

Татьяна П. Карпилович (Минск)

Грамматические преобразования в научном тексте

Ключевые слова:
научная статья, композиционная структура, первичный текст, вторичный текст, резюме, кодирование информации, морфологические преобразования, синтаксические преобразования.

У раду се описују типови граматичких трансформација приликом кодирања информација текста научног чланка у његов резиме. Морфолошке трансформације се односе на замену у погледу вида и времена глагола у оквиру парадигме, док се синтаксичке трансформације анализирају на нивоу реченичних чланова и на плану реченичних конструкција. Посебно се разматрају услови за трансформације глаголског рода.

Научный текст носит ярко выраженный когнитивный характер, поскольку в нем не только фиксируется, но и создается научное знание. Эти функции научного текста обуславливают возможность применения к его анализу когнитивно-коммуникативного подхода, предполагающего рассмотрение его собственно языковых свойств и релевантных когнитивных и прагматических факторов, определяющих коммуникативные стратегии субъектов и выбор конкретных языковых форм в процессе речевой деятельности (Кубрякова 2004; Цурикова 2002).

Одним из видов речевой деятельности является смысловая компрессия текста, т.е. построение вторичного текста (резюме, реферата, аннотации) на основе некоторого первичного текста. Закономерности построения, или кодирования, вторичного текста можно выявить, например, путем сопоставительного анализа текста научной статьи и ее авторского резюме. Такой анализ был проведен нами на материале научных статей по вычислительной технике и компьютерной лингвистике из англоязычной журнальной периодики и сборников материалов международных конференций. В этих

Система 2006

изданиях научные публикации имеют четкую композиционную структуру, включающую с небольшими вариациями следующие части: *Abstract/Summary, Introduction, Previous Research, Methods, Results, Discussion, Conclusion*. Такая каноническая структура предоставляет широкие возможности для изучения не только процессов декодирования, но и кодирования научного знания. С одной стороны, организация научной информации по определенной композиционной модели, отражающей динамику познавательной деятельности с акцентуацией нового знания на фоне уже известного, является коммуникативно эффективной, поскольку способствует ориентации читателя в сложном интеллектуальном содержании научного текста. С другой стороны, четкая композиция научной статьи дает возможность исследователю выявить закономерности кодирования информации в ее различных структурных частях, установить их связи в рамках целого текста.

В соответствии с данными когнитивной психологии в процессе кодирования информации из первичного текста во вторичный можно условно выделить два основных этапа, целями которых являются: а) выбор наиболее важной информации из исходного текста, б) преобразование этой информации во вторичный текст. В нашем исследовании моделирование первого этапа осуществляется на основе данных, полученных в результате сопоставительного анализа авторских предтек-

стов¹ и других композиционных частей научных статей на английском языке. Этот анализ показал, что релевантная информация для кодирования во вторичном тексте сконцентрирована в сильных позициях научной статьи – в ее вводной и заключительной частях, что соответствует когнитивной интерпретации этих периферийных частей как *фигур на фоне* всего пространства научного текста. Введение в исследуемых предметных областях, как правило, содержит данные для индикативного резюме (своего рода аннотации), а именно такие смысловые компоненты, как *анализ предшествующих исследований* и *сообщение о теме собственного исследования (гипертеме)*. Заключение, являясь обобщением проделанной работы, обязательно включает основные научные *результаты и выводы* и, таким образом, может служить основой для построения информативного резюме, или краткого реферата (Карпилович 2000).

Для моделирования второго этапа процесса кодирования вторичного текста каждое предложение авторского резюме последовательно сопоставлялось с идентичными или близкими по содержанию фрагментами введений и/или заключений научных статей. В результате этого анализа было установлено, что, помимо простого переноса отдельных предложений в резюме, исходный текст чаще всего подвергается таким преобразованиям, как исключение (опущение элементов) и перифразирование. Рассмотрим подроб-

1) В исследуемом нами материале такие тексты называются *Abstract* 'Реферат', *Summary* 'Резюме' или используются без названия. По этой причине, а также в связи с тенденцией в современных лингвистических исследованиях использовать термин *резюме* для обозначения вторичного текста любой стилиевой принадлежности, составленного автором или другим лицом, в дальнейшем изложении мы будем пользоваться именно этим термином для обозначения изучаемых вторичных текстов.

нее *грамматическое перифразирование* как один из наиболее употребительных типов преобразований, которые можно наблюдать при сопоставлении основного текста научной статьи и ее резюме.

Теоретической основой грамматического перифразирования является наличие в языке *грамматических синонимов*, которые традиционно классифицируются на *морфологические* и *синтаксические*. К первым обычно относят синонимичные словоформы, ко вторым – синонимичные модели словосочетаний, предложений, придаточных предложений и оборотов. К *морфологическим* синонимам в исследуемых научных текстах принадлежат, главным образом, глагольные видо-временные формы: *Future Indefinite* → *Present Indefinite*, *Past Indefinite (Present Perfect)* → *Present Indefinite*. Такие преобразования можно рассматривать как транспозицию грамматических категорий, под которой понимается замена форм в пределах парадигмы, когда одна из форм заменяется на какой-либо ее парадигматический противочлен (Гак 1998). Так, при сопоставлении введения научной статьи и резюме можно наблюдать транспозицию форм, представляющих глагольную категорию *проспекции*, при которой маркированный член оппозиции заменяется на немаркированный (*Future Indefinite* → *Present Indefinite*). Например: *After explaining Application Center 2000 clusters, its applications and architecture, I'll explain how Application Center 2000 handles load balancing...* → *This article explains network load balancing...*

Транспозиция грамматических категорий времени (*Past Indefinite* → *Present Indefinite*) и временной отнесенности (*Present Perfect* → *Present Indefinite*) была обнаружена нами при сопоставлении предложений заключения и резюме. На-

пример: *I described the first experiment...*
→ *In this paper, I describe...*; *Two different problems have been selected for experiments...*
→ *... two problems are selected...*

Приведенные выше морфологические преобразования, которые сводятся к замене на форму *Present Indefinite*, можно объяснить прежде всего тем, что в научном стиле английского языка эта форма преобладает, поскольку она фиксирует вневременной, обобщенный характер действия (Рейман 1977). Кроме того, можно предположить, что при создании резюме, которое строится после написания основного текста, его автор сознательно или интуитивно учитывает возможность его функционирования в качестве самостоятельного текста (например, в компьютерных базах данных или в реферативных изданиях). Поэтому в таком тексте неуместны, например, формы *Future Indefinite* при описании основных результатов исследования, которые уже проведены и изложены в опубликованной научной статье.

Наличие же иных форм, помимо форм *Present Indefinite*, во введении и заключении научной статьи можно объяснить последовательным изложением информации в этих композиционных частях. Во введении при формулировании гипертемы, подтемы и основных результатов исследования, которые *будут* излагаться в данной статье, наряду с формами *Present Indefinite* автор может использовать и формы *Future Indefinite*. В заключении, где подводятся итоги проведенного исследования, закономерны формы *Past Indefinite* и *Present Perfect*.

Среди *синтаксических* синонимов, которые являются основой грамматических преобразований научной статьи в резюме, нами выявлены одноуровневые корреляты (словосочетание – словосо-

четание, предложение – предложение) и разноуровневые корреляты (придаточное предложение – словосочетание). При этом под синтаксическими синонимами понимаются разнооформленные конструкции со сходным грамматическим значением. Среди одноуровневых синтаксических синонимов в функции второстепенных членов предложения преобладают определительные и обстоятельственные структуры. Например: *circle detection algorithm* → *algorithm for detecting circles*; *for the automatic acquisition of statistical models* → *for automatically acquiring stochastic models*; *for instability prevention* → *to prevent instability*; *for producing comprehensive text summaries* → *to produce comprehensive text summaries*.

Как показывают примеры, необходимым условием преобразований субстантивных словосочетаний в герундиальные и инфинитивные обороты является наличие отглагольного имени существительного в качестве ядра исходного сегмента и соответственно возможность образования неличной формы глагола (*acquisition* – *acquiring*, *prevention* – *to prevent*, *detection* – *detecting*).

Перифразирование второстепенных членов предложения может носить разноуровневый характер: придаточное предложение → оборот с неличной формой глагола, придаточное предложение → субстантивное словосочетание. Например: *realistic images that contain circular objects with occluded circles and missing edges* → *realistic images containing circular objects with occluded circles and missing edges*; *to show that the algorithm works correctly* → *for the algorithm's correctness*. Следует отметить, что обратные преобразования наблюдаются значительно реже, так как их результатом являются более объемные фразы

по количеству не только служебных, но и знаменательных слов.

К синонимичным моделям на уровне предложения традиционно относят активные и соответствующие им пассивные конструкции, поскольку общий смысл предложения в них не меняется (Гак 1998: 410). Рассматривая ниже залоговые преобразования, мы будем исходить из предположения о том, что синонимия – это относительное явление и что абсолютные синонимы в языке среди синтаксических конструкций, как и в лексике, являются исключением. Синонимия активных и пассивных структур состоит в том, что они описывают одну и ту же денотативную ситуацию. Основное различие заключается в том, что в активных конструкциях предмет или лицо, выраженное подлежащим, производит действие (является агенсом); в пассивных конструкциях действие глагола-сказуемого направлено на лицо или предмет, выраженное подлежащим, которое выполняет семантическую функцию пациенса. С позиций когнитивного подхода считается, что выбор подлежащего в английском языке зависит от фокусного внимания говорящего (Tomlin 1994). Если предикация переходна и фокусный референт – агенс, выбирается форма глагола в действительном залоге; если в фокусе внимания находится пациенс, то выбирается форма в страдательном залоге.

Понятие фокуса внимания связывает проблему выбора активной или пассивной конструкции с вопросом выделенности информации в тексте. В когнитивной лингвистике предлагается различать естественную и неестественную (или маркированную) выделенность. Естественная выделенность имеет место при иконическом принципе построения

текста: в соответствии с порядком вещей во внешнем мире вначале вербализуется источник или агенс, а новая информация располагается ближе к концу предложения. Маркированная выделенность наблюдается при выдвигании значимой в том или ином отношении информации в приоритетную синтаксическую позицию в предложении (для английского языка это начало предложения).

Исследователи научных текстов отмечают, что пассивные конструкции в них употребляются гораздо чаще, чем активные, что обычно объясняют объективностью и беспристрастностью научного изложения. Поэтому логично было бы предположить, что в залоговых преобразованиях при построении резюме ведущее место занимают трансформации активных конструкций в пассивные. Однако количественные подсчеты залоговых преобразований в нашем материале показали, что частотность преобразований «актив → пассив» не намного превышает количество трансформаций «пассив → актив» (56% против 44%). Это можно объяснить тем, что ситуация научно-познавательной деятельности, реализованная в научном тексте, включает не только объектный (предметный), но и субъектный компонент, непосредственно связанный с личностью автора. Хотя «автор научного текста стремится к самоустранению в пользу предмета речи, он никогда не достигает этой цели в полной мере» (Матвеева 1990: 35). Поэтому непосредственно в тексте статьи, а также в резюме можно встретить активные конструкции с подлежащим, выраженным личным местоимением *we* или *I*. Например: *We obtain up to 70% accuracy of unrestricted text, ...; I end the paper with the evaluation of...* . Широкое употребление личного местоимения во множественном

числе можно объяснить сложившейся традицией, а также тем, что большинство статей в исследуемых предметных областях представляют собой результаты коллективного труда.

Как показали наши наблюдения, своеобразной чертой научных текстов на английском языке является то, что безличность изложения достигается, помимо пассивных конструкций, и активными структурами, в которых позицию подлежащего занимает одно из неодушевленных имен существительных: *article, paper, work, report, section* и др. Например: *This paper explores the use of results of linguistic research into the lexicon...* Это нетипичная активная структура, поскольку в ней, несмотря на наличие глагола в действительном залоге, подлежащее выполняет роль локатива, а не агенса. Хотя агенс вербально не выражен, он подразумевается, что в ряде случаев подтверждается текстовыми преобразованиями, в которых неодушевленное имя существительное заменяется на личное местоимение, и наоборот: *In this paper we argue that...* → *This paper argues that...* В таких предложениях авторы отождествляют себя со своей работой и наделяют ее свойствами агентивности, и поэтому преобразования в пассивные конструкции наблюдаются и для таких структур.

Рассмотрим условия, при которых становятся возможными преобразования активных структур в пассивные при кодировании информации в резюме научного текста. Эти условия можно расклассифицировать на внутренние (характеристики исходной активной структуры) и внешние (когнитивно-коммуникативные факторы). Рассмотрим вначале первую группу факторов. Естественно, что главным условием пассивизации является употребительность глагола исходного

предложения в страдательном залоге. Сопоставительный анализ первичных и вторичных научных текстов показал, что пассивным трансформациям чаще всего подвергаются предложения, построенные на основе следующих групп переходных глаголов: глаголы познавательной деятельности – *analyze, base, characterize, propose, discuss, integrate*; глаголы демонстрации – *demonstrate, show*; глаголы сообщения информации – *describe, present, report* и др. Если исходное предложение содержит непереходный глагол, то оно не может быть подвергнуто пассивной трансформации, например: *Our system participated in the official TREC Q&A evaluation.*

Второе условие для проведения трансформации пассивизации предполагает проверку семантической роли подлежащего: для проведения пассивной трансформации оно должно обозначать агенс (*we, author* и др.) или локатив (*article, paper, work, report, system* и др.). В случае иных семантических функций подлежащего предложение не может быть преобразовано в пассивную конструкцию, как в следующем примере: **Composition of the elementary trees in a TAG derivation uses the two operations of substitution and adjoining.**

Третье условие предполагает анализ группы дополнения в исходной структуре. Если дополнение имеет громоздкое распространенное определение, то пассивная трансформация нарушит один из основных принципов английского словоупотребления (*end-weight*) – «не помещать распространенные структуры в начало предложения» (Quirk et al.: 1973). Такое объемное определение присутствует, к примеру, в следующем предложении: *This paper gives methods for constructing text relation maps and for using text re-*

lations to access and use text databases.

Однако существует способ соблюсти указанный принцип и провести пассивную трансформацию для случая с распространенной группой дополнения в активной структуре. Это использование прерывистых структур, т.е. постановка распространенного определения в конец предложения за сказуемым: *Methods are given for constructing text relation maps and for using text relations to access and use text databases.*

При проведении конкретных текстовых преобразований в целях построения вторичного текста учитываются все три рассмотренных выше условия. Так, например, следующее предложение из первичного текста преобразовано в пассивную конструкцию в тексте резюме, поскольку подлежащее *we* – агенс, глагол-сказуемое *separate* может употребляться в страдательном залоге, а дополнение состоит из трех полных слов: *After binarization, we separate a graphic image part from a text region by considering chain codes of connected components* → *After binarization, a graphic image part is separated from a text region by considering chain codes...* Следует отметить, что подлежащее активной конструкции не преобразуется в дополнение с предлогом *by*, поскольку в научном тексте агенс в пассивных конструкциях, как правило, опускается.

Особым является случай, когда пассивная конструкция в резюме, образованная из активной в исходном тексте, является главной частью сложноподчиненного предложения не с пациенсом или локативом в позиции подлежащего, а с формальным *it*, вводящим придаточное предложение. Как известно, норма языка запрещает продвижение придаточного в позицию подлежащего в таких случаях. Например:

The paper proposes that a linguistically-motivated template entry should be designed for each type of word, ... → It is proposed that a linguistically-motivated template entry should be designed for each type of word, ...

Способ выделенности при таких преобразованиях не меняется: новая информация располагается в конце предложения и выражена придаточным предложением, что свойственно нейтральной речи. Изменяется главная часть сложноподчиненного предложения, которая приобретает безличный характер и является одним из клише, характерным для научного стиля речи: *it is shown, it is proposed, it is demonstrated* и др.

Перейдем к рассмотрению внешних (когнитивно-коммуникативных) факторов, вызывающих употребление пассивных конструкций во вторичном тексте. С позиций когнитивного подхода, как уже указывалось, преобразование актива в пассив связано с продвижением в начальную позицию или ближе к началу предложения паиенса, который находится в фокусном внимании адресанта. Этот фактор, будучи ведущим для употребления пассива вообще, особенно актуален для инициальных предложений резюме. Именно первое предложение резюме заостряет внимание адресата на новом знании, полученном в результате конкретного исследования или на современных направлениях в данной научной области. Например: *A new neural paradigm called diagonal recurrent neural network is presented; In recent years, information retrieval techniques have been used for automatic generation of semantic hyper-text links.*

Однако распределение информации в тексте обусловлено не только когнитивными принципами, но и нормами коммуникации, в частности, необходимостью

обеспечить непрерывность и связность потока информации, что выражается в известных закономерностях развертывания текста – тема-рематических прогрессиях (Daneš 1974). Наблюдения показали, что пассивные конструкции в резюме научного текста реализуют в основном линейную прогрессию и прогрессию с производными темами. При линейной модели рема предыдущего предложения (или ее часть) становится темой пассивной конструкции, последовательно и логично продвигая описание вперед. Например: *This paper presents some computationally efficient predictors of the cost-effectiveness of the two main classes of selective regression testing approaches. These predictors are computed from data about the coverage relationship...*

В прогрессии с производными темами последние вытекают из общей темы (гипертемы) резюме и всего текста научной статьи. Подобное развертывание вторичного текста обусловлено его компрессионным характером, стремлением авторов изложить максимум информации минимальными средствами. В свою очередь, маркированность тем в последовательности параллельных структур обеспечивает лаконичность, четкость изложения и, в результате, их оперативное восприятие адресатом. Например: *Arbitrary application data structures are allowed. Units of measurement are converted as needed.*

Как было отмечено выше, сопоставительный анализ первичных и вторичных текстов выявил также достаточно высокую употребительность преобразований «пассив → актив» при построении вторичного текста. Если выбор пассивной конструкции обусловлен прежде всего необходимостью маркированного выделения новой или наиболее значимой

информации, то активные конструкции с агенсом в позиции подлежащего, выделяя новую информацию нейтральными средствами, позволяют подчеркнуть личный вклад авторов. Например: *In our approach, a summary is assembled out of passages extracted from the original text, ...* → *In our approach, we assemble a summary out of passages extracted from the original text...*; *The measures are based on ...* → *We base these measures on...* В исследуемом массиве резюме научных статей употребительны последовательности активных структур с константными темами, выраженными личным местоимением во множественном числе: *We compute aggregate ROC curves based on 1 set of 50 object images and another set of 10 aerial images. We analyze the performance of 11 different edge detectors reported in the literature.* В таких случаях повторяемость тематических элементов усиливается синтаксическим параллелизмом, позволяя говорить о наличии залогового согласования между предложениями научного текста, который создает своеобразный ритм изложения. Вместе с тем при выборе способа акцентуации информации в каждом отдельном предложении учитываются и закономерности развертывания связно-

го текста, которые зависят не только от сложившихся в языке тема-рематических моделей, но и от интенций авторов тем или иным образом маркировать свои коммуникативные и информативные намерения (Лузина 1996).

Безусловно, средствами грамматической синонимии такой тип текстового преобразования, как перифразирование, не исчерпывается. Преобразование текста научной статьи в резюме осуществляется также с помощью лексической, лексико-грамматической синонимии и путем логико-смыслового перифразирования. Все эти типы преобразований способствуют варьированию средств вербализации смысловых компонентов вторичного научного текста (резюме) по сравнению с первичным (собственно текстом научной статьи). Вместе с тем выбор того или иного типа преобразования исходного текста обусловлен необходимостью компрессированного изложения информации в резюме, когнитивными и коммуникативными принципами распределения и акцентуации информации в связном тексте, а также системно-структурными свойствами языковых единиц, вовлеченных в процесс кодирования научной информации.

summary

Σ Grammatical Transformations in a Scientific Text

This paper considers the problem of grammatical transformations in scientific texts written in the English language. By comparing the peripheral parts of a scientific article (introduction and conclusion) and its summary (abstract) two main types of grammatical transformations have been singled out: morphological and syntactical. Morphological transformations mean transposition of the forms representing the grammatical categories

of prospect, tense and correlation. The resulting grammatical form for all the cases is the Present Indefinite, which is attributed to the generalized character of scientific prose and the possibility for a summary to function as a separate text.

Syntactical transformations occur on the levels of parts of the sentence and sentence structures. Special attention is paid to the factors which allow passive transformations of sentences. They are classified into internal factors describing the features of the source structure and external factors connected with the cognitive principles of foregrounding, thematical progressions in a connected text and the author's intentions. The types of transformations studied help to reveal the regularities of coding information from the primary scientific text into its summary.

Источники исследования

- Advances in Automatic Text Summarization. – Cambridge (Mass), London (England), 1999.
 AT&T Technical Journal. – 1995. N1, 2, 5; 1996. N1, 2, 5.
 Computational Intelligence. – 1994. N4.
 Computational Linguistics. – 1991. N4; 1995. N2. 1996. N1.
 Computer Vision and Image Understanding. – 2001. Vol. 81–84.
 International Journal of Lexicography. – 1991. Vol.4. N1.
 IEEE Transactions on Software Engineering. – 1995. N8; 1997. N1, 2, 3, 4.
 IV International Conference on Language Resources and Evaluation. – Lisboa, 2004.
 MSDN Magazine. – 2001. N5, 7.
 Machine Translation. – 1992/1993. N7. 1995. N9.
 Software Practice and Experience. – 1995. N8, 10.

Литература

- Гак 1998: **Гак, В. Г.** Языковые преобразования. – Москва: Языки русской культуры, 1998. – 768 с.
 Карпилович 2000: **Карпилович, Т. П.** Сравнительный анализ текстов резюме и заключений научных статей. – In: Слово в языке и речи. – Минск.: МГЛУ, 2000. – С. 97–106.
 Кубрякова 2004: **Кубрякова, Е. С.** Язык и знание. – Москва: Языки славянской культуры, 2004. – 560 с.
 Лузина 1996: **Лузина, Л. Л.** Распределение информации в тексте: когнитивный и прагматилистический аспекты. – Москва: ИНИОН РАН. – 1996. – 155 с.
 Матвеева 1990: **Матвеева, Т. В.** Функциональные стили в аспекте текстовых категорий. – Свердловск: Изд-во Урал. ун-та, 1990. – 255 с.
 Рейман 1997: **Рейман, Е. А.** Обусловленность употребления личных глагольных форм характером содержания научного текста. – In: Язык и стиль научной литературы. – Москва: Наука, 1977. – С. 255–269.
 Цурикова 2002: **Цурикова, Л. В.** Естественность дискурса как когнитивно-прагматический феномен. – In: С любовью к языку: Сб. науч. тр., посвящ. Е. С. Кубряковой. – Москва; Воронеж, 2002. – С. 419–429.

ТАТЬЯНА П. КАРПИЛОВИЧ

Danes 1974: **Danes, F.** Functional Sentence Perspective and the Organization of the Text. – In: Papers on Functional Sentence Perspective. – Prague, 1974.

Quirk et al. 1973: **Quirk, R., Greenbaum S., Leech G., Svartvik J.** A University Grammar of English. – London: Longman, 1973. – 484 p.

Tomlin 1994: **Tomlin, R.** Focal Attention, Voice and Word Order: An Independent Cross Linguistic Study. – In: Word Order in Discourse. – Amsterdam, 1994.