

ЧТО ТАКОЕ ТРАНСФОРМАЦИОННЫЙ АНАЛИЗ?

Одним из приемов описания синтаксических структур является метод так называемого синтаксического анализа по «непосредственно составляющим», при котором некоторое распространенное предложение сводится к сочетанию нескольких основных единиц, определяющих общий «скелет» всего предложения. Так, например, предложение *Маленький мальчик поймал большой мячик* может быть представлено следующим образом¹:

[illegible]

Из ядерных предложений можно построить, добавляя, согласно правилам данной грамматики, определенные классы слов, развернутые предложения достаточно большой длины. Например, предложение *Сестра читает книгу* можно развернуть в предложение *Моя младшая сестра весь день с удовольствием читает очень интересную книгу*. Такой процесс развития некоторого окончательного ядра называется *де ривацией*. В данном случае структура предложения не меняется. Однако это же предложение без ущерба для общего смысла может быть преобразовано следующим образом: *Сестра читает книгу — Книга читается сестрой*. Преобразования, совершенные в предложении, могут быть соответствующим образом записаны:

 $C^1_{\text{им. Г}} C^2_{\text{вин.}}$
$$x_1 - y - x_2 \rightarrow x_2 - y - x_1$$

(т. е. члены ядерной струк-

$$x_2 - y - x_1 \rightarrow C^2_{\text{им.}} - \Gamma - (\text{ся}) - C^1_{\text{тв.}} \quad \text{туры меняются местами).}$$

В дальнейшем используются следующие обозначения для частей речи в русском языке: С — существительное, Г — глагол, П — прилагательное, Н — наречие, Пр — предлог, Сз — союз, М — местоимение, Св — связочные глаголы, Мд — модальные глаголы.

тической логики, где в качестве основных категорий принимаются, с одной стороны, некоторые аксиомы, а с другой — правила их порождения и преобразования.

Если принять в качестве исходных данных некоторое характерное для данного языка количество ядерных структур, то все предложения иной структуры можно рассматривать как производные от этих основных структур. Зная правила выведения подобных производных из первичных ядер, можно построить трансформационную грамматику данного языка, способную порождать синтаксические структуры любой сложности. Таким образом, «все многообразие различных видов предложений внутри одного языка может быть описано в терминах небольшого числа ядерных структур, небольшого числа основных трансформаций и небольшого числа рекурсивных правил для сочетания предложений и для словаря»².

Наиболее подробно теория трансформационной грамматики была разработана в трудах З. Харриса и его ученика Н. Хомского. Согласно Хомскому, вся система построения грамматики трансформаций, порождающих разные типы предложений реального языка, представляется состоящей из трех последовательных этапов: 1) определение структуры предложения, 2) трансформационные преобразования, 3) морфологический уровень³. На последнем уровне происходит превращение некоторой единицы с определенным грамматическим признаком в реальное слово данного языка. Так, например, для русского языка на этом уровне произойдет превращение слова со значением «идти» с признаком «прошедшее время» в слово *шел* и т. д.

Типы трансформаций, выделяемых Харрисом и Хомским, классифицируются следующим образом:

1. Прежде всего различаются обязательные (obligatory) и желательные (optional) трансформации. Обязательная трансформация для русского языка производится, например, при построении сложного предложения с определительным словом *который* в род. падеже. В этом случае слово *который* должно быть обязательно помещено за определяемым словом в им. падеже (*Я видел человека, лицо которого показало мне знакомым*). Желательной трансформацией для русского языка могут считаться некоторые случаи инверсии (*Зима наступила* → *Наступила зима*).

2. Трансформации могут различаться и по тому, изменяется ли в данном случае вся структура предложения (как в приведенном выше примере трансформации активной формы в пассивную или при построении вопросительной конструкции в английском языке), или к неизменяемому ядру добавляются по определенным правилам некие новые элементы. Последнее происходит при трансформации отрицания (Totр.):

Я хочу читать → *Я не хочу читать*

$(x_1 - x_2 - x_3 \rightarrow x - \text{не} - x_2 - x_3)$.

3. В основу различения трансформаций может лечь и иной признак — производятся ли преобразования внутри одной структуры или объединяются и трансформируются две (или более) структуры. Примером последнего случая может служить трансформация объединения: *Он читает эту книгу* + *Я тоже читаю эту книгу* → *Он читает эту книгу, и я тоже*.

4. Далее трансформации классифицируются и на том основании, меняют ли члены ядерной структуры после трансформации принадлежность к определенному классу слов или нет. Например, при так называемой «трансформации номинализации» из ядерной глагольной структуры выводится группа существительного: *Он читает книгу* → *Его чтение книги* или *Чтение им книги*. При активно-пассивной трансформации принадлежность всех членов ядерной структуры к различным классам слов не меняется.

Всего Н. Хомский выделяет двадцать четыре типа различных трансформаций. Сюда входит трансформация номинализации, трансформация адъективации, трансформация связи слов, трансформация активной формы в пассивную, трансформация вопроса, отрицания, инверсии и др. Над одной ядерной структурой может производиться не одно преобразование, а несколько преобразований. В данном случае последовательность этих преобразований далеко не безразлична, а они должны осуществляться по заранее обусловленным правилам. Эта последовательность трансформаций, каждая из которых предполагает уже совершенной предыдущую, образует своего рода алгебру трансформаций. Например, изменения глаголов, в частности преобразования в числе, могут быть произведены лишь после изменения активной формы в пассивную, так как в данном случае число глагола может измениться. Для построения предложения *Не хочу ли я читать?* необходимо произвести три трансформации: трансформацию отрицания (Totр.), трансформацию инверсии (Тинв.) и ли-трансформацию (Тли):

² Z. S. Harris, Linguistic transformations for information retrieval. Preprint of papers for the International conference on scientific information, Washington, 1958.

³ N. Chomsky, Syntactic structures, 's-Gravenhage, 1957.

Я хочу читать

$x_1 x_2 x_3$

Тотр. $x_1 - x_2 - x_3 \rightarrow x_1 - \text{не} - x_2 - x_3$

Тинв. $x_1 - \text{не} - x_2 - x_3 \rightarrow \text{не} - x_2 - x_1 - x_3$

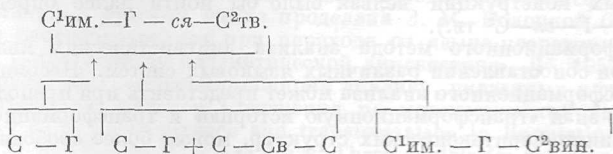
Тли $\text{не} - x_2 - x_1 - x_3 \rightarrow \text{не} - x_2 - \text{ли} - x_1 - x_3$

Алгебра трансформаций, их внутренняя структура, равно как и максимальное число трансформаций, которые могут производиться над одной структурой, различаются для каждого языка, хотя, разумеется, некоторые трансформации могут (овпадать для разных языков. При определении различных трансформаций, производимых над ядерными структурами, большое значение имеет качество самих структур, считающихся основными для данного языка. При выборе между двумя эквивалентными по смыслу предложениями большое значение приобретает вопрос о том, какую структуру из них считать основной, а какую — производной. Например, для русского языка можно было бы считать основной пассивную конструкцию $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$, а активную форму считать производной от нее. Однако в этом случае возможны были бы квазитрансформации для предложений типа *Он смеется веселым смехом* (также $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$), которые могут по общей формуле быть трансформированы в *Веселый смех смеет его*. Потому удобнее в качестве основного ядра взять структуру активной формы $C^1 \text{им.} - \Gamma - C^2 \text{вин.}$, где Γ — переходный глагол.

Классы, над которыми производятся определенные трансформации, могут отличаться от обычных частей речи. Например, прилагательные *железный* и *хороший* принадлежат к разным классам, так как второе из них может трансформироваться в конструкцию со словом *очень*, а первое — нет.

Применение трансформационного метода анализа лингвистического материала в некоторых случаях дает в руки лингвисту возможность более глубоко понять специфику того или иного языкового явления. В этом отношении трансформационный анализ существует как метод выделения различных классов слов или некоторых сочетаний слов. Трансформационный анализ помогает в этом случае разбить формально совпадающие языковые единицы на более мелкие (благодаря выяснению трансформационных возможностей для отдельных членов этих классов). Трансформационный анализ является здесь как бы вспомогательным методом для составления некоторой точной грамматики. Так, например, совпадающие по форме предложения *Человек причисляется гребенкой*, *Он вернулся моряком* и *Поле обрабатывается трактором* (все три в форме $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$) представляют на самом деле совершенно различные структуры. В первых двух предложениях глагол находится в активной форме, в третьем — в пассивной. Если же попытаться определить «трансформационную историю» этих предложений, то первое из них восходит к предложению *Человек причисывается* (СГ), второе — к двум предложениям *Он вернулся* — *Он стал моряком* [СГ/С(Св)С]. Третье предложение порождено из структуры *Трактор обрабатывает поле* ($C^1 \text{им.} - \Gamma - C^2 \text{вин.}$).

Итак, в данном случае форма $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$ восходит к трем различным структурам:



Попытку классификации различных преобразований, возможных для конструкций с творительным падежом в русском языке, сделала Д. С. Уорф в статье «Трансформационный анализ конструкции с творительным падежом в русском языке»⁴. Она выделяет 6 видов конструкций с творительным падежом и показывает, в какие конструкции могут трансформироваться различные предложения с творительным падежом. Так, например, внутри формы $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$ Уорф различает семь разных структур в зависимости от того, в какие структуры может трансформироваться эта конструкция. В частности, предложение *Учреждение руководится директором* может трансформироваться лишь в конструкцию $C^2 \text{им.} - \Gamma - C^1 \text{тв.}$ (*Директор руководит учреждением*), а предложение *Лука заливались водой* трансформируется в конструкцию $C^2 \text{им.} - \Gamma - C^2 \text{вин.}$ (*Вода залила луга*), конструкцию $C^1 \text{вин.} - \Gamma - C^2 \text{тв.}$ (*Лука залило водой*) и в конструкцию $\Gamma - \text{Пр} - C^2$.

Различного типа трансформации для формы $C^1 \text{им.} - \Gamma - \text{ся} - C^2 \text{тв.}$ представлены Д. С. Уорф в виде приводимой ниже таблицы, из которой следует, какие «совпадающие», но по существу различные конструкции скрываются за внешне тождественной формулой:

⁴ D. S. Worth, Transform analysis of Russian instrumental constructions, «Word», vol. 14, № 2-3, 1958.

C^2 им. — Г — C^1 вин.	C^2 им. — Г — C^1 тв.	Г — ся — Пр	C^1 вин. — Г безл. — C^2 тв.	C^1 им. — Г — C^2 вин. — C^2 тв.	C^1 им. — бмл — C^2 тв.	C^1 им. — Г — ся — Пр — C^2	
+	—	+	—	—	—	—	Комната наполнялась толпой
+	—	+	+	—	—	—	Лука заливались водой
+	—	—	—	—	—	—	Счет составляется бухгалтером
—	+	—	—	—	—	—	Учреждение руководится директором
(+)	—	+	—	+	—	—	Студент порезался бритвой
(—)	—	+	—	—	+	—	Иван вернулся стариком
(—)	—	—	—	—	—	+	Баржи тянулись рядами

Анализ «трансформационной истории» каждого типа обусловил возможность построения таблицы трансформационных соответствий для каждого из семи типов. В данном случае трансформационный анализ как метод оказался гораздо продуктивнее метода описания по «непосредственно составляющим», применяя который при анализе подобных конструкций нельзя было бы пойти далее определения общей формулы (C^1 им.—Г—ся— C^2 тв.).

Идея трансформационного метода анализа лингвистических явлений находит применение и при сопоставлении различных языковых систем. Несомненный интерес применение трансформационного анализа может представить при преподавании языка, когда учащийся, зная «трансформационную историю» и трансформационные возможности для большинства синтаксических структур, может более свободно конструировать фразы изучаемого языка.

Принципы трансформационного анализа используются для создания особого вида алгоритма машинного перевода, когда производится «свертывание» синтаксической структуры при анализе переводимого языка и соответственное «развертывание» при синтезе переводящего языка⁵. При синтезе в машинном переводе трансформационный анализ как таковой именно и является одним из ведущих методов машинного перевода. Если же понимать трансформационный метод более широко, как общий анализ синтаксических структур одного языка и их трансформацию в другом языке, то этот метод, как полагает И. Бар-Хиллел, может служить основой для машинного перевода в целом⁶.

Трансформационный метод может применяться при преобразовании не только целых предложений, но и отдельных словосочетаний как глагольного, так и именного характера. Например, словосочетание *стеклянная посуда* трансформируется в конструкцию *посуда из стекла* (ПС→СПрС), конструкция *говорить громким голосом*

⁵ Описание такой структуры алгоритма машинного перевода см. в статье Т. Н. Молюшниковой «Некоторые вопросы синтаксиса в связи с машинным переводом с английского языка на русский», ВЯ, 1957, № 4.

⁶ J. Bar-Hillel, Report on the state of machine translation in the United States and Great Britain, Jerusalem (Israel), 1959 (mimeogr. изд.).

(ГПС) — в конструкцию *громко говорить* (НГ)⁷. Подобные эквивалентные конструкции есть трансформационные преобразования одна другой. По своему употреблению такие синонимические эквиваленты, однако, отличаются от синонимических отношений на лексическом уровне. В последнем случае одна лексическая единица заменяется в речи другой при помощи простой подстановки: *Я увидел врага — Я увидел неприятеля*. При трансформационной же замене обязательно необходимы внутренние преобразования среди компонентов исходной структуры (*виноградное вино* → *вино из винограда*: $P - S \text{ им.} \rightarrow x_1 - x_2 \rightarrow x_3 - \text{из} - x_3 \rightarrow S^1 \text{ им.} - \text{из} - S^2 \text{ род.}$).

Несмотря на все более распространяющееся в последнее время применение трансформационного анализа как одного из продуктивных и новых методов исследования лингвистических явлений, в целом методика подобного подхода к языку нельзя считать единой и в достаточной степени установившейся. Прежде всего бросается в глаза неоднородность трансформаций, предлагаемых Хомским и Харрисом. Такая трансформация, как преобразование активной формы в пассивную, несомненно, имеет место в реальной речи и совершается каждым говорящим в каждом языке, где существуют две эти конструкции. Другие же трансформации, такие, как трансформация противоположности, уподобления и др., носят несколько искусственный характер. Не существует достаточно четких критериев для определения того, какую конструкцию считать для данного языка исходной, а какую производной от нее. Некоторые структуры могут при речевом процессе функционировать параллельно для носителя языка. Среди трансформаций, выделяемых Хомским, есть трансформации и слишком общего характера (трансформация отрицания, трансформация вопроса), и весьма частные, факультативные трансформации (типа трансформации сепарации для английского языка).

Наиболее сложный и спорный вопрос, возникающий в связи с изучением и применением трансформационного анализа, — это проблема значения синтаксической структуры. Если при ряде трансформаций (активно-пассивной, соединения и т. д.) общее значение (смысл) высказывания действительно остается неизменным и происходит лишь внешнее преобразование компонентов ядерной структуры, то при трансформации отрицания, трансформации вопроса и тому подобных трудно говорить о полной тождественности «исходной» и «преобразованной» структур. В ряде случаев приходится говорить о трансформации с добавлением некоторого элемента ($x + a$), как, например, в трансформации *стеклянная посуда* → *посуда из стекла*, или о трансформациях с полным сохранением числа преобразуемых элементов. В связи с этой неопределенностью в понимании «значения» неясным остается вопрос о границах применения трансформационного метода. Излишне широкое понимание этого приема анализа может привести к универсализации самого понятия трансформаций и трансформ, которые, перестав быть точными и единообразными, потеряют свое значение для научного исследования⁸.

Т. М. Николаева

⁷ Классификация подобных эквивалентных переходов и трансформаций для разного рода сочетаний в русском языке проделана З. М. Волоцкой (З. М. В о л о ц к а я, Вопросы словообразования при переходе от языка-посредника к выходному языку, Тезисы Совещания по математической лингвистике, Л., 1959).

⁸ Помимо работ по трансформационному анализу, указанных в статье, см. также следующую литературу: «Тезисы Совещания по математической лингвистике», Л., 1959; Z. S. H a r r i s, Co-occurrence and transformation in linguistic structure, «Language», vol. 33, № 3, 1957; N. C h o m s k y, Three models for the description of language, «I.R.E. Transactions on information theory», vol. IT-2, № 3, Sept., 1956; N. C h o m s k y, Linguistics, logic, psychology and computers. Lectures given at the University of Michigan, 1958; R. B. L e e s, [печ. на кн.:] N. Chomsky, Syntactic structures..., «Language», vol. 33, № 3, 1957.