

© 2006 г. О. В. ДРАГОЙ

РАЗРЕШЕНИЕ СИНТАКСИЧЕСКОЙ НЕОДНОЗНАЧНОСТИ: ПРАВИЛА И ВЕРОЯТНОСТИ

Естественному языку свойственны предложения, в которых существует более одного способа синтаксического анализа, и носитель языка должен преодолеть эту неоднозначность, чтобы адекватно проинтерпретировать сообщение. Данные психолингвистических экспериментов свидетельствуют о существовании межъязыковых различий в понимании конструкций типа *Кто-то застрелил служанку актрисы, которая стояла на балконе*. Для объяснения этого факта выдвинут ряд гипотез: одни лишь модифицируют принцип универсального синтаксического анализа, другие предлагают новые правила, специфичные для конкретного языка, третьи ориентированы на вероятностный подход. В настоящей статье приводятся результаты экспериментального и корпусного исследований на материале русского языка, свидетельствующие в пользу языковой специфичности механизмов анализа и использования в процессе разрешения указанного вида многозначности как правил, так и информации о лингвистических вероятностностях.

0. ВВЕДЕНИЕ

Неоднозначность присуща любому естественному языку на всех его уровнях. Когда человек воспринимает (читает, слышит) некий текст, ему нужно одновременно принимать множество решений относительно его интерпретации. В частности, это касается синтаксического анализа: как слова группируются во фразы и каким образом новая фраза должна присоединяться к предыдущей. Проблема неоднозначности, как правило, не актуальна для говорящего или пишущего – в нормальной коммуникации люди не вкладывают в высказывание одновременно несколько смыслов. Выбор между различными вариантами интерпретации осуществляется только в процессе понимания (у слушающего или читателя).

Так, предложения типа (1) с определительным придаточным, которое модифицирует одно из трех существительных, входящих в состав сложной именной группы (ИГ), демонстрируют неоднозначность интерпретации:

- (1) *Теракты изменили философию системы безопасности, которая раньше требовала от пилотов выполнения приказов террористов: посадить самолет в любом аэропорту, а затем начать переговоры...* [Известия, 2002.10.27]

При совпадении рода и числа существительных придаточное *которая раньше требовала от пилотов выполнения приказов террористов...* может относиться к любому из трех имен: *философия* (ИГ1), *система* (ИГ2) или *безопасность* (ИГ3), а все предложение имеет три прочтения:

- (1а) *выполнения приказов террористов требовала философия* (придаточное модифицирует ИГ1),
 (1б) *выполнения приказов террористов требовала система* (придаточное модифицирует ИГ2),
 (1в) *выполнения приказов террористов требовала безопасность* (придаточное модифицирует ИГ3).

Очевидно, в данном случае неоднозначность возникает из-за невозможности установить референциальные отношения в предложении вне более широкого контекста: рассматривая высказывание изолированно, мы можем по-разному построить синтаксические связи между частями предложения. Тем не менее, в условиях реальной коммуникации мы успешно преодолеваем эту неоднозначность. Вопрос заключается в том, какие принципы лежат в основе интерпретации синтаксически неоднозначных предложений подобного типа.

Далее, в первом разделе статьи, мы рассмотрим первые разработки проблемы синтаксической неоднозначности в отечественной и западной традициях, а также обоснуем выбор конструкции, находящейся в фокусе нашего исследования, – предложения с определительным придаточным, которое модифицирует сложную ИГ. Проблема синтаксической неоднозначности рассматривается в рамках двух основных направлений: одно предлагает анализ, основанный на правилах, второе – вероятностное прогнозирование разрешения неоднозначности. Описанию соответствующих моделей посвящен второй раздел. В третьем представлены проведенные нами на материале русского языка психолингвистическое и корпусное исследования, направленные на выявление факторов, которые определяют анализ неоднозначных предложений с придаточным, модифицирующим сложную ИГ. Наконец, в последней части статьи обобщены результаты выполненной работы.

1. ИСТОРИЯ ВОПРОСА

Работы отечественных авторов, посвященные вопросу разрешения синтаксической неоднозначности, начали появляться в 1960-х годах и были связаны с задачей автоматической обработки текстов естественного языка [Иорданская 1967; Гладкий 1985; Дрейзин 1988]. Ранее данная проблема исследовалась, главным образом, в стилистике. При этом, как отмечает Л.Н. Иорданская [Иорданская 1967], рассматривались наиболее явные случаи синтаксической неоднозначности. За прошедшие сорок лет проблема не потеряла актуальности, и приходится признать, что в настоящий момент “полностью автоматизированные процедуры, даже самые эффективные, не могут обеспечить надежного разрешения лингвистической неоднозначности” [Иомдин и др. 2005].

При разработке систем автоматического анализа и перевода текстов используется несколько способов преодоления многозначности: формулируются и уточняются правила обработки языковых данных; разрабатываются вероятностные анализаторы, учитывающие статистику употребления языковых единиц; создаются базы знаний, позволяющие учитывать экстралингвистическую информацию [Лазурский и др. 2005]. Представляется продуктивным использование гибридного подхода: например, учет и правил, и вероятностей, что реализовано, например, в лингвистическом процессоре ЭТАП-3 [Богуславский и др. 2003]. Однако пока все эти методы ограничены по эффективности и, безусловно, проигрывают механизмам, которые задействует человек при понимании предложений, имеющих более одного способа анализа. В связи с этим возникает вопрос, насколько указанные подходы соответствуют тем процессам, которые происходят в сознании носителя языка при разрешении многозначности.

Проблема отыскания принципов, лежащих в основе принятия таких решений человеком, получила широкое освещение в работах западных психолингвистов. Первые модели анализа предложения опирались на теорию Хомского [Chomsky 1965]: они утверждали, что механизм анализа подчиняется строгим правилам и является универсальным. Так, предложенная в 70-х годах XX в. Модель заблуждения (Garden Path model) предполагала, что разрешение синтаксической неоднозначности рассматривается как процесс, подчиняющийся двум принципам: принципу Минимального Присоединения (“при присоединении нового материала к конструируемому синтаксическому дереву необходимо использовать как можно меньше узлов”) и принципу Позднего Закрытия (“по возможности присоединять новые лексические единицы к той фразовой группе, которая анализируется в данный момент”) [Секерина 2002: 242]. Таким образом, первые разработки

проблемы разрешения неоднозначности предложений опирались на действие строгих синтаксических правил, которые считались базовыми и универсальными, то есть действующими при понимании предложений любого языка.

Последнее утверждение подверглось сомнению после публикации работы [Cuetos, Mitchell 1988], где исследовалось присоединение определительных придаточных к сложной ИГ. На материале английского и испанского языков авторы рассмотрели конструкции типа:

- (2) *Someone shot the maid of the actress who was on the balcony.*
Alguien disparó contra la criada de la actriz que estaba en el balcón.
Кто-то выстрелил в служанку актрисы, которая стояла на балконе.

Как и в случае примера (1), данное предложение синтаксически неоднозначно и имеет прочтения:

(2a) *На балконе стояла служанка* (придаточное модифицирует ИГ1).

(2б) *На балконе стояла актриса* (придаточное модифицирует ИГ2).

В соответствии с принципом Позднего Закрытия предложения (1) и (2) должны быть проинтерпретированы во всех языках соответственно как: *выполнения приказов террористов требовала безопасность и на балконе стояла актриса*. Однако Ф. Куэтос и Д. Митчелл показали, что предпочтения носителей разных языков различаются. Англоговорящие испытуемые действительно следовали принципу Позднего Закрытия и выбирали интерпретацию (2б). Испаноговорящие же предпочитали интерпретацию (2а), и это явление получило название раннего закрытия. В ходе дальнейших исследований группу языков позднего закрытия дополнили шведский, норвежский, румынский, бразильский португальский, арабский, а раннее закрытие обнаружилось в нидерландском, французском, немецком, хорватском, африкаанс [Fernández 2000].

В связи с выявлением межъязыковых различий в интерпретации предложений типа (1) и (2) исследования неоднозначных конструкций с определительным придаточным приобрели чрезвычайную важность, поскольку этот тип неоднозначности был единственным свидетельством того, что в разных языках одна и та же конструкция анализируется в соответствии с различными принципами. По этой же причине обратились к исследованию понимания предложений типа (1) и мы, чтобы проверить, какие из существующих психолингвистических гипотез подтверждаются материалом русского языка.

2. ПРАВИЛА ИЛИ ВЕРОЯТНОСТИ, УНИВЕРСАЛЬНЫЙ АНАЛИЗ ИЛИ ЯЗЫКОВАЯ СПЕЦИФИЧНОСТЬ

Вопрос о характере операций, которые используются носителем языка в процессе понимания конструкций типа (1) и (2), получил широкое освещение в психолингвистических работах на материале все возрастающего числа языков. Помимо новых гипотез, оставшихся в рамках Модели Заблуждения, появились идеи объяснять процесс разрешения неоднозначности в процессе понимания совершенно новыми структурными принципами, а также вероятностными механизмами. Рассмотрим основные направления, в рамках которых разрабатывается указанная проблема.

2.1. Универсальные правила синтаксического анализа

Пытаясь реабилитировать Модель Заблуждения после обнаруженного в исследовании [Cuetos, Mitchell 1988] факта межъязыковых различий в разрешении синтаксической неоднозначности, последователи идеи об универсальности синтаксического анализа предложили ряд модифицирующих Модель Заблуждения гипотез. Межъязыковые

различия в отношении придаточного к тому или иному члену сложной ИГ объяснялись разницей не в самом синтаксическом анализе, который по-прежнему предполагался универсальным, а в принципах, действующих на других этапах обработки лингвистического материала.

Так, Л. Фразье предложила принцип Относительной Релевантности (Relativized Relevance Principle): все интерпретации предложения должны быть грамматичны, информативны и соответствовать актуальному дискурсу [Frazier 1990]. В частности, была выдвинута идея о действии дополнительного дискурсивного фактора на постсинтаксическом этапе анализа неоднозначных предложений типа (2). Дело в том, что в английском языке, в отличие от испанского, существует способ однозначно указать, *кто стоял на балконе* – так называемый саксонский родительный падеж:

(3) *Someone shot the actress's maid who was on the balcony.*

В соответствии с постулатом Грайса [Grice 1975] – говорящий должен воспроизводить ровно такое количество языковой информации, которое необходимо, чтобы быть понятым, – если англоязычный собеседник будет иметь в виду, что *на балконе стояла служанка*, он выразит свою мысль в форме (3), а не (2), то есть использует саксонский родительный. В испанском такой возможности нет, поэтому в нем, утверждает Л. Фразье, действует принцип Относительной Релевантности: предпочтение с ИГ2 переходит на ИГ1.

Тем не менее, принцип Относительной Релевантности противоречил данным ряда языков: в нидерландском и хорватском, где есть морфологический родительный падеж, наблюдается предпочтение раннего закрытия [Brysbaert, Mitchell 1996; Lovrić 2002], а в румынском без такой возможности – позднее [Ehrlich et al. 1999].

Подобная ситуация повторялась неоднократно. Для объяснения механизмов, лежащих в основе интерпретации неоднозначных предложений с определительным придаточным, предлагалось некое “универсальное” объяснение: определяющими анализ предлагалось считать прагматическую информацию [De Vincenzi, Job 1993; 1995], дискурсивные факторы [Frazier, Clifton 1996], морфосинтаксические характеристики относительного местоимения [Hemforth et al. 1996]; для определительных придаточных даже выделялся особый тип анализа [Frazier, Clifton 1997]. Тем не менее, впоследствии, с привлечением данных новых языков, каждому из этих объяснений находился контрпример [Fernández 2000].

Пожалуй, из всех моделей, опирающихся на универсальные принципы анализа, дальше всех от Модели Заблуждения ушла теория Просодической Сегментации Дж. Фодор [Fodor 1998]. Дж. Фодор предположила, что межъязыковые различия в анализе предложений с определительным придаточным обусловлены просодическими характеристиками языков: такие предложения имеют разное просодическое оформление в разных языках. В работе [Fodor 2001] выдвинута гипотеза Имплицитной Просодии, в соответствии с которой просодический контур проецируется на предложение даже при чтении “про себя”: анализатор выбирает интерпретацию, связанную с наиболее естественным в данном языке просодическим оформлением конструкции, который эксплицирован при говорении. Также Дж. Фодор сформулировала Закон Антигравитации: для более “тяжелых” составляющих предпочтительнее присоединение к более “высоким” узлам синтаксического дерева, для более “легких” – к более “низким” [Fodor 1998]. Причем “вес” составляющих в различных языках неодинаков, что и определяет анализ. Если составляющая является короткой (“легкой”), то наблюдается позднее закрытие.

Отметим, что Закон Антигравитации был подтвержден на материале английского [Fodor 1998], арабского [Abdelghany, Fodor 1999], хорватского [Lovrić, Fodor 2000], французского [Quinn et al. 2000], немецкого [Walter et al. 1999], испанского [Igoa 1999] и русского [Федорова, Янович 2005] языков. Во всех указанных языках наблюдается одна и та же закономерность: короткие придаточные чаще, чем длинные, присоединяются к ИГ2. Однако сама по себе справедливость Закона Антигравитации не приближает нас к

пониманию, почему существуют межъязыковые различия в предпочтении того или иного вида закрытия. Для этого необходимы эксперименты, в которых проверялось бы влияние просодических характеристик языка на анализ.

Итак, в модификациях Модели Заблуждения остается неизменным базовое правило работы синтаксического анализатора – принцип Позднего Закрытия. А межъязыковые различия в отнесении придаточного к тому или иному имени сложной ИГ обусловлены разницей не в самом синтаксическом анализе, а в принципах, действующих на других этапах. Тем не менее, большинство гипотез, предложенных в рамках этого подхода, не может объяснить все разнообразие накопленных языковых фактов. Критика принципа Позднего Закрытия привела к возникновению альтернативного подхода, ориентированного на языковую специфичность анализа. Сохранилась идея о том, что анализ синтаксически неоднозначных предложений подчиняется неким правилам, но предлагаются совершенно новые структурные принципы работы анализатора. Кроме того, в рамках данного направления была предпринята попытка не только определить, как происходит анализ, но и по чему он осуществляется в конкретном языке именно таким образом.

2.2. Специфичные для конкретного языка правила

Модель, ориентированная на языковую специфичность анализа, была предложена Э. Гибсоном и коллегами и получила название **Двухфакторная Модель Гибсона** [Gibson et al. 1996]. Она предполагает, что принцип, благоприятствующий позднему закрытию, действительно активен при синтаксическом анализе, но он взаимодействует, по крайней мере, с одним другим фактором – принципом Близости к Предикату, в соответствии с которым предпочтительно отнесение придаточного к такому имени, которое в синтаксической структуре расположено ближе к вершине предикации.

В качестве фактора, создающего тенденцию к позднему закрытию, Э. Гибсон и коллеги рассматривают Предпочтение Последнего (выбирать в качестве вершины придаточного наиболее близкую к нему ИГ) – вариант принципа Позднего Закрытия, который применим ко всем потенциальным вершинам придаточного и может взаимодействовать с другими влияющими на предпочтение факторами.

Поскольку действие принципа Предпочтения Последнего является мотивированным с точки зрения общих характеристик памяти (последняя поступившая в обработку информация требует меньших усилий для активации), Э. Гибсон и коллеги полагают, что этот фактор имеет одинаковую силу в разных языках. В отношении же фактора Близости к Предикату предлагается считать, что языки отличаются по этому параметру. Так, в ряде языков принцип Близости к Предикату оказывается сильнее принципа Предпочтения Последнего и создает предпочтение к раннему закрытию в случае, когда определительное придаточное модифицирует двучленную ИГ. С другой стороны, в работе [Gibson et al. 1996] утверждается, что удлинение конструкции благоприятствует усилению фактора Предпочтения Последнего. Иными словами, если при интерпретации предложений с определительным придаточным, которое может быть отнесено к одной из двух ИГ, случаи раннего закрытия преобладают (что соответствует активации фактора Близости к Предикату), то в случае трехчленной конструкции фактор Предпочтения Последнего, тем не менее, оказывается сильнее.

Э. Гибсон и коллеги сосредоточились на исследовании неоднозначных конструкций типа (4), в которых определительное придаточное может модифицировать любое имя из трехчленной ИГ:

- (4) *La lámpara cerca de la pintura de la casa que fue dañada en la inundacion.*
The lamp near the painting of the house that was damaged in the flood.
Светильник рядом с картиной дома который(ая) пострадал(а) при наводнении.

Результаты экспериментов показали, что для испано- и англоговорящих испытуемых ИГ2 (*картина*) является наименее предпочтительной вершиной придаточного. В большинстве случаев выбиралась ИГ3 (*дом*), во вторую очередь ИГ1 (*светильник*). Полученные данные не согласуются с действием единственного фактора при разрешении неоднозначности, поскольку кривая распределения предпочтений немонотонна (предпочтения убывают по схеме ИГ3>ИГ1>ИГ2, а не ИГ3>ИГ2>ИГ1). Таким образом, результаты говорят в пользу взаимодействия двух факторов, поддерживающих противоположные тенденции – относить придаточное в первую очередь то к ИГ1, то к ИГ3.

Подобное распределение предпочтений при разрешении неоднозначности в предложениях типа (4), как выяснилось в ходе дальнейших исследований, наблюдается и в японском [Miyamoto et al. 1999], немецком [Walter, Hemforth 2001], нидерландском [Wijnen et al. 1999] языках.

Обращаясь к Двухфакторной Модели Гибсона для интерпретации полученных экспериментальных данных, мы неизбежно сталкиваемся с вопросом: от чего зависит оценка фактора Близости к Предикату в конкретном языке? Почему в одних языках (испанский) он оказывается сильнее, чем в других (английский), когда мы рассматриваем возможность присоединения придаточного к двум ИГ?

С одной стороны, как предположили Э. Гибсон и коллеги [Gibson et al. 1996], предпочтение того или иного типа закрытия может определяться параметрами Универсальной Грамматики: грамматические различия между языками порождают различия в анализе. Следовательно, ожидается, что определенный набор лингвистических параметров будет коррелировать с поведением этого фактора. Причем связь между параметрами Универсальной Грамматики и действием фактора, влияющего на интерпретацию неоднозначных предложений, может быть и непрямой. Иными словами, грамматические характеристики языка не жестко определяют механизм анализа, а влияют на оценку приемлемости различных конструкций в этом языке. Разные языки будут иметь разную частоту встречаемости неоднозначностей конкретного типа и разную частоту того или иного типа закрытия.

Так, Ф. Куэтос и Д. Митчелл [Cuetos, Mitchell 1988] объясняют различия в предпочтениях англо- и испаноязычных носителей в случае с двучленными конструкциями типа (2) допустимостью постановки определения (в частности, прилагательного) перед существительным в этих языках. В испанском определение может стоять между существительным и другим модификатором (предложной группой или определительным придаточным), то есть для вершины придаточного нормально находиться на расстоянии от присоединяемой составляющей. Поэтому в испанском раннее закрытие более ожидаемо и связано с меньшими трудностями понимания, чем в английском.

В свою очередь, Э. Гибсон и коллеги [Gibson et al. 1996] предположили, что для объяснения выбора типа закрытия в конкретном языке нужно обратиться к понятию среднего расстояния между предикатом и его аргументами, которое и определяет силу фактора Близости к Предикации. Чем больше это расстояние, тем более активен должен быть фактор, чтобы позволить синтаксические отношения между линейно неблизкими единицами. С этой точки зрения языки со строгим порядком слов SVO, как английский, будут характеризоваться небольшим значением среднего расстояния и, соответственно, малой силой фактора Близости к Предикату, поскольку аргументы глагола линейно близки к предикату. Наоборот, в испанском, в языке с тем же базовым порядком слов SVO, допускается и порядок VOS, при котором субъект удален от предиката, – следовательно, сила фактора Близости к Предикату увеличивается. На основе предложенной идеи Э. Гибсон и коллеги делают следующее предсказание: языки с порядком слов VOS, VSO, SOV и OSV будут характеризоваться сильной активацией фактора Близости к Предикату, что отразится в предпочтении раннего закрытия. В языках же SVO и OVS ожидается позднее закрытие из-за небольшого допустимого расстояния между предикатом и его аргументами.

Обе предложенные гипотезы [Cuetos, Mitchell 1988] и [Gibson et al. 1996] были опровергнуты фактическим материалом: нидерландский язык, в котором порядок существи-

тельного и прилагательного такой же, как в английском, и бразильский португальский с порядком слов SVO – оба демонстрируют раннее закрытие [Brysbaert, Mitchell 1996; Finger, Zimmer 2000], в то время как для них предсказывалось позднее. Следовательно, расстояние между составляющими ИГ или предикатом и его аргументами не является решающим фактором, на основе которого мы могли бы объяснять существующие межъязыковые различия в интерпретации неоднозначных предложений с определительным придаточным.

Описывая методы, с помощью которых исследователи пытаются решить проблему неоднозначности в системах автоматического анализа текстов на естественном языке, Лазурский и коллеги [Лазурский и др. 2005] отмечают, что эти же методы используются и человеком при естественном анализе: “правильный подход моделирует знание носителем правил и законов родного языка; подход, использующий базы знаний, – не извлекаемые непосредственно из текста знания; статистический подход призван заменить имеющуюся у каждого человека языковую интуицию”. Что касается экстралингвистической информации, которая довольно плохо поддается формализации и экспериментальным исследованиям, то этот компонент не получил пока воплощения в моделях разрешения синтаксической неоднозначности. Идея же влияния вероятностных характеристик языковых единиц на интерпретацию предложений нашла свое место в психолингвистической теории.

2.3. Вероятностный подход

Как альтернатива идее существования правил, определяющих анализ, к исследованию естественного языка был применен вероятностный подход: анализатор независим от грамматики и чувствителен к частоте употребления конструкций определенного типа в языке. Иными словами, анализатор использует только вероятностные отношения между единицами языка, предлагая синтаксическую интерпретацию на основе того, с какой вероятностью анализируемые единицы связывались раньше. При таком подходе мы должны рассматривать принцип Предпочтения Последнего [Gibson et al. 1996] как частноязыковую стратегию, не имеющую ничего общего с базовыми принципами языка или ограничениями на память. Стратегия, благоприятствующая раннему закрытию, имеет с этой точки зрения не меньшую вероятность быть использованной. Ф. Куэтос и Д. Митчелл [Cuetos, Mitchell 1988] обозначают данную гипотезу как **Лингвистический Тюнинг**.

Гипотеза Лингвистического Тюнинга в формулировке Д. Митчелла [Mitchell et al. 1995] предполагает, что решающее воздействие на анализ оказывает прошлый лингвистический опыт носителя конкретного языка. Это означает, что выбор интерпретации предложения определяется конструкциями, наиболее частотными в языке.

Так, например, имеется неоднозначное предложение:

- (5) *Я задаю вопрос, исходя из редакции статьи Конституции, которая требует установления этого обстоятельства для признания неконституционной деятельности партии, и, естественно, исхожу не из программных документов, а из фактов деятельности* [А. Яковлев. Омут памяти].

Если в языке однозначные предложения с ИГ1 в роли вершины придаточного более частотны (пример (6)), чем те, где вершина – ИГ2 или ИГ3 (примеры (7) и (8) соответственно), то и в случае неоднозначности интерпретация в пользу ИГ1 будет предпочтительнее.

- (6) *Под ее руководством была разработана грандиозная программа покорения космоса, которая необычайно высоко подняла престиж державы* [В. Быков. Оборонка].

- (7) *Не останавливаясь здесь более подробно на **понятиях физики живого**, которая будет подробно рассмотрена во второй части, обсудим кратко идеи симметрии-асимметрии применительно к проблемам объектов живой и неживой природы* [В.В. Горбачев. Концепции современного естествознания].
- (8) *Ван Гог **вынашивает идею товарищества художников**, которые превратят в конце концов обывательский городишко Арль в художественную столицу мира или в крайнем случае Франции* [Г. Бурков. Хроника сердца].

Следовательно, в предложении (5) придаточное будет восприниматься как относящееся к ИГ1 и получит интерпретацию: *установления этого обстоятельства требует редакция*.

Непосредственно после того, как была предложена гипотеза Лингвистического Тюнинга, в ее подтверждение нашлись, казалось бы, убедительные свидетельства во французском [Baltazart, Kister 1995], испанском и английском [Cuetos et al. 1996] языках. В корпусе письменных текстов, исследованном в указанных работах, общеязыковое распределение частотности конструкций с той или иной ИГ в роли вершины придаточного соответствовало предпочтениям, демонстрируемым испытуемыми в ходе психолингвистических экспериментов. Это означало, что носители языка действительно могут использовать стратегию обращения к своему прошлому языковому опыту (который отражен в корпусе) при разрешении синтаксической неоднозначности. Однако впоследствии были обнаружены несоответствия между корпусными и экспериментальными частотами, в частности, на материале нидерландского языка [Mitchell, Brysbaert 1998]: в то время, как экспериментальные данные свидетельствовали о предпочтении раннего закрытия, в корпусе выявилось доминирование количества случаев позднего. Спустя несколько лет было найдено объяснение данному противоречию: согласно работам [Desmet et al. 2002; Desmet, Gibson 2003], на распределение раннего и позднего закрытия повлияла одушевленность имен, составляющих сложную ИГ. В экспериментах, описанных в статье [Mitchell, Brysbaert 1998], использовался стимульный материал (одушевленное существительное + одушевленное существительное), наименее представленный в корпусе, что и стало причиной различий результатов корпусного исследования и эксперимента.

3. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

Одной из задач, которые стояли перед нами в настоящем исследовании, было выяснить, какие правила анализа используют (и используют ли вообще) носители русского языка в процессе понимания неоднозначных предложений, в которых придаточное модифицирует одно из существительных сложной ИГ. Отметим, что в большинстве психолингвистических работ, посвященных данной проблематике, рассматривается конструкция с двучленной ИГ, как в предложениях типа (2). Так, для русского языка была выявлена тенденция к раннему закрытию [Fedorova, Yanovich 2005]. Как показано в работе [Gibson et al. 1996], более сложная структура предложений (с трехчленной ИГ) требует, с одной стороны, усложнения моделей, объясняющих их анализ, но, с другой, позволяет обнаружить факторы, эффект которых маскируется действием других в случае двучленной ИГ. Поэтому в настоящем исследовании мы рассматриваем конструкцию с тремя потенциальными вершинами придаточного:

- (9) *Я задаю вопрос, исходя из **редакции статьи Конституции**, которая требует установления этого обстоятельства для признания неконституционной деятельности партии, и, естественно, исхожу не из программных документов, а из фактов деятельности* [А. Яковлев. Омут памяти].

Как и предложения типа (1), данная конструкция (9) потенциально неоднозначна и имеет три возможных прочтения:

- (9а) Установления этого обстоятельства требует редакция.
(9б) Установления этого обстоятельства требует статья.
(9в) Установления этого обстоятельства требует Конституция.

С другой стороны, мы решили проверить, правомерно ли применение вероятностного подхода для объяснения понимания подобных предложений: мы обратились к проверке гипотезы Лингвистического Тюнинга на материале русского языка. Для этого потребовалось сравнить данные проведенного экспериментального исследования, направленного на выявление предпочтений в интерпретации предложений типа (9), и анализа корпуса письменных текстов, в котором отражены частотные показатели встречаемости того или иного вида закрытия в русском языке.

Одним из пунктов, вызвавших спор сторонников и противников гипотезы Лингвистического Тюнинга, стал вопрос о корректности сопоставления данных корпуса (результата порождения речи) и экспериментов на понимание, которые привлекались для доказательства гипотезы наряду с исследованиями порождения экспериментальных конструкций. Дело в том, что в работе [Gibson, Schütze 1999] у англоговорящих испытуемых было выявлено предпочтение раннего закрытия. Однако корпусное исследование показало, что в текстах частота встречаемости позднего закрытия значительно превосходит частоту раннего. Авторы сделали вывод: при анализе неоднозначных предложений не используются частоты, представленные в корпусе; следовательно, принципы, лежащие в основе порождения и понимания речи, должны различаться. Иными словами, существует фактор, работающий при понимании и не активный при порождении. Впоследствии Т. Десмет и Э. Гибсон [Desmet, Gibson 2003] показали, что выявленные в работе [Gibson, Schütze 1999] противоречия между корпусными и экспериментальными данными оказались связаны не столько с различными механизмами, лежащими в основе понимания и порождения, а скорее с самим стимульным материалом, и обосновали корректность сравнения материала, представленного в корпусе, с данными экспериментов на понимание речи. Наиболее строго аналогичность процессов понимания и порождения в случае синтаксической неоднозначности показана в работе [Desmet et al. 2005], где приведены данные эксперимента, выполненного по методике записи движений глаз, и подтверждено, что если учитывать характеристики существенных – потенциальных вершин придаточного (одушевленность и конкретность), то корпусные частоты согласуются с результатами исследования понимания неоднозначных предложений с определенным придаточным.

3.1. Эксперимент-опросник

При планировании исследования мы исходили из нескольких возможных вариантов механизма анализа неоднозначных предложений типа (9). Во-первых, предпочтения испытуемых могли распределяться следующим образом: как и в случае двучленной ИГ, наиболее предпочтительным было бы отнесение придаточного к ИГ1, следующим по степени предпочтительности стала бы ИГ2, и, наконец, в последнюю очередь испытуемые выбирали бы ИГ3. Такая картина предпочтений свидетельствовала бы о том, что тенденция к раннему закрытию распространяется на все составляющие сложной ИГ. Во-вторых, если в русском языке действует только фактор, благоприятствующий раннему закрытию, тогда разницы в процентном выражении выбора ИГ2 и ИГ3 не будет. Оба случая поддерживают гипотезу о наличии единственного активного фактора при разрешении данного типа неоднозначности. Если же в анализ вовлечены как минимум два фактора, тогда кривая распределения предпочтений будет немонотонной. В соответствии с идеями Э. Гибсона и коллег, можно предполагать, что принцип Предпочтения Последнего действительно взаимодействует с фактором Близости к Предикату, и ИГ1 с ИГ3 будут иметь преимущество перед ИГ2 в процентном выражении.

Стимульный материал состоял из 51 предложения, среди которых экспериментальных было 16, отвлекающих – 32, тренировочных – 3. Отвлекающие и тренировочные

предложения были однозначными, экспериментальные – многозначными и оставляли возможность выбора любой из трех интерпретаций.

- (10) *Почтальон попросил жену разнести письма, потому что его помощник заболел.* (отвлекающее предложение)
(11) *Нам позвонил брат напарника водителя, который вчера видел ограбление.* (экспериментальное предложение)

Испытуемые. В эксперименте приняли участие 36 русскоязычных испытуемых в возрасте от двадцати до двадцати пяти лет, студенты и аспиранты МГУ им. М.В. Ломоносова.

Процедура проведения эксперимента. Испытуемые тестировались индивидуально, каждый эксперимент продолжался примерно 15 минут. На экране компьютера появлялось предложение, например (10) или (11), испытуемый читал его вслух и затем, нажав кнопку “Дальше”, переходил к следующей странице, где предлагалась перифраза этого предложения с тремя вариантами интерпретации, порядок которых был сбалансирован:

- (10а) *Разносить письма пришлось...*
Почтальону Жене Помощнику
(11а) *Ограбление видел...*
Брат Напарник Водитель

Испытуемый выбирал один из вариантов, произнося его вслух и одновременно нажимая на соответствующую клавишу, и переходил к следующему предложению.

Обсуждение результатов. Анализ результатов эксперимента показал, что случаи отнесения придаточного к ИГ1 составили 39,2%, к ИГ2 – 27,3%, к ИГ3 – 33,5%. То есть наиболее предпочтительной была ИГ1, далее следовала ИГ3, и реже всего вершиной становилась ИГ2. При этом зависимость между способом отнесения придаточного и местом, которое вершина занимает в составе сложной ИГ, является статистически значимой: $F(2,11)=7,83$, $p<0,001$; $F(2,39)=3,37$, $p<0,04$ (однофакторный анализ АНОВА).

Полученные результаты частично могут быть объяснены двухфакторной моделью, предложенной в работе [Gibson et al. 1996]. В терминах модели мы наблюдаем в русском языке активацию фактора Близости к Предикату, который оказывается сильнее фактора Предпочтения Последнего. Это проявляется в немонотонности распределения предпочтений испытуемых:

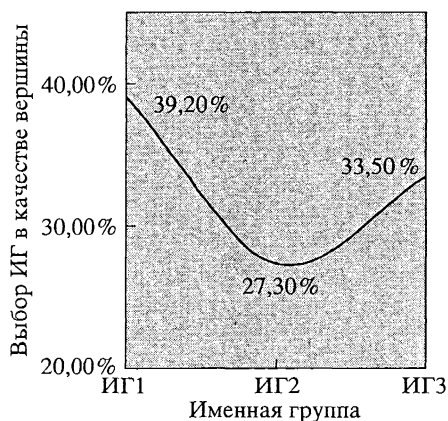


Диаграмма 1. U-образное распределение предпочтений при интерпретации неоднозначных предложений с определительным придаточным и трехчленной ИГ в русском языке

Однако модель Э. Гибсона и коллег предполагает, что удлинение конструкции благоприятствует усилению фактора Предпочтения Последнего. Действительно, это положение нашло подтверждение в испанском [Gibson et al. 1996], немецком [Walter, Hemforth 2001], японском [Miyamoto et al. 1999] и нидерландском языках [Wijnen et al. 1999], в которых наблюдается раннее закрытие, когда выбор заключается между двумя ИГ, но случаи позднего более частотны, если сложная ИГ состоит из трех имен. Однако данные русского языка отражают другое распределение предпочтений: интерпретация в пользу ИГ1 при двучленной конструкции [Fedorova, Yanovich 2005] сохраняется, как показал проведенный эксперимент, и при трехчленной. Причем разница между процентным соотношением выбора ИГ1 и ИГ2 в русском и, например, в нидерландском языках оказывается одинаковой: 39,2% ИГ1 и 33,5% ИГ3 в нашем эксперименте и 38% и 44% соответственно в исследовании [Wijnen et al. 1999].

Таким образом, подвергается сомнению положение авторов работы [Gibson et al. 1996] о факторе Предпочтения Последнего как универсальном и первичном, только модулирующемся фактором Близости к Предикату, сила которого, в свою очередь, зависит от длины сложной ИГ, содержащей вершину определительного придаточного. Большее число случаев выбора ИГ1 русскоговорящими испытуемыми свидетельствует в пользу того, что фактор Близости к Предикату (или любой другой фактор, благоприятствующий раннему закрытию) оказывается достаточно сильно активирован и в случае удлинения сложной ИГ. Следовательно, именно он должен рассматриваться как первичный, оказывающий наибольшее влияние на интерпретацию неоднозначных конструкций с определительным придаточным в русском языке.

С другой стороны, методика, использованная в проведенном эксперименте (опросник), относится к типу так называемых опосредованных (off-line), когда исследованию подвергаются более поздние этапы анализа предложения, на которых решающими могут оказаться факторы постсинтаксической природы. Наоборот, непосредственные (on-line) методики рассматриваются как отражающие ранние этапы принятия решений относительно синтаксического анализа. Так, например, в работе [Kamide, Mitchell 1997] описаны два эксперимента на материале японского языка: в ходе первого (опросник, опосредованная методика) выявлено предпочтение раннего закрытия для неоднозначных предложений с определительным придаточным, модифицирующим двучленную ИГ; второй же эксперимент (саморегуляция чтения, непосредственная методика) показал, что изначально, в процессе анализа, испытуемые относят придаточное к линейно наиболее близкой ИГ, но по завершении синтаксического анализа целого предложения предпочтение изменяется в пользу раннего закрытия. Чтобы исключить возможность того, что полученные нами результаты зависят от метода исследования, мы планируем провести новый эксперимент – с использованием методики саморегуляции чтения – на том же стимульном материале. Если эксперимент покажет то же распределение предпочтений, это станет более сильным аргументом в пользу Двухфакторной Модели Гибсона, поскольку пока мы не можем исключить возможность того, что фактор, благоприятствующий раннему закрытию, начинает действовать на постсинтаксических стадиях анализа.

3.2. Анализ корпуса

Для проверки гипотезы Лингвистического Тюнинга нам потребовалось сравнить полученные экспериментальные данные с частотностью того или иного вида закрытия в предложениях с определительным придаточным в русском языке. Чтобы определить частоту встречаемости таких конструкций, мы предприняли корпусное исследование.

Описание корпуса. В исследовании был использован Национальный корпус русского языка (www.ruscorpora.ru). Корпус представлен письменными текстами различных жанров и стилей в пропорции, отражающей современный русский литературный язык. В момент проведения исследования (весна 2005 года) Национальный корпус русского языка содержал 13 246 текстов общим объемом 35 238 929 словоупотреблений.

Отбор и классификация контекстов. Поиск осуществлялся по лексико-грамматическим признакам слов, входящих в конструкцию, и параметру расстояния между составляющими конструкции: нас интересовала последовательность трех существительных (причем два последних должны быть в форме родительного падежа и без предлогов), за которыми следует местоимение-прилагательное “который”, являющееся началом определительного придаточного. Поиск по корпусу выдал наличие около трех тысяч описанных контекстов. Часть полученных контекстов не соответствовала предъявляемым к конструкции требованиям и была удалена из совокупности примеров, которые подверглись исследованию. Так, из рассмотрения удалялись предложения, в которых члены сложной ИГ не являлись зависимыми:

- (12) *Карьерной была **речь маршала Конева, которого Жуков спас от расстрела, заступившись за него перед Сталиным*** [А. Яковлев. Омут памяти].

Значительную часть нерелевантных для исследования контекстов составляли примеры из корпуса с неснятой омонимией:

- (13) *И приятная новость: **нашелся “отец” милой зебры, которая в метро постоянно звонит своим родственникам*** [Вечерняя Москва, 2002.04.11].

После отбора интересующих нас примеров “вручную”, мы получили выборку из 822 контекстов. В отношении каждого из оставшихся контекстов было определено, к какой из трех ИГ относится придаточное. Предложения, для которых нельзя было принять такое решение в силу их неоднозначности, были исключены из выборки, например:

- (14) *Нет, у нее явное **изобилие пространства лица, которое ничто не может спасти*** [Г. Щербакова. Мальчик и девочка].

В результате поиска по корпусу было обнаружено 584 контекста, релевантных для анализа, – однозначных предложений, содержащих определительное придаточное, относящееся к трехчленной ИГ, как пример (15):

- (15) *Пространство между пластинками несегментированной части мезодермы превращается в **полость тела зародыша, которая в организме человека подразделяется на брюшную, плевральную и перикардальную полости*** [Р. Самусев, Ю. Селин. Анатомия человека].

Полученные контексты подверглись дальнейшей классификации по количеству имен – потенциальных вершин определительного придаточного, которые в силу своих морфосинтаксических характеристик могут возглавлять придаточное. Так, в предложении (16) возможен единственный выбор вершины в пользу существительного *группа*, поскольку местоимение *которая* согласуется по роду и числу только с ним. В предложениях (17) и (18) возможно соответственно два и три формальных “претендента” на вершину придаточного.

- (16) *Клуб обещал помочь Детскому фонду, и они действительно многое сделали, например взяли на себя **спонсорство группы врачей, которая** отправилась на две недели в Тбилиси со своим оборудованием и провела там серию сложных операций* [В. Фетисов. Овертайм].
- (17) *Все это я говорю для того, чтобы подвести аудиторию к **пониманию источников юмора, которым** пронизаны “Двенадцать стульев”* [В. Катаев. Алмазный мой венец].

(18) ...Н. Томским, З. Азгуром и др. создали **образец монумента вождя, которому и следовали скульпторы рангом ниже** [М. Чегодаева. Соцреализм. Мифы и реальность].

Мы предположили, что для предложений типа (16), (17) и (18) соотношение предложений с различными типами закрытия также может различаться. Говорящий или пишущий, порождая высказывания (17) и (18), может и не осознавать, что в процессе понимания формальные характеристики имен сложной ИГ предоставляют альтернативу в выборе вершины придаточного. Эта альтернатива успешно разрешается, поскольку в понимании задействованы операции над смыслами: в предложении (17) мы относим придаточное к ИГ3 (*юмор*), а в (18) – к ИГ1 (*образец*). Однако без проверки нельзя было сбрасывать со счетов фактор наличия в сложной ИГ формальных вершин – имен с морфосинтаксическим оформлением, удовлетворяющим требованиям к вершине придаточного.

Обсуждение результатов. Процентное соотношение ИГ1, ИГ2 и ИГ3, являющихся вершинами определительного придаточного, отражено в табл. 1. Как видно, наиболее частотны ИГ1-вершины (259 из 584 контекстов). То есть в 44,3% случаев порождения конструкции со сложной ИГ и определительным придаточным вершиной является ИГ1. Реже встречаются предложения, где придаточное модифицирует ИГ2 и ИГ3 (28,1% и 27,6% соответственно). Статистический анализ (однофакторная АНОВА) показал, что позиция вершины в составе сложной ИГ значимо коррелирует с наблюдаемым способом отнесения придаточного: $F(2,1749)=24,57$, $p<0,00001$. Причем соотношение имен, возглавляющих придаточное, приблизительно одинаково для всех трех типов конструкций – с одной, двумя и тремя формальными вершинами. Поэтому мы отбрасываем выдвинутую ранее гипотезу о возможном влиянии фактора количества формальных вершин на распределение и далее рассматриваем выборку контекстов, не классифицируя их по данному признаку.

Присоединение определительного придаточного к ИГ1, ИГ2 и ИГ3 по данным корпуса

Кол-во формальных вершин придаточного	ИГ1		ИГ2		ИГ3		Всего
	Кол-во	%	Кол-во	%	Кол-во	%	
1 (предложения типа (16))	143	42,3	97	28,7	98	29	338
2 (предложения типа (17))	100	46,7	60	28	54	25,2	214
3 (предложения типа (18))	16	50	7	21,9	9	28,1	32
Всего	259	44,3	164	28,1	161	27,6	584

Таким образом, данные проведенного эксперимента и корпусного исследования свидетельствуют в целом в пользу гипотезы Лингвистического Тюнинга. В эксперименте испытуемые чаще относили придаточное к ИГ1, в корпусе текстов в большинстве случаев придаточное модифицирует ту же ИГ. Это означает, что носители русского языка действительно могут использовать стратегию обращения к своему прошлому лингвистическому опыту при разрешении синтаксической неоднозначности.

При планировании корпусного исследования мы опирались на исходную и получившую наибольшее распространение формулировку гипотезы Лингвистического Тюнинга [Mitchell et al. 1995]: для объяснения работы синтаксического анализатора необходимо учитывать частотность целой синтаксической конструкции. В широком же смысле вероятностный подход к разрешению синтаксической неоднозначности основан на учете частотности того или иного анализа в прошлом языковом опыте, на предположении, что именно анализ “по аналогии” с большой вероятностью будет правильным. При

этом остается открытым вопрос, частотность каких языковых единиц мы учитываем. Теоретически можно рассматривать языковые вероятности любого уровня: частотность лексем, синтаксических конструкций, дискурсивного контекста. Более того, возможен учет характеристик слов (например, одушевленность существительного), экстралингвистического контекста или индивидуальных предпочтений носителя языка [Pickering et al. 2000].

Что касается моделей, учитывающих только лексические вероятности, то и в их пользу свидетельствуют данные ряда исследований [MacDonald et al. 1994; Spivey-Knowlton, Sedivy 1995; Trueswell et al. 1993], однако этот подход не объясняет предпочтения относительно конструкции с определительным придаточным: как обсуждалось выше, Э. Гибсон и коллеги [Gibson et al. 1996] показали, что в случае конструкции с трехчленной ИГ в испанском языке наблюдается предпочтение позднего закрытия, а для двухчленной – раннее. Поскольку в конструкциях с двумя ИГ употреблялись две из трех лексем, использованных в экспериментальном материале с тремя ИГ, в данном случае нельзя объяснить выбор того или иного предпочтения лексическими частотами.

Хотя в работе [Mitchell et al. 1995] обозначалась возможность использования смешанного подхода (учета частот синтаксических конструкций и лексики), первый из авторов гипотезы Лингвистического Тюнинга, Д. Митчелл, подчеркивал, что исключительно синтаксическая версия гипотезы согласовывалась со всеми данными, полученными на тот момент. Тем не менее, сейчас уже ясно, что такая версия не является обоснованной и что смешанный вариант имеет большую объяснительную силу. Так, например, Т. Десмет и коллеги [Desmet et al. 2005] предлагают рассматривать категорию одушевленности как дискурсивную переменную, полагая, что это скорее приведет к прогрессу в понимании разрешения исследуемого типа неоднозначности, чем структурная модель. К такому же выводу пришли и авторы работ [Gilboy et al. 1995; Frazier, Clifton 1996; Traxler et al. 1998; Hemforth et al. 2000]. Конечно, как замечено в исследовании [Gilboy et al. 1995], смещение фокуса от синтаксиса к дискурсу не освобождает нас от необходимости искать фактор, управляющий присоединением определительного придаточного к сложной ИГ. Это только значит, что подобные структуры могут нам сказать гораздо меньше о работе анализатора, нежели раньше думалось.

Как показали Т. Десмет и коллеги [Desmet et al. 2002; 2005] и на корпусных, и на экспериментальных данных, при анализе неоднозначных предложений с определительным придаточным необходимо учитывать характеристики существительных – потенциальных вершин придаточного (их одушевленность и конкретность). На материале нидерландского языка было выявлено, что в конструкциях, содержащих одушевленную ИГ1, раннее закрытие встречается чаще, и в большинстве случаев раннего закрытия одушевленные ИГ1 являются конкретными существительными.

В проведенном эксперименте (см. раздел 3.1) мы не контролировали специально такие характеристики ИГ, как одушевленность. В числе шестнадцати экспериментальных единиц были представлены все восемь логически возможных сочетаний одушевленных (О) и неодушевленных (Н) ИГ1, ИГ2 и ИГ3: ООО, ННН, ОНО, НОН, ООН, ОНН, НОО, ННО. Однако примеров сочетаний ННН было пять, ННО – три, ОНО и НОН – по два, остальных – по одному. Несмотря на то, что экспериментальный материал нельзя в строгом смысле назвать сбалансированным, не наблюдалось такой явной неравномерности, как в эксперименте, описанном в работе [Mitchell, Brysbaert 1998], где большинство примеров содержало ИГ с двумя одушевленными существительными. К тому же в эксперименте Т. Десмета и коллег [Desmet et al. 2002] показано, что ИГ1-одушевленная увеличивает процент раннего закрытия, а в нашем случае более двух раз (если бы сочетание каждого вида встречалось два раза, материал был бы сбалансирован) представлены сочетания ННН, ННО, которые не содержат ни на месте ИГ1, ни на месте ИГ2 одушевленного имени. Следовательно, можно предположить, что в проведенном эксперименте влияние категории одушевленности было минимальным.

С другой стороны, мы проанализировали корпусный материал на предмет распределения одушевленных существительных. Оказалось, что одушевленность ИГ действи-

тельно тесно связана с модификацией этой ИГ придаточным: в случае, когда одна из ИГ является одушевленной (ОНН, НОН, ННО), придаточное относится именно к ней. Если в конструкции имеется несколько одушевленных ИГ (ООО, ООН, ОНО, НОО), то вершиной являлась та из них, которая расположена ближе к вершине предикации (для ООО, ООН, ОНО – ИГ1, для НОО – ИГ2). Наконец, в случае ННН придаточное чаще относится к ИГ1. Иными словами, наблюдается следующая закономерность: в корпусе чаще всего вершинами определительных придаточных являются ИГ1, при этом если позицию ИГ1 занимает одушевленное существительное, то с вероятностью 75% она является вершиной придаточного. Следовательно, материал русского языка свидетельствует о том, что, рассматривая гипотезу Лингвистического Тюнинга в ее “синтаксическом” варианте (по Д. Митчеллу) и обнаруживая данные в ее пользу, необходимо учитывать и характеристики составляющих исследуемой конструкции, поскольку вопрос о языковых единицах, на которые “настраивается” анализатор, еще не получил окончательного разрешения.

Более того, обращаясь к порождению речи, мы обнаруживаем тенденцию связывать одушевленные сущности с позицией субъекта. В работе [Bock et al. 1992] описан эксперимент, в ходе которого испытуемых просили описать картинку, содержащую одушевленного участника в роли пациента и неодушевленного в роли агенса. И испытуемые демонстрировали тенденцию к образованию пассивных конструкций, где одушевленный пациент занимал позицию субъекта. Аналогично, когда испытуемых просили оценить степень приемлемости предложения, они приписывали больше баллов в случае одушевленного субъекта, чем неодушевленного [Corrigan 1986]. Таким образом, как отмечено в работе [Кибрик 2004], одушевленность является немаркированным когнитивным свойством участника ситуации, занимающего позицию субъекта. Та же идея нашла отражение в шкале одушевленности, в соответствии с которой выстраивается иерархия реализации тех или иных синтаксических средств [Silverstein 1976].

Описанная тенденция связывать одушевленность с позицией субъекта действия и может выражаться в наблюдаемом предпочтении присоединять определительное придаточное к одушевленному имени. В определительном придаточном местоимение, замещающее соответствующее полнозначное существительное, чаще всего находится в позиции подлежащего [Desmet et al. 2005], а появление одушевленного подлежащего наиболее ожидаемо. Следовательно, местоимение скорее будет замещать одушевленное, чем неодушевленное имя, а все придаточное будет скорее отнесено к одушевленному существительному.

Таким образом, очевидно, что выбор вершины определительного придаточного не определяется только настройкой на синтаксические конструкции, которые составляют лингвистический опыт носителя языка. Модель, ориентированная на объяснение явления позднего / раннего закрытия в различных языках, не может игнорировать характеристики конкретных слов, заполняющих конструкцию, и должна предусматривать использование этой информации при анализе. В частности, это касается одушевленности имен, входящих в сложную ИГ, которая модифицируется определительным придаточным. Для того, чтобы строго показать, насколько тип закрытия коррелирует с оформлением ИГ по категории одушевленности, требуется новое исследование, в котором экспериментальный материал был бы сбалансирован по этому признаку и содержал блоки однотипных предложений, различающихся только сочетанием одушевленных и неодушевленных имен в пределах сложной ИГ. Такой эксперимент позволит говорить о влиянии фактора одушевленности на разрешение синтаксической неоднозначности в процессе понимания.

4. ЗАКЛЮЧЕНИЕ

Наибольшее распространение в исследованиях синтаксической неоднозначности получили два подхода:

- предлагающий некие правила, в соответствии с которыми происходит понимание предложений с несколькими возможными вариантами интерпретации,

- ориентированный на вероятность того или иного вида анализа на основе предшествующего опыта взаимодействия с подобными конструкциями.

В рамках направления, связанного с постулированием правил, спорят, одинаковы ли эти правила для всех языков или специфичны для каждого. Подобно тому, как специалисты в области автоматической обработки текстов опытным путем пришли к тому, что использование смешанной модели “правила + вероятности” является наиболее эффективным, в рамках психолингвистической теории, как нам представляется, гибридный подход также имеет преимущества.

Проведенный нами эксперимент показал, что при понимании неоднозначных предложений с определительным придаточным, модифицирующим трехчленную ИГ, испытуемые демонстрируют одновременное действие противоположных тенденций – относить придаточное к ИГ1 и к ИГ3. Наблюдаемый факт соответствует Двухфакторной Модели Гибсона, которая предлагает правило действия двух факторов: Предпочтения Последнего и Близости к Предикату, взаимодействие которых и отражается на картине предпочтений.

С другой стороны, во второй части предпринятого исследования мы изучили корпусные данные с целью подвергнуть проверке гипотезу Лингвистического Тюнинга, связывающую указанные межъязыковые различия с частотностью того или иного анализа в прошлом лингвистическом опыте. Как и экспериментальный материал, корпус демонстрирует наибольшее процентное выражение случаев отнесения придаточного к ИГ1. Однако, ориентируясь только на анализ вероятностей, мы не смогли бы предсказать, что второй по степени предпочтительности в роли вершины придаточного выступает наиболее близкая к нему ИГ3, поскольку по данным корпуса ИГ–2 и ИГ–3 вершины представлены в русском языке приблизительно в равном процентном соотношении.

Кроме того, более детальный анализ корпуса показал, что тип закрытия зависит не только от того, какое место занимает имя в составе сложной ИГ, но и от характеристики этого имени по категории одушевленности. Это свидетельствует о недостаточности рассмотрения языковой настройки только на синтаксические конструкции.

Таким образом, накопленные к данному моменту экспериментальные и корпусные данные говорят о том, что для адекватного описания механизмов разрешения синтаксической неоднозначностью в процессе понимания необходимо привлекать модели, учитывающие и вероятности употребления языковых единиц (причем не только синтаксического уровня), и правила, составляющие языковую компетенцию носителя языка.

СПИСОК ЛИТЕРАТУРЫ

- Богуславский и др. 2003 – *И.М. Богуславский, Л.Л. Иомдин, В.Г. Сизов, И.С. Чардин*. Использование размеченного корпуса текстов при автоматическом синтаксическом анализе // Труды международной конференции “Когнитивное моделирование в лингвистике–2003”. Варна, 2003.
- Дрейзин 1988 – *Ф.А. Дрейзин*. Синтаксическая омонимия // Машинный перевод и прикладная лингвистика. М., 1988.
- Гладкий 1985 – *А.В. Гладкий*. Синтаксические структуры естественного языка в автоматизированных системах общения. М., 1985.
- Иомдин и др. 2005 – *Л.Л. Иомдин, И.М. Богуславский, А.В. Лазурский, Л.Г. Митюшин, В.Г. Сизов, Л.Г. Крейдлин, А.С. Бердичевский*. Интерактивное разрешение неоднозначности в процессе автоматической обработки текстов: проблемы и перспективы // Труды Международного семинара Диалог. М., 2005.
- Иорданская 1967 – *Л.Н. Иорданская*. Синтаксическая омонимия в РЯ (с точки зрения автоматического анализа и синтеза) // НТИ. 1967. № 5.
- Кибрик 2004 – *А.Е. Кибрик*. Лингвистическая реконструкция когнитивной структуры // Первая Российская конференция по когнитивной науке. Тезисы докладов. Казань, 2004.
- Лазурский и др. 2005 – *А.В. Лазурский, А.С. Бердичевский, Л.Г. Крейдлин, Л.Г. Митюшин, В.Г. Сизов*. Интерактивное разрешение лексической и синтаксической неоднозначности в

- системах автоматической обработки естественного языка // Сборник работ стипендиатов компании Яндекс (http://company.yandex.ru/grant/2005/01_Lazursky_102921.pdf).
- Секерина 2002 – И.А. Секерина. Психолингвистика // Современная американская лингвистика: фундаментальные направления. М., 2002.
- Федорова, Янович 2005 – О.В. Федорова, И.С. Янович. Разрешение синтаксической многозначности в русском языке: роль длины и структуры придаточного // Труды Международного семинара Диалог. М., 2005.
- Abdelghany, Fodor 1999 – H. Abdelghany, J.D. Fodor. Low attachment of relative clauses in Arabic // Poster presented at AMLaP'99, Edinburgh, UK, September 23–25, 1999.
- Baltazart, Kister 1995 – D. Baltazart, L. Kister. Corréation entre détermination et sélection d'un anaphorisé dans ne structure N de N // Paper presented at the Séminaire "Anaphore et Référence", Nancy (CRIN), 20–22 September, 1995.
- Bock et al. 1992 – J.K. Bock, H. Loebell, R. Morey. From conceptual roles to structural relations: Bridging the syntactic cleft // Psychological review. 1992. 99.
- Brysbaert, Mitchell 1996 – M. Brysbaert, D.C. Mitchell. Modifier attachment in sentence parsing: Evidence from Dutch // Quarterly journal of experimental psychology. 49A. 1996. 3.
- Chomsky 1965 – N. Chomsky. Aspects of the theory of syntax. Cambridge (Mass.), 1965.
- Corrigan 1986 – R. Corrigan. The internal structure of English transitive sentences // Memory and cognition. 1986. 14.
- Cuetos, Mitchell 1988 – F. Cuetos, D.C. Mitchell. Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish // Cognition. 1988. 30.
- Cuetos, Mitchell, Corley 1996 – F. Cuetos, D.C. Mitchell, M.M.B. Corley. Parsing in different languages // M. Carreiras, J.E. García-Albea, N. Sebastián-Gallés (eds.). Language processing in Spanish. Mahwah (NJ), 1996.
- Desmet et al. 2002 – T. Desmet, M. Brysbaert, C. De Baecke. The correspondence between sentence production and corpus frequencies in modifier attachment // Quarterly journal of experimental psychology. 2002. 55A (3).
- Desmet et al. 2005 – T. Desmet, C. De Baecke, D. Drieghe, M. Brysbaert, W. Vonk. Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account // Language and cognitive processes. 2005. 20.
- Desmet, Gibson 2003 – T. Desmet, E. Gibson. Disambiguation preferences and corpus frequencies in noun phrase conjunction // Journal of memory and language. 2003. 49(3).
- De Vincenzi, Job 1993 – M. De Vincenzi, R. Job. Some observations on the universality of the Late-Closure strategy // Journal of psycholinguistic research. 22. 1993. 2.
- De Vincenzi, Job 1995 – M. De Vincenzi, R. Job. An investigation of Late Closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation // Journal of experimental psychology: Learning, memory, and cognition. 21. 1995. 5.
- Ehrlich et al. 1999 – K. Ehrlich, E. Fernández, J.D. Fodor, E. Stenshoel, M. Vinereanu. Low attachment of relative clauses: New data from Swedish, Norwegian and Romanian // Poster presented at the 12th Annual CUNY conference on human sentence processing. March 18–20. New York, 1999.
- Fedorova, Yanovich 2005 – O. Fedorova, I. Yanovich. Early preferences in RC attachment in Russian: The effect of working memory differences // Paper presented at FASL-14. May 6–8. Princeton, 2005.
- Fernández 2000 – E. Fernández. Bilingual sentence processing: relative clause attachment in English and Spanish. Doctoral dissertation. CUNY graduate center. New York, 2000.
- Finger, Zimmer 2000 – I. Finger, M. Zimmer. Relative clause attachment preference in Brazilian Portuguese. Unpublished manuscript, 2000.
- Fodor 1998 – J.D. Fodor. Learning to parse? // Journal of psycholinguistic research. 27. 1998. 2.
- Fodor 2001 – J.D. Fodor. Prosodic disambiguation in silent reading // Paper presented at the 32nd meeting of the North East linguistic society (NELS). New York, 2001.
- Frazier 1990 – L. Frazier. Parsing modifiers: Special purpose routines in the human sentence processing mechanism? // D.A. Balota, G.B. Flores d'Arcais, K. Rayner (eds.). Comprehension processes in reading. Hillsdale (NJ), 1990.
- Frazier, Clifton 1996 – L. Frazier, C.J. Clifton. Construal. Cambridge (Mass.), 1996.
- Frazier, Clifton 1997 – L. Frazier, C.J. Clifton. Construal: Overview, motivation, and some new evidence // Journal of psycholinguistic research. 26. 1997. 3.
- Gibson et al. 1996 – E. Gibson, N. Pearlmutter, E. Canseco-Gonzalez, G. Hickok. Cross-linguistic attachment preferences: Evidence from English and Spanish // Cognition. 1996. 59.

- Gibson, Schütze 1999 – E. Gibson, C.T. Schütze. Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency // *Journal of memory and language*. 1999. 40.
- Gilboy et al. 1995 – E. Gilboy, J.M. Sopena, C. Clifton, L. Frazier. Argument structure and preferences in the processing of Spanish and English complex NPs // *Cognition*. 1995. 54.
- Grice 1975 – H.P. Grice. *Logic and conversation* // P. Cole, J.L. Morgan (eds.). *Speech acts*. New York, 1975.
- Hemforth et al. 1996 – B. Hemforth, L. Konieczny, C. Scheepers. Syntactic and anaphoric processes in modifier attachment // Poster presented at the 9th Annual CUNY conference on human sentence processing. March 21–23. New York, 1996.
- Hemforth et al. 2000 – B. Hemforth, L. Konieczny, H. Seelig, M. Walter. Case matching and relative clause attachment // *Journal of psycholinguistic research*. 29. 2000. 1.
- Igoa, Carreiras, Meseguer 1998 – J.M. Igoa, M. Carreiras, E. Meseguer. A study on Late Closure in Spanish: Principle-grounded vs. frequency-based accounts of attachment preferences // *Quarterly journal of experimental psychology*. 51A. 1998. 3.
- Kamide, Mitchell 1997 – Y. Kamide, D.C. Mitchell. Relative clause attachment: Non-determinism in Japanese parsing // *Journal of psycholinguistic research*. 1997. 26.
- Lovrić 2002 – N. Lovrić. It's the prosody that matters in Croatian // Poster presented at the 15th Annual CUNY conference on human sentence processing. March 21–23. New York, 2002.
- Lovrić, Fodor 2000 – N. Lovrić, J.D. Fodor. Relative clause attachment in sentence parsing: Evidence from Croatian // Poster presented at the 13th Annual CUNY conference on human sentence processing. March 30 – April 1. La Jolla (CA), 2000.
- MacDonald et al. 1994 – J.L. MacDonald, N.J. Pearlmutter, M.S. Seidenberg. Lexical nature of syntactic ambiguity resolution // *Psychological review*. 1994. 101.
- Mitchell, Brysbaert 1998 – D.C. Mitchell, M. Brysbaert. Challenges to recent theories of cross linguistic variation in parsing: Evidence from Dutch // D. Hillert (ed.). *Sentence processing: A cross linguistic perspective*. San Diego (CA), 1998.
- Mitchell et al. 1995 – D.C. Mitchell, F. Cuetos, M.M.B. Corley, M. Brysbaert. Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records // *Journal of psycholinguistic research*. 24. 1995. 6.
- Miyamoto et al. 1999 – E.T. Miyamoto, E. Gibson, N.J. Pearlmutter, T. Aikawa, S. Miyagawa. A U-shaped Relative Clause attachment preference in Japanese // *Language and cognitive processes*. 1999. 14 (5/6).
- Pickering et al. 2000 – M.J. Pickering, M.J. Traxler, M.W. Crocker. Ambiguity resolution in sentence processing: evidence against frequency-based accounts // *Journal of memory and language*. 2000. 43.
- Quinn et al. 2000 – D. Quinn, H. Abdelghany, J.D. Fodor. More evidence of implicit prosody in reading: French and Arabic relative clauses // Poster presented at the 13th Annual CUNY Conference on human sentence processing. March 30–April 1. La Jolla (CA), 2000.
- Silverstein 1976 – M. Silverstein. Hierarchy of features and ergativity // R.M.W. Dixon (ed.). *Grammatical categories in Australian languages*. Canberra, 1976.
- Spivey-Knowlton, Sedivy 1995 – M. Spivey-Knowlton, J.C. Sedivy. Resolving attachment ambiguities with multiple constraints // *Cognition*. 1995. 55.
- Traxler et al. 1998 – M.J. Traxler, M.J. Pickering, C. Clifton. Adjunct attachment is not a form of lexical ambiguity resolution // *Journal of memory and language*. 1998. 39.
- Trueswell et al. 1993 – J.C. Trueswell, M.K. Tanenhaus, C. Kello. Verb-specific constraints in sentence processing – separating effects of lexical preference from garden-path // *Journal of experimental psychology: Learning, memory, and cognition*. 1993. 19.
- Walter, Hemforth 2001 – M. Walter, B. Hemforth. Attachment preferences of extraposed and adjacent relative clauses following three-site NPs in German // Poster presented at the 11th Annual CUNY conference on human sentence processing. March 19–21. New Brunswick (NJ), 2001.
- Walter et al. 1999 – M. Walter, B. Hemforth, L. Konieczny, H. Seelig. Same size sisters in German? // Poster presented at the 12th Annual CUNY conference on human sentence processing. March 18–20. New York, 1999.
- Wijnen et al. 1999 – F. Wijnen, C. Troos, H. Quené. Prosodic phrasing and relative clause attachment in a three-site context // Poster presented at the 12th Annual CUNY conference on human sentence processing. March 18–20. New York, 1999.