

СИНТАКСИЧЕСКАЯ НЕОДНОЗНАЧНОСТЬ И РОЛЬ ПУНКТУАЦИИ В ЕЕ РАЗРЕШЕНИИ: ВЫБОР РАННЕГО ИЛИ ПОЗДНЕГО ЗАКРЫТИЯ¹

М.С. Власов, О.В. Торопчина

Ключевые слова: пунктуация, синтаксическая неоднозначность, присоединение определительного придаточного предложения, раннее закрытие, позднее закрытие, прайминг.

Keywords: punctuation, syntactical ambiguity, relative clause attachment, early closure, late closure, priming.

Введение

В настоящий момент известно, что лингвистическими детерминантами синтаксической неоднозначности на уровне предложения являются:

- грамматические формы слов;
- варианты лексической сочетаемости слов;
- отсутствие необходимых компонентов в предложении;
- порядок слов;
- пунктуационное оформление и другие факторы².

Б.Л. Иомдин и А.С. Бердичевский в своей статье [Бердичевский, Иомдин, 2007] обозначают проблему учета пунктуации в системах автоматической обработки текста (АОТ) при разрешении неоднозначности разного вида. Признавая не только синтаксическую, но и семантическую функцию знаков препинания для систем АОТ, авторы считают, что конкретные алгоритмы реализации пунктуационного разрешения неоднозначности для систем АОТ сильно зависят от общего устройства конкретной системы. По их мнению, пунктуация может использоваться (и используется) в системах, основанных как на правиле, так и на статистическом подходе, но способы использования должны различаться. В работе продемонстрировано, что пунктуация заключает в себе ценную и порой не так уж сложно извлекаемую информацию, которой не стоит пренебрегать,

¹ Исследование выполнено при поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – кандидатов наук (проект № МК-2986.2014.6 «Норма и стандарт в русской пунктуации в России и США»).

² Последний фактор был рассмотрен М.С. Власовым в предыдущих работах на материале русского и английского языков [Власов, 2008].

а системы АОТ следует последовательно обучать пунктуационным правилам и соблюдению этих правил. Нисколько не упрощая проблемы обучения пунктуации не только систем АОТ, но и самих носителей языка, авторы статьи предлагают переводить знаки препинания лексическими средствами (русское двоеточие дополнить *а именно* или заменить на *потому что*, запятую дополнить союзом *и* или заменить на *него* и т.п.). Данные проблемы представляют сегодня особый интерес для лингвистов и разработчиков систем АОТ.

Не меньшую актуальность приобретают сегодня психолингвистические исследования разрешения глобальной синтаксической неоднозначности в языках разной структуры. Так, некоторые типы такой неоднозначности (например, синтаксическая неоднозначность сложноподчиненных предложений с определительными придаточными типа *Кто-то застрелил служанку актрисы, которая стояла на балконе*) являются специальными предметами исследований на материале разных языков [Федорова и др., 2007], [Анисимов, 2013], [Fodor, 1998, p. 285-319]. Сегодня доказано, что одна группа языков демонстрирует предпочтение раннего закрытия (РЗ): например, носители русского, польского, японского, африкаанса, греческого, голландского, немецкого, французского языков присоединяют к придаточному первый член сложной именной группы (И1 – вершина придаточного: *на балконе стояла служанка*). Носители английского, арабского, норвежского, румынского, шведского языков, напротив, предпочитают принцип позднего закрытия (ПЗ): в качестве вершины придаточного они выделяют второй член именной группы (И2 – вершина придаточного: *на балконе стояла актриса*).

Вопрос о том, как происходит выбор вершины, изначально решался в рамках теории универсального анализа: считалось, что синтаксически неоднозначные предложения интерпретируются носителями разных языков в соответствии с одной и той же стратегией [Fodor et al., 1974]. Однако известная работа [Cuetos, Mitchell, 1988] скорректировала направления психолингвистических изысканий в этой области, и неоднозначные предложения с определительными придаточными сыграли в этом решающую роль: на их материале было обнаружено, что носители разных языков используют не одну и ту же, а разные стратегии. Универсальность анализа подверглась сомнению, и возникла идея о том, что ментальный механизм понимания предложения имеет частноязыковую природу. Впоследствии, в дополнение к двум указанным точкам зрения (универсальной и частноязыковой), была высказана третья: выбор стратегии анализа также зависит от индивидуальных характеристик носителя языка [Pearlmutter, MacDonald, 1995]. Таким образом, на данный момент не существует единой точки зрения на рассматриваемую проблему.

Большая экспериментальная работа проведена на данном материале М.В. Юдиной. Автор допускает ряд факторов, влияющих на выбор РЗ или ПЗ: 1) длина придаточного предложения (короткие придаточные чаще модифицируют ПЗ, а в случае длинных придаточных распределение зависит от других факторов); 2) лингвистическая настройка (при разрешении неоднозначности человек склонен настраиваться на ту стратегию, которая является в языке наиболее частотной); 3) одушевленность существительных (при наличии в ИГ одушевленного и неодушевленного существительного одушевленное существительное выбирается чаще в любой позиции); 4) влияние контекста (после контекста, склоняющего к РЗ, будет больше РЗ, и наоборот)¹; 5) *синтаксический прайминг*² (отмечена «преднастройка» на РЗ у 57% испытуемых после РЗ-прайма и у 46% испытуемых также «преднастройка» на РЗ после ПЗ-прайма); 6) класс глагольной вершины (то есть некоторые предложения во всех экспериментах показали схожие результаты по критерию присоединения придаточного, в независимости от экспериментального условия); 7) тип предиката (предложения с глаголами чувственного восприятия и глаголы движения имеют большой процент РЗ, глаголы же, связанные с мышлением и интеллектом, – ПЗ) [Юдина, 2010].

Кроме этого, в работе О.В. Федоровой установлено, что помимо этих факторов, определяющих тип выбора (РЗ или ПЗ), следует учитывать осознаваемость / неосознаваемость испытуемым потенциальной неоднозначности [Федорова и др., 2007].

Исследование пунктуационного фактора при разрешении синтаксической неоднозначности данного вида на материале русского языка ранее не проводилось.

Методика проведения эксперимента

¹ Согласно данным М.В. Юдиной, после контекста, склоняющего к РЗ, 91% испытуемых выбрали вариант РЗ, а после контекста, склоняющего к ПЗ – 60% реакций в пользу РЗ.

² В психологии *праймингом* называется изменение скорости или точности решения задачи после предъявления информации, связанной с содержанием или контекстом этой задачи, но не соотносящейся прямо с ее целью, а также повышение вероятности спонтанного воспроизведения этой информации в подходящих условиях (по М.В. Фаликману и А.А. Койфману). В психолингвистике *синтаксический прайминг* – это тенденция говорящего повторять синтаксическую конструкцию высказывания, произнесенного незадолго до этого. О.В. Федорова трактует синтаксический прайминг как «явление, при котором обработка одного высказывания упрощает обработку другого высказывания за счет повторения его синтаксической структуры» [Федорова, 2009]. Явление синтаксического прайминга заключается в следующем: при ответной реакции на какой-либо стимул говорящий склонен использовать те синтаксические конструкции, которые он в недавнем прошлом каким-либо образом обработал (услышал, прочитал, произнес). Высказывание, осуществляющее преднастройку, называется «праймом», а высказывание, на порождение или понимание которого, как предполагается, окажет влияние прайм, называют «целью» (цит. по: [Юдина, 2009]).

Учитывая накопленный опыт исследований вариантов раннего / позднего закрытия, мы поставили перед собой специфическую задачу – раскрыть роль пунктуации в процессе идентификации и разрешения синтаксической неоднозначности носителями русского языка. Для выяснения того, влияет ли пунктуационное оформление предложения на выбор варианта РЗ или ПЗ, мы составили список из шести предложений сходной синтаксической структуры, три из которых были неоднозначными (в них допускалось присоединение к определительному придаточному и первого, и второго члена ИГ), а три других предложения – с единственной возможной вершиной придаточного. Введение отвлекающих предложений обусловлено тем, что в ходе эксперимента возможен эффект «самопрайминга» (подсознательной самонастройки) на определенный тип закрытия (описано в [Юдина, Федорова, 2009]).

В качестве исследуемого нами фактора выступало наличие или отсутствие запятой перед определительным придаточным. Нулевая гипотеза состояла в отсутствии различий в выборе вариантов РЗ и ПЗ в зависимости от данного фактора. Тестовые предложения №№ 2, 3, 5 содержали синтаксическую неоднозначность.

Предложения с расставленными запятыми:

1. *На полу было много воды, которая постоянно прибывала и прибывала.*

2. *Преступник застрелил служанку актрисы, которая отказалась пустить его в квартиру.*

3. *Мальчик швырял в воробьев камнями, которые разлетались во все стороны.*

4. *Это была деревня, в которой я так любил бывать в детстве.*

5. *Заседание было посвящено рассмотрению заявлений рабочих, которые поступили за последний месяц.*

6. *На перрон прибыл поезд, который ждали уже двое суток.*

Предложения с отсутствующими запятыми:

1. *На полу было много воды которая постоянно прибывала и прибывала.*

2. *Преступник застрелил служанку актрисы которая отказалась пустить его в квартиру.*

3. *Мальчик швырял в воробьев камнями которые разлетались во все стороны.*

4. *Это была деревня в которой я так любил бывать в детстве.*

5. *Заседание было посвящено рассмотрению заявлений рабочих которые поступили за последний месяц.*

6. *На перрон прибыл поезд который ждали уже двое суток.*

В пилотном эксперименте приняли участие 20 носителей русского языка (10 мужчин и 10 женщин) в возрасте от 18 до 25 лет, не имеющие филологического образования, от которых было получено 120 реакций на предложенные стимулы, в том числе 60 реакций – на тестовые предложения. Испытуемые были разбиты на две группы: 10 человек получали предложения с расставленными запятыми, 10 – с пропущенными. Фактор наличия / отсутствия запятой перед определительным придаточным носил имплицитный характер: второй группе не сообщалось, что предложения содержат пунктуационные ошибки. Обеим группам предлагались 2 задания: 1) прочитать вслух указанные предложения (запись фиксировалась на диктофон); 2) выбрать один из двух вариантов ответа по трем предложениям с синтаксической неоднозначностью (с целью выявления предпочтения раннего или позднего закрытия). Вопросы были следующие:

1. Преступника отказалась впустить в дом: а) актриса (И2), б) служанка (И1).
2. В разные стороны разлетались: а) воробы (И1), б) камни (И2).
3. За последние месяц поступили: а) рабочие (И2), б) заявления (И1).

Предъявляемые три неоднозначных предложения имели разные условия семантического присоединения придаточного, обладали различным *прайм-эффектом* (в нашем случае – это преднастройка на определенный тип закрытия). Данная методика эксперимента широко известна в современной психолингвистике [Юдина, Федорова, 2009], [Scheepers, 2003].

Так, в предложении № 2, являющемся хрестоматийным, наблюдается тенденция к выбору РЗ в 62% случаев [Юдина, 2010], при этом каждое из двух имен (И1, И2) теоретически могло оказаться вершиной придаточного. В предложении № 3 мы допускаем эффект ПЗ-прайма: глагол *разлетаться* в русском языке чаще сочетается с неодушевленными существительными (по данным Национального корпуса русского языка наблюдается 1879 вхождений подобных сочетаний с заданным расстоянием от 1 до 3 слов, а при сочетании с одушевленными существительными – только 504 вхождения). Такое распределение предпочтений подтверждается и нашим интроспективным предположением (существительное *камни* сочетается с глаголом *разлетаться* чаще, чем существительное *воробы*). При ПЗ-прайме предпочтение РЗ в русском языке составляет 60% [Юдина, 2010]. Наконец, предложение №5 обладало четко выраженным эффектом РЗ-прайма, с приоритетным выбором И1 (*заявления*) в качестве вершины придаточного. В подобных случаях для русского языка характерно РЗ в 91% случаев [Юдина, 2010].

Помимо фиксирования ответов испытуемых (относительно типа закрытия) с помощью программы *WaveSurfer* регистрировались следующие акустические данные:

- 1) общий темп произнесения предложений (слов в секунду);
- 2) абсолютный и относительный темп произнесения И1 и И2, разность между данными переменными;
- 3) длительность пауз между предложениями-стимулами и перед определительными придаточными.

Результаты эксперимента.

При постановке запятой перед союзным словом в предложении № 2 испытуемые выбрали РЗ в 60% случаев и ПЗ – в 40%; при отсутствии запятой – РЗ – в 90% и ПЗ в – 10% соответственно. По предложению № 3 реакции распределились следующим образом: при наличии запятой – ПЗ – 60%, РЗ – 40%, при отсутствии – ПЗ – 70%, РЗ – 30%. На предложение № 5 получены следующие варианты: при наличии запятой – РЗ – 100%, при отсутствии – РЗ – 80%, ПЗ – 20%.

Стоит отметить, что по гендерному признаку предпочтение вида закрытия распределилось абсолютно в равной пропорции: и РЗ, и ПЗ предпочли по 50% мужчин и 50% женщин.

Исследование фактора наличия или отсутствия запятой проводилось нами с использованием непараметрических процедур средствами SPSS. Независимой переменной являлось условие наличия или отсутствия запятой перед придаточным предложением. Основной зависимой переменной являлся тип закрытия.

В качестве других зависимых переменных выступали следующие акустические характеристики реакций испытуемых: 1) абсолютный темп¹ произнесения И1; 2) абсолютный темп произнесения И2; 3) разница между абсолютным темпом И1 и И2; 4) относительный темп произнесения И1; 5) относительный темп произнесения И2; 6) разница между относительным темпом И1 и И2; 7) длительность паузы перед началом произнесения предложения; 8) длительность паузы перед определительным придаточным.

По U-критерию Манна-Уитни нами не было установлено статистически значимых различий в выборе варианта закрытия в зависимости от наличия или отсутствия запятой во всех трех тестовых предложениях. При этом наблюдаемые отклонения средних значений в предпочтениях испытуемых не подтверждаются какими-либо значимыми различиями на уровне акустики в реакциях на предложения № 2, № 3. Возможно, полученные результаты связаны с небольшим объемом выборки.

¹ Измерялся в количестве слогов, произнесенных за 1 секунду (далее – слог / сек).

Единообразие ответов и их изохронность темповым стратегиям произнесения прослеживалось только в реакциях на предложение № 5, хотя данный эффект лишь подтвердил нашу гипотезу о семантическом прайминге: при РЗ-прайме с учетом приоритета РЗ в русском языке результат выбора РЗ очевиден и легко прогнозируем. Однако было установлено влияние фактора наличия / отсутствия запятой на темп произнесения первого члена сложной именной группы: И1 читается быстрее при наличии запятой перед относительным придаточным (в среднем 8,65 слог / сек), при ее отсутствии темп произнесения И1 снижается (в среднем 7,33 слог / сек), причем это различие является квазизначимым ($Z = -1,89$, $p = 0,059$):

Параметр	Наличие ЗП n=10 (среднее значение)	Отсутствие ЗП n=10 (среднее значение)	Различия по критериям			
			<i>U Манна-Уитни</i>	<i>W Вилкоксона</i>	<i>Z</i>	<i>p</i>
Абсолютный темп произнесения И1 ^{абс} , слог / сек.	8,65	7,33	25,0	80,0	-1,89	0,059

Таблица 1. Влияние фактора наличия запятой на абсолютное значение темпа произнесения И1 в предложении № 5.

Отмечая то, что испытуемые отдали явное предпочтение И1 в данном предложении, а двум группам оно предъявлялось с разным пунктуационным оформлением, мы предполагаем, что снижение темпа произнесения И1 при отсутствии запятой говорит о большем затруднении в синтаксическом анализе данного предложения.

В ходе эксперимента обнаружены гендерные различия, связанные с действием пунктуационного фактора. Так, было установлено, что отсутствие запятой перед придаточными слабо влияет на разницу в темпе произнесения И1 мужчинами (6,37 слог / сек) и женщинами (5,88 слог / сек). При наличии запятой наблюдается квазизначимое различие: у мужчин – 7,18 слог/сек, у женщин – 5,73 слог / сек ($Z = -1,888$, $p = 0,059$). То есть если запятая перед придаточным предложением поставлена, то мужчины произносят И1 быстрее, чем женщины в среднем в 1,25 раза.

Выявлены гендерные различия в среднем темпе произнесения тестовых предложений при наличии запятой: мужчины склонны произносить их быстрее, в среднем со скоростью 7,1 слог / сек, женщины – со скоростью 6,35 слог/сек, то есть предпочитают более медленное прочтение. Данный эффект имеет высокую статистическую достоверность ($Z = -2,136$, $p = 0,033$). Если же запятая отсутствует, то такого эффекта не наблюдается.

Полученные результаты представлены в таблицах 2 и 3:

Параметр	Отсутствие ЗП		Различия по критериям			
	Мужчины n=15 (среднее значение)	Женщины n=15 (среднее значение)	<i>U</i> Манна-Уитни	<i>W</i> Вилкоксона	<i>Z</i>	<i>p</i>
Абсолютный темп произнесения И1 ^{абс} , слог/сек.	6,37	5,88	90,00	210,00	-0,933	0,351
Абсолютный темп произнесения И2 ^{абс} , слог/сек.	6,40	5,78	59,00	179,00	-2,221	0,026
И1 ^{абс} –И2 ^{абс}	-0,03	0,10	110,50	230,50	-0,083	0,934
Общий темп произнесения предложений, слог/сек. (Т ^{общ})	6,92	6,65	87,50	207,50	-1,037	0,300
Относительный темп произнесения И1 ^{отн} (И1 ^{абс} /Т ^{общ})	0,91	0,89	99,50	219,50	-0,540	0,589
Относительный темп произнесения И2 ^{отн} (И2 ^{абс} /Т ^{общ})	0,94	0,87	89,50	209,50	-0,955	0,340
И1 ^{отн} –И2 ^{отн}	-0,02	0,01	108,50	228,50	-0,166	0,868
Длительность паузы перед прочтением предложения, сек	0,73	0,44	83,50	203,50	-1,204	0,229

Длительность паузы перед придаточным, сек.	0,10	0,03	108,00	228,00	-0,287	0,774
--	------	------	--------	--------	--------	-------

Таблица 2. Гендерные различия в темповых стратегиях произнесения тестовых предложений при отсутствии запятой.

Параметр	Наличие ЗП		Различия по критериям			
	Мужчины n=15 (среднее значение)	Женщины n=15 (среднее значение)	U Манна-Уитни	W Вилкоксона	Z	p
Абсолютный темп произнесения И1 ^{абс} , слог/сек.	7,18	5,73	67,00	187,00	-1,888	0,059
Абсолютный темп произнесения И2 ^{абс} , слог/сек.	6,32	5,71	69,00	189,00	-1,804	0,071
И1 ^{абс} –И2 ^{абс}	0,86	0,03	85,50	205,50	-1,120	0,263
Общий темп произнесения предложений, слог/сек. (Т ^{общ})	7,10	6,35	61,00	181,00	-2,136	0,033
Относительный темп произнесения И1 ^{отн} (И1 ^{абс} /Т ^{общ})	1,00	0,89	85,50	205,50	-1,122	0,262
Относительный темп произнесения И2 ^{отн} (И2 ^{абс} /Т ^{общ})	0,90	0,91	112,50	232,50	0,000	1,000
И1 ^{отн} –И2 ^{отн}	0,10	-0,02	80,50	200,50	-1,328	0,184
Длительность паузы перед прочтением предложения, сек	0,54	0,51	96,00	216,00	-0,684	0,494
Длительность паузы перед придаточным, сек.	0,07	0,06	98,00	218,00	-0,773	0,440

Таблица 3. Гендерные различия в темповых стратегиях произнесения тестовых предложений при наличии запятой.

Выводы

Запятая в сложноподчиненном предложении с определительным придаточным с двумя возможными вершинами (И1 / И2) при определенных условиях может оказывать влияние на процесс его дизамбигуации испытуемыми.

Эксперимент показал, что роль пунктуационного фактора в предложении № 3 (с ПЗ-праймом) оказалась незначительной. При анализе предложения № 2 этот фактор в целом повлиял на усиление приоритета РЗ, однако значимых темповых различий в произнесении отдельных сегментов предложения мы не обнаружили. Кроме этого, не прослеживалась изохронность паузы и наличия запятой перед придаточным предложением. Иначе говоря, ни один из «просодических ключей» разрешения данного вида неоднозначности с опорой на пунктуацию нами обнаружен не был.

Принимая во внимание приоритет РЗ в русском языке (вслед за О.В. Федоровой), мы зафиксировали эффект влияния отсутствия запятой в предложении с РЗ-праймом (№ 5) на замедление темпа произнесения И1 и снижение уровня предпочтения РЗ со 100% до 80%. Такого эффекта мы не получили по двум другим тестовым предложениям, в которых пунктуационный фактор не играл такой значимой роли.

Отмечены гендерные различия в темпе произнесения первого члена сложной именной группы в зависимости от пунктуационного фактора: женщины склонны к более медленному прочтению И1, чем мужчины. Наблюдается усиление данного эффекта при наличии запятой перед придаточным предложением.

Изучение пунктуационного фактора в разрешении глобальной синтаксической неоднозначности при разных условиях прайминга является предметом нашего ближайшего исследования.

Литература

Анисимов В.Н. Движения глаз при чтении предложений с синтаксической неоднозначностью в русском языке: автореф. дис. ... канд. биол. наук. М., 2013.

Бердичевский А.С., Иомдин Б.Л. Роль пунктуации в разрешении неоднозначности // Компьютерная лингвистика и интеллектуальные технологии. М., 2007.

Власов М.С. Теоретико-экспериментальное исследование процессов порождения и восприятия «естественной» пунктуации (на материале русского и английского языков): дис. ... канд. филол. наук. Бийск, 2008.

Федорова О.В., Янович И.С., Юдина М.В. Синтаксическая неоднозначность в эксперименте и в жизни // Компьютерная лингвистика и интеллектуальные технологии. М., 2007.

Юдина М.В. Референциальный контекст и синтаксическое наведение на факторы, влияющие на разрешение синтаксической неоднозначности: автореф. дис. ... канд. филол. наук. М., 2010.

Юдина М.В., Федорова О.В. Разрешение синтаксической неоднозначности: эффекты прайминга и самопрайминга // Компьютерная лингвистика и интеллектуальные технологии. М., 2009. Вып. 8 (15).

Cuetos F., Mitchell D.C. Cross-linguistic differences in parsing: restrictions on the use of the late closure strategy in Spanish // *Cognition*. 1988. № 30.

Fodor J.A., Bever T., Garrett M. *The Psychology of Language: an Introduction to Psycholinguistics and Generative Grammar*. New York, 1974.

Fodor J.D. Learning to parse? // *Journal of Psycholinguistic Research*. 1998. № 27.

Pearlmutter N.J., MacDonald M.C. Individual differences and probabilistic constraints in syntactic ambiguity resolution // *Journal of Memory and Language*. 1995. № 34.

Scheepers C. Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production // *Cognition*. 2003. № 89.