ВОПРОСЫ ЯЗЫКОЗНАНИЯ

№ 3 1996

КРИТИКА И БИБЛИОГРАФИЯ

ОБЗОРЫ

© 1996 г. И.А. СЕКЕРИНА

АМЕРИКАНСКИЕ ТЕОРИИ СИНТАКСИЧЕСКОГО АНАЛИЗА ПРЕДЛОЖЕНИЯ В ПРОЦЕССЕ ПОНИМАНИЯ

Синтаксический анализ предложения в процессе понимания (syntactic sentence processing, syntactic parsing) (далее "синтаксический анализ предложения") представляет собой составную часть процесса понимания предложения в целом. Считается, что с психолингвистической точки зрения аппарат понимания предложения человеком (анализатор, процессор) включает в себя фонетико-фонологический анализ каждого слова, лексическую и морфологическую информацию, построение синтаксической структуры предложения и, наконец, его семантическую интерпретацию. Естественный интерес представляют попытки смоделировать каждую из названных выше частей процесса понимания предложения человеком. Очевидно, что подобное разграничение функций модулей аппарата понимания не является абсолютным — модули взаимодействуют между собой на разных этапах анализа предложения.

В современной американской психолингвистике синтаксический анализ предложения занимает одно из ведущих мест по числу публикаций и конференций. Этот повышенный интерес связан не только с тем, что синтаксический анализ предложения в общем находится в центре внимания теоретической лингвистики, но и с тем фактом, что именно эту часть процесса понимания предложения оказывается возможным изучать с помощью экспериментальных методов. Поскольку целью данного обзора является описание и сравнение нескольких ведущих в американской психолингвистике моделей синтаксического анализа предложения с теоретической точки зрения, экспериментальные исследования в рамках той или иной модели будут по большей части опущены.

Предлагаемый обзор состоит из двух разделов. Раздел 1 представляет собой краткое описание терминов, понятий и языкового материала, используемых в теории синтаксического анализа предложения. Раздел 2 посвящен рассмотрению и сравнению трех ведущих моделей синтаксического анализа предложения в английском языке. В конце обзора кратко суммируются положения теории синтаксического анализа предложения и основные характеристики трех главных моделей.

1. ОСНОВНЫЕ ПОНЯТИЯ И МЕТОДЫ СИНТАКСИЧЕСКОГО АНАЛИЗА ПРЕДЛОЖЕНИЯ

Синтаксический анализ предложения представляет собой процесс определения синтаксической структуры предложения. Традиционно, в соответствии с наиболее распространенной в Америке теорией трансформационно-порождающей грамматики, разработанной Хомским и его последователями [Chomsky 1981; 1986; Секерина 1994], — теории управления и связывания (government and binding theory), синтаксическая структура предложения представляется в виде фразового дерева (phrase-marker), состоящего из фразовых групп и терминальных символов — слов. Синтаксическое дерево позволяет определять структурные зависимости его частей и описывать их свойства. Помимо построения дерева синтаксический анализ предложения должен включать в себя процедуры проверки построенного дерева и его корректи-

ровки (если на каком-нибудь последующем этапе понимания предложения окажется, что первоначальный синтаксический анализ предложения был неверным).

Центральной проблемой для психолингвистических моделей синтаксического анализа предложения является вопрос о том, как люди преобразуют цепочку слов в синтаксическое дерево для представления его структуры. Очевидно, что напрямую ответить на этот вопрос не так-то просто, главным образом из-за того, что синтаксический анализ предложения осуществляется необыкновенно быстро. В результате пока представляется невероятно сложным охарактеризовать аналитические операции (parsing strategies) непосредственно тогда, когда они осуществляются в процессе синтаксического анализа предложения. Речь идет о возможности измерения времени (в миллисекундах), которое уходит на построение того или иного синтаксического дерева и, тем самым, о возможности сравнения нескольких деревьев с точки зрения того, насколько быстрее можно проанализировать разные синтаксические структуры. Особенно ценными в этом плане оказываются многозначные предложения (ambiguous sentences), т.е. такие предложения, которые могут иметь два или более разных синтаксических представления. Информанты слушают (или читают) многозначные предложения, и на основании того, как они справляются с разрешением многозначности, можно опосредованно судить о характере аналитических операций, производимых синтаксическим анализатором в том месте, где эта многозначность возникает.

Синтаксически многозначные предложения бывают двух типов: временно многозначные (temporarily ambiguous) и глобально многозначные (globally ambiguous). Особое внимание в теории синтаксического анализа предложения в процессе понимания отводится временно многозначным предложениям, т.е. таким предложениям, в которых до определенного момента возможны два (или больше) представления, а в одном конкретном месте эта многозначность разрешается в пользу только одного представления. Это связано с тем, что только такие предложения позволяют судить о том, как проходит анализ предложения, какие возникают трудности и как происходит корректировка. Глобально многозначные предложения в этом смысле представляют собой материал, который дает возможность судить о характере анализа только после того, как он полностью завершен. Это, конечно, отнюдь не означает, что теория синтаксического анализа предложения не интересуется однозначными (unambiguous sentences) и глобально многозначными предложениями. Однако именно временно многозначные предложения дают наилучшую возможность исследовать аналитические операции, характеризующие синтаксический анализ в процессе развертывания предложения.

Интересно отметить, что временно многозначные предложения бывают двух типов: такие предложения, которые не вызывают осознанных трудностей при анализе возникшей многозначности (1.3) и такие, которые представляют большие трудности даже тогда, когда известно, что предложение требует особых усилий для понимания (1.4). Предложения второго типа считаются неприемлемыми (unacceptable). В психолингвистике неприемлемые предложения занимают такое же центральное место, как грамматически неправильные (ungrammatical) предложения в теоретическом синтаксисе. В то время как грамматически неправильные предложения помогают установить и определить принципы грамматики, неприемлемые предложения являются практически единственным способом на сегодняшний день, который позволяет пролить свет на характер аналитических операций, осуществляемых человеком в процессе синтаксического анализа предложения. Очевидно, однако, что грамматически неправильные предложения совсем не обязательно должны быть неприемлемыми и наоборот. Таким образом, мы имеем дело с четырехчленной матрицей, представленной в Таблице 1[1].

[1] В теории синтаксиса значком * принято обозначать грамматически неправильные предложения. В теории синтаксического анализа предложения значком # принято обозначать такие временно многозначные предложения, которые представляют значительные трудности для понимания.

Таблица 1

Типы предложений

	Грамматически неправильные	Грамматические правильные
Неприемлемые	**The horse **race** past the barn fell* Лошадь, гнать мимо сарая, упала	*# The horse raced past the barn fell* Лошадь, прогнанная мимо сарая, упала
Приемлемые	**The horse **ride** past the barn fell* Лошадь, проскакать мимо сарая, упала	*The horse ridden past the barn fell* Лошадь, проскакавшая мимо сарая, упала

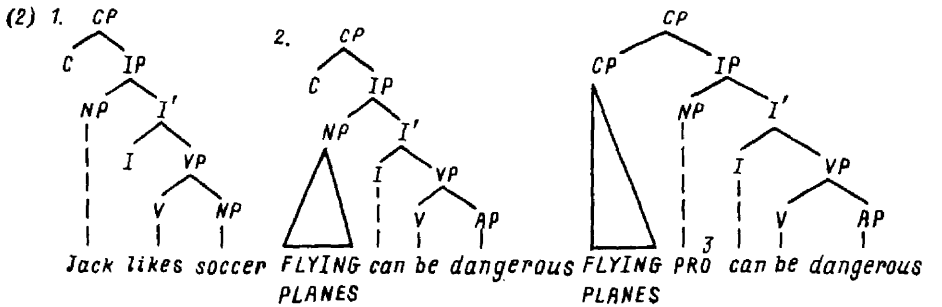
Классические примеры классификации предложений английского языка с точки зрения многозначности представлены в Таблице 2.

Таблица 2

Классификация английских предложений с точки зрения многозначности

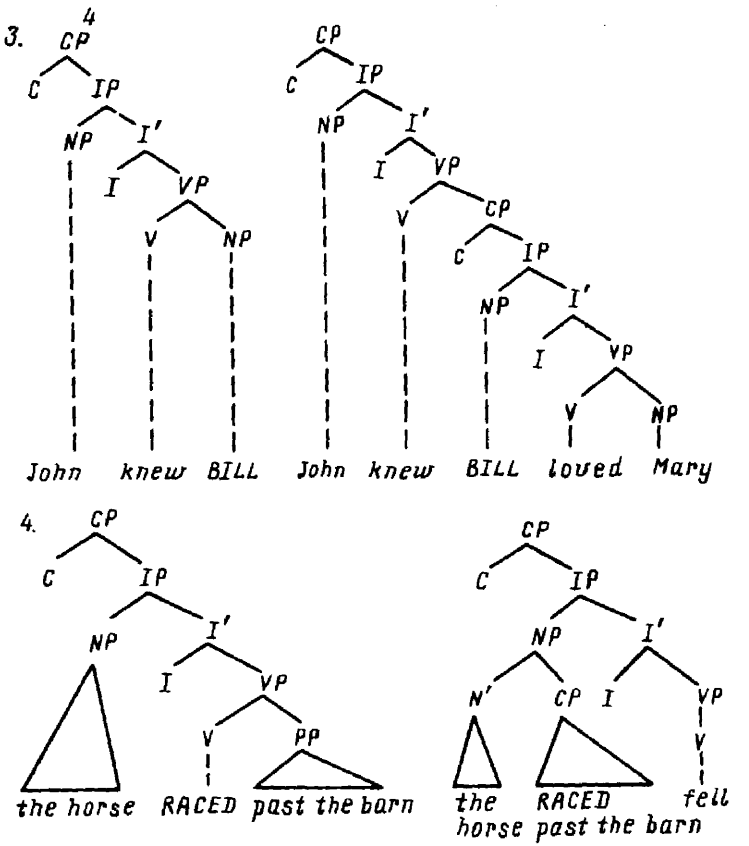
(1)	1. *Jack likes soccer* "Джек любит футбол"	немногозначное предложение
2. *FLYING PLANES*[2] *can be dangerous* "Летающие самолеты могут быть опасны" "Летать самолетами может быть опасно"		глобально многозначное предложение: первый возможный анализ второй возможный анализ
3. *John knew BILL loves Mary* "Джон знал БИЛЛА ... любит Мэри"? "Джон знал, что БИЛЛ любит Мэри"		временно многозначное предложение: первоначальный неправильный анализ последующий скорректированный анализ
4. *# The horse RACED past the barn fell* "Лошадь пробежала мимо сарая... упала"? "Лошадь, ПРОГНАННАЯ мимо сарая, упала"		временно многозначное предложение: первоначальный неправильный анализ последующий скорректированный анализ

Пример (1.1) является предложением, который имеет однозначное синтаксическое представление (2.1) (см. ниже); пример (1.2) может иметь два синтаксических представления (2.2) — в первом словосочетание *FLYING PLANES* является подлежащим простого предложения, а во втором оно представляет собой герундивный оборот, являющийся подчиненным предложением:



[2] Здесь и далее заглавными буквами выделены фразовые группы, которые могут иметь два и более различных синтаксических представления.

[3] Символом PRO в теории управления и связывания обозначается пустой, местоименный элемент (empty pronominal), имеющий неопределенный референт и занимающий позицию подлежащего в инфинитивных и герундивных предложениях.



Та часть предложения, которая может иметь два или более синтаксических представления, называется областью многозначности (ambiguous region), а все, что следует за областью многозначности, — областью снятия многозначности (disambiguating region). Таким образом, в примере (1.4) слово *RACED* является многозначным с точки зрения части речи: оно может пониматься как глагол "ПРОБЕЖАЛА", являющегося сказуемым простого предложения, или в качестве причастия "ПРОГНАННАЯ" в составе причастного оборота, определяющего существительное *horse* "лошадь" (2.4). Область многозначности состоит из всего причастного оборота *RACED past the barn* "ПРОБЕЖАЛА/ПРОГНАННАЯ мимо сарая", а глагольная группа *fell* "упала" представляет собой область снятия многозначности.

---

[4] При построении синтаксических представлений (деревьев) здесь и далее используются символы фразовых групп, принятые в теории управления и связывания [Chomsky 1981; 1986]: CP (complementizer phrase) эквивалентна S', то есть предложению; IP (inflectional phrase) — группа глагольного словоизменения; NP (noun phrase) — именная группа; VP (verbal phrase) — глагольная группа; AP (adjectival phrase) — адъективная группа; PP (prepositional phrase) — предложная группа; Det — детерминатор, то есть артикль. Фразовые группы проецируются от главного слова, возглавляющего фразовую группу — (C, I, V, N, A, P), и иначе называются максимальными проекциями (maximal projections). Между максимальной проекцией и главным словом может находиться промежуточная категория (C', I', V', N', A', P'). Все они являются понятиями X-штрих модуля (X-bar theory).

Исторически сложилось так, что большая часть работ в области теории синтаксического анализа предложения основывалась на материале английского языка, что связано, в первую очередь, с преобладанием до недавнего времени исследований по синтаксису именно этого языка[5]. Вследствие этого представляется важным кратко рассмотреть типологию многозначных предложений, выявленных в английском языке. Эта типология, которая в том или ином виде ныне фигурирует в большинстве работ, посвященных вопросам синтаксического анализа предложения [Frazier, Clifton в печати; Gibson 1991; Pritchett 1992], представлена в (3) (1—5, 10—11 — временно многозначные предложения; 6—9 — глобально многозначные):

(3)1. главное предложение/сокращенное[6] придаточное определительное предложение: *# The horse RACED past the barn fell* "Лошадь, ПРОГНАННАЯ мимо сарая, упала";

2. прямое дополнение/подлежащее относительного придаточного предложения: *BILL knew JOHN liked Mary* "Билл знал [,что][7] Джону нравится Мэри";

3. прямое дополнение/подлежащее: *# While Mary was sewing THE DRESS fell on the floor* "Пока Мэри шила [,] платье упало на пол";

4. придаточное дополнительное/определительное: *# The doctor persuaded the patient THAT was having trouble with to leave* "Доктор уговорил пациента, с которым у него не ладились отношения [,] уйти";

5. конструкции с двойным дополнением: *# I gave the man who was eating THE ICE-CREAM* "Я дал человеку, который ел [,] мороженое";

6. присоединение предложной группы к именной группе/глагольной группе: *The cop saw the spy WITH THE BINOCULARS* "Полицейский увидел шпиона с биноклем/в бинокль";

7. присоединение наречия к высшему/низшему предложению: *Bill thought John died YESTERDAY* "Билл думал [,] что Джон умер вчера";

8. присоединение предложной группы к высшему/низшему предложению: *I put the book that you were reading IN THE LIBRARY* "Я оставил книгу, которую ты читал [,] в библиотеке";

9. придаточное определительное предложение внутри сложной именной группы: *The son of the colonel WHO died in the car accident* "Сын полковника, который погиб в автомобильной катастрофе";

10. сочинительные конструкции: *Bill kissed Mary AND her sister laughed* "Билл поцеловал Мэри [,] и ее сестра засмеялась";

11. лексическая многозначность: *Without HER contributions failed to come* "Без нее/ее пожертвования перестали приходить".

Одним из главных вопросов, возникающих в процессе синтаксического анализа временно многозначных предложений, является вопрос о характере аналитических операций, применяемых человеком в области многозначности. Во время анализа примера (2.4), услышав слово *RACED* "ПРОБЕЖАЛА/ПРОГНАННАЯ", человеку надо принять решение, какое синтаксическое представление построить. В принципе, существуют три возможных способа продолжить анализ предложения в области многозначности. Во-первых, можно построить только одно синтаксическое представление,

---

[5] В последнее время положение стало меняться в лучшую сторону, так как появилось несколько интересных работ на материале других языков. Синтаксический анализ предложения в других языках представляет собой самое перспективное и быстро развивающееся направление в психолингвистике, и объем настоящего обзора не позволяет остановиться на этой теме. Она должна стать предметом отдельной работы.

[6] Сокращенными такие придаточные определительные предложения называются потому, что в них опущены определительное слово *that* "который" и вспомогательный глагол в пассивной конструкции *be* "быть". Ср.: *The horse that was raced past the barn fell.*

[7] Квадратные скобки означают, что в английском оригинале данные слова опущены, однако в таком виде дословный перевод на русский язык оказывается невозможным.

Таблица 3

Классификация моделей синтаксического анализа предложения

Модель	Тип аналитических операций снятия многозначности	Тип используемой информации	Компонент корректировки
Модель заблуждения и ее разновидности	Последовательный анализ	Только синтаксическая информация	Фразер и Райнер: выборочная корректировка; Инуе и Фодор: те же операции, что и при первоначальном анализе (стирание); Фодор и Инуе: особые операции (исправление);
Модель главного слова	Последовательный анализ	Только синтаксическая информация	[Pritchett 1991]: Тэта корректировка [Prichett 1992]: Принцип непосредственной локальности
Модель структурного детерминизма / Модель минимальных обязательств	Последовательный анализ	Только синтаксическая информация	нет отдельного компонента корректировки
Модели Гибсона, Горрелла [Gorell 1987; 1989], Николь и Пикеринга, Хикока	Параллельный анализ	Только синтаксическая информация	нет
Д-Модель Маркуса	Анализ при помощи отсрочки	Только синтаксическая информация	нет

не обращая внимания на другие возможные альтернативы, а затем скорректировать первоначальный анализ, если таковой окажется неправильным. Во-вторых, можно построить все возможные альтернативные синтаксические представления. В-третьих, можно продолжить анализ, отложив разрешение многозначности на некоторое время по крайней мере до тех пор, пока не появится информация, позволяющая снять многозначность. Первый тип аналитических операций снятия многозначности называется последовательным анализом (serial parsing). Этот тип аналитических операций характерен для работ Л. Фразер [Frazier 1978; 1987; 1982; 1990; Frazier 1993] и ее коллег — Дж.Д. Фодор [Foder, Inoue 1993], А. Инуе [Inoue, Fodor 1993], Феррейры и Хендерсона [Ferreira, Henderson 1991]. Второй тип называется параллельным анализом (parallel parsing) и является ведущим принципом в моделях, разработанных Гибсоном [Gibson 1991], Бервиком и Фонгом [Berwick, Fong в печати], Горреллом [Gorrell 1992; 1992 в печати], Николь и Пикерингом [Nicol, Pickering 1993], Хикоком [Hickok 1993]. Третий тип называется анализом при помощи отсрочки (delay parsing). Он характерен для детерминистских моделей Маркуса [Marcus 1980; Marcus, Hindle, Fleck 1983] и частично его последователей Горрелла [Gorrell 1992; 1992 в печати; 1993 рукопись; 1993; 1994 в печати] и Вайнберг [Weinberg 1993; 1994 в печати].

Характер и тип аналитических операций снятия многозначности является одним из трех главных параметров, различающих модели синтаксического анализа предложения в процессе его понимания человеком. Второй параметр определяет виды информации, которые могут оказывать влияние на процесс разрешения многозначности. Здесь

существуют две возможности: только структурная (синтаксическая) информация может быть принята во внимание при разрешении многозначности [Cuetos, Mitchell 1988; De Vincenzi 1989; Ferreira, Henderson 1991; Fodor, Inoue 1993; Frazier 1978; 1987; 1990; Frazier, Clifton 1993 в печати; Frazier, Fodor 1978; Frazier, Rayner 1982; Gorrell 1992; 1992 в печати; 1993; 1993a; 1994 в печати]; или же не только синтаксическая, но и информация, относящаяся к более высоким уровням иерархии, может влиять на процесс разрешения многозначности, а именно, семантическая, прагматическая, контекстная информация и знание реального мира [Altmann, Steadman 1988; Carlson, Tanenhaus 1988; Crain, Steadman 1985; Sells 1985]. И, наконец, для моделей последовательного синтаксического анализа предложения центральным компонентом является процесс корректировки, а также степень сложности корректировки неправильного первоначального анализа [Ferreira, Henderson 1991; Fodor, Inoue 1993; Frazier 1978a; 1982; 1990; 1993; Gorrell 1992; 1992a; 1993; 1993a; Gorrell в печати; Inoue, Fodor 1994; Pritchett 1987; 1987a; 1991; 1992].

В разделе 2 предлагается обзор трех самых влиятельных американских моделей синтаксического анализа предложения на настоящий день: модели заблуждения (Garden-Path parsing model), модели главного слова (ad-Driven parsing model) и модели структурного детерминизма (Structural Determinism parsing model).

В Таблице 3 в обобщенном виде представлена классификация существующих в американской психолингвистике моделей синтаксического анализа предложения.

2. ТРИ ВЕДУЩИЕ МОДЕЛИ СИНТАКСИЧЕСКОГО АНАЛИЗА ПРЕДЛОЖЕНИЯ В АМЕРИКАНСКОЙ ПСИХОЛИНГВИСТИКЕ

Три ведущие модели синтаксического анализа предложения в американской психолингвистике — модель заблуждения, модель главного слова и модель структурного детерминизма основаны на принципе последовательного анализа. Их также объединяет и тот факт, что во всех трех моделях только структурная (синтаксическая) информация может оказывать влияние на способ разрешения многозначности при анализе временно многозначных предложений. Таким образом, в соответствии с классификацией факторов, определяющих характеристики той или иной модели, изложенной в разделе 1, все три модели попадают в один класс. Однако они различаются по третьему параметру, а именно, что из себя представляет компонент корректировки (reanalysis component), то есть аналитические операции, используемые при корректировке первоначального неправильного анализа, и степень сложности корректировки (cost of reanalysis).

2.1. Модель заблуждения (перцептивная модель) Л. Фразер и модификации этой модели

Модель заблуждения была самой первой моделью синтаксического анализа предложения в американской психолингвистике. Интересно происхождение названия этой модели. В английском языке существует фразеологический оборот *to lead down the garden-path*, означающий "вводить кого-то в заблуждение, заводить не туда". Принято считать, что в предложениях типа (1.3—4) человек первоначально избирает неправильный синтаксический анализ предложения, т.к. его вводят в заблуждение специфические аналитические стратегии, которые применяются автоматически, не принимая во внимание другие факторы. Таким образом, временно многозначные предложения типа (1.3—4) получили название "предложений заблуждения" (garden-path sentences)[8]. Как уже было упомянуто выше, предложения заблуждения бывают двух видов — (1.3) и (1.4). Некоторые исследователи используют термин "предложения заблуждения" только для предложений второго типа, представляющих осознанную

[8] Для того чтобы передать разговорный характер этого английского фразеологического оборота, нужно специфическое русское выражение. Например, можно назвать предложения такого типа "сусанинскими" предложениями. Пока же за неимением ничего лучшего я ограничусь прямым переводом с английского и впредь буду называть многозначные предложения "предложениями заблуждения".

трудность для аппарата понимания человеком, иначе называемых каноническими предложениями заблуждения (canonical garden-path sentences). Другие исследователи относят оба вида таких предложений к предложениям заблуждения и делят их на легкие и сложные предложения заблуждения. Как будет видно из последующего обсуждения, разница между этими двумя типами играет роль в теории синтаксического анализа предложения.

Часто модель заблуждения называют еще и перцептивной моделью [Pritchett 1992], так как аналитические операции синтаксического анализа предложения в этой модели мотивированы с функциональной точки зрения их зависимостью от общих когнитивных процессов, например, ограниченностью краткосрочной памяти, и при этом считаются независимыми и никак не связанными с аспектами грамматической теории.

Разработка модели заблуждения началась в середине 1970-х годов с опубликованием статей Т. Бевера [Bever 1970] "Когнитивные основы языковых структур" и Дж. Кимболла "Семь принципов анализа поверхностной структуры предложения в естественном языке" [Kimball 1973] и книги Дж.А. Фодора, Т. Бевера и М. Гарретта "Психология языка" [Fodor, Bever, Garrett 1974], ставших классическими в теории синтаксического анализа предложения.

Каноническая Сентоидальная Стратегия Фодора, Бевера и Гарретта. Фодор, Бевер и Гарретт [Fodor, Bever, Garrett 1974] выдвинули предположение, что модуль синтаксического анализа предложения (human parser) оперирует несколькими эвристическими операциями. Одна из таких операций, называемая канонической сентоидальной стратегией (Canonical sentoid strategy), является причиной того, почему люди первоначально анализируют неправильно предложения типа (1.4): *# The horse RACED past the barn fell* "Лошадь, ПРОГНАННАЯ мимо сарая, упала". Каноническая сентоидальная стратегия состоит в том, что первый фрагмент вида N...V...(N) анализируется как главное предложение, если при этом глагол не помечен как глагол подчиненного предложения. Таким образом, каноническая сентоидальная стратегия объясняет, почему в (1.4) первоначальный анализ фрагмента "Лошадь ПРОБЕЖАЛА мимо сарая..." как главного предложения оказывается неправильным. Еще один класс предложений — предложения с несколькими внутренними придаточными предложениями (center-embedded clauses) — является чрезвычайно проблематичным для синтаксического анализатора: *# The editor authors the newspaper hired liked laughed* "Редактор, который авторам, которых наняла газета, нравится, рассмеялся". И в этом случае каноническая сентоидальная стратегия объясняет трудности анализа данного предложения, связанного со сложностью вычленения последовательности типа N...V в этом предложении.

Семь принципов анализа поверхностной структуры предложения Кимболла [Kimball 1973]. Дж. Кимболл пересмотрел каноническую сентоидальную стратегию, поскольку она была не в состоянии объяснить многие явления, наблюдаемые при синтаксическом анализе предложений, и предложил семь принципов анализа поверхностной структуры предложения[9] (4):

(4) 1. Принцип "сверху вниз" (top down): Анализ естественного языка осуществляется по принципу "сверху вниз" [Kimball 1973 : 20].

2. Принцип ассоциации справа (Right association): Терминальные символы присоединяются к самому нижнему нетерминальному символу [Kimball 1973 : 24].

3. Принцип новых узлов (New nodes): Сигналом для создания нового узла является появление слова незнаменательной части речи [Kimball 1973 : 29].

4. Принцип двух предложений (Two sentences): Могут быть одно-

[9] Более подробное обсуждение классической модели синтаксического анализа предложения, предложенной Кимболлом, см. [Pritchett 1992].

временно проанализированы составляющие не более чем двух предложений [Kimball 1973 : 33].

5. Принцип закрытия (closure): Фразовая группа закрывается как можно раньше, кроме тех случаев, когда следующий анализируемый символ является непосредственно составляющим символом этой группы [Kimball 1973 : 36].

6. Принцип фиксации структуры (Fixed structure): После того, как последний непосредственно составляющий символ фразовой группы сформирован и она закрыта, возвращение назад к этой группе с целью реорганизации ее составляющих символов создает перцептивную нагрузку (perceptual complexity) для анализатора [Kimball 1973 : 37].

7. Принцип анализа (processing): Когда фразовая группа проанализирована, она поступает дальше на этап синтаксического (возможно, семантического) анализа и стирается из краткосрочной памяти [Kimball 1973 : 38].

Принципы 1, 2, 3 и 5 определяют непосредственно эвристические операции построения анализа предложения, в то время как принципы 4, 6 и 7 характеризуют перцептивную нагрузку на анализатор.

Принцип 2 (Ассоциация справа) является одним из двух центральных принципов при объяснении трудностей, связанных с анализом предложений заблуждения. Например, этот принцип объясняет разницу в приемлемости предложений (5а) и (5б), где придаточное определительное присоединяется к самой низшей (правой) фразовой группе, придавая ей нормальное значение: *attractive job* "привлекательная должность" — в (5а) и аномальное — *attractive New York* "привлекательный Нью-Йорк" в (5б). Следует учесть, что прилагательное "привлекательная" может также относиться и к именной группе *the girl* "девушка":

(5) а. *The girl took the job that was attractive*
"Девушка заняла привлекательную должность"
"Привлекательная девушка заняла должность"
б. *The girl went to New York that was attractive*
"Девушка поехала в привлекательный Нью-Йорк"
"Привлекательная девушка поехала в Нью-Йорк".

Принцип 4 (два предложения) объясняет аномальный характер предложений с несколькими внутренними встроенными придаточными (см. пример # *The editor authors the newspaper hired liked laughed*), сложность анализа которых, как видно из возможности их переформулировки, заключается не в многозначности. Принцип 5 (закрытие) – второй по важности в модели Кимболла, в сочетании с принципом 6 (фиксация структуры) он является причиной трудностей при анализе канонических предложений заблуждения типа (1.4).

Сосисочный аппарат Л. Фразер и Дж.Д. Фодор. Именно принципы ассоциации справа и закрытия легли в основу модели заблуждения как таковой, представленной в статье Л. Фразер и Дж.Д. Фодор "Сосисочный аппарат: Новая двухступенчатая модель синтаксического анализа предложения" [Frazier, Fodor 1978] и в диссертации Л. Фразер "О понимании предложений: стратегии синтаксического анализа предложения" [Frazier 1978]. Синтаксический анализ предложения в "сосисочном аппарате" (the sausage machine) осуществляется двухступенчато. На первом этапе, называемом этапом "Предварительной упаковки фраз" (ПУФ) (the preliminary phrase packager), терминальные и фразовые символы приписываются части предложения, состоящей из примерно шести слов. Такое ограничение размера буфера связано с характеристиками краткосрочной памяти. На втором этапе, называемом этапом "управления структурой предложения" (УСП) (the sentence structure supervisor), фразовые "упаковки", поступившие с первого этапа, связываются друг с другом при помощи вышестоящих узлов. Таким образом, ПУФ первоначально пропускает части предложения через свой шестисимвольный буфер, а затем УСП строит цельное синтаксическое представление.

Аналитические операции, связанные с разрешением многозначности в сосисочном аппарате, определяются главным образом через взаимодействие двух эвристических принципов: принципа "минимального присоединения" (МП) (minimal attachment) и принципа "позднего закрытия" (ПЗ) (late closure). Несмотря на то, что многие характеристики сосисочного аппарата существенно устарели и не используются в современных версиях модели Заблуждения, принципы МП и ПЗ до сих пор являются ведущими и своего значения не утратили.

Принцип минимального присоединения: При присоединении нового материала к конструируемому синтаксическому дереву используйте как можно меньше узлов, соблюдая при этом принципы правильности построения в грамматике [Frazier 1987].

Принцип позднего закрытия: По возможности присоединяйте новые лексические единицы к той фразовой группе или тому предложению, которые анализируются в данный момент [Frazier 1987].

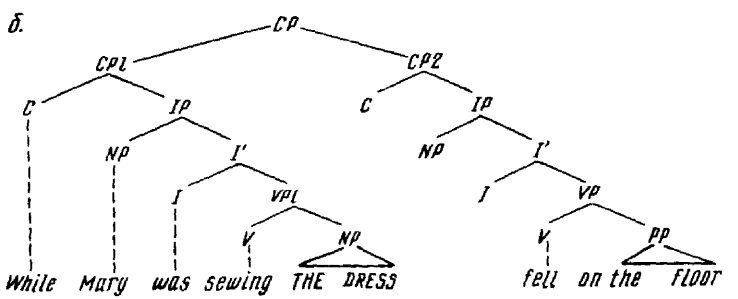
Этих двух принципов оказалось достаточно для объяснения многих легких и сложных предложений заблуждения, и поэтому оказалось возможным отказаться от семи принципов, предложенных в модели Кимболла. В частности, каноническое предложение заблуждения типа (1.4) первоначально анализируется как простое предложение благодаря МП (6):

(6) <[*NP the horse*] [*VP* [*RACED*] [*pp past the barn*]]>
1 2 3 4 5 6

ПУФ группирует шесть терминальных символов в две фразовые группы — NP и VP; затем УСП строит вышестоящий узел. Этот узел может быть или узлом IP (предложение), или узлом NP, распространенным придаточным предложением. Поскольку при построении IP потребуется меньше нетерминальных узлов (см. схему (2.4)), как того требует МП, именно этот вариант и будет выбран. Затем, когда появится слово *fell* "упала", этот анализ окажется неправильным, и потребуется корректировка.

Принцип ПЗ нужен для объяснения другого типа канонических предложений заблуждения:

(7) а. # *While Mary was sewing THE DRESS fell on the floor*
"Пока Мэри шила, [,] ПЛАТЬЕ упало на пол"



Когда появится слово *THE DRESS* "ПЛАТЬЕ", его можно будет присоединить к уже частично построенному синтаксическому представлению либо в качестве прямого дополнения в составе VP1, *либо* в качестве подлежащего в составе CP2, о существовании которого на этот момент пока еще не известно. Как того требует ПЗ, NP "ПЛАТЬЕ" будет присоединена в качестве прямого дополнения в составе VP1, так как именно VP1 анализировалась в данный момент.

Несмотря на то огромное влияние, которое оказала модель сосисочного аппарата на все последующие исследования в области теории синтаксического анализа предложения, в ней не хватало очень серьезного компонента — компонента корректировки.

Как уже было отмечено в разделе 1, модель заблуждения в целом основывается на принципе последовательного анализа, то есть при анализе временно многозначных предложений анализируется только один из возможных вариантов синтаксического представления данного предложения. Если же этот первоначальный анализ окажется впоследствии неверным (а именно тогда, когда появится информация, разрешающая многозначность), то приходится отказываться от этого анализа и корректировать его или строить другой анализ. Таким образом, для того, чтобы модель заблуждения в том виде, в каком она была предложена, могла претендовать на психолингвистическую достоверность, она должна суметь объяснить, каким образом анализатор корректирует первоначальный неправильный анализ и почему сложные предложения заблуждения практически не поддаются корректировке, в то время как легкие предложения заблуждения даже не осознаются человеком как таковые.

Компонент корректировки в модели заблуждения. Л. Фразер и К. Райнер в статье "Ошибки и их исправление в процессе понимания предложения: движения глаз при анализе структурно многозначных предложений" [Frazier, Rayner 1982] предложили три гипотезы о характере возможной процедуры корректировки: корректировка с начала (forward reanalysis), корректировка с конца (backward reanalysis) и выборочная корректировка (selective reanalysis).

Корректировка с начала предполагает, что после обнаружения ошибки человек возвращается к самому началу предложения и по-новому анализирует все предложение целиком в поисках альтернативных вариантов. Корректировка с конца означает, что человек начинает пересматривать слово за словом в обратном порядке, пытаясь найти новый альтернативный вариант. И, наконец, выборочная корректировка предполагает, что анализатор использует любую доступную ему информацию с целью достижения правильного анализа, и поэтому корректировка может быть успешно ограничена одной конкретной фразовой группой. Корректировка будет выявлена при помощи диагностического анализа ошибки, что позволит выборочно сконцентрировать внимание именно на той части предложения, которая послужила причиной первоначального неправильного анализа. Таким образом, процедура корректировки может быть очень конкретной и зависеть от той информации, которой располагает анализатор и которая может быть использована в этой процедуре.

Проведя серию психолингвистических экспериментов на материале легких и сложных предложений заблуждения, Фразер и Райнер пришли к выводу, что результаты этих экспериментов противоречат как гипотезе корректировки с конца, так и гипотезе корректировки с начала. Измерения движений глаз информантов показали, что люди возвращаются только к той части предложения, которая составляет область многозначности, и ни к каким другим частям. Таким образом, был сделан вывод, что движения глаз подтверждают гипотезу выборочной корректировки, так как регрессивные движения глаз информантов были зафиксированы именно назад к области многозначности предложения. Гипотеза выборочной корректировки, по мнению Фразер и Райнера, также отвечает на вопрос, почему существует такая разница в корректировке разных типов предложений заблуждения, когда оказывается, что легко скорректировать одни предложения и почти невозможно другие. Суть дела заключается в том, что при попытке корректировки сложных предложений заблуждения аналитические операции синтаксического анализа опять оказываются не в состоянии правильно проанализировать предложение, в то время как легкие предложения заблуждения успешно поддаются корректировке. Однако, как замечают авторы, гипотеза выборочной корректировки нуждается в многократной проверке и более детальной разработке, так как пока представляется возможным описать только главную идею, лежащую в ее основе.

Наступивший после опубликования статьи Фразер и Райнера [Frazier, Rayner 1982] период в разработке и опробировании модели заблуждения принес много интересных модификаций и дополнений. В частности, две модели, разработанные Инуе и Фодор — ранжированный флажковый последовательный анализатор (ranked flagged serial

parser) и модель диагностики (diagnosis model) [Fodor, Inoue 1993] — являются самыми новыми, и их появление связано со все более возрастающим интересом к анализу фактов других языков [Inoue, Fodor 1994].

Ранжированный флажковый последовательный анализатор (ranked flagged serial parser). Необходимость появления этой модели была вызвана попыткой приложить модель заблуждения к синтаксическому анализу предложения в японском языке. Особенность японского языка заключается в том, что он имеет огромное количество многозначных предложений, так как в его фразовых группах главное слово[10] (ядро, вершина) (head) является последним (head-final). Первоначальная гипотеза состоит в том, что анализатор делает очень мало ошибок, так как он принимает решения о снятии многозначности только тогда, когда появляется информация достаточная для того, чтобы данные решения были наверняка правильными. Такой анализатор может быть назван "информационно ведóмым" (information-paced parsing). Однако идея такого чисто информационно ведóмого анализатора не подтверждается экспериментальными исследованиями на материале японского языка, так как они показывают, что при синтаксическом анализе предложения японский анализатор принимает решения о снятии многозначности немедленно, задолго до того, как появляется информация, разрешающая многозначность. Чисто информационно ведóмый анализатор не годится даже и для английского языка, так как в таком случае в английском языке должны полностью отсутствовать предложения заблуждения. Таким образом, процедура раннего принятия решений о снятии многозначности должна быть встроена в модель синтаксического анализа предложения для обоих языков.

Модель, которую предлагают Инуе и Фодор, принимает обязательные решения во всех случаях многозначности, не дожидаясь ее разрешения, и следует при этом принципу минимального присоединения. Таким образом, анализатор, используя любую доступную ему информацию, принимает решения и присоединяет каждую новую фразовую группу к уже частично построенному синтаксическому представлению даже тогда, когда имеющаяся информация неполна. Такой способ функционирования анализатора связан с его стремлением как можно более последовательно следовать принципу минимального присоединения. Тем не менее, уверенность, с которой анализатор принимает решения о снятии многозначности, пропорциональна качеству информации, имеющейся в его распоряжении для подобных решений, и поэтому в этом смысле такой анализатор может быть назван информационно ведóмым.

Процесс корректировки в такой модели, или анализ по второму разу (second-pass processing, reanalysis), должен удовлетворять двум условиям: 1) выборочная легкость, и 2) универсальность. Анализатор, стремящийся наиболее последовательно следовать принципу минимального присоединения, обречен на то, что он будет часто анализировать японские предложения как легкие предложения заблуждения, и как следствие этого ему придется часто их корректировать. Чем более многозначна синтаксическая конструкция, тем больше существует альтернативных структур, среди которых анализатору предстоит выбрать одну. Однако как уже было отмечено выше, при первоначальном анализе анализатор вычисляет ранжированный коэффициент уверенности в правильности выбора в зависимости от того, на основе какой информации выбирается то или иное решение. Уверенность будет обратно пропорциональна количеству возможных альтернативных структур в области многозначности. Кроме того, операции корректировки сами по себе психологически не перегружают перцептивный аппарат человека, поэтому самым главным фактором, определяющим легкость корректировки является то, насколько упрямо анализатор будет держаться за свой первоначальный, неправильный, анализ предложения.

---

[10] Существуют несколько способов перевода термина *head* на русский язык: "главное слово", "ядро", "вершина". Я буду называть модель Притчетта "head-driven model" моделью главного слова и использовать термины "ядро" и "главное слово".

Нежелание анализатора корректировать первоначальный неправильный анализ предложения связан с тем, что нет гарантии, что существует лучший анализ и что анализатор его найдет, даже если такой анализ существует. Поэтому отказ от первоначального анализа всегда сопряжен с риском, особенно для последовательного анализатора, который в отличие от параллельного анализатора конструирует только одну возможную структуру в области многозначности. Хотя анализатор строит только одну из возможных структур в области многозначности, тем не менее он в этом месте "ставит флажок", указывающий, что существуют другие альтернативные структуры, и поэтому модель называется "флажковой". Если в дальнейшем окажется, что придется отказаться от первоначального анализа, то анализатор возвращается к флажку и выбирает одну из возможных альтернативных структур. Набор флажков, сигнализирующий область многозначности, также может определять степень уверенности анализатора в правильности выбранного анализа. Чем больше флажков, тем меньше будет уверенность анализатора в том, что тот или иной выбранный вариант — правильный, и тем легче будет выбрать альтернативный вариант в случае, если первоначальный окажется неправильным. Важно отметить, что в данной модели при корректировке используются те же аналитические операции, что и при первоначальном анализе. Анализатор начинает корректировку с области многозначности, выбирая второй альтернативный вариант из набора возможных и просто анализирует предложение снова, используя стандартные принципы синтаксического анализа.

Каким образом определяется легкость или сложность каждого конкретного альтернативного варианта, делая предложение тем самым легким или сложным предложением заблуждения? Одного показателя – количества флажков в области многозначности – не только недостаточно для этого, но к тому же это часто может привести к нежелательным результатам, поскольку можно прийти к заключению, что все предложения заблуждения в английском языке должны быть сложными, так как в английском языке редко бывают возможными более двух альтернативных вариантов. И наоборот, в японском языке все предложения заблуждения должны быть легкими, так как количество возможных альтернативных вариантов очень большое. Инуе и Фодор предлагают ввести понятие "ранжированности" альтернативных вариантов, и таким образом, не количество альтернатив, а их ранг, важность по сравнению с первоначальным вариантом будет определять выбор второго подходящего варианта. Например, в предложении (1.4) *#The horse RACED past the barn fell* "Лошадь, ПРОГНАННАЯ мимо сарая, упала" второй возможный вариант анализа значительно менее предпочтителен, чем первый, а в предложении *He put the book on the table into his pocket* "Он положил книгу со стола в карман" второй вариант почти так же предпочтителен, как и первый. Таким образом, когда первоначальный анализ оказывается неправильным, анализатор возвращается назад в область многозначности и выбирает альтернативный вариант, но только при условии, если он находит этот второй вариант почти в такой же степени предпочтительным, как и первоначальный. Если такового не окажется, анализатор предпочтет придерживаться первоначального варианта, даже если таковой проблематичен[11].

Как было описано выше, принципы минимального присоединения и позднего закрытия характеризуют аналитические операции анализатора при первоначальном синтаксическом анализе предложения во всех моделях, относящихся к классу моделей заблуждения. Не составляет исключение и ранжированный флажковый последовательный анализатор, предложенный Инуе и Фодор. Что, однако, отличает одну модель этого класса от другой, так это принципы корректировки, в частности, каким

[11] Поскольку модель ранжированного флажкового последовательного анализатора пока еще находится в стадии разработки и по-настоящему экспериментально не проверена, техническая сторона модели может быть описана только в относительно общих терминах. Более конкретные детали, как то: на основе каких факторов происходит ранжировка; очевидность связи между первоначальным неправильным и последующим правильным анализами предложения – нуждаются в доработке и проверке.

образом различаются легкие и сложные предложения заблуждения. Инуе и Фодор предложили принцип эффекта тематического наложения (thematic overlay effect). Когда начинается процесс корректировки первоначального неправильного синтаксического анализа предложения, очень важно, чтобы при этом семантическое представление, основанное на первоначальном синтаксическом анализе, было стерто, уничтожено. Однако оказывается, что анализатор часто "забывает" это сделать: как только часть семантического представления построена, она обычно существует до тех пор, пока ее не стирает наложившаяся на нее новая часть. Эффект тематического наложения, тем не менее, предполагает, что разные семантические корректировки не одинаковы: когда старое семантическое представление стирается автоматически, корректировка оказывается легкой, когда оно остается, корректировка сильно затруднена, если вообще оказывается возможной. Например, в предложении *Mary gave her books to Susan* "Мэри дала ее книги Сюзан" эффект тематического наложения предсказывает, что корректировка первоначального анализа" Мэри дала ей книги..." будет легкой, потому что новая тематическая структура дает возможность реализации обеих ролей в конструкции с двойным дополнением – темы и цели, и таким образом, ни одна из приписанных ролей не исчезает.

Модель ранжированного флажкового последовательного анализатора в том виде, как она кратко описана выше, основана на том предположении, что аналитические операции, используемые при корректировке, ничем не отличаются от тех, что действуют при первоначальном анализе предложения. Однако пока ни подтвердить, ни опровергнуть это предположение не удается, так как экспериментальные исследования процесса корректировки находятся в зачаточном состоянии. Это дает возможность выдвигать различные теоретические гипотезы о характере процесса корректировки, каждая из которых равно заслуживает изучения. Поэтому Фодор и Инуе [Fodor, Inoue 1993] предложили другую модель синтаксического анализа предложения, которая отличается от предыдущей модели именно характером компонента корректировки.

Модель диагностики Фодор и Инуе [Fodor, Inoue 1993][12]. Эта модель представляет собой версию ранжированного флажкового последовательного анализатора, которая характеризуется следующими свойствами:

1) путем угадывания выбирается и строится один последовательный синтаксический анализ предложения, в результате чего совершается много ошибок, и предложения часто анализируются как предложения заблуждения;

2) анализатор справляется с анализом предложений заблуждения, не прибегая к помощи никаких несинтаксических механизмов;

3) при помощи специальных аналитических операций корректировки анализатор перестраивает неправильное синтаксическое представление. Набор таких операций очень ограничен и является универсальным для всех языков. Некоторые из этих операций легкие, а некоторые – сложные, хотя возможно, что одна и та же операция может иногда быть легкой, а иногда сложной. Принцип, который определяет легкость/сложность корректировки, зависит от того, насколько очевидна связь между первоначальным неправильным и последующим правильным анализами предложения;

4) анализатор может принимать стратегическое решение, стоит ли предпринять корректировку, так как первоначальный анализ предложения, хоть и несовершенный, может все равно оказаться самым лучшим;

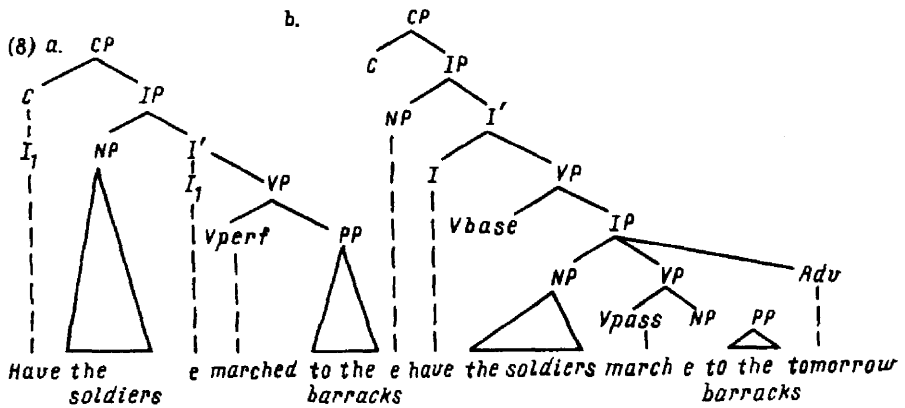
5) анализатор принимает такое стратегическое решение, основываясь на том насколько ценность следующего возможного анализа соотносится с ценностью первоначального.

Характеристики 1, 2, 4 и 5 в модели диагностики остались теми же, что и в модели ранжированного флажкового последовательного анализатора. Однако коренным обра-

[12] Полное описание и теоретическое обоснование этой модели еще далеки от окончания, поэтому в настоящем обзоре приведены лишь ее первоначальные положения, что делается с разрешения авторов.

зом изменился характер процедуры корректировки. Если в модели ранжированного флажкового последовательного анализатора корректировка состояла из стирания первоначального и проведения повторного синтаксического анализа предложения в области многозначности с использованием стандартных аналитических операций, то в модели диагностики используются специальные процедуры, фактически исправляющие первоначальный неправильный анализ. Важно то, насколько легко анализатору "заметить" правильный альтернативный анализ. Вспомним, что в каждом предложении заблуждения всегда есть слово, которое вдруг оказывается невозможным присоединить к уже частично построенному синтаксическому представлению. Это слово является "симптомом", указывающим на возникшую проблему. Природа этого слова-симптома является информацией, которую может использовать анализатор, чтобы установить "диагноз" проблемы и затем предложить способ избавления от симптома. Если правильный диагноз оказывается очевидным в результате симптома, корректировка предложения заблуждения будет быстрой и безболезненной. Но если симптом недостаточно информативен или, больше того, сам вводит в заблуждение по поводу характера проблемы, тогда правильная корректировка будет сильно затруднена, если вообще окажется возможной[13].

В качестве примера для иллюстрации модели диагностики на материале английского языка Фодор и Инуе предлагают анализ предложения *HAVE the soldiers marched to the barracks tomorrow*[14]; а) Промаршировали ли солдаты к казармам *?завтра?* (8а); б) Отправьте солдат маршем к казармам завтра (8б). Наречие *завтра* семантически несовместимо с первоначально предпочтительным синтаксическим анализом предложения (8а), в котором *HAVE the soldiers marched* представлены в качестве инверсивной (вопросительной) конструкции со вспомогательным глаголом и глаголом в форме перфекта. Эту проблему, однако, легко заметить, поэтому анализатор, достигнув слова *завтра* в предложении, знает, что его первоначальный анализ неправильный. Несмотря на это, оказывается совсем не просто выбрать правильный альтернативный анализ (8б), который представляет собой каузативную конструкцию в форме повелительного наклонения. Корректировка синтаксического представления (8а) для того, чтобы получить (8б), включает целый набор структурных и семантических операций, одного объема которых достаточно, чтобы затруднить корректировку:



С каноническими предложениями заблуждения дело обстоит еще сложнее. Вернемся еще раз к предложению (1.4) *#The horse RACED past the barn fell* "Лошадь ПРОБЕЖАЛА мимо сарая ?упала". Сначала процесс диагностики идет успешно: анализатор узнает, что слово *упала* нуждается в подлежащем, однако в первона-

[13] Интересно отметить, что еще раньше Фразер и Райнер [Frazier, Rayner 1982] выдвинули гипотезу "выборочного анализа" (см. выше), которую можно считать прямым предшественником модели диагностики.

[14] В данном примере слово-симптом – *tomorrow*.

чально сконструированном синтаксическом представлении для него подлежащего нет. Анализатор вынужден предпринять одну из аналитических процедур корректировки, а именно ему следует "создать именную группу" (create a NP), которую он затем при помощи другой операции может "украсть" (steal) для того, чтобы сделать ее подлежащим глагола *упала*. Таким образом, необходимо так перестроить уже частично построенное синтаксическое представление до слова *упала*, чтобы оно состояло из X + NP, где X – любой набор слов, который может предшествовать NP + VP. При такой перестройке, как и всегда, анализатор будет стремиться следовать принципу "минимум всего" (minimal everything), поэтому первым вариантом будет предположение, что набор X пуст. Теперь нужно найти способ превратить предложение "Лошадь ПРОБЕЖАЛА мимо сарая" в именную группу, и эту процедуру схематично можно представить следующим образом:

(9) ЗАДАЧА: НАЙТИ ПОДЛЕЖАЩЕЕ ДЛЯ *УПАЛА*

- достать лексическую информацию для слова *RACED;*
- предположить, что подлежащее–именная группа для глагольной формы *упала* включает все слова слева от *упала;*
- в английском языке артикль и главное существительное предшествуют всем остальным составляющим именной группы. Следовательно, артикль подлежащего скорее всего первое слово *the*, и наиболее вероятное главное слово – первое, т.е. *horse;*
- это существительное не может иметь высказывание в качестве дополнения, поэтому та часть предложения, которая следует за существительным *horse*, является не аргументом, а адъюнктом (adjunct)[15];
- адъюнкт представляет из себя глагольную группу без подлежащего, поэтому, следуя расширенному принципу проекции (extended projection principle) [Chomsky 1981; 1986; Секерина 1994], у него должно быть подлежащее, выраженное пустой категорией. Кроме того, отсутствует комплементайзер (complementizer), возглавляющий CP. Все эти факты вместе приводят к выводу, что ядро IP тоже пусто;
- таким образом, *RACED* не является спрягаемой глагольной формой и не может быть глаголом в прошедшем времени;
- вновь вернуться в лексикон и проверить, что *RACED* может не являться спрягаемой глагольной формой, если это слово – причастие, образованное не от переходного глагола *RACE*, а от непереходного глагола *RACE;*
- анализ в виде сокращенного определительного придаточного с причастием соответствует всему остальному, что есть в предложении-адъюнкте, если только вставить "след" (trace) именной группы и проиндексировать его так же, как и пустое подлежащее.

Аналитические операции корректировки в (9) состоят из упаковки *RACED past the barn* в предыдущую именную группу, т.е. опускания ее вниз синтаксического дерева и упаковки вовнутрь уже построенного узла именной группы. Операция опускания не влечет за собой последствий в виде усложнения анализа, а вот сама упаковка непосредственно представляет из себя сложную операцию. Таким образом, сложность процедуры корректировки в данном примере, как и в других канонических предложениях заблуждения, заключается в том, что анализатору приходится осуществлять большое количество дедуктивных операций и много гадать, прежде чем становится очевидным, что нужно вернуться в лексикон, чтобы найти правильную синтаксическую категорию для многозначного слова.

Модель ассоциации Л. Фразер и Ч. Клифтона [Frazer, Clifton 1993]. Все версии модели заблуждения, кратко описанные выше, при синтаксическом анализе предложения следовали двум принципам построения синтаксической струк-

[15] См. более подробное обсуждение разницы между аргументами и адъюнктами ниже.

туры – принципу минимального присоединения и принципу позднего закрытия, предложенных Л. Фразер в ее диссертации [Frazier 1978] еще 20 лет назад. Основной характеристикой обоих принципов считался их универсальный характер, т.е. оба они должны систематически направлять анализатор при анализе многозначных синтаксических структур любых типов во всех языках мира. Абсолютное большинство экспериментальных и теоретических исследований в области синтаксического анализа предложения до конца 1980-х годов было посвящено рассмотрению первого принципа, принципа минимального присоединения, его характеристик и универсальности. Первыми психолингвистами, проведшими серию экспериментов с целью определить характер действия принципа позднего закрытия, были Ф. Куэтос и Д. Митчелл, опубликовавшие в 1988 г. статью "Языковая вариативность в синтаксическом анализе предложения: ограничения при использовании стратегии позднего закрытия в испанском языке" [Cuetos, Mitchell 1988].

Куэтос и Митчелл исследовали характер применения ПЗ в испанском и английском языке, используя опросники и последовательное чтение предложений с фиксацией времени, уходящего на определение значения частей предложения[16]. Материалом для этих экспериментов послужили сложноподчиненные предложения с определительным придаточным, которое модифицирует существительное, входящее в состав сложной именной группы (10):

(10) а. *Alguien disparó contra* [NP*el criado de la actriz*] *que estaba en el balcón*

б. *Someone shot* [NP*the servant of the actress*] *who was on the balcony* "Кто-то застрелил [NPслугу актрисы], которая/который был/а на балконе"

В испанском, как и в английском языке, предложения такого типа являются многозначными: речь может идти или о застреленном слуге актрисы, который был на балконе, или о застреленной актрисе, которая была на балконе. Если люди при восприятии этого предложения следуют ПЗ, то они должны присоединить придаточное предложение к той фразовой группе, которая на данный момент была сформирована последней, т.е., к именной группе, непосредственно предшествующей *que*. По результатам опросника оказалось, что испаноговорящие в 62% предложений предпочли первую интерпретацию ("на балконе был слуга"), т.е. избрали раннее закрытие.

Три последующих эксперимента с фиксацией времени, понадобившегося на чтение предложений, подтвердили гипотезу, что испаноговорящие информанты не следуют принципу ПЗ. Предложения из первого эксперимента были модифицированы таким образом, чтобы заставить информантов выбрать второе прочтение (на балконе была актриса). Оказалось, что на понимание таких модифицированных предложений уходит значительно больше времени, чем на понимание контрольных (однозначных) предложений. Следовательно, информанты вначале присоединяли придаточное предложение к более высокой составляющей, следуя раннему закрытию, а когда появились слова, требующие скорректировать эту интерпретацию и "опустить" придаточное предложение ниже в дереве, на эту процедуру уходило дополнительное время.

По результатам этих экспериментов Куэтос и Митчелл пришли к выводу, что испаноговорящие нарушают универсальный принцип ПЗ при интерпретации предложений такого типа. Похожие эксперименты были проведены на материале некоторых других романских языков, в частности, итальянского и французского, в которых

[16] Существует несколько хорошо разработанных экспериментальных методов, которые в разной степени позволяют исследовать аналитические операции, используемые синтаксическим анализатором при анализе предложения. Некоторые из этих методов являются постоперационными (off-line), такие, как, например, опросники, заканчивание предложения, устное определение грамматической правильности или приемлемости предложения и т.п. Другие, более важные, в большей или меньшей степени позволяют непосредственно исследовать операции в процессе их применения (on-line): последовательное чтение частей или всего предложения, при котором скорость чтения регулируется самим читающим, многомодальное зондирование, запись движения глаз и т.п. Объем данной работы не позволяет остановиться более подробно на характеристике экспериментальной части теории синтаксического анализа предложения.

результаты тоже свидетельствовали о некотором предпочтении раннего закрытия. Теория универсальной грамматики, однако, предполагает, что если принцип универсален, то для него не может быть исключений. Поэтому после опубликования результатов испанских экспериментов Куэтоса и Митчелла в теории синтаксического анализа предложения начались долгие дебаты по поводу универсальности характера принципа позднего закрытия. Критике подверглись как экспериментальные предложения, составленные Куэтосом и Митчеллом, так и методика самих экспериментов. Даже после нескольких лет постоянных дискуссий, заполненных повторениями испанских экспериментов с использованием других методик, модификацией предложений и обращением к материалу других языков вопрос продолжал оставаться спорным.

И вот, наконец, появилось новое, неожиданное объяснение результатам экспериментов Куэтоса и Митчелла, а также результатам экспериментов на материале других языков, в которых исследовались предложения этого типа, а именно, сложноподчиненные предложения с определительным придаточным, модифицирующим одно из существительных в составе последней именной группы главного предложения. Это объяснение предложено Л. Фразер и Ч. Клифтоном в рамках "модели ассоциации" (construal) [Frazier, Clifton 1993]. Именно различные способы присоединения определительного придаточного послужили отправной точкой создания этой модели. Фразер и Клифтон заметили, что семантика главного слова в именной группе, которую модифицирует определительное придаточное, оказывает влияние на способ присоединения этого придаточного. Это главное слово может означать материал, количество, родственные отношения, принадлежность и т.п., и поэтому существует некая иерархия предпочтительности присоединения определительного придаточного или к главному слову, или к зависимому слову в составе сложной именной группы. Авторы предположили, что синтаксический анализ таких определительных придаточных не подчиняется принципу позднего закрытия, как считалось раньше. А именно, они не присоединяются к последней предшествующей им сложной именной группе, а "ассоциируются" (are construed) с ней.

Ассоциация, по определению авторов, состоит из двух частей: синтаксическая операция ассоциации с некоторой областью предложения и процесс интерпретации. Ассоциация отличается от присоединения тем, что она не подчиняется общим принципам присоединения. Таким образом, при ассоциации структурные отношения между некоторыми узлами синтаксического дерева (являются ли два узла сестрами или находятся в отношении зависимости мать–дочь) остаются неопределенными. Такие структурные принципы, как принцип минимального присоединения и принцип позднего закрытия не применяются при синтаксическом анализе некоторых конструкций, в частности, при анализе сложноподчиненных предложений с определительным придаточным, модифицирующим одно из существительных в составе последней именной группы главного предложения. В случае временной многозначности в таких предложениях любой из возможных вариантов может быть выбран в зависимости от того, какому из них отдается предпочтение благодаря разным несинтаксическим факторам, но с одним ограничением: все предпочтительные варианты анализа должны оставить эту часть предложения внутри "анализируемой в настоящий момент тематической области" (current processing domain).

Схематично модель Ассоциации можно представить следующим образом:

(11) (1) Применение ассоциации: только непервичные синтаксические отношения подчиняются принципу ассоциации, а те фразовые группы, которые могут выражать первичные синтаксические отношения, ей не подчиняются.

(2) Принцип ассоциации: (а) ассоциировать XP с анализируемой в настоящий момент тематической областью; (б) проинтерпретировать XP внутри этой анализируемой тематической области, используя одновременно несинтаксические принципы интерпретации и структурные принципы грамматической правильности.

(3) Анализируемая в настоящий момент тематическая область представляет собой расширенную фразовую группу XP (XP = X_{max}), управляющую последним терминальным элементом X, приписывающим "тематические роли" (theta-assigner).
(4) Принципы интерпретации (не ранжированы):
(а) максимы Грайса;
(б) принцип немедленной интерпретации;
(в) минимальная ревизия;
(г) априорная вероятность;
(д) принцип референциальности;
(е) фокус и контекстная значимость.

Как определено в (11.1), только непервичные синтаксические отношения подчиняются принципу ассоциации. Первичные отношения представляют собой отношение между главными подлежащим и сказуемым вместе со всеми остальными обязательными позициями, определяемыми этим отношением. Первичные отношения включают в себя следующие случаи: 1) подлежащее и сказуемое; 2) их дополнения и другие обязательные составляющие; 3) дополнения и другие обязательные составляющие первичных фразовых групп. Вторичные отношения определяются как непервичные. Примерный набор вторичных отношений для английского языка, по мнению Фразер и Клифтона, включает следующие конструкции (Таблица 4) [Fodor, Bever, Garrett 1974]:

Таблица 4

Вторичные отношения в английском языке

(12)	Структура	Пример	Перевод
(а) Сложноподчиненные предложения с определительным придаточным, модифицирующим одно из существительных в составе последней (сложной) NP главного предложения	[Nl (of) N2 Relative Clause]	*a table of wood that was from Galicia...*	"Стол из дерева, которое из Галиции"
(б) Сложноподчиненные предложения с определительным придаточным, вынесенным вперед	[NP V NP Adv] Relative Clause	*Some girl hit some boy last night who was...*	"Какая-то девочка ударила какого-то мальчика, который (ая)"
(в) Вторичная предикативность	[NP V NP] AP	*Mary brought Fido home wet/dripping*	"Мэри привела Файдо домой мокрым/взмокшая"
(г) Сочинительные конструкции	[XP]and XP	*The nurse weighed John and Mary (both/together)*	"Медсестра взвесила Джона и Мэри (обоих/вместе)"
(д) Присоединение модифицирующего подчинительного предложения	[NP VP] because S	*The doctor didn't leave because he was angry*[17]	"Доктор не ушел, потому что он рассердился"

[17] Источник многозначности в примере (12д) – сфера действия отрицания:
(а) "Доктор ушел не потому, что он рассердился"
(б) "Доктор не ушел, потому что он рассердился".

Рассмотрим на примере, иллюстрирующем предложения со вторичной предикативностью (тип (12в)), как применяется принцип Ассоциации. Предложение (13) может иметь три разных значения:

(13) *Sam hammered the metal flat*

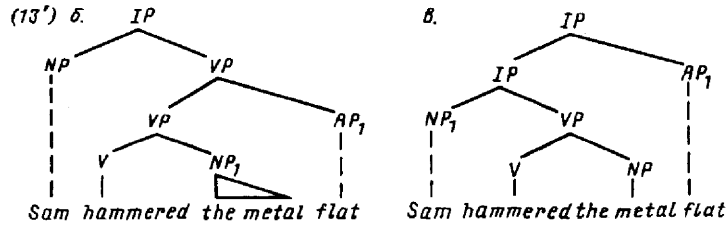
"Сэм бил молотком металл плоский"

а. результативное состояние: *Сэм расплющил металл при помощи молотка;*

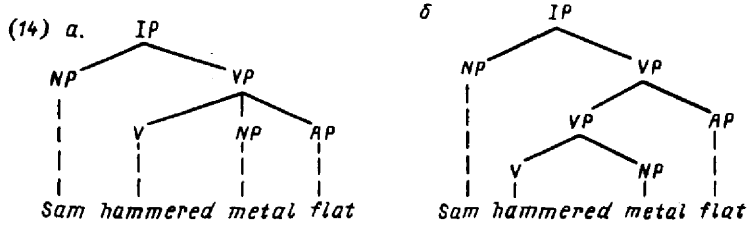
б. прямое дополнение: *Сэм бил молотком (уже) расплющенный*$_1$ *металл*$_1$;

в. подлежащее: *Сэм*$_1$, *расплющенный*$_1$, *бил металл молотком*[18]

Когда значение (13а) возможно, обычно именно оно является предпочтительным. Семантика результирующего состояния требует, чтобы причастие/прилагательное "расплющенный" модифицировало прямое дополнение. При таком прочтении "расплющенный" представляет собой обязательную часть дополнения при глаголе. В отличие от значения (13а), при прочтениях (13б, в) "расплющенный" функционирует в качестве вторичного предиката. Если представить себе, что значение (13а) не существует, становится неясно, можно ли доказать, что значение (13б), при котором "расплющенный" модифицирует прямое дополнение, предпочтительнее, чем (13в), при котором оно модифицирует подлежащее. Если мы предположим, что синтаксические представления для (13б) и (13в) выглядят следующим образом:



то, следуя структурному принципу позднего закрытия, (13'б) должно быть предпочтительным значением, а прочтение (13'в) должно представлять собой по крайней мере легкое предложение заблуждения, в результате чего на его понимание при эксперименте должно уходить больше времени из-за необходимости корректировки. Однако, как показывают предварительные экспериментальные исследования, семантически однозначные предложения с модифицированным подлежащим не дают оснований думать, что такая корректировка имеет место. Таким образом, мы видим, что обычного структурного предпочтения для более низкого присоединения вторичного предиката в таких предложениях не существует. Однако не существует предпочтения и для более высокого присоединения. Вместо этого такие неструктурные факторы, как семантика и вид глагола, характеризуют перцептивную сложность и предпочтительное значение этих многозначных предложений. Фразер и Клифтон предлагают следующие синтаксические представления для (13'б) и (13'в), в которых анализатор ассоциирует вторичный предикат, модифицирующий прямое дополнение или подлежащее, с некоей неопределенной синтаксической структурой (14а) или (14б):



[18] Интерпретация (13в) возможна, но в данном примере является семантически аномальной. Ср. *Sam*$_1$ *hammered the metal naked*$_1$ "Сэм, голый, бил металл молотком".

Вторичные синтаксические отношения, вроде вторичной предикативности, следуя принципу ассоциации, соотносятся с анализируемой в настоящий момент тематической областью, определенной в (11.3). Например, рассмотрим сложноподчиненные предложения с определительным придаточным, модифицирующим одно из существительных в составе сложной именной группы главного предложения. Это определительное придаточное обычно ассоциируется со второй, нижней, именной группой NP2, когда та является адъюнктом имени N1 с приписанной ей тематической ролью (эту роль приписывает предшествующий предлог); однако оно ассоциируется со всей сложной именной группой, если NP2 является дополнением N1. Определительное придаточное интерпретируется в составе нижней именной группы NP2 в первом случае, а во втором случае оно потенциально может ассоциироваться с любой из двух возможных именных групп в составе сложной именной группы[19].

Рассмотрим теперь вкратце принципы интерпретации, представленные в (11.4). Максима Грайса (11.4 а) [Grice 1989] объясняет разницу в предпочтительности присоединения определительного придаточного в английском языке к нижней из двух именных групп в составе сложной именной группы, содержащей именную группу в родительном падеже. Принцип немедленной интерпретации (11.4 б) необходим для объяснения предложений с "потому что" (12д). Этот принцип означает, что только обязательные семантические решения принимаются при отсутствии доказательств. Каждая именная группа (кроме именных предикатов) должна быть привязана к какому-нибудь объекту в дискурсе (связном тексте) либо при помощи постулирования нового объекта для этой именной группы, либо путем ассоциации с уже существующим объектом. Предикат должен быть индексно связан с каким-нибудь объектом в дискурсе. В случае предложений с "потому что" предикат немедленно ассоциируется с подлежащим предложения, что приводит к предпочтению того значения, когда придаточное причины модифицирует все предложение целиком, а не VP.

Принципы минимальной ревизии (11.4 в) и априорной вероятности (11.4 г) заимствованы без обсуждения из статей Л. Фразер [Frazier 1990] и Крейна и Стидмана [Crain, Steedman 1985] соответственно. Принцип референциальности (11.4 д) необходим для объяснения того факта, что ограничительные определительные придаточные имеют тенденцию присоединяться к тому главному слову, которое соответствует новому или старому объекту в дискурсе, так как такое слово считается референциальным. Только такие именные группы, которые имеют в своем составе артикль или квантор, считаются референциальными. И, наконец, фокус и контекстная значимость (11.4 е) указывают на то, что предпочтение отдается той фразовой группе, которая находится в фокусе или иным другим способом выделена в контексте.

В качестве заключения Фразер и Клифтон приводят причины, побудившие их сформулировать принцип ассоциации. Некоторые синтаксические конструкции не поддавались многочисленным попыткам разных исследователей анализировать их при помощи принципов минимального присоединения и позднего закрытия. Постоянное предпочтение одного структурного варианта при анализе первичных (синтаксических) отношений, как предсказывают эти принципы, не наблюдается при анализе вторичных ассоциаций. Таким образом, с добавлением принципа Ассоциации к арсеналу универсальных принципов появилась, наконец, возможность последовательно объяснить как структурные предпочтения в одних случаях, так и их отсутствие в других.

---

[19] Когда в состав сложной именной группы входит предлог, например, *кофта из шерсти, которую привезли из Испании*, то в зависимости от типа предлога он может либо приписывать либо не приписывать своему дополнению тета-роль, что существенным образом влияет на выделение тематической области.

* * *

В кратком описании модели заблуждения и ее модификаций, изложенном выше, наше внимание было сосредоточено только на одном аспекте синтаксического анализа предложения, а именно, на непосредственном "построении синтаксического представления (дерева)" (structure building). Это хоть и центральное, но не единственное направление, в котором идет разработка и экспериментальное апробирование данной модели. Другими важными направлениями можно считать следующие исследования: 1) описание и анализ "предложений с грамматическими зависимостями на расстоянии" (long-distance dependencies); 2) возможность параметризации синтаксического анализа предложения, связанная с появлением теории принципов и параметров в рамках синтаксической теории управления и связывания; 3) применение модели заблуждения (и других моделей тоже) для синтаксического анализа предложения в других языках[20]. Ниже в качестве иллюстрации предлагается краткое описание первых двух направлений.

Предложения с грамматическими зависимостями на расстоянии. Обсуждение синтаксического анализа предложения до настоящего момента велось с точки зрения построения синтаксического представления предложения, то есть, дерева, где все терминальные символы-слова представляли собой полнозначные лексические единицы. Однако, как это хорошо известно из теории синтаксиса, существуют так называемые пустые категории (empty categories), которые являются характерной чертой многих типов синтаксических структур. В синтаксической теории управления и связывания Хомского [Chomsky 1981; 1986; Sells 1985] выделяют четыре типа пустых категорий: след именной группы (NP-trace), след вопросительного слова (Wh-trace), большое ПРО (PRO) и маленькое про (pro)[21]. Пустые категории отличаются от других синтаксических объектов типа лексических ядер и их проекций тем, что они не имеют фонолого-фонетического представления, то есть, не произносятся, являясь при этом полноправными членами синтаксического представления предложения. Однако в синтаксической теории не существует полностью убедительных аргументов, доказывающих существование пустых категорий; поэтому оказываются так важны доказательства, если таковые возможны, из области психолингвистики, в частности, синтаксического анализа предложения в процессе понимания.

Вследствие чрезвычайной важности этого вопроса в психолингвистических исследованиях в области синтаксического анализа предложения особое внимание стало уделяться предложениям с заполнителем/пропуском (filler-gap dependencies). Естественно, разработчики модели заблуждения не остались в стороне от этой проблемы. Было первоначально выдвинуто две гипотезы: активного заполнителя (active filler hypothesis), сформулированная Л. Фразер [Frazier 1987], и "пропуск-только-как-второе-средство" (gap-as-second-resort hypothesis), предложенная Дж.Д. Фодор [De Vincenzi 1989; Frazier 1987]. Гипотеза Фразер, основанная на материале голландского языка, предполагает, что наличие обязательного заполнителя заставляет действовать ана-

---

[20] См. сноску 5.

[21] Примеры, характеризующие все четыре пустые категории, представлены ниже, где индекс i означает кореференцию в интерпретации:

-NP-trace: *John$_i$ was arrested t_i* "Джон арестован" (след именной группы).

-Wh-trace: *Who$_i$ did the police arrest t_i* "Кого арестовала полиция" (след вопросительного слова).

-PRO: *John$_i$ wants PRO$_i$ to study* "Джон хочет учиться".

-pro: *pro$_i$ ha chiamato Giovanni$_i$* (итальянский) "Позвонил Джованни".

Основы теории пустых категорий и определение понятия цепи см. в [Sells 1985]. Важно отметить, тем не менее, что другие современные американские синтаксические теории имеют свои представления о пустых категориях. В частности, грамматика главного слова и фразовых структур (head-driven phrase structure grammar) признает лишь след вопросительного слова, в то время как лексико-функциональная грамматика (lexical-functional grammar) не признает никаких пустых категорий вообще.

лизатор в особом режиме, в котором заполнитель "активно ищет свой пропуск". Гипотеза Фодор, основанная на материале английского языка, разрешает наличие пропуска только в том случае, если анализатор заранее знает, что фраза определенной категории с лексическим содержанием (то есть, не пустая) должна встретиться в анализируемом предложении, но отсутствует. Более поздние исследования на материале итальянского языка, проведенные М.Де Винченци в ее диссертации "Стратегии синтаксического анализа в языке с нулевыми подлежащими" [De Vincenzi 1989], показали, что возможно объединение этих двух гипотез в одну стратегию, которой пользуется синтаксический анализатор при анализе предложений с грамматическими зависимостями на расстоянии. Ею был предложен "принцип минимальной цепи" (minimal chain principle):

> Принцип минимальной цепи (minimal chain principle): избегайте постулировать лишние звенья цепи в поверхностной структуре предложения, но не откладывайте постулирование обязательных звеньев [De Vincenzi 1989: 99].

Возможность параметризации синтаксического анализа предложения, связанная с появлением теории принципов и параметров в рамках синтаксической теории управления и связывания. Какое отношение имеет теория грамматики, будь то теория принципов и параметров в рамках теории управления и связывания или любая другая, к теории синтаксического анализа предложения? Оказывается, что взаимосвязь между двумя теориями является принципиальной, так как ответ на этот вопрос дает возможность ответить на более общий вопрос теории лингвистики, а именно, как связаны между собой знание языка и процесс пользования языком. Принципиальное различие между этими двумя явлениями нам хорошо знакомо из истории лингвистики, и наиболее известна формулировка, предложенная Ф. де Соссюром [Saussure 1932]: противопоставление "языка" и "речи". Хомский определил это как противопоставление языковой компетенции (competence) и языковой деятельности (performance). Лингвистика пытается пролить свет на взаимоотношение между грамматикой и синтаксическим анализатором, чтобы ответить на эти более общие вопросы.

В начале развития теории синтаксического анализа предложения предполагалось, что грамматика и синтаксический анализатор представляют собой разные явления (см. [Pritchett 1992] по истории вопроса). Были, однако, предложены и такие психолингвистические модели синтаксического анализа предложения, где грамматика и синтаксический анализатор были взаимосвязаны в том смысле, что последний в основном зависел от первой. Хорошо известна модель синтаксического анализа предложения, предложенная в рамках "деривационной теории сложности" (the derivational theory of complexity) [Pritchett 1992], которая утверждала, что деривационная (трансформационная) история предложения эквивалентна его поверхностной перцептивной сложности. Эта теория была следствием гипотезы, что синтаксический анализатор является "грамматикой наоборот". Эту гипотезу, тем не менее, не только удалось доказать экспериментально, но были найдены аргументы, утверждающие прямо противоположное. Тогда большинство исследователей пришло к выводу, что трансформационно-порождающая грамматика не может быть "психологически реальной" (psychologically real), и начался поиск независимых аналитических принципов, которые отражали бы общие когнитивные процессы и не были никак связаны ни с какой конкретной теорией грамматики. Таковы, например, принципы минимального присоединения и позднего закрытия в модели заблуждения.

Однако в последнее время появились новые модели синтаксического анализа предложения, бросившие вызов влиятельной модели заблуждения, например, модель главного слова Притчетта [Pritchett 1987; 1987a; 1991; 1991a] и модель минимальных обязательств Вайнберг [Taraban 1988; Weinberg 1993; 1994]. Они утверждают, что грамматика может быть зависима от синтаксического анализатора, а не наоборот.

Конечно, очевидно, что полностью синтаксический анализатор не может определять характер грамматики, так как это осложнит объяснение языкового разнообразия. Ведь считается, что если аналитические стратегии отражают фундаментальные и автоматические когнитивные процессы человека, то эти процессы должны быть едиными. Таким образом, очевидно, что не все грамматические принципы являются аналитическими; тем не менее, структуры обоих компонентов обязательно являются совместимыми, хотя вопрос о конкретной природе этой взаимосвязи пока остается открытым.

Объяснение характера взаимосвязи между грамматикой и синтаксическим анализатором осложняется еще и тем обстоятельством, что идет бурное развитие самой теории грамматики. Все современные теории грамматики, ведущие свое происхождение от "расширенной стандартной теории трансформационно-порождающей грамматики" (extended standard theory), – "теория управления и связывания", грамматика главного слова и фразовых структур (head-driven phrase structure grammar) и "лексико-функциональная грамматика" (lexical-functional grammar) [Sells 1985], – больше не используют правила порождения синтаксических представлений, а пользуются всеми возможными обобщенными представлениями, из которых при помощи грамматических ограничений (constraints) выбираются только те, которые подчиняются универсальным грамматическим принципам.

Этот переход от системы с правилами порождения к системе с ограничениями особенно важен в синтаксической теории управления и связывания и в ее новой версии, теории принципов и параметров (principles and parameters theory). Центральным для этой теории с середины 1980-х годов стала разработка универсальных принципов и конкретно языковых параметров. Универсальные принципы являются частью универсальной грамматики (УГ) (universal grammar), и они соблюдаются во всех языках без исключения. В качестве примера универсальных принципов можно привести модуль Х-штриха, принцип пустых категорий (the empty category principle), "теорию связывания" (binding theory), тета-критерий и т.п. [Chomsky 1981; 1986; Секерина 1994]. С другой стороны, в УГ существует набор бинарных параметров, каждый из которых может иметь значение "да" или "нет". Любой язык в таком случае будет характеризоваться значениями определенного набора параметров, и это будет являться уникальной комбинацией, отличающей данный язык от любого другого языка. В качестве примеров параметров можно привести "параметр нулевого подлежащего" (null subject parameter), опущение *про* (pro-drop parameter), движение вопросительного слова в синтаксисе (Wh-movement in syntax), главное слово (ядро) – первое (head-initial), главное слово (ядро) – последнее (head-final) и т.п. Например, используя параметры, можно сравнить два языка – английский и японский (только частично, в качестве иллюстрации) [Frazier 1990]:

Таблица 5

Некоторые типологические различия английского и японского языков

	Английский язык	Японский язык
Нулевое подлежащее	нет	да
Опускание *про*	нет	да
Движение вопросительного слова в синтаксисе	да	нет
Главное слово – первое	да	нет
Главное слово – последнее	нет	да

Взаимосвязь между такой грамматикой, основанной на принципах, и синтаксическим анализатором не является очевидной. Однако если показать, что глубинные и абстрактные грамматические принципы оказывают существенное влияние на поверх-

ностный синтаксический анализ предложения, то можно утверждать, что основой синтаксического анализа предложения в процессе понимания является локальное применение глобальных грамматических принципов.

Две интересные модели синтаксического анализа предложения основаны на этом утверждении: модель главного слова Притчетта [Pritchett 1987; 1987a; 1991; 1992; 1992a] и модель минимальных обязательств Вайнберг [Taraban 1988; Weinberg 1993; 1994]. Более подробно модель минимальных обязательств будет описана в разделе 2.3, а сейчас обратимся к основным положениям модели главного слова.

2.2. Модель главного слова Б. Притчетта

Модель главного слова была первоначально изложена Б. Притчеттом в его неопубликованной диссертации "Феномен предложений заблуждения и грамматические основы синтаксического анализа предложения" [Pritchett 1987]. Вслед за диссертацией последовала серия статей [Pritchett 1987a; 1991; 1992a] и, наконец, в 1992 г. вышла монография "Грамматическая компетенция и синтактико-аналитическая деятельность" [Pritchett 1992].

В качестве теоретических основ своей модели Притчетт предлагает поиск ответа на два фундаментальных вопроса: (1) Каким образом людям удается так быстро, автоматически строить грамматически обоснованное синтаксическое представление предложения? (2) Какова взаимосвязь между анализатором и грамматической теорией, благодаря чему это возможно? Исследование предложений заблуждения, на основании которых можно строить предположения о характере деятельности анализатора, как предлагает автор, дает единый ответ на оба вопроса: ядро синтаксического анализа предложения состоит из локального применения глобальных синтаксических принципов.

Притчетт придерживается той точки зрения, что между грамматической теорией и анализатором существует тесная взаимосвязь. Можно даже предположить, считает он, что грамматическая теория является производной от теории анализа предложения, а не наоборот, хотя, конечно, такая точка зрения спорна. Наличие огромной вариативности в составе естественных языков указывает на то, что грамматическая теория имеет право на самостоятельное существование вне зависимости от характеристик анализатора. Тем не менее, глубинные и абстрактные грамматические принципы оказывают сильное влияние на (поверхностную) синтактико-аналитическую деятельность анализатора, указывая на их тесную взаимосвязь.

Автор строит свою модель синтаксического анализа предложения, исходя из принципа присоединения, которому следует анализатор [Pritchett 1992: 12]:

> Тета присоединение (theta attachment): тета критерий должен удовлетворяться на каждом этапе синтаксического анализа предложения, принимая во внимание максимальную тематическую (тета) сетку (набор валентностей) (theta grid).

Тета критерий, один из главных компонентов теории управления и связывания [Chomsky 1981; 1986; Секерина 1994], означает, что каждый аргумент должен иметь только одну тематическую роль и каждая тематическая роль должна быть приписана только одному аргументу. Принцип тета присоединения должен следить за тем, чтобы анализатор получал из лексикона максимальную тематическую сетку для каждого слова, обладающего такой сеткой, и обеспечивал сведéние тематических ролей и слов в соответствии с тета критерием в каждый данный момент синтаксического анализа предложения. Таким образом, каждая фразовая группа будет присоединяться к уже частично построенному синтаксическому представлению предложения в то место, в котором она может получить тематическую роль в соответствии с независимыми принципами грамматики.

Для иллюстрации действия принципа тета присоединения рассмотрим в слегка упрощенном виде анализ следующего простого предложения: *Without her contributions the orphanage closed* "Без ее пожертвований детский дом закрылся". Вначале *without*

"без" при помощи информации из лексикона идентифицируется как предлог, имеющий тематическую сетку, т.е. валентности. Хотя этот предлог может приписать какой-нибудь составляющей одну тематическую роль, на настоящий момент именная группа, которая может служить аргументом, еще не доступна для анализа. В следующий момент появляется слово *her* "(н)ее", которое является либо именной группой, могущей служить дополнением к предлогу, либо местоименным определением для следующего за ним имени существительного. Таким образом, уже здесь возникает временная многозначность. Однако предлог *без* имеет при себе тематическую роль, а *(н)ее* может быть аргументом, поэтому следуя принципу тета присоединения, "нее" так присоединяется к уже частично построенному синтаксическому представлению, чтобы получить эту роль, а именно становится сестрой предлога в синтаксическом дереве. Второй возможный вариант присоединения *(н)ее* при этом не проходит, потому что в таком случае это слово будет определением, а определения, не являясь аргументами, не требуют тематической роли, и следовательно, имеющаяся у предлога тематическая роль останется невостребованной. Следующее слово — *contributions* "пожертвований" — идентифицируется как существительное. Анализатор опять оказывается перед выбором: можно принять это существительное за главное слово новой именной группы, у которой еще нет тематической роли, нарушая тем самым принцип тета присоединения, или можно построить распространенную именную группу [*her contributions*]NP "ее пожертвований". В соответствии с принципом тета присоединения выбирается второй вариант. Распространенная именная группа *her contributions* как единое целое получает тематическую роль от предлога, которому теперь удалось ее удачно приписать. Вслед за определенным артиклем появляется существительное *orphanage* "детский дом", но на данный момент нет свободных тематических ролей и, кроме того, это слово нельзя включить в состав предыдущей именной группы. Но дальше следует слово *closed* "закрылся", которое имеет тематическую роль, и именная группа "детский дом" присоединяется к нему в качестве подлежащего, соответственно получив имеющуюся у него тематическую роль.

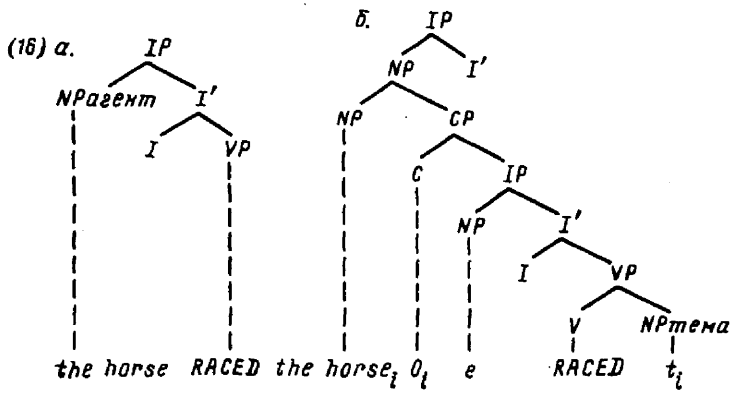
В модели главного слова синтаксический анализ предложения осуществляется последовательно, поэтому в ней так же, как и в модели заблуждения, компонент корректировки занимает центральное место. В качестве компонента корректировки Притчетт предлагает принцип тета корректировки (theta reanalysis constraint):

> Принцип тета корректировки: синтаксическая корректировка, при которой составляющий предложение элемент с приписанной ему тематической ролью выносится за пределы текущей тематической области (current theta domain) и присоединяется к новой тематической области, является невозможной для автоматического аппарата понимания человека (анализатора) (Pritchett 1992: 94).

По мнению автора, принцип тета корректировки объясняет все предложения заблуждения, представленные в (3) (как легкие, так и сложные). Рассмотрим, как модель главного слова при помощи принципа тета корректировки объясняет канонические предложения заблуждения на примере предложения (1.4): *#The horse RACED past the barn fell* "Лошадь, ПРОГНАННАЯ мимо сарая, упала".

The horse "лошадь" идентифицируется как именная группа, но никакие действия по присоединению не предпринимаются в данный момент, так как пока отсутствует слово, приписывающее тематические роли. Вслед за этим анализируется слово *RACED* "ПРОБЕЖАЛА/ПРОГНАННАЯ", которое в лексиконе обозначено как глагол с тематической сеткой (валентной структурой) <АГЕНТ, ТЕМА>. Поскольку это слово омонимично с морфологической точки зрения — прошедшее время глагола или страдательное причастие — на данном этапе могут быть построены два разных синтаксических представления, связанные с двумя разными наборами тематических ролей. Тета присоединение выбирает первый вариант: синтаксическое представление первых

двух слов в этом предложении представляет собой главное предложение (16а), а не сокращенное придаточное определительное (16б):



В случае (16а) именная группа *the horse* будет присоединена в качестве подлежащего и получит тематическую роль АГЕНТА от глагола *RACED*, а роль ТЕМЫ останется локально не приписанной. В случае (16б) роль ТЕМЫ может быть приписана пустой категории e_i, имеющей одинаковый индекс с нулевым оператором 0_i [Chomsky 1981; 1986; Sells 1985]. Хотя в обоих случаях приписывается только одна роль, в конструкции (16б) нужно строить еще одну, определительную, именную группу, в которой главное слово — *the horse*, а для нее, как видно из синтаксического представления (16б), нет тематической роли, что явно нарушает принцип тета присоединения. Вследствие этого анализатор выбирает первое синтаксическое представление.

Следующей анализируется предложная группа *past the barn*, и она присоединяется к уже частично построенному синтаксическому представлению. Далее идентифицируется глагол *fell*, который в лексиконе имеет тематическую сетку (набор валентностей), состоящей из одной роли <ТЕМА>. Теперь для того, чтобы полностью удовлетворить грамматические принципы, нужно перестроить *the horse RACED past the barn* так, чтобы часть предложения *RACED past the barn* стала именной группой, могущей получить от глагола *fell* его обязательную тематическую роль. Для этого нужно, чтобы существительное *the horse*, которое вначале было ядром (вершиной) именной группы с приписанной ему ролью глагола *RACED*, было скорректировано так, чтобы стать ядром подлежащего глагола *fell* и получить от него его роль ТЕМЫ. Структурная позиция подлежащего *fell* лежит вне пределов тематической области *RACED*, и таким образом, принцип тета корректировки объясняет, почему такая корректировка является невозможной для анализатора.

Тета присоединение и принцип тета корректировки составляют основу модели главного слова, первоначально разработанной Причеттом в его статье "Феномен предложений заблуждения и грамматические основы анализа предложения" [Pritchett 1987], основанной на его диссертации, и в статье "Позиция главного слова и многозначность при синтаксическом анализе предложения" [Pritchett 1991].

Последующие исследования показали, однако, что следуя принципу тета корректировки, конструкции с двойным дополнением (3.5) должны быть сложными предложениями заблуждения, но слушатели не считают эти предложения таковыми. Например, сравним предложение типа *They gave her gifts on Christmas* "Они подарили ей подарки на Рождество" с предложением *They gave her gifts to John* "Они отдали ее подарки Джону". Во втором предложении именная группа "подарки" должна быть скорректирована так, чтобы она стала главным словом первой именной группы "ее", а не возглавляла свою собственную отдельную, вторую именную группу. Это очевид-

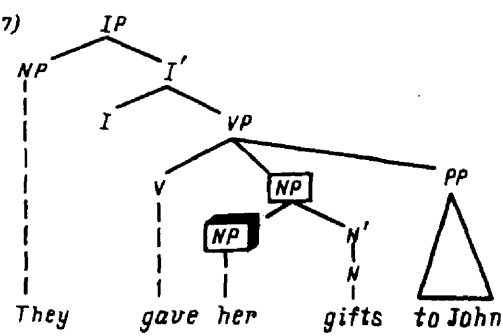
ное нарушение принципа тета корректировки, тем не менее, это предложение не представляет никаких трудностей для анализатора. Таким образом, принцип тета корректировки нуждался в модификации. Такая модификация модели была предложена Притчеттом в его книге [Pritchett 1992].

Как было отмечено выше, особенность модели главного слова заключается в той тесной связи, которая постулируется между теорией грамматики и характеристиками анализатора. Именно грамматические принципы как таковые являются неотъемлемой составной частью теории синтаксического анализа предложения. Однако принцип тета корректировки является нововведением; такого принципа не существует в теории управления и связывания, на которой Притчетт основывает свою модель анализа предложения. Кроме того, формулировка принципа тета корректировки характеризует процесс корректировки в терминах соотношения позиции в синтаксическом представлении предложения, которую первоначально занимает элемент, и позиции, на которую он перемещается после корректировки. Первоначальная позиция называется исходной (source), а окончательная называется результирующей (target). Следовательно, принцип тета корректировки можно переформулировать как условие, характеризующее взаимоотношение между исходной и результирующей позициями в синтаксическом представлении предложения, используя при этом чисто структурные синтаксические понятия управления и доминирования (dominance). Переформулированный таким образом принцип тета корректировки был назван Причеттом принципом непосредственной локальности[22] (on-line locality constraint):

> Принцип непосредственной локальности: Исходная позиция (если таковая имеется), должна управлять или доминировать над результирующей позицией (если таковая имеется), которую займет элемент предложения после корректировки, иначе присоединение окажется невозможным для автоматического аппарата понимания человека [Pritchett 1992: 101].

Теперь, по утверждению автора, конструкции с двойным дополнением перестают представлять проблему для модели главного слова. Предложение *They gave her gifts to John* "Они отдали ее подарки Джону" после корректировки будет иметь следующее представление (17)[23]:

(17)



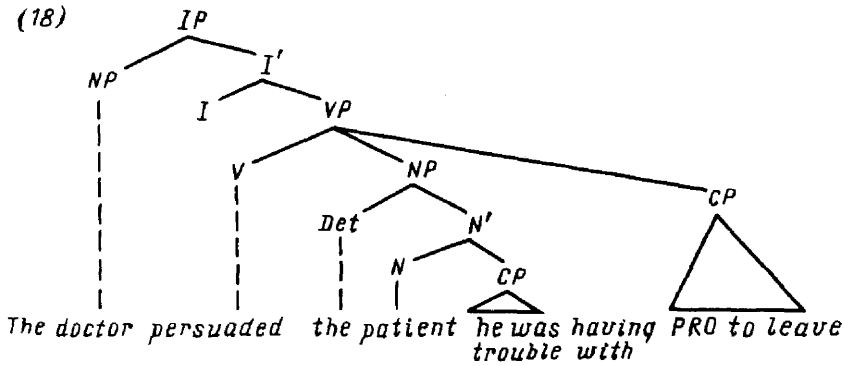
Как мы видим, принцип непосредственной локальности дает объяснение тому факту, что корректировка в (17) не является проблематичной. В окончательном синтаксическом представлении предложения (17) позиция второго дополнения, которую первоначально занимала именная группа *gifts* "подарки", управляет главным

[22] Термин "on-line" противопоставляется по своим характеристикам "off-line". Процедура, происходящая "on-line", имеет место непосредственно, прямо во время какого-либо процесса (в данном случае речь идет о мгновенном, автоматическом, имеющем место в сознании человека процессе синтаксического анализа предложения при его понимании). "Off-line" процедура происходит сразу после того, как этот процесс завершился (см. также сноску 16).

[23] Простым квадратом обозначена исходная позиция, объемным — результирующая позиция.

словом N результирующей позиции, поскольку она управляет максимальной проекцией NP этого существительного.

Нужно теперь доказать, что принцип непосредственной локальности, как и принцип тета корректировки, объясняет все типы многозначности. Притчетт применяет принцип непосредственной локальности при анализе шести типов многозначности из (3) и утверждает, что этот принцип работает во всех случаях. Рассмотрим для примера, как с помощью тета присоединения, которое продолжает занимать центральное место в модели главного слова, и принципа непосредственной локальности проходит анализ предложения типа (3.4) — придаточное дополнительное/определительное: *#The doctor persuaded the patient THAT he was having trouble with to leave* "Доктор уговорил пациента, с которым у него не ладились отношения [,] уйти":



В синтаксическом представлении [Frazier 1982] вторая позиция аргумента внутри VP, которую первоначально занимало придаточное дополнительное, не управляет результирующей позицией, так как управление блокировано промежуточной максимальной проекцией NP, и элемент предложения, подвергающийся корректировке, является не аргументом, а адъюнктом. Мы видим, что так же, как раньше трудность этого предложения заблуждения объяснялась с помощью принципа тета корректировки, сейчас она объясняется с помощью принципа непосредственной локальности.

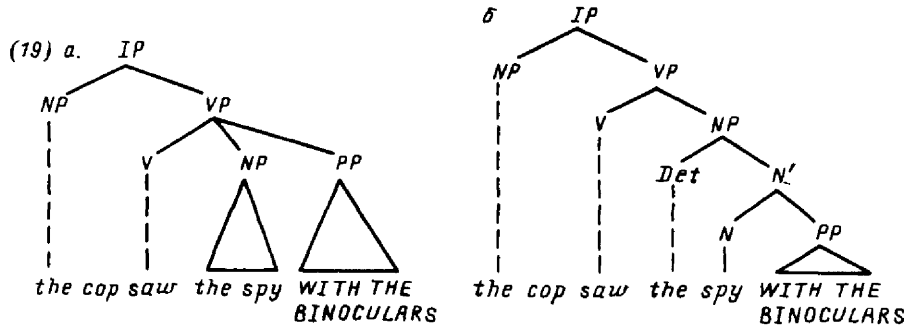
Тета присоединение, основанное на грамматическом принципе тета критерия, — универсальный принцип, по мнению Притчетта. Можно также предположить, что и другие глобальные грамматические принципы применяются подобным же способом. Таким образом, автор предлагает следующую формулировку принципа обобщенного тета присоединения (generalized theta attachment):

> Принцип обобщенного тета присоединения: Каждый синтаксический принцип во время анализа предложения стремится к тому, чтобы быть максимально удовлетворенным [Pritchett 1992: 138].

Синтаксические принципы удовлетворяются тем более, чем более мотивировано (licensed) появление каждого элемента в данной конкретной части предложения, и соответственно, чем больше мотивирующих признаков (падежей, тематических ролей) приписано.

Модель главного слова в том виде, как она описана выше, в основном отвечает на вопрос, каким образом анализируются и как присоединяются к синтаксическому дереву аргументы. Противопоставление аргументов и адъюнктов является одним из центральных вопросов для теории управления и связывания и теории синтаксиса в целом. Было давно отмечено, что аргументы и адъюнкты ведут себя по-разному по отношению к самым различным синтаксическим модулям, а именно, существует принципиальная асимметрия в их поведении. В частности, некоторые типы синтаксического передвижения (syntactic movement) применимы к аргументам, но не приме-

нимы к адъюнктам[24]. Кроме того, аргументы занимают такие позиции в предложении, где они могут получить падеж[25] и тематическую роль; адъюнкты же не имеют ни падежа, ни тематической роли. Таким образом, присоединение той или иной фразовой группы к уже частично построенному синтаксическому представлению должно отражать разницу между аргументами и адъюнктами. Например, в предложении типа (3.6) — присоединение предложной группы к именной группе/глагольной группе: *The cop saw the spy WITH THE BINOCULARS* "Полицейский увидел шпиона с биноклем/в бинокль" — предложная группа *WITH THE BINOCULARS* может быть присоединена либо как квази-аргумент глагола (19а), либо как адъюнкт второй именной группы (19б):



Многочисленные эксперименты доказывают, что информанты предпочитают присоединять многозначные предложные и именные группы в качестве аргументов, а не адъюнктов. Поэтому любая модель синтаксического анализа предложения, стремящаяся адекватно смоделировать психологическую реальность, должна объяснять это предпочтение. Притчетт отмечает, что разграничение аргументов и адъюнктов само по себе недостаточно. Как известно, существует определенный класс наречий и обстоятельств других частей речи, которые можно охарактеризовать как "квази-аргументы". В частности, обстоятельства места, времени и "инструментальные" обстоятельства ведут себя по отношению к определенным синтаксическим процессам как обязательные дополнения (квази-аргументы), в то время как обстоятельства образа действия и причины являются прототипическими адъюнктами. Модель главного слова объясняет особенности присоединения обстоятельств первого типа, "квази-аргументов", при помощи принципа обобщенного тета присоединения. С другой стороны, эта модель ничего не предсказывает в случае обстоятельств второго типа. В предложении *Frank visited the man WITH THE SUITCASE* "Франк навестил человека с ЧЕМОДАНОМ" предложная группа может быть адъюнктом как именной, так и глагольной группы, однако обобщенное тета присоединение не отдает предпочтения ни одному из этих двух возможных вариантов. В то же время очевидно, что модификация именной группы более уместна. К сожалению, замечает автор, вопросы, связанные с грамматическим характером адъюнктов и их синтаксическим анализом, практически не изучены, поэтому на настоящем этапе развития лингвистики можно только строить догадки[26] на эту тему. Модель главного слова в явном виде признает тот факт, что

---

[24] Принципиальные различия между аргументами и адъюнктами — слишком широкая тема для того, чтобы ее можно было кратко изложить в обзоре. Эту разницу можно приблизительно сравнить с традиционным делением на дополнения (прямые, косвенные) и обстоятельства. В теории управления и связывания существуют классические тесты, помогающие различить аргументы и адъюнкты.

[25] Речь здесь идет о синтаксических падежах, а не о морфологических падежах, поэтому теория падежа в теории управления и связывания является универсальной и применяется даже в тех языках, в которых морфологических падежей нет.

[26] Книга Притчетта [Pritchett 1992] была опубликована в 1992 г., когда модель ассоциации Фразер и Клифтона [Frazier, Clifton 1993], в центре внимания которой находятся именно адъюнкты и модифицирующие конструкции, была еще неизвестна.

некоторые глагольные адъюнкты ведут себя как дополнения, и этим объясняется их аргументоподобный статус. Но поскольку на настоящий момент не существует адекватной грамматической теории адъюнктов, объяснение характера их поведения в процессе синтаксического анализа предложения придется отложить до тех пор, пока эта теория будет разработана в достаточной мере.

В качестве завершающего штриха к своей модели Притчетт очень бегло рассматривает вопрос, каким образом модель главного слова объясняет характерные особенности синтаксического анализа предложения в других языках, а именно, в китайском, иврите, немецком, японском и корейском. Однако, как уже было отмечено выше, синтаксический анализ предложения в других языках следует рассмотреть в отдельном обзоре.

2.3. Модель структурного детерминизма Горрелла [Gorrell 1992; 1992a; 1993; 1994] / минимальных обязательств Вайнберга [Weinberg 1993; 1994].

Детерминистская модель М. Маркуса [Marcus 1980; 1983]. Предшественницей модели структурного детерминизма можно считать модель детерминистского анализа предложения, предложенную Маркусом в 1980 г. Маркус выдвинул гипотезу, что естественный язык анализируется детерминированно слева направо, то есть на каждом отдельном этапе анализа при возникновении временной многозначности принимается одно-единственное структурное решение (никакого параллелизма), и это принятое решение не может быть изменено никогда (никакой корректировки). Таким образом, анализатор всегда действует детерминистски: его действия всегда однозначны. Детерминистский анализатор имеет не только доступ к лексической информации и информации субкатегоризации, но и обладает специальным буфером, позволяющим ему заглядывать вперед, чтобы иметь доступ к той части предложения, которая еще не проанализирована. Объем этого буфера, по мнению автора, равен трем составляющим предложения. Важно отметить, что объем буфера измеряется составляющими, то есть, фразовыми группами, а не отдельными словами.

Маркус также предложил компьютерное воплощение своей теоретической модели синтаксического анализа предложения, получившее название программы ПАРСИФАЛЬ. ПАРСИФАЛЬ имитирует действия анализатора и поэтому должен так же, как и люди, испытывать трудности. Однако трудности представляют только те предложения с временной многозначностью, в которых область снятия многозначности лежит за пределами трехчленного буфера. Так, в каноническом предложении заблуждения (1.4) как только подлежащее the horse "лошадь" оказывается проанализированным, буфер будет содержать в себе глагол, предлог и именную группу: <RACED>$_V$ <past>$_P$ <the barn>$_{NP}$. Последующий глагол *fell* еще не видим, а когда он появляется, первоначально построенный анализ оказывается неправильным. Однако поскольку анализ является детерминистским, раз построенное синтаксическое представление нельзя автоматически изменить, требуются уже сознательные усилия для повторного анализа. Маркус считает, что тот факт, что ПАРСИФАЛЬ не в состоянии успешно проанализировать с первого раза сложные предложения заблуждения, составляет аргумент в пользу психологической реальности его модели.

Однако модель Маркуса в том виде, как она была изложена изначально, нуждалась в доработке, так как было очевидно, что существуют очень короткие сложные предложения заблуждения, целиком помещающиеся в буфере, и такие предложения (исходя из модели Маркуса) не должны представлять никаких трудностей. Тем не менее, предложения типа *#The boat FLOATED quickly sank* "Лодка, ПРОПЛЫВШАЯ/ПРОГНАННАЯ быстро, утонула" так же сложно анализировать, как и более длинные варианты канонических предложений заблуждения. Маркус, Хиндл и Флек [Marcus, Hindle, Fleck 1983] опубликовали сильно доработанную версию детерминистской модели, получившую название дескриптивной модели, или Д-модели.

Синтаксическое представление предложения, получившееся в результате анализа, в этой модели представляет собой не фразовое дерево, а набор синтаксических описаний (descriptions), которые содержат принципиально меньше информации, чем фразовое дерево. Синтаксическое представление состоит из описания отношений просто доминирования, но не прямого, непосредственного доминирования[27]. Такие отношения доминирования дают набор составляющих предложения, но не обязательно непосредственно составляющих, что позволяет описаниям фразовой структуры оставаться не до конца определенными. Например, предложение *I drove my aunt from Peoria's car* "Я вел машину моей тети из Пеории" представляет трудности для ПАРСИФАЛЯ (но не для человека) потому, что он присоединяет именную группу *my aunt* в качестве непосредственного прямого дополнения, то есть слишком высоко в дереве, вместо того, чтобы присоединить ее в качестве притяжательного адъюнкта прямого дополнения. Таким образом, принципиальное отличие Д-теории от детерминистской модели заключается в том, что при анализе подобных предложений фиксируется только информация о том, что глагольная группа доминирует над многозначной именной группой, а непосредственное доминирование не указывается. По мере того, как анализируется новый лексический материал в предложении, к уже существующим описаниям отношений доминирования добавляются новые и, в конце концов, точное расположение именной группы в синтаксическом представлении предложения можно определить по сумме таких описаний. В этом и состоит суть синтаксического анализа предложения в Д-теории.

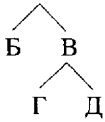
Модель структурного детерминизма Горрелла. Как было отмечено выше, основная идея Д-теории, разработанной Маркусом и его коллегами, заключается в том, что при описании отношений между фразовыми группами в синтаксическом представлении предложения фиксируются только отношения просто доминирования, но не отношения непосредственного доминирования. Эта идея легла в основу модели структурного детерминизма П. Горрелла [Gorrell 1992; 1992a; 1993; 1994] и модели минимальных обязательств А. Вайнберг [Weinberg 1993; 1994]. Эти две модели весьма тесно связаны и похожи.

Для начала мы рассмотрим модель структурного детерминизма. Горрелл предлагает описание этой модели в целом ряде своих работ [Gorrell 1992; 1993; 1993a; 1994] и в новой книге под названием "Синтаксис и восприятие" [Gorrell 1992a]. Поскольку книга суммирует большинство положений статей, остановимся подробнее на теории в том виде, как она изложена в книге.

Центральным конструктом своей модели Горрелл считает синтаксическое представление предложения в виде фразового дерева. Фразовое дерево несет в себе следующую информацию: а) иерархическую группировку частей предложения в составляющие; б) последовательность составляющих слева направо; в) грамматические категории составляющих. Фразовое дерево состоит из узлов, представляющих грамматические категории и ветвей, которые соединяют эти узлы между собой. Грамматические отношения между узлами описываются при помощи отношения доминирования. Отношение доминирования определяется следующим образом: для двух узлов x и y, x доминирует над y тогда и только тогда, когда поддерево, связывающее эти два узла в дереве, состоит исключительно из нисходящих ветвей. Отношение 'x доминирует над y' записывается в виде d (x, y). Если между x и y нет промежуточных узлов, x непосредственно доминирует над y. Последовательность узлов слева направо определяется при помощи отношения предшествования (precedence). Если x предшествует y, это отношение обозначается в виде p (x, y). Так, в дереве (20a) узел А доминирует над всеми остальными узлами дерева, Б предшествует В и всем узлам, над которыми доминирует В, т.е. Г и Д. Все возможные отношения доминирования и предшествования в этом дереве представлены в

[27] Просто доминирование означает, что узел А выше в дереве, чем узел Б; непосредственное доминирование означает, что узел А является матерью узла Б.

списке (20б):

(20) а. 

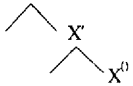
б. d (А, Б) d (А, В) d (А, Г) d (А, Д) p (Б, В) p (Б, Г) p (Б, Д) d (В, Г) d (В, Д) p (Г, Д)

Помимо отношений доминирования и предшествования, считающихся первичными, между узлами дерева существуют и другие, вторичные, отношения, речь о которых пойдет ниже.

Существуют несколько условий, которые характеризуют правильность построения дерева. Одно из них имеет более общий характер. Оно предполагает, что два узла могут находиться в отношении предшествования, только если они не находятся в отношении доминирования. Другое условие имеет более частный характер и применяется только к синтаксическому дереву. Это условие определяет способ приписывания имен узлам и следует при этом положениям X-штрих модуля (X-bar theory) в теории Управления и Связывания. X-штрих модуль постулирует, что каждая категория проецирует (projects) узлы, доминирующие над ней, т.е. N → NP, V → VP и так далее. Промежуточная проекция главного слова X — X' (X-штрих)[28] — появляется только тогда, когда она ветвится. Вся необходимая синтаксическая информация будет представлена следующим образом: грамматические категории главного и зависимых слов и типы возможных дополнений главного слова определены в лексиконе, грамматические категории проекций и отношения доминирования определены при помощи X-штрих модуля, а отношения предшествования — при помощи параметра следования главного слова.

Фразовое дерево в том виде, как оно описано выше, является результатом деятельности анализатора, следовательно, анализатор представляет собой механизм, который строит структурное представление для набора символов на входе, основанное на синтаксически значимой лексической информации и общих принципах синтаксиса. Главными операциями анализатора являются операции создания узлов (node creation) и присоединения. Операция создания узлов вызывается либо при помощи X-штрих модуля, либо при помощи субкатегоризации. Операция присоединения представляет собой процедуру установления либо отношения доминирования, либо отношения предшествования между узлами x и y. Подобное определение присоединения представляет анализатор в виде механизма для построения дерева, в котором отношения доминирования и предшествования являются первичными структурными отношениями в синтаксическом представлении. Все остальные виды отношений между узлами в дереве, например, управление[29], являются вторичными и всегда определяются через посредство первичных отношений. Какое представление построит такой анализатор, когда он анализирует многозначное предложение? Горрелл предлагает следующий принцип, принцип простоты, который ограничивает действия операций создания узлов и присоединения:

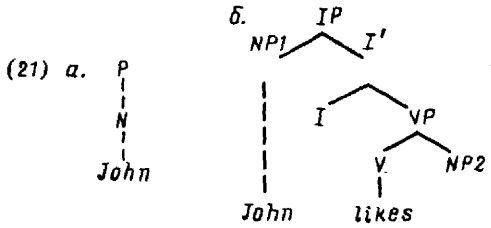
---

[28] В теории управления и связывания приняты следующие обозначения для элементов X-штрих модуля: X^0 — главное слово, терминальный символ; X'— категория, промежуточная между главным словом и его максимальной проекцией; X" — (X-двойной штрих) — максимальная проекция главного слова, XP, фразовая группа X' и X" (= XP) — нетерминальные символы: 

[29] Под управлением здесь и далее в описании модели структурного детерминизма имеется в виду стандартное определение этого отношения в теории управления и связывания, основанное на понятии "сикоманды" (c-command).

Принцип простоты (simplicity): Строить самое простое дерево, насколько это возможно, исходя из характера материала на входе и принципов грамматики [Gorrell 1992: 124].

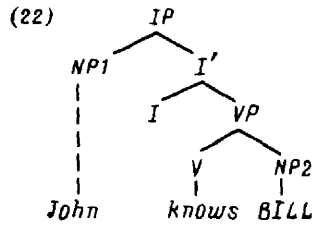
Рассмотрим на нескольких примерах, как действует модель структурного детерминизма. Начнем с простого однозначного предложения типа *John likes basketball* "Джон любит баскетбол". Принято считать, что такие предложения однозначны, однако в тот момент, когда на вход анализатора поступает слово *John* "Джон", возможны несколько вариантов его анализа: а) Джон и Мэри любят баскетбол, б) Джон, я его люблю, в) Джон и Мэри, г) Джон. Ведь в модели структурного детерминизма анализатор не обладает буфером, позволяющим ему заглянуть вперед, как это предполагалось в модели Маркуса. Анализатор строит именную группу NP для существительного "Джон", т.е. создает максимальную проекцию (21а). Это самая простая из возможных структур. Следующее слово "любит" является спрягаемой формой глагола, и поэтому анализатор строит VP и IP, и построенное на данный момент синтаксическое представление может быть изображено следующим образом (21б):



в. d (IP, NP1) d (IP, I') d (I', I) d (VP, V)
d (VP, NP2) d (I', VP) p (I, VP) p (V, NP2)

В (21в) выборочно представлены некоторые отношения доминирования и предшествования в дереве (21б). Пустой узел NP2 построен анализатором, поскольку того требует субкатегоризация глагола "любит", который должен иметь при себе обязательное прямое дополнение. Когда будет проанализирована именная группа "баскетбол", она будет присоединена в качестве этого прямого дополнения.

По точно такой же схеме происходит анализ легкого предложения заблуждения[30] типа *John knows BILL likes basketball* "Джон знает, [что] Билл любит баскетбол". Однако в этом предложении есть многозначное с точки зрения валентностей слова *knows* "знает". Синтаксическим представлением для этого предложения будет (22) после того, как проанализирована пост-глагольная именная группа NP2:

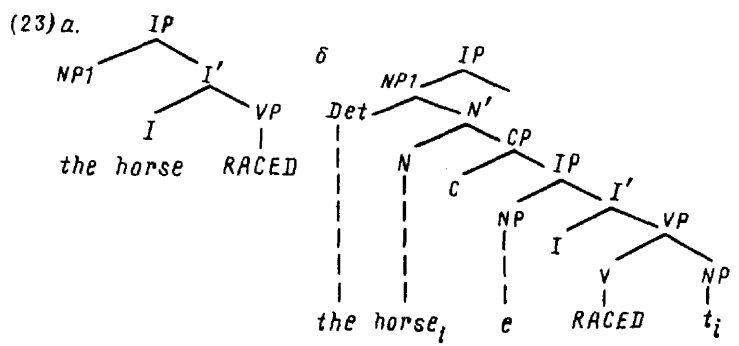


Анализатор следует тому предположению, что VP доминирует над пост-глагольной именной группой NP2 и в этой позиции V также доминирует над NP2. Кроме того, V

[30] Горрелл, так же как и Притчетт, считает только сложные предложения заблуждения предложениями заблуждения, поскольку в легких предложениях заблуждения корректировка происходит автоматически и незаметно для человека.

предшествует NP2, а следовательно, NP2 может получить и тематическую роль, и синтаксический падеж. Когда же анализируется глагол "любит", то строится дополнительная структура, в которой между VP и NP2 вставляется CP. Никакого эффекта заблуждения не происходит, поскольку в модели структурного детерминизма только операции уничтожения или изменения структуры, не являющиеся автоматическими, влекут за собой эффект заблуждения, в то время как добавление структуры, которое всегда происходит автоматически, не имеет последствий в виде осознанной корректировки предложения. Уничтожение и изменение структуры представляют собой операции по изменению первичных отношений доминирования и предшествования. В примере (22) глагол в главном предложении продолжает предшествовать NP2, поскольку он предшествует узлу CP, который теперь доминирует над именной группой NP2. Но если первичные отношения в данном примере не меняются после появления слова "любит", этого нельзя сказать про вторичные отношения. Глагол V в главном предложении больше не управляет именной группой NP2 как раньше, поскольку управление блокируется двумя новыми максимальными проекциями, CP и IP, а если нет отношения управления между двумя фразовыми группами, то невозможно ни приписывание тематической роли, ни синтаксического падежа. Операции приписывания тематической роли и синтаксического падежа и есть вторичные отношения. Они представляет собой интерпретационные процессы, существующие внутри синтаксиса. Эти процессы корректируют интерпретацию синтаксического представления предложения по мере того, как анализируется каждая лексическая единица.

Обратимся теперь к анализу сложных предложений заблуждения в том виде, как он осуществляется в модели структурного детерминизма. Каноническое предложение заблуждения (1.4) — *#The horse RACED past the barn fell* "Лошадь, ПРОГНАННАЯ мимо сарая, упала". Первоначальная структура, следуя принципу простоты, изображена в (23а), а правильная, более сложная, структура — в (23б) (см. (2.4)):



Следует обратить внимание на то, что в обоих случаях (23а) и (23б) NP1 доминирует над VP, и поэтому эффект заблуждения не может возникнуть в результате изменения отношений доминирования. Он возникает в результате изменения отношений предшествования, так как максимальной проекции VP причастия "ПРОГНАННАЯ" в (23а) предшествует максимальная проекция NP "лошадь". В (23б), однако, это не так, потому что NP "лошадь" доминирует над VP "ПРОГНАННАЯ". Как мы видим, происходит изменение одного из первичных отношений предшествования, и именно это вызывает эффект заблуждения.

Но отношения предшествования представляют собой только один из типов первичных отношений в модели структурного детерминизма, второй тип, как мы помним, составляют отношения доминирования. Рассмотрим пример предложения, в котором необходимо изменить именно первоначально построенные отношения доминирования, чтобы проверить, действительно ли в таком случае возникает эффект заблуждения. В предложении (7) *#While Mary was sewing THE DRESS fell on the floor* "Пока Мэри шила,

ПЛАТЬЕ упало на пол" следующее за глаголом слово присоединяется к уже частично построенному синтаксическому представлению как часть глагольной группы, потому что это самое простое представление (как того требует принцип простоты), к тому же оправданное материалом на входе и соответствующее принципам грамматики. За именной группой "ПЛАТЬЕ" следует глагол "упало", и возникает эффект заблуждения, потому что скорректированный анализ требует изменения одного из первоначально построенных отношений доминирования, а именно отношения доминирования между глагольной группой, возглавляемой глаголом "шила", и именной группой "ПЛАТЬЕ".

Таким образом, Горрелл предлагает модель структурного детерминизма, являющуюся последовательной (serial) моделью синтаксического анализа предложения, в которой нет необходимости в отдельном компоненте корректировки. Она концептуально проста, так как осуществляет операции только структурного характера: операции создания узлов и присоединения. Эмпирически модель в состоянии объяснить широкий спектр как однозначных предложений, так и легких и сложных предложений заблуждения. Она органически включает в себя важные аспекты других моделей (например, принцип простоты основывается на принципе минимального присоединения модели заблуждения), однако идет дальше в том, что придает не меньше значения отношениям предшествования, чем отношениям доминирования. Таким образом, заключает автор, модель структурного детерминизма мотивирована как с теоретической, так и с эмпирической точки зрения.

Модель минимальных обязательств. Модель минимальных обязательств описана А. Вайнберг в нескольких статьях [Weinberg 1993; 1994]. Наиболее подробно особенности этой модели представлены в статье [Weinberg 1993].

Модель минимальных обязательств называется так потому, что, по мнению автора, анализатор будет строить минимальное синтаксическое представление, необходимое для приписывания всех имеющихся в нем грамматических признаков. Формально это требование воплощено в принципе быстрой интерпретации (the principle of quick interpretation): Анализатор присоединяет аргументы, используя минимальное количество отношений доминирования и грамматических принципов, необходимых для приписывания значимых характеристик.

Модель имеет следующие характеристики:

1) анализатор строит не полностью определенное (underspecified) синтаксическое представление, представленное в виде отношений доминирования и предшествования. Это представление должно быть максимально мотивировано различными грамматическими модулями (X-штрих модуль, модуль синтаксического падежа, тета-критерий и т.д.);

2) анализатор не обладает буфером, позволяющим ему заглядывать вперед на несколько шагов. На вход анализатора поступает единовременно по одному слову, и анализатор пытается присоединить это слово к уже частично построенному синтаксическому представлению;

3) синтаксический анализ предложения осуществляется последовательно;

4) после того как появляется информация, снимающая многозначность, анализатор может добавить к уже построенному им набору не полностью определенных отношений доминирования и предшествования новые отношения;

5) при синтаксическом анализе предложения анализатор следует принципу монотонности, который означает, что анализатор может добавлять новые отношения к уже частично построенному представлению в любой момент, но не может менять уже построенные отношения.

Центральным понятием модели минимальных обязательств является понятие мотивирования (licensing) и мотивирующего слова (licenser). Каждая фразовая группа для того, чтобы иметь возможность присоединиться к уже частично построенному синтаксическому представлению, должна быть тем или иным способом мотивирована.

Вайнберг выделяет два типа отношений мотивированности: внутреннее (internal) и внешнее (external).

Внутреннее мотивирование определяется некоторым признаком, которым обладает определенная категория и который является сигналом для анализатора, чтобы построить данную категорию, так как для того, чтобы быть построенной, каждая категория должна быть внутренне мотивирована. Каждая максимальная проекция должна быть мотивирована отдельно. Признаками, служащими для внутреннего мотивирования категории, являются признаки, характеризующие X-штрих модуль и падежи. Внешнее мотивирование представляет собой приписывание некоторых грамматических признаков категории со стороны мотивирующего слова, находящегося за пределами фразовой группы, в состав которой входит данное слово. Чаще всего это означает, что одна фразовая группа доминирует над другой. Вайнберг предлагает следующий принцип мотивированности: "Когда внутренний состав категории полностью определен, она должна быть присоединена при помощи отношения доминирования к какой-нибудь другой категории; если невозможно построить такое мотивированное отношение доминирования, присоединения не получается" [Weinberg 1993: 348]. Мотивированное отношение доминирования возникает тогда, когда обе категории, мотивирующая и мотивированная, внутренне мотивированы, а отношение доминирования между ними — внешне мотивировано. Выделяются следующие типы внешней мотивированности: а) приписывание тематической роли (со стороны глагола своим дополнениям и подлежащему); б) приписывание синтаксического падежа (предлог приписывает синтаксический падеж своему объекту); в) функциональная селекция (например, артикль функционально выбирает именную группу); г) модификация (наречие модифицирует глагол); д) предикация (отношение между определительным придаточным предложением и его главным словом); е) согласование между спецификатором (specifier) и главным словом (ядро IP приписывает падеж подлежащему, находящемуся в позиции спецификатора.

Принцип быстрой интерпретации объясняет, почему легкие предложения заблуждения не вызывают сложностей для анализатора при их анализе. В предложении *John knew BILL liked Mary* "Джон знал, [что] Билл любит Мэри", принцип быстрой интерпретации указывает анализатору на то, что именная группа "БИЛЛ" является дополнением первого глагола. Когда на входе появляется следующее слово, то есть "любит", анализатор должен "опустить" эту именную группу в придаточное предложение при помощи дополнительных отношений доминирования. Первоначальное отношение доминирования, при котором первый глагол доминирует над именной группой, осталось неизменным, и осознанного эффекта заблуждения не возникает.

Тот факт, что анализатор при синтаксическом анализе предложения должен следовать принципу монотонности (характеристика 5), объясняет канонические, то есть сложные предложения заблуждения. Рассмотрим опять классический пример (1.4) #*The horse RACED past the barn fell* "Лошадь, ПРОБЕЖАЛА/ПРОГНАННАЯ мимо сарая, упала". Как мы помним, при анализе "ПРОБЕЖАЛА/ПРОГНАННАЯ" синтаксическое представление является многозначным: либо это может быть простое предложение, либо сокращенное определительное придаточное предложение. Первый вариант представляет собой минимальную структуру, поскольку в таком случае нет нужды строить сложное подлежащее, распространенное придаточным предложением. Поэтому, следуя принципу быстрой интерпретации, анализатор строит следующие отношения доминирования и предшествования, соответствующие первому варианту (24):

(24) d (IP NP I) d (NP, N (the horse)) d (I, VP) d (VP, V (*RACED*))
p ([NP the [N, horse]], *RACED*)

Однако глагол "упала" требует пересмотреть этот первоначальный анализ. Второй вариант — сокращенное определительное придаточное — несовместим с тем набором первичных отношений, который был выстроен при первоначальном анализе, поскольку

категория N-штрих (N'), доминирующая над глагольной фразой *RACED*, больше не будет ей предшествовать. Одна и та же категория, как уже было отмечено выше, не может одновременно доминировать и предшествовать по отношению к другой категории (см. синтаксическое представление в (16б)).

Таким образом, модель минимальных обязательств в состоянии объяснить широкий круг предложений заблуждения разных видов, используя при этом набор простых принципов. Тем не менее, эта модель ограничена в том смысле, что минимальное синтаксическое представление следует из принципа монотонности в том смысле, что разрешается добавлять структуру, но не разрешается ее менять. Анализатор ведет себя очень консервативно при постройке синтаксического представления предложения. Монотонность является характерной особенностью детерминистского анализа, и предпочтение присоединения фразовой группы как аргумента, а не адъюнкта, следует из того факта, что детерминистские системы осуществляют анализ снизу вверх. Обладающий такими характеристиками анализатор не нуждается в особом компоненте корректировки.

СПИСОК ЛИТЕРАТУРЫ

Секерина И.А. 1994 — Модели синтаксического анализа предложения в процессе понимания. Обзор американских психолингвистических теорий // Реферативный журнал. Социальные и гуманитарные науки. Серия 6: Языкознание. № 3. М., 1994.

Berwick R., Fong S. 1994 — Madam, Butterfly redux: parsing English and Japanese with a Principles and parameters approach // Japanese sentence processing / Ed. by R. Mazuka, N. Nagai. New Jersey, 1994.

Bever T.G. 1970 — The cognitive basis for linguistic structures // Cognition and development of language / Ed. by J.R. Hayes. N.Y., 1970.

Carlson G., Tanenhaus M. 1988 — Thematic roles and language comprehension // Syntax and semantics / Ed. by W. Wilkins. San Diego, 1988.

Chomsky N. 1981 — Lectures on government and binding. Dordrecht, 1981.

Chomsky N. 1986 — Barriers. Cambridge, 1986.

Crain S., Steedman M. 1985 — On not being led up the garden path: the use of context by the psychological parser // Natural language parsing / Ed. by D. Dowty, L. Kartunnen, A. Zwicky. Cambridge (Mass.), 1985.

Cuetos F., Mitchell D. 1988 — Cross-linguistic differences in parsing: restrictions on the use of the late closure in Spanish // Cognition, Lausanne, 1988.

De Vincenzi M. 1989 — Syntactic parsing strategies in a null-subject language. Ph. D. Dissertation. University of Massachusetts, 1989.

Ferreira F., Henderson J. 1991 — How is verb information used during syntactic parsing // Understanding word and sentence / Ed. by G.B. Simpson. Amsterdam, 1991.

Fodor J.A., Bever T.G., Garrett M.F. 1974 — The psychology of language: an introduction to psycholinguistics and generative grammar. N.Y., 1974.

Fodor J.D., Inoue A. 1993 — The diagnosis and cure of garden paths. School and University Center, City University of New York, 1993 (рукопись).

Frazier L. 1978 — On comprehending sentences: syntactic parsing strategies. Ph. D. dissertation. University of Connecticut. 1978.

Frazier L. 1987 — Syntactic processing: evidence from Dutch // Natural language and linguistic processing. 1987. V. 5. № 4.

Frazier L. 1990 — Exploring the architecture of the language system // Cognitive models of speech processing: psycholinguistic and computational perspectives / Ed. by G. Altmann. Cambridge (Mass.), 1990.

Frazier L., Clifton C. 1993 — Construal. University of Massachusetts, 1993.

Frazier L., Fodor J.D. 1978 — The sausage machine: a new two-stage parsing model // Cognition. 1978. V. 6.

Frazier L., Rayner K. 1982 — Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences // Cognition psychology. 1982. V. 14.

Gibson E.A. 1991 — A computational theory of human linguistic processing: memory limitations and processing breakdown. Ph. D. dissertation. Carnegie Mellon University. Pittsburg, 1991.

Gorrell P. 1987 — Studies of human syntactic processing: ranked parallel vs. serial models. Ph. D. dissertation. University of Connecticut, 1987.

Gorrell P. 1989 — Establishing the loci of serial and parallel effects in syntactic processing // Journal of psycholinguistic research. 1989. V. 18. № 1.

Gorrell P. 1992 — The parser as tree builder. University of Maryland. College Park, 1992 (рукопись).

Gorrell P. 1992a — Syntax and perception. N.Y., 1992.
Gorrell P. 1993 — Form- and context-based parsing strategies. University of Maryland. College Park, 1993 (рукопись).
Gorrell P. 1993a — Structural relations in the grammar and the parser. University of Maryland. College Park, 1993 (рукопись).
Gorrell P. 1994 — Japanese trees and the garden path // Japanese sentence processing / Ed. by R. Mazuka, N. Nagai. New Jersey, 1994.
Grice P.H. 1989 — Studies in the way of words. Cambridge (Mass.), 1989.
Hickok G. 1993 — Parallel parsing: evidence from reactivation in garden-path sentences // Journal of psycholinguistic research. 1993. V. 22. № 3.
Inoue A., Fodor J.D. 1994 — Information-paced parsing of Japanese // Japanese sentence processing / Ed. by R. Mazuka, N. Nagai. New Jersey, 1994.
Kimball J. 1973 — Seven principles of surface structure parsing in natural languages // Cognition. 1973. V. 2. № 1.
Marcus M. 1980 — A theory of syntactic recognition for natural language. Cambridge (Mass.), 1980.
Marcus M., Hinndle D., Fleck M. 1983 — D-theory: talking about trees // Proceedings of the 21-st annual meeting of the Association for computational linguistics. Cambridge, 1983.
Nicol J., Pickering M. 1993 — Processing syntactically ambiguous sentences: evidence from semantic priming // Journal of psycholinguistic research. 1993. V. 22. № 3.
Pritchett B. 1987 — Garden path phenomena and the grammatical basis of language processing. Ph. D. dissertation. Harvard University. Cambridge (Mass.), 1987.
Pritchett B. 1987a — Garden path phenomena and the grammatical basis of language processing // Language. 1987. V. 64. № 4.
Pritchett B. 1991 — Head position and parsong ambiguity // Journal of psycholinguistic research. 1991. V. 20. № 2.
Pritchett B. 1992 — Grammatical competence and parsing performance. Chicago; London, 1992.
Pritchett B. 1992 a — Parsing with grammar: islands, heads and garden paths // Islands constraints / Ed. by H. Goodlick, M. Rochemont. N.Y., 1992.
Saussure F. 1932 — Cours de linguistique générale. Paris, 1932.
Sells P. 1985 — Lectures on contemporary syntactic theories: an introduction to government-binding theory, generalized phrase structure grammar, and lexical-functional grammar. Stanford, 1985.
Taraban R., McClelland J. 1988 — Constituent attachment and thematic role assignment in sentence processing: influences of context-based expectations // Journal of memory and language. 1988. V. 27. № 6.
Weinberg A. 1993 — Parameters in the theory of sentence processing: minimal commitment theory goes east // Journal of psycholinguistic research. 1993. V. 22. № 3.
Weinberg A. 1994 — Licensing constraints and the theory of language processing // Japanese sentence processing / Ed. by R. Mazuka, N. Nagai. New Jersey, 1994.