

многие подходы основаны на вариантах моделей марковских цепей, условных случайных полей или на классификаторах [1, с. 218]. Многие модели позволяют ответить на следующие вопросы.

1. Имеет ли данная именная группа своим референтом некую именованную сущность?

2. Какая категория из списка лучше всего соответствует именной группе в данном контексте?

3. Кто или что является референтом данной именной группы?

Во всех приведенных выше примерах у имен собственных имелся конкретный референт. Однако в следующем предложении вряд ли имеется в виду столица Франции: *Дмитрий Rogozin: «Генералы в окопах должны быть, а не в парижсах».*

Эффективность параметров для систем машинного обучения зависит от языка и жанра текста. Так, во многих языках важным признаком именованных сущностей является их написание с заглавной буквы. Однако, например, в немецком языке, где все существительные пишутся с заглавной буквы, и в текстах, полученных при автоматическом распознавании речи, этот параметр оказывается бесполезным. Важным фактором является также наличие леммы слова в списке известных именованных сущностей. Для слов и именных групп, отсутствующих в списке, решение о том, идет ли речь об именованной сущности и, если да, то какого типа, принимается на основе анализа микроконтекста. Например, слово *Mrs.* в английском языке, *Frau* в немецком языке или *госпожа* в русском языке с большой долей вероятности сигнализируют о том, что за ним следует имя человека.

## ЛИТЕРАТУРА

1. Прикладная и компьютерная лингвистика / под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. – М. : ЛЕНАНД, 2016. – 320 с.
2. *Fader, A. Identifying relations for open information extraction / A. Fader, S. Doderland, O. Etzioni // Proc. of EMNLP-11. – 2011. – P. 1535–1545.*

The article deals with the concept of the named entity and the categories it may be referred to. The main problems of automatic named entities detection and extraction from written texts are under consideration.

**А. М. Савич**  
Минск, МГЛУ

## ВИДЫ ТРАНСФОРМАЦИЙ ПРИ АВТОМАТИЧЕСКОМ ПЕРЕВОДЕ АНГЛОЯЗЫЧНОГО ОФИЦИАЛЬНО-ДЕЛОВОГО ТЕКСТА НА РУССКИЙ ЯЗЫК

В статье представлены результаты и основные выводы исследования лексических трансформаций при автоматическом переводе англоязычных текстов 17 целей ООН в области устойчивого развития на русский язык.

Предполагается, что отношение между отрезком оригинала и соответствующим отрезком перевода, можно представить как преобразование (трансформацию) первого во второй по определенным правилам. Подобные переводческие трансформации могут рассматриваться как приемы перевода, которые переводчик использует для преодоления типичных трудностей. Основопологающей задачей является уточнение смысла, вкладываемого в понятие *переводческая трансформация*. Переводческая трансформация – это такой процесс перевода, в ходе которого система смыслов, заключенная в речевых формах исходного текста, воспринятая и понятая переводчиком в силу его компетентности, трансформируется естественным образом вследствие языковой асимметрии в более или менее аналогичную систему смыслов, облакаемую в формы языка перевода [1].

Чтобы лучше понять само явление переводческих трансформаций, стоит разобраться в причинах их возникновения [2].

1. Несовпадение систем исходного языка (ИЯ) и переводного языка (ПЯ). Возможности русского словосложения не позволяют использовать немецкую, например, или английскую конструкцию, поэтому переводчик применяет иные по отношению оригинала лексико-грамматические конструкции.

2. Несовпадение норм ИЯ и ПЯ. Языковая норма выражается в общественно санкционированных формах языкового выражения. Норма выступает в функции фильтра языковой системы: либо задерживает, либо пропускает языковые продукты системы. Неправильное в норме неправильно всегда и везде.

3. Несовпадение речевых норм (узусов) носителей ИЯ и ПЯ. Узус отделяет одно, что принято в одной ситуации и не принято в другой. Узус – (некоторые ситуативные правила выбора средств языкового выражения) – это языковые привычки носителей языка.

4. Расхождения преинформационных запасов носителей ИЯ и ПЯ культурно-исторического или актуально-событийного характера.

Переходя к проблеме разделения трансформаций на виды, существует множество различных точек зрения, однако большинство лингвистов в своих работах отмечают, что все переводческие трансформации подразделяются на лексические, грамматические и смешанные (или комплексные).

В настоящее время особо актуальным является получение информации в скорейшие сроки, ее общедоступность, а также важность ускорения перевода этой информации на другие языки. Основные трудности, возникающие при переводе официально-деловых текстов, связаны с необходимостью выбора переводных эквивалентов для оригинала. Суть возникающей задачи сводится к поиску такого содержательного эквивалента, при котором соблюдаются и смысл, и нормы переводящего языка.

Материалом для данного исследования послужили англоязычные тексты 17 целей в области устойчивого развития, принятых ООН и их параллельные переводы на русский язык, взятые из интернет-сайта [www.un.org/](http://www.un.org/)

sustainabledevelopment/sustainable-development-goals/. В данных текстах были выделены лексические трансформации главных опорных слов. Например, для текста «Goal 13: Take urgent action to combat climate change and its impacts» были выделены следующие замены: в большинстве случаев это словарные переводные эквиваленты (*climate* – ‘климат’, *change* – ‘изменение’, *country* – ‘страна’, *temperature* – ‘температура’, *level* – ‘уровень’), помимо этого, адекватная замена (*global* – ‘мировой’), логическое развитие (*emissions* – ‘объем выбросов’). Стоит отметить, что у лексической единицы *global* в данном тексте два перевода.

Все лексические единицы, относящиеся к основному содержанию текстов, были переведены с помощью программы машинного перевода PROMT Professional 10. Для данного текста было выделено минимальное количество замен при автоматическом переводе: адекватная замена (*global* – ‘мировой’). В абсолютном большинстве преобладают словарные переводные эквиваленты и дословный перевод.

А, например, текст «Goal 10: Reduce inequality within and among countries» включает следующие трансформации: словарный эквивалент (*inequality* – ‘неравенство’, *country* – ‘страна’, *developing* – ‘развивающийся’, *reduce* – ‘уменьшать, сокращать’), логическое развитие (*income* – ‘уровень дохода’).

Далее были подсчитаны среднестатистические данные по всем встретившимся в анализируемых текстах лексическим трансформациям. Эти данные мы свели в таблицу.

Среднестатистические данные по лексическим трансформациям  
в исследуемых англоязычных официально-деловых текстах

Виды лексических трансформаций	Среднестатистические данные, выполненные автоматически		Среднестатистические данные, выполненные переводчиком	
	количественный показатель	относительная частота	количественный показатель	относительная частота
Словарные переводные эквиваленты	39	0,78	46	0,62
Адекватная замена	7	0,14	8	0,11
Логическое развитие	2	0,04	8	0,11
Конкретизация	1	0,02	6	0,08
Замена частей речи	0	0	4	0,05
Генерализация	1	0,02	2	0,03

Рассмотрев переводческие лексические трансформации в исследуемых текстах, мы пришли к выводу, что чаще всего при переводе англоязычных официально-деловых текстов используются словарные переводные эквиваленты. Иногда встречаются генерализация, конкретизация, замена частей речи и смысловое развитие. При автоматическом переводе с помощью программы машинного перевода наиболее характерными для данных текстов являются адекватная замена и лексическое развитие, правда встречается они редко. Особую трудность в системе машинного перевода вызывают многозначные слова (*people, food, water*), значение которых зависит от контекста и может значительно различаться. Например, лексическая единица *water* переводится как ‘водный ресурс’, ‘водоснабжение’.

## ЛИТЕРАТУРА

1. *Комиссаров, В. Н.* Современное переводоведение : курс лекций / В. Н. Комиссаров. – М. : ЭТС, 1999. – 192 с.
2. *Бархударов, Л. С.* Язык и перевод. Вопросы общей и частной теории перевода / Л. С. Бархударов. – 2-е изд. – М. : ЛКИ, 2007. – 240 с.

In the article the results and main conclusions of the research of lexical transformations in automatic translation of English-language texts of 17 UN goals for sustainable development into Russian are presented.

**А. В. Скрябина**  
Минск, МГЛУ

## АВТОМАТИЧЕСКОЕ ИНДЕКСИРОВАНИЕ АНГЛОЯЗЫЧНЫХ ПУБЛИЦИСТИЧЕСКИХ ТЕКСТОВ

В статье рассматривается проблематика автоматического индексирования англоязычных новостных текстов на основе выделения ключевых слов.

В настоящее время системы, автоматизирующие обработку текста с целью выявления его смысла и структуры, крайне немногочисленны или находятся в экспериментальной стадии разработки. В основном такие системы выделяют лишь смысл отдельных слов либо словосочетаний, но структура целого предложения целиком не анализируется. Сам процесс описания смыслового содержания текста средствами информационно-поискового языка (ИПЯ) называется индексированием. А именно это процесс присвоения документам и запросам или их частям индексов. Итоговая продукция этих процессов содержится в поисковом образе документа.

В нашем исследовании материалом послужили англоязычные новостные тексты по теме «Культура». Все публицистические тексты взяты с интернет-ресурсов [www.bbc.com](http://www.bbc.com) и [www.theguardian.com](http://www.theguardian.com). В нашей работе устанавли-