C_0) =Sbl(Есть M1); СНЯТЬ Усл(C_0)= =Sbl(Есть M2); p.323. Есть M3; p.321. M3=S,A,$\langle$quot$\rightarrow C_1$, Ф;	(18)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$IV1(sg.mf) (19)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VI1(partpass1) (20)III1(partpass1) $\xrightarrow{3}$VII1 (21)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIIIjam1(ingpt) (22)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIII[jam1+ING1](sg.mf) (23)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIIIjamming1(sg.mf)
				p.321(partpass1).1. СНЯТЬ Усл(C_0)= =Sbl(M1=X); СНЯТЬ Усл (C_0)= =Sbl(M2=X); p.321(partpass1).2. M2=$\langle$by4 $\xrightarrow{2}$S, $\langle$by4 $\xrightarrow{2}$V.ingger;	

1	2	3	4	5	6
II	2	V.copul.aux	fin.3sg.pres	p.321(fin). Gm=S; p.312(fin).2. Если C_1=Gm(C_0) и C_1≠(what,whatever,whoever)и нет C_2<C_1 и C_2=(whatever,whichever) то C_1<C_3<C_0 и C_3=S;	(12)II2(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIIIjamming1(sg.mf) [Нет I1(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres), I2(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres)]
	3	V.aux	fin.3sg.pres	p.323(fin). Если нет C_1∩C_0 и C_1=(than,as,whatever,whoever) и нет C_2<C_0 и C_2=(ord,what,whatever,whichever) то есть M1(C_0); p.321. M1=S;	(13)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ I1(sg.mf) (14)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ I2(sg.mf) (15)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ IV1(sg.mf) (16)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIII[jam1+ING1](sg.mf) (17)II3(fin.3sg.pres)$\xrightarrow{1}$ $\xrightarrow{1}$ VIIIjamming1(sg.mf)
				p.321(fin). Gm=S; p.312(fin).2. Если C_1=Gm(C_0) и C_1≠(what,whatever,whoever) и нет C_2=(whatever,whichever) то C_1<C_3<C_0 и C_3=S;	[Нет I1(sg.mf)$\xrightarrow{m}$ $\xrightarrow{m}$ II2(fin.3sg.pres), I2(sg.mf)$\xrightarrow{m}$ II2(fin.3sg.pres)]

1	2	3	4	5	6
II	4	V.aux	fin.3sg.pres	p.323. Есть M2; p.321. M2=Prep,Adv;	Нет II4(fin.3sg.pres)
III TERM	1	V	fin.past	p.312(fin).10. Если $C_1 \cap C_0$ и C_1= =fin и нет C_2(parent)$\langle C_0$ то есть $C_3 \langle C_4 \langle C_1$ и C_3=S и C_4=S;	Нет III1(fin.past)
		V.A.nomin. adverb	partpass1	p.323(partpass1).1. СНЯТЬ Усл (C_0) =Sbl(Есть M1); СНЯТЬ Усл (C_0)= =Sbl(Есть M2); p.323: Есть M3; p.321. M3=S,A,$\langle$quot$\rightarrow C_1$, Ф;	(18)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$IV1(sg.mf) (19)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VI1(partpass1) (20)III1(partpass1) $\xrightarrow{3}$VII1 (21)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIIIjam1(ingpt) (22)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIII[jam1+ING1](sg.mf) (23)III1(partpass1) $\xrightarrow{3}$ $\xrightarrow{3}$VIIIjamming1(sg.mf)
				p.321(partpass1).1. СНЯТЬ Усл(C_0)= =Sbl(M1=X); СНЯТЬ Усл (C_0)= =Sbl(M2=X); p.321(partpass1).2. M2=$\langle$by4 $\xrightarrow{2}$S, $\langle$by4 $\xrightarrow{2}$V.ingger;	

1	2	3	4	5	6
III	1	V.A.nomin advert	partpass1	p.32 (partpass1).3. Gm=S,V и не aux;	(24)IV1(sg.mf) m→ m→ III1(partpass1) (25)VIII[jam1+ING1](sg.mf) m→ m→ III1(partpass1) (26)VIIIjamming1(sg.mf) m→ m→ III1(partpass1)
				p.312(partpass1).2. Если C_1=Gm(C_0) и $C_0<C_1$ и C_1=V то есть $C_0<C_3<C_1$ и C_3=comm;	[Нет VI1(fin.past) m→ m→ III1(partpass1) VI1(partpass1) m→ m→ III1(partpass1), VI1(partperf) m→ m→ III1(partpass1) VIIIjam1(ingger) m→ m→ III1(partpass1), VIIIjam1(ingpt) m→ m→ III1(partpass1), VIIIjam1(ingabs) m→ m→ III1(partpass1)
				p.312(partpass1).1. Если C_1=G(C_0) и C_1=S и $C_1<C_2<C_0$ то C_2≠V.fin;	[Нет I1(sg.mf) m→ m→ III1(partpass1), I2(sg.mf) m→ III1(partpass1)]
		V	partperf	p.323(partperf).1. Есть $C_1<C_0$ и C_1=have;	нет III1(partperf)

1	2	3	4	5	6
IV AMPLI-TUDE	1	S	sg.mf	p.321. M1=S.attr, ⟨of7 $\xrightarrow{2}$⟩S; p.312(attr).1. Если C_1=G(C_0) то C_0<C_1; p.312(attr).2. Если C_1=G(C_0) и C_0<C_2<C_1 то C_2=(S,A,crd,hyphen, quot) и C_2≠Art;	[нет VIII[jam1+ING1](sg.mf) $\xrightarrow{m}$ IV1(sg.mf), VIIIjamming1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ IV1(sg.mf)] [нет I1(sg.mf) $\xrightarrow{m}$IV1(sg.mf), I2(sg.mf) $\xrightarrow{m}$IV1(sg.mf)]
				p.321. M2=⟨of7 $\xrightarrow{2}$⟩S;	
				p.321(sg.mf). Gm=S,V.inf, V.partpass1; p.312(sg.mf).1. Если C_1=Gm(C_0) и C_0≠temp то C_0=attr p.312(attr).1. Если C_1=G(C_0) то C_0<C_1;	(27)VI1(partpass1) $\xrightarrow{m}$ $\xrightarrow{m}$ IV1(sg.mf) (28)VIII[jam1+ING1](sg.mf) $\xrightarrow{m}$IV1(sg.mf) (29)VIIIjamming1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$IV1(sg.mf) [нет I1(sg.mf) $\xrightarrow{m}$IV1(sg.mf), I2(sg.mf) $\xrightarrow{m}$IV1(sg.mf), III1(partpass1) $\xrightarrow{m}$IV1(sg.mf)]
V HYPHEN	1			p.321. Gm=S; p.312. Если C_1=Gm(C_0) то C_1^C_0	(30)IV1(sg.mf) $\xrightarrow{m}$ V
VI MODULATE	1	V	fin.past	p.323(fin). Если нет C_1^C_0 и C_1=(than,as,whatever,whoever) и	(31)VI1(fin.past) $\xrightarrow{1}$ $\xrightarrow{1}$ IV1(sg.mf)

I	2	3	4	5	6

			fin.past	нет $C_2<C_0$ и C_2=(crd,what,whatever,whoever) то есть M1(C_0);
				p.321. M1=S;
				p.312(fin).6. Если C_1=M1(C_0) и $C_1<C_2<C_0$ и C_2=fin и нет $C_1<C_3$(parent)$<C_0$ и $C_1\neq$(what,whatever,whoever) и нет $C_4<C_0$ и C_4=(whatever,whichever) то есть $C_1<C_5<C_2$ и C_5=S;
				p.323. Есть M2;
				p.321. M2=S;
				p.312(fin).8. Если C_1=Mi(C_0) и $C_1\neq$M1(C_0) и C_1=S и $C_1\neq$wh то $C_0<C_1$;
					(32)VI1(fin.past) $\xrightarrow{1}$ $\xrightarrow{1}$VIII[jam1+ING1](sg.mf)
					(33)VI1(fin.past) $\xrightarrow{1}$ $\xrightarrow{1}$VIIIjamming1(sg.mf)
					[нет VI1(fin.past) $\xrightarrow{1}$ $\xrightarrow{1}$ I1(sg.mf)
					VI1(fin.past) $\xrightarrow{1}$ I2(sg.mf)]
					(34)VI1(fin.past) $\xrightarrow{2}$ $\xrightarrow{2}$ VIII[jam1+ING1](sg.mf)
					(35)VI1(fin.past) $\xrightarrow{2}$ $\xrightarrow{2}$ VIIIjamming1(sg.mf)
					[нет VI1(fin.past) $\xrightarrow{2}$ $\xrightarrow{2}$ I1(sg.mf),
					VI1(fin.past) $\xrightarrow{2}$ I2(sg.mf)
					VI1(fin.past) $\xrightarrow{2}$ IV1(sg.mf)]

				p.321(fin). Gm=S; p.312(fin).2. Если C_1=Gm(C_0) и $C_1\neq$(what,whatever,whoever) и нет $C_2<C_0$ и C_2=(whatever,whichever) то $C_1<C_3<C_0$ и C_3=S; p.312(fin).3.Если C_1=Gm(C_0) и	[нет IV1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(fin.past)]

1	2	3	4	5	6
VI	1	V	fin.past	$C_1 < C_2 < C_o$ то C_2=(S,A,Adv,crd, hyphen,quot)	I1(sg.mf) $\xrightarrow{m}$ VI1(fin.past), I2(sg.mf) $\xrightarrow{m}$ VI1(fin.past)]
		V.A.nomin. adverb	partpass1	p.323(partpass1).1. СНЯТЬ Усл(C_o)= =Sbl(Есть M1); СНЯТЬ Усл (C_o)= =Sbl(Есть M2); p.321(partpass1).1. СНЯТЬ Усл(C_o)= =Sbl(M1=X); СНЯТЬ Усл (C_o)= =Sbl(M2=X); p.321(partpass1).2. M2=<by4 $\xrightarrow{2}$>S, <by4 $\xrightarrow{2}$>V.ingger; p.321(partpass1).3. Gm=S,V и не aux;	(36)III1(fin.past) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1) (37)III1(partpass1) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1) (38)III1(partperf) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1) (39)IV1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1) (40)VIII[jam1+ING1](sg.mf) $\xrightarrow{m}$ VI1(partpass1) (41)VIIIjamming1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1)

I	2	3	4	5	6
VI	1	V.A.nomin adverb	partpass1	p.312(partpass1).2. Если C_1=Gm(C_o) и C_1=V и C_o<C_1 то есть C_o<C_3<C_1 и C_3=c-m;	[Нет VIIIjam1(ingger) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1), VIIIjam1(ingpt) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1) VIIIjam1(ingabs) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1)]
				p.312(partpass1).1. Если C_1=G(C_o) и C_1=S и C_1<C_2<C_o то C_2≠V.fin;	[Нет I1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VI1(partpass1), I2(sg.mf) $\xrightarrow{m}$ VI1(partpass1)]
		V	partperf	p.323(partperf).1. Есть C_1<C_o и C_1=have;	Нет VI1(partperf)
VII	1	A		p.321(A). Gm=S;	(42)VIII[jam1+ING1](sg.mf) $\xrightarrow{m}$ VII1 (43)VIIIjamming1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VII1
				p.312(A).7. Если C_1=G(C_o) и C_1=S и C_o≠partpass1 и C_1<C_2<C_o то C_2=Adv;	[Нет I1(sg.mf) $\xrightarrow{m}$ VII1, I2(sg.mf) $\xrightarrow{m}$ VII1, IV(sg.mf) $\xrightarrow{m}$ VII1]
VIII JAM	1	V	ingger	p.323. Есть M2; p.321. M2=S; p.312(ingger).2. Если C_1=Mi(C_o) и C_1≠M1(C_o) то C_o<C_1;	Нет VIIIjam1(ingger)

1	2	3	4	5	6
VIII	1	V.A.Adverb	ingpt	p.323(ingpt).1. СНЯТЬ Усл (C_0)= =Sbl(Есть Mi); p.323(ingpt).2. Нет MI; p.321. M2=S; p.312(ingpt).1. Если C_1=Mi(C_0) и C_1=S то C_1=attr; p.312(attr).2. Если C_1=G(C_0) и $C_0<C_2<C_1$ то C_2=(S,A,crd,hyphen, quot) и C_2≠Art;	[1~~4~~]VIIIjam1(ingpt) $\xrightarrow{2}$ $\xrightarrow{2}$ IV1(sg.mf) [Нет VIIIjam1(ingpt) $\xrightarrow{2}$ $\xrightarrow{2}$ I1(sg.mf), VIIIjam1(ingpt) $\xrightarrow{2}$ I2(sg.mf)]
				p.321(A). Gm=S и не (it3,wh); p.312(A).7. Если C_1=G(C_0) и C_1= =S и $C_1<C_2<C_0$ и C_0≠partpass1 то C_1=Adv;	[Нет I1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt), I2(sg.mf) $\xrightarrow{m}$ VIIIjam1(ingpt) IV1(sg.mf) $\xrightarrow{m}$ VIIIjam1(ingpt)]
				p.321(adverb). Gm=V и не aux;	[1~~45~~]IIII1(fin.past) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt) [1~~46~~]IIII1(partpass1) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt) [1~~47~~]IIII1(partperf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt) [1~~48~~]VI1(fin.past) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt)

1	2	3	4	5	6
VIII	1	V.A.Adverb	ingpt		(49)VI1(partpass1) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt) ~~(50)~~VI1(partperf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjam1(ingpt)
		V	ingabs	p.323. Есть M2; p.321. M2=S; p.312(ingabs).4. Если C_1=Mi(C_o) и $C_1 \neq$M1(C_o) то $C_o < C_1$;	Нет VIIIjam1(ingabs)

VIII JAM1+ +ING1		S	sg.mf	p.323(ING1). СНЯТЬ Усл (C_o)= =Sbl(Есть Mi);
				p.321(ING1).1. СНЯТЬ Усл (C_o)= =Sbl(M1=X); СНЯТЬ Усл (C_o)=Sbl (M2=X);
				p.321(ING1).2. M1=<of7 $\xrightarrow{2}$>S, <by4 $\xrightarrow{2}$>S,S.poss;
				p.321(ING1).3. M2=<of7 $\xrightarrow{2}$>S,S.attr;
				p.312(attr).2. Если C_1=G(C_o) и $C_o < C_2 < C_1$ то C_2=S,A,crd,hyphen, quot) и $C_2 \neq$Art;
					~~(51)~~VIII[jam1+ING1](sg.mf) $\xrightarrow{2}$ IV1(sg.mf)
					[Нет VIII[jam1+ING1](sg. mf) $\xrightarrow{2}$ I1(sg.mf), VIII[jam1+ING1](sg.mf) $\xrightarrow{2}$ $\xrightarrow{2}$ I2(sg.mf)]

I	2	3	4	5	6
VIII JAM1+ +ING1	1	S	sg.mf	p.321(sg.mf). Gm=S,V.partpass1, V.inf; p.312(sg.mf).1. Если C_1=Gm(C_0) и C_0≠temp то C_0=attr; p.312(attr).1. Если C_1=G(C_0) то C_0<C_1;	[Нет I1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIII[jam1+ING1](sg.mf), I2(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIII[jam1+ING1](sg.mf), IV1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIII[jam1+ING1](sg.mf)]
VIII JAMMING	1	S	sg.mf	p.321. M1=<by4 $\xrightarrow{2}$ S; M2=S.attr,<of7 $\xrightarrow{2}$ S; p.312(attr).2. Если C_1=G(C_0) и C_0<C_2<C_1 то C_2=(S,A,crd,hyphen, quot) и C_2≠Art;	(52)VIIIjamming1(sg.mf) $\xrightarrow{2}$ $\xrightarrow{2}$ IV1(sg.mf) [Нет VIIIjamming1(sg.mf) $\xrightarrow{2}$ $\xrightarrow{2}$ I1(sg.mf) VIIIjamming1(sg.mf) $\xrightarrow{2}$ $\xrightarrow{2}$ I2(sg.mf)]
				p.321(sg.mf). Gm=S,V.partpass1, V.inf; p.312(sg.mf).1. Если C_1=Gm(C_0) и C_0≠temp то C_0=attr; p.312(attr).1. Если C_1=G(C_0) то C_0<C_1;	[Нет I1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjamming1(sg.mf), I2(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjamming1(sg.mf), IV1(sg.mf) $\xrightarrow{m}$ $\xrightarrow{m}$ VIIIjamming1(sg.mf)]

Как видно из таблицы I, уже на первом этапе анализа происходит снятие значительной части омонимии: отвергаются единицы IIpa1(fin.3sg pres), IIbe4(fin.3sg pres), IIIterm1(fin.past), IIIterm1(partperf), VImodulate1(partperf), VIIIjam1(ingger), VIIIjam1(ingabs) . Соответственно, снимаются и все установленные с этими единицами потенциальные связи: (8), (10), (36), (38), (45), (47), (50). Таким образом, всего на первом этапе устанавливается 45 разных потенциальных связей.

Таблица 2.

1	2	3	4	5	6
VI MODU-LATE	1	V	fin.past	p.322. M1= "прибор", "энергия"; M2= "прибор", "энергия", "величина";	
				p.213(amplitude1)."величина";	нет (31)
				p.213(jam1)."воздействие" и "отрицательность";	нет (32), (34)
				p.213(jamming1) "воздействие" и "отрицательность"	нет (33), (35)
				p.323(fin). Если нет $C_1 \cap C_0$ и C_1=(than,as,whatever,whoever) и нет $C_2 \subset C_0$ и C_2=(crd,what,whatever,whichever) то есть M1(C_0)	нет VI1(fin.past)

1	2	3	4	5	6
VI	1	V	fin.past	p.323. Есть M2	
VIII JAM1	1	V	ingpt	p.322. M2="прибор", "предикат"; p.213(amplitude1). "величина";	нет (44)
JAM1+ +ING1		S	sg.mf	- " -	нет (51)
JAMMING	1	S	sg.mf	- " -	нет (52)

Таблица 3.

I THIS	2	S.pronpers ant	sg.mf	p.324. G≠copul	1. нет (8)
II BE	2	V.copul.aux	fin.3sg.pres	p.324(fin).1. Если C_1=Mi(C_o) и C_1≠M1(C_o) и C_1≠Mm(C_o) и C_o<C_1<C_2 то C_2≠Mi(C_o)	2.Если есть (I) то нет (9) и нет (IO) и нет (II) и нет (I2)
				p.325.7. Если C_1=M1.1 и C_2=M2.2 то $T(C_1) \leq T(C_2)$; p.213(amplitude1) "величина"; p.213(this1). "факт" и "местоименность" и "определенность";	3.Если (2) то нет (7) и нет (II) и нет (I2) 4.Если (4) то нет (7) и нет (9) 5.Если (5) то нет (7) и нет (9)

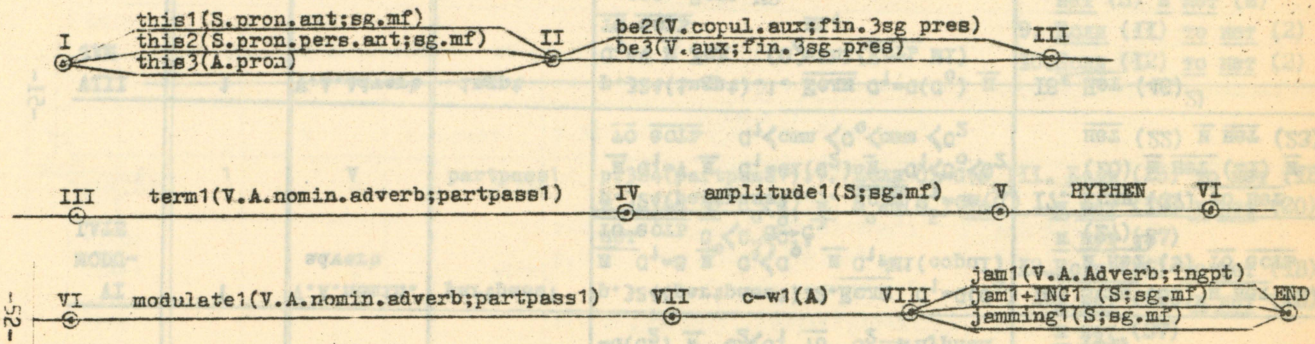
	2	3	4	5	6
	2	V.copul.aux.fin.3sg.pres		p.213(jam1). "воздействие" и "отрицательность"; p.213(jamming1). "воздействие" и "отрицательность";	6. Если (6) то нет (7) и нет (9) 7. Если (7) то нет (2) и нет (4) и нет (5) и нет (6) 8. Если (9) то нет (4) и нет (5) и нет (6) 9. Если (II) то нет (2) I0. Если (I2) то нет (2)
-ed	1	V	partpass1	p.324(partpass1).1. Если C_1=G(C_0) и C_1=S и $C_0 < C_1$ и C_0=G(C_2) то нет $C_0 < C_2 < C_1$;	II. Если (25) то нет (I8) и нет (I9) и нет (20) и нет (37) I2. Если (26) то нет (I8) и нет (I9) и нет (20) и нет (37)
-ing HYPHEN	1			p.324.1. Если C_1=G(C_0) и C_2=G(C_1) и $C_1 < C_3 < C_2$ и $C_3 \neq$hyphen то $C_3 \xrightarrow{m}$ hyphen	I3. Если (30) то нет (28) и нет (29)
				p.324.2. Если C_1=G(C_0) и C_2=G(C_1)	I4. Если (30) то нет (I8)

I	2	3	4	5	6
V HYPHEN	1			и $C_2 \prec C_1$ то $C_2 \xrightarrow{m}$ hyphen	
				p.324.3. Если $C_1=G(C_o)$ и $C_1=G(C_2)$ и $C_2 \prec C_1$ то $C_2 \xrightarrow{m}$ hyphen	I5. Если (30) то нет (24)
VI MODU-LATE	1	V.A.nomin. adverb	partpass1	p.324(partpass1).2.Если $C_1=G(C_o)$ и C_1=S и $C_1 \prec C_o$ и $C_1 \neq$Mi(copul) то есть $C_o \rightarrow C_3$	I6. Если (39) и нет (2) и нет (9) то есть (27)[1]
				p.324(partpass1).3. Если $C_1=Gm(C_o)$ и C_1=V и $C_1=Gi(C_2)$ и $C_1 \prec C_o \prec C_2$ то есть $C_1 \prec cmm \prec C_o \prec cmm \prec C_2$	I7. Если (37) то нет (20) и нет (21) и нет (22) и нет (23)
VIII JAM	1	V.A.Adverb	ingpt	p.324(ingpt).1. Если $C_1=G(C_o)$ и C_1=V и Усл (C_o=Sbl(Есть Mi) то есть Mi;	I8. Нет (46)[2]
				p.323. Есть M2	

1) Связи (39) и (27) несовместимы, поскольку обе связывают одни и те же слова, а между двумя словами разрешается устанавливать не более одной синтаксической связи. Поэтому условие I6 фактически означает, что связь (39) не может быть включена в дерево фразы без одной из связей: (2) или (9).

2) Если сохранить связь (46), где jam1(ingpt) имеет глагольного "хозяина" term1(partpass1), не будет выполнено требование обязательности M2(jam1), т.к. единственная потенциальная связь, реализующая эту валентность, отвергается по соображениям семантики (см. таблицу 2).

Вспомогательное линейное представление 2.



Комментарий к обработке матрицы потенциальных связей.

Как видно из матрицы, слова this 3, be 2 и be3 не имеют потенциальных "хозяев" и, следовательно, могут быть включены в синтаксическую структуру анализируемой фразы только в качестве вершины этой структуры. This 3, однако, будучи прилагательным, не может стоять в вершине дерева зависимостей, в связи с чем это значение имеющейся в нашей фразе словоформы this следует отбросить. Предположим, что в качестве вершины нашего дерева выступает be3 (тем самым из рассмотрения исключаются все связи, в которых участвует be 2).

Согласно требованиям единственности связей сильного управления, а также обязательности заполнения валентности MI(be 3), в дереве должна иметься ровно одна из перечисленных в строке be 3 связей: (13), (14), (15), (16), (17). Связи (13) и (14) соединяют be 3 с this 1 или this 2, соответственно;при этом ни this 1, ни this 2 сами не участвуют ни в каких других связях, так что получить связное дерево на этом пути невозможно. Если же выбрать одну из связей (15), (16) или (17), то тем самым окажутся запрещены связи (13) и (14) - единственные, с помощью которых можно включить в дерево фразы, имеющее своей вершиной be 3, одно из ГМП словоформы this . Таким образом, и здесь связное дерево не будет получено, и, следовательно, единицу be 3 следует исключить из дальнейшего рассмотрения, точно так же, как и все установленные с ней связи. Теперь на роль вершины фразы остается только один претендент: be 2. При этом оказывается, что из двух альтернативных ГМП словоформы this в дерево фразы можно включить только this 1 (у this 2 нет потенциальных хозяев) и притом ровно одним способом - по связи (7). Отсюда, по требованию единственности связей сильного управления, связи (9), (11) и (12) следует считать недопустимыми. По условию 7 таблицы 3 недопустимы также связи (2), (4), (5) и (6). Вычеркнув их из матрицы, мы получим, что в строке be 2

G		M	I			II	
			This1 (sg.mf)	This2 (sg.mf)	This3	Be2 (fin.3sg. pres)	Be3 (fin3sg. pres)
I	This1 (sg.mf)						
	This2 (sg.mf)						
	This3						
II	Be2 (fin.3sg.pres)		(7):1				
	Be3 (fin.3sg.pres)		(13):1	(14):1			
III	Term1 (partpass1)						
IV	Amplitude1 (sg.mf)						
V	HYPHEN						
VI	Modulate1 (partpass1)						
VII	c-w1						
VIII	Jam1 (ingpt)						
	jam1+ING1 (sg.mf)						
	Jamming1 (sg.mf)						

III	IV	V	VI	VII	VIII		
Term1 (partpass1)	Amplitude (sg.mf)	HYPHEN	modulate1 (partpass1)	c-w1	jam1 (ing_{pt})	jam1+ING (sg.mf)	jamming1 (sg.mf)
(1) :2	(2) :2 (9) :1				(4) :2	(5) :2 (11) :1	(6) :2 (12) :1
	(15) :1					(16) :1	(17) :1
	(18) :3		(19) :3 (37) :m	(20) :3	(21) :3	(22) :3	(23) :3
(24) :m		(30) :m	(39) :m				
	(27) :m				(49) :m		
(25) :m	(28) :m		(40) :m	(42) :m			
(26) :m	(29) :m		(41) :m	(43) :m			

остается только одна потенциальная связь, реализующая обязательную для be 2 валентность M2: (1). Следовательно, связь (1) должна быть обязательно включена в дерево фразы, а связи (24), (25) и (26) - вычеркнуты из матрицы (по требованию единственности хозяина любого слова).

В рассматриваемой матрице имеется еще одна единица, которая может быть включена в дерево только одним способом: это HYPHEN с единственным потенциальным хозяином amplitude1 (связь (30)). Таким образом, в любой вариант синтаксической структуры данной фразы должна входить, кроме связей (1) и (7), также связь (30). Это значит, что (по условиям 13 и 14 таблицы 3) из матрицы должны быть вычеркнуты связи (18), (28) и (29), что дает нам единственного потенциального хозяина для amplitude1 (связь (27)) и заставляет вычеркнуть связь (39), симметричную связи (27) (т.е. установленную между теми же двумя словами).

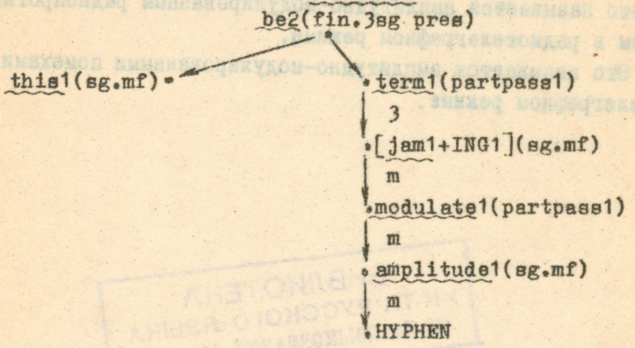
Из еще не рассмотренных связей обязательна одна из связей (19), (20), (21), (22) или (23). Из них связь (19) неприемлема, т.к. при этом в дерево не может быть включена единица c-w1 (она может быть связана либо с [jam1+ING1], либо с jamming, но ни одна из этих двух единиц при наличии в дереве связи (19) в него не включается). Если взять связь (20), то для словоформы jamming в дерево можно включить только ГМП jam1(ingpt) - по связи (49); два других альтернативных ГМП этой словоформы: [jam1+ING1](sg.mf) и jamming1(sg.mf) - в этом варианте структуры фразы участвовать не могут. Но тогда единственный потенциальный хозяин для modulate1 должен устанавливаться по связи (37), а эта связь, согласно условию (17) таблицы 3, запрещает выбранную нами связь (20). Таким образом, рассматриваемый вариант анализа оказывается внутренне противоречивым и должен быть отвергнут. Если мы, далее, возьмем связь (21), т.е. опять обратимся к ГМП jam1(ingpt) словоформы jamming, мы не сможем включить в дерево фразы словоформу c-w. Следовательно, должна быть выбрана либо связь

(22), либо связь (23): а единица jam1(ingpt) - отвергнута как недопустимая в контексте анализируемой фразы.

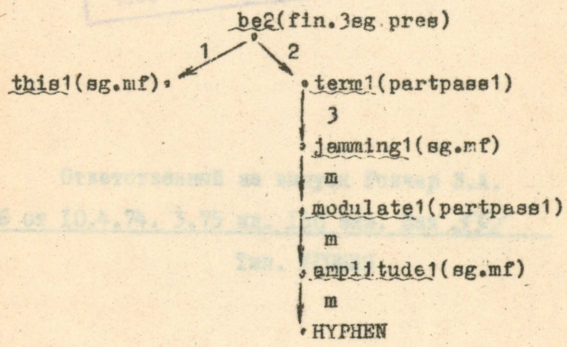
Если будет выбрана связь (22), то в дерево должна быть включена также связь (42), дающая единственного в этом случае хозяина для c-w1 , и связь (40), устанавливающая хозяина для modulate1 (связь (37), как и для ранее рассмотренного случая, невозможна в силу требования 17 таблицы 3). Если же мы выберем связь (23), то из тех же соображений в дерево должны быть включены связи (43) и (41).

Таким образом, для анализируемой фразы могут быть получены две альтернативные синтаксические структуры, отличающиеся только выбором ГМП словоформы jamming:

(I)



(II)



Стремиться к однозначному выбору той или иной из этих структур нам представляется нецелесообразным, так как из английской фразы неясно, что именно имеется в виду: процесс помехообразования ([jam1+ING1]) или сами помехи (jemming1).

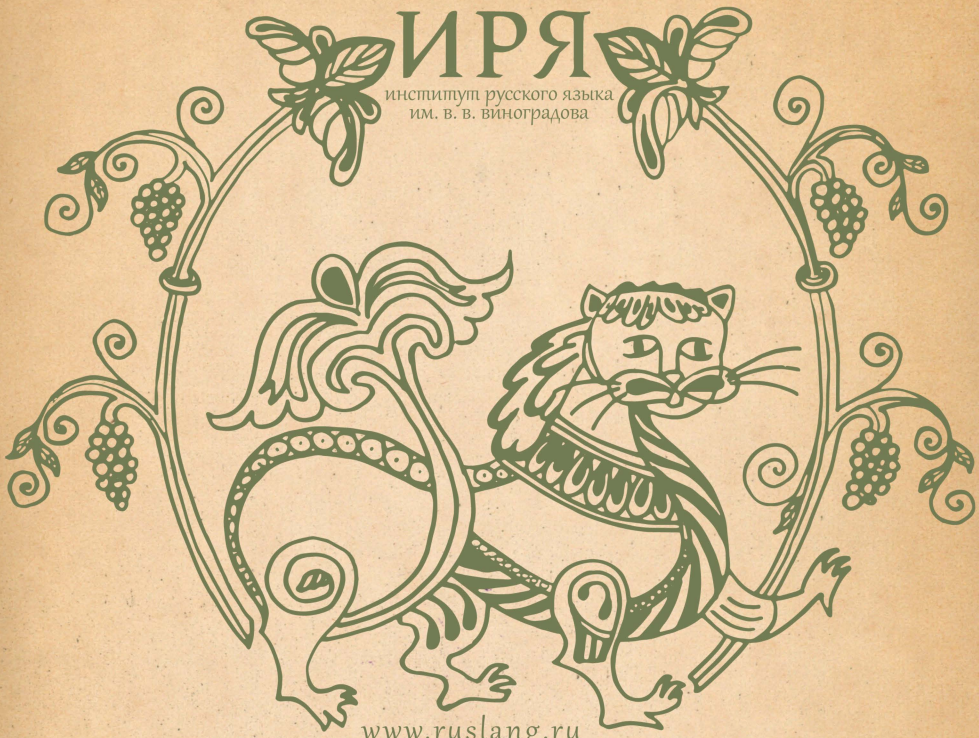
Если теперь заменить аналитическую форму be2(fin.3sg.pres)- $\xrightarrow{2}$ term1(partpass1) одним узлом term1(fin.ind.pass. imperf.improgr.pres) и подставить на место всех английских единиц, имеющихся в дереве, русские эквиваленты, предусмотренные для них в словаре АРМАС, то при адекватно организованном русском синтезе для рассматриваемой фразы будут получены переводы:

(I) Это называется амплитудно-модулированным радиопротиводействием в радиотелеграфном режиме.

(II) Это называется амплитудно-модулированными помехами в радиотелеграфном режиме.

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР

ИРЯ
институт русского языка
им. в. в. виноградова



www.ruslang.ru

Ответственный за выпуск Гончар Л.А.

№ А02086 от 10.4.74. 3,75 пл. 150 экз. Зак. 591

Тип. МГПИИЯ

Бесплатно.