

# Проект типологической базы данных по синтаксическим ограничениям на передвижения

## Project of the typological database on syntactic constraints on movement

**Сердобольская Н. В.** (serdobolskaya@gmail.com)

Российский государственный гуманитарный университет

**Циммерлинг А. В.** (meinmat@yahoo.com)

Московский государственный педагогический университет им. М. А. Шолохова

**Аркадьев П. М.** (alpgurev@gmail.com)

Институт славяноведения

Работа представляет собой пилотный проект типологической базы данных по синтаксическим передвижениям и ограничениям на передвижения. Цель настоящей базы данных — обобщение и фиксация опыта, накопленного различными исследователями в ходе изучения синтаксических ограничений в языках, принадлежащих различным языковым семьям и ареалам.

### 1. Описание проекта

Работа представляет собой пилотный проект типологической базы данных по порядку слов и синтаксическим ограничениям (далее: БД). Данный проект<sup>1</sup> реализуется группой исследователей под руководством А. В. Циммерлинга (см. сайт <http://antonzimmerling.wordpress.com/current-projects/>). Участники проекта: А. В. Циммерлинг, Н. В. Сердобольская, Н. Б. Пименова, П. М. Аркадьев и А. А. Перекрестенко.

БД создается как аналог типологическим базам данных, посвященных подробному описанию отдельных явлений в неродственных языках (см., например, базу данных по реципроку <http://languagelet.uu.nl/burs/>, по согласованию <http://www.smg.surrey.ac.uk/Agreement/index.aspx>, супплетивизму <http://www.smg.surrey.ac.uk/Suppletion/index.aspx> и др.). Цель настоящей БД — обобщение и фиксация опыта, накопленного различными исследователями в ходе изучения синтаксических передвижений и ограничений на передвижения в языках, принад-

лежащих различным языковым семьям и ареалам. В частности, использованы параметры, которые обсуждаются в работах Гринберг 1963; 1970; Dryer 1985; 1992; Hawkins 1994; Циммерлинг 2006; 2007; Franks, King 2000; Тестелец 1999; 2001. Хорошо известно, что синтаксические параметры изучены в различных языках неравномерно. В частности, в грамматиках многих языков можно найти информацию о базовом порядке слов в именной группе или о наличии *wh*-передвижения, однако лишь в немногих источниках может содержаться указание на наличие в языке синтаксических островов или последовательное перечисление ограничений на топиализацию. Мы ставим задачу разработать признаковую базу, которая может быть использована как основа для анкеты по описанию синтаксических передвижений и ограничений в произвольном языке.

Аналогичным образом, неравномерно исследовано взаимодействие различных параметров порядка слов и синтаксических передвижений и ограничений. В типологических работах по порядку слов, в частности, подробно изучены вопросы взаимодействия порядков элементов в различных типах составляющих (именные группы, предложные/последовательные группы и т. п., см. Гринберг 1970; Dryer

<sup>1</sup> Работа выполнена при поддержке гранта РГНФ 09-04-00297а. Авторы выражают благодарность О. И. Беляеву за помощь в технических вопросах.

1992, а также типологический проект по порядку слов <http://linguistics.buffalo.edu/people/faculty/dryer/dryer/database>). Однако о взаимодействии порядка слов и различных типов синтаксических передвижений на настоящий момент гораздо меньше. Планируемая БД призвана служить отправной точкой для такого рода исследований.

БД создается в формате Microsoft Access 2003. Признаковая база, лежащая в основе БД, разработана на основе параметров, обсуждаемых в (Циммерлинг 2006; 2007 и др.) и содержит следующие три группы параметров: признаки порядка слов, синтаксические ограничения и передвижения. В отличие от типологической базы данных по порядку слов, разработанной М. Драйером (<http://linguistics.buffalo.edu/people/faculty/dryer/dryer/database>), в настоящей базе данных сведения о базовом порядке слов даются не в качестве основного, а лишь в качестве вспомогательного параметра (необходимого для учета синтаксических ограничений на передвижения). На настоящий момент БД существует в пилотной версии, которую планируется в дальнейшем совершенствовать за счет увеличения количества признаков, уточнения и улучшения интерфейса, а также за счет привлечения материала других языков. На настоящий момент БД содержит информацию лишь о трех языках (русский, английский, датский). В дальнейшем планируется расширение языковой выборки за счет имеющегося у группы материала. БД реализуется на английском языке.

## 2. Признаковая база

Сложность построения типологической базы данных состоит в отграничении непосредственно наблюдаемых фактов от теоретических конструктов. В особенности, это касается баз данных, фиксирующих синтаксические или семантические явления. Например, хорошо известное отклонение от базового порядка слов в русском языке (см., напр., Ковтунова 1976), проиллюстрированное в (1), формально может быть описано тремя способами: как перестановка подлежащего и глагольной группы, вынос глагольной группы влево или вынос подлежащего вправо:

(31) *Кто пришел? — Пришел отец.*

Многие синтаксические явления трактуются исследователями различным образом — в зависимости от теоретической парадигмы или от традиции описания, принятой при анализе языков данного ареала или языковой семьи. Например, вынос топиального элемента в крайнюю правую позицию в романских языках (2) описывается, как Right Detachment, Right Dislocation и т. п. (см. Lambrecht 1995; Erteschik-Shir 2007; Fernandez-Vest 2005 и др.).

### Французский

- (2) *Je l' ai vue, ta voiture.*  
я он иметь видеть твой машина  
Твою машину я видел. (букв. Я ее видел, твою машину.)

В целях устранения неоднозначности, возникающей за счет использования различной терминологии, в настоящей БД делается попытка отделить непосредственно наблюдаемые факты языка от их теоретической интерпретации, принятой в литературе. Для этого все параметры делятся на два типа: параметры дескриптивного и теоретического характера. Параметры первого типа фиксируют непосредственно наблюдаемые факты (например, базовый порядок слов в именной группе, возможность разрыва именной группы глаголом: рус. *Белую принеси миску*); параметры второго типа отражают теоретические выводы и приводятся с опорой на соответствующие теоретические статьи (например, наличие в языке локативной инверсии, топиализации и т. п.). Параметры первого типа приводятся в листаемых вкладках, параметры второго типа следуют в прямоугольнике в нижней части формы (см. рис. 1). Левая верхняя часть записи содержит базовую информацию о синтаксисе языка: название (включая сведения о генетической принадлежности), морфологический тип (агглютинативный, флективный, изолирующий, наличие инкорпорации), преобладание суффиксации или префиксации в словоизменении, базовый тип кодирования семантических ролей (аккузативный, эргативный, активный или нейтральный). Кроме того, приводится список лексических категорий языка и проблемы, возникающие при их выделении.

Перечислим используемые нами параметры. Синтаксические явления, непосредственно наблюдаемые при анализе языковых данных и требующих минимального теоретического обоснования, разделены на три группы и представлены в трех вкладках, соответственно. Используются следующие вкладки:

### 2.1. Базовый порядок слов и разрыв составляющих (Basic word order)

- базовый порядок слов во всех типах фразовых составляющих (например, Basic word order: NP);
- допустимость перестановок (задается в том же окошке, что и базовый порядок);
- разрывность всех типов фразовых составляющих элементами внешней составляющей (задается наличием галочки в соответствующем окошке, например, Split: Gen N; в поле комментариев содержатся указания на типы элементов, способных разрывать составляющие).

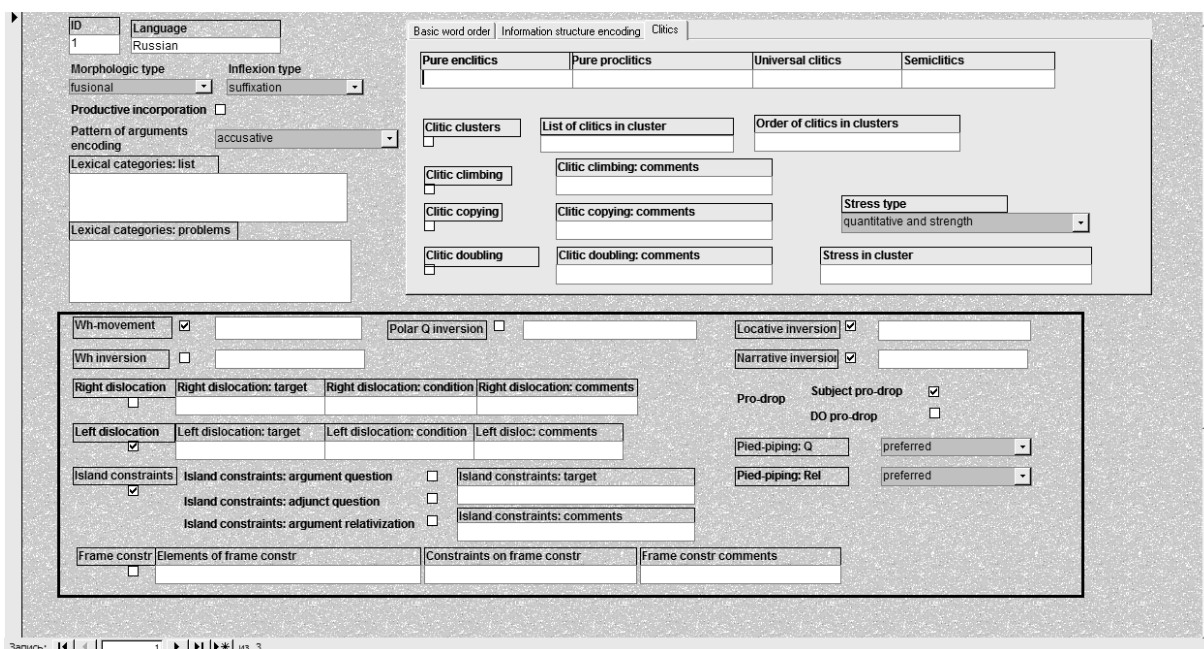


Рис. 1. База данных по порядку слов и синтаксическим ограничениям в языках мира: вкладка «клитики»

## 2.2. Коммуникативные категории (Information structure encoding)

- базовые средства выражения различных типов коммуникативного членения высказывания: предикатный фокус, узкий фокус, контрастивный фокус, топик и т. п. (см. Lambrecht 1995; Ertheschik-Shir 2007; Янко 2001; 2008);
- пример, глоссы и перевод на основные средства кодирования информационной структуры.

## 2.3. Клитики

- список клитик в анализируемом языке (если таковые имеются)
  - а) строгие энклитики (pure enclitics);
  - б) строгие проклитики (pure proclitics);
  - в) универсальные клитики (universal clitics), или клитики, способные употребляться как энклитики или проклитики;
  - г) полуклитики (semiclitics), или элементы, ведущие себя, как клитики, не во всех контекстах;
- наличие цепочек клитик (clitic clusters);
- список клитик, способных выступать в цепочке;
- порядок клитик в цепочке;
- наличие явления clitic climbing (в конструкциях с сентенциальными актантами употребление при главном глаголе клитики, семантически относящейся к зависимому глаголу, см. Rizzi 1982);
- наличие явления clitic copying (повтор клитики при одном полноударном слове);
- наличие явления clitic doubling (или местоименная реприза, т. е. употребление полной

именной группы и кореферентного ей местоимения, см. (2));

- тип ударения в языке (динамическое, силовое, силовое и динамическое, музыкальное);
- взаимодействие ударения с клитиками и цепочками клитик.

Каждый параметр сопровождается полем, содержащим комментарии и ссылки на имеющуюся литературу.

В прямоугольнике в нижней части формы даны теоретические выводы о наблюдаемых синтаксических передвижениях и ограничениях. Приводимые выводы базируются на имеющихся теоретических работах либо на исследованиях, осуществляемых участниками проекта. БД содержит информацию о следующих синтаксических явлениях:

- 1) Pro-drop, т. е. опущение местоимений в актантных ИГ;
- 2) Pied-Piping, или «эффект крысоллова» (необходимость передвижения вершины составляющей при передвижении зависимого);
- 3) Wh-передвижение;
- 4) Wh-инверсия (наличие инверсии в специальных вопросах);
- 5) Q-инверсия (наличие инверсии в общих вопросах);
- 6) Локативная инверсия;
- 7) Нарративная инверсия;
- 8) Коммуникативно нагруженные синтаксические передвижения (Right Dislocation и Left Dislocation);
- 9) Синтаксические острова;
- 10) Рамочные конструкции.

Для каждого из перечисленных параметров предусмотрено поле комментариев для более подробной информации и ссылок на имеющуюся литературу.

Параметр Pro-Drop (опущение местоимений, обозначающие актаны предиката) фиксируется отдельно для подлежащих ИГ, отдельно для ИГ, обозначающих прямое дополнение. (См. Huang 1984 относительно необходимости разграничения данных явлений.) Для фиксации явления Pied-Piping (иначе, эффект крысолова: передвижение вершины составляющей при передвижении его зависимого) предусмотрено четыре варианта заполнения: 1) Pied-Piping обязателен (напр., английский язык); 2) предпочтителен, но в большинстве случаев необязателен (напр., русский язык); 3) запрещен (авторам неизвестны такие примеры, однако они предполагаются теоретически возможными); 4) параметр неприменим, т. к. в языке нет синтаксических передвижений на поверхностном уровне (т. н. *in situ* языки; по-видимому, к таким языкам относится китайский). В БД отдельным образом отражается информация о явлении Pied-Piping при частном вопросе и релятивизации.

Коммуникативно нагруженные синтаксические передвижения (параметры Right Dislocation и Left Dislocation) объединяют все передвижения, кодирующие различные коммуникативные категории — включая топикализацию, передвижение фокуса влево, передвижение фокуса вправо, передвижение топика вправо (Right Detachment, ср. (2)) и т. п. (см. Циммерлинг 2007). Такое объединение перечисленных типов передвижений обосновано следующим. Не вызывает сомнений, что, например, различные типы передвижения фокусных составляющих влево существенным образом различаются в языках мира. Поэтому было бы неправомерно объединять их под одним термином. С другой стороны, в силу традиции определенный тип явлений тем не менее принято объединять одним термином — речь идет о топикализации. Несмотря на очевидные различия в морфосинтаксическом и интонационном оформлении топика и ограничениях на топикализацию в разноструктурных языках, термин «топикализация» активно используется различных лингвистических парадигмах и описательных традициях. Такое объединение мотивировано сходством описываемых явлений в разных языках.

Таким образом, используемая исследователями терминология не дает равномерного описания для различных коммуникативно нагруженных типов передвижений: некоторые типы передвижений

объединяются в большие группы, а некоторые получают отдельные наименования. Данные наименования в большинстве случаев не являются общепринятыми, а в значительной мере зависят от теоретической парадигмы исследователя, описывающего язык. При сохранении терминологии в настоящей БД пришлось бы тем не менее для каждого случая приводить подробные пояснения относительно конкретного типа передвижения и ограничений на его использования в данном конкретном языке. Учитывая данные соображения, в рамках данной БД было решено полностью отказаться от принятой терминологии, разделяя все имеющиеся в языке передвижения на две основные группы: передвижения вправо и передвижения влево. Далее, созданы поля для уточнения, о каком типе передвижения идет речь: поле Right/Left Dislocation target служит для указания на коммуникативную категорию, вызывающую передвижение (топик, фокус, узкий фокус, контрастный топик и т. п.), поле Right/Left Dislocation condition содержит информацию об ограничениях на данное передвижение (например, в даргинском языке очень редко допускается передвижение генитива из именной группы и т. п.). Наконец, поле Right/Left Dislocation comments предназначено для остальных комментариев и ссылок на имеющуюся литературу. Таким образом, отказ от использования терминов, принятых для различных типов передвижений, позволяет выявить сходства и различия в системах топикальных и фокусных передвижений различных языков, что необходимо для систематизации имеющейся информации и унификации терминологии в данной области.

Параметр «Синтаксические острова» включает в себя следующую информацию. Фиксируется наличие следующих типов островных ограничений в языке: ограничения на вопрос и релятивизацию из различных типов составляющих. В поле Island constraints: target перечисляются те типы составляющих, для которых действуют данные ограничения. Поле Island constraints: comments содержит комментарии и ссылки на имеющуюся литературу.

Рамочные конструкции характеризуются по следующим признакам. Поле Elements of frame constructions содержит указание на элементы, способные образовывать рамочную конструкцию. В поле Constraints on frame constructions приводятся данные об условиях, при которых реализуется рамочная конструкция. Данные сведения снабжены комментариями и ссылками на источники (поле Frame constructions comments).

## Литература

1. *Dryer M. S.* Greenbergian word order correlations // *Language*. 1992. Vol. 68. N 1. 81–138.
2. *Dryer M. S.* Word Order // T. Shopen (ed.) *Language Typology and Syntactic Description*, vol. 1: Clause Structure. Cambridge University Press. Cambridge: 1985.
3. *Erteschik-Shir N.* Information structure. The syntax — discourse interface // Oxford University Press: 2007.
4. *Fernandez-Vest J.* Vers une typologie linguistique du détachement à fondement ouralien d'Europe // *Bulletin de la Société de Linguistique de Paris*, BSL, C1, 1. Paris: 2006. Pp. 173–224.
5. *Franks S., King T. H.* A Handbook of Slavic clitics. Oxford University Press. Oxford: 2000.
6. *Hawkins J. A.* A performance theory of order and constituency // Cambridge University Press. Cambridge: 1994.
7. *Huang C., James T.* On the Distribution and Reference of Empty Pronouns // *Linguistic Inquiry* 15: 1984. Pp. 531–574.
8. *Lambrecht K.* Information Structure and Sentence Form. Topic, Focus and the Mental Representation of Discourse Referents // Cambridge University Press: 1995.
9. *Rizzi L.* Issues in Italian Syntax // Foris: 1982.
10. *Ross J. R.* Constraints on Variables in Syntax. Ph.D. thesis // Massachusetts Institute of Technology, Cambridge: 1967.
11. *Гринберг Дж.* Квантитативный подход к морфологической типологии языков // *Новое в лингвистике* // М., 1963. Вып. III. С. 60–95.
12. *Гринберг Дж.* Некоторые грамматические универсалии, преимущественно касающиеся порядка значимых элементов // *Новое в лингвистике* // М.: 1970. Вып. V. С. 114–162.
13. *Гринберг Дж., Осгуд Ч., Дженкинс Дж.* Меморандум о языковых универсалиях // *Новое в лингвистике* // М.: 1970. Вып. V. С. 31–44.
14. *Ковтунова И. И.* Современный русский язык. Порядок слов и актуальное членение предложения // Просвещение. М.: 1976.
15. *Тестелец Я. Г.* Порядок слов и структура составляющих // А. Е. Кибрик (ред.) *Элементы цахурского языка в типологическом освещении* // «Наследие». М.: 1999. С. 293–346.
16. *Тестелец Я. Г.* Введение в общий синтаксис // РГГУ. М.: 2001.
17. *Циммерлинг А. В.* Отношение свободного порядка слов и модели инверсии // *Труды международного семинара «Диалог 2006: компьютерная лингвистика и информационные технологии»* // М.: 2006. С. 541–544.
18. *Циммерлинг А. В.* Порядок слов в русском языке // *Текст, Структура и семантика. Доклады XI международной конференции* // МГУ. М.: 2007. С. 138–151.
19. *Янко Т. Е.* Коммуникативные стратегии русской речи. М.: 2001.
20. *Янко Т. Е.* Интонационные стратегии русской речи в сопоставительном аспекте. М.: 2008.