

Котельникова Евгения Владимировна

ПРОЕКЦИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА КОГНИТИВНЫЕ МОДЕЛИ СТИЛИСТИКИ В СМЕШАННОМ ПЕРЕВОДЕ НАУЧНО-ИННОВАЦИОННОГО ДИСКУРСА

В процессах смешанного перевода межкультурной коммуникации при репрезентации стиля и идиостиля исходного текста в переводном тексте стилистика текста учитывается в настройках исходных параметров машинного перевода опосредствованно и частично в отличие от лексики и семантики. На этапе постредактирования смешанного перевода стиль рассматривается как принцип построения целостной формы исходного текста и реализуется обратный процесс - от содержания переводного текста к его форме и стилю, на основе отбора средств выражения среди возможных синонимических средств языка.

Адрес статьи: www.gramota.net/materials/2/2011/4/27.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2011. № 4 (11). С. 102-105. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2011/4/

© Издательство "Грамота"

Информацию о том, как опубликовать статью в журнале, можно получить на Интернет сайте издательства: www.gramota.net
Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: voprosy_phil@gramota.net

УДК 81'322.4:81'42

В процессах смешанного перевода межкультурной коммуникации при репрезентации стиля и идиостиля исходного текста в переводном тексте стилистика текста учитывается в настройках исходных параметров машинного перевода опосредствованно и частично в отличие от лексики и семантики. На этапе постредактирования смешанного перевода стиль рассматривается как принцип построения целостной формы исходного текста и реализуется обратный процесс - от содержания переводного текста к его форме и стилю, на основе отбора средств выражения среди возможных синонимических средств языка.

Ключевые слова и фразы: смешанный перевод; идиостиль; машинный перевод; исходный текст; переводной текст; научно-инновационный дискурс.

Евгения Владимировна Котельникова

Кафедра английского языка для гуманитарных специальностей

Факультет лингвистики и журналистики

Ростовский государственный экономический университет (РИНХ)

jenni.cot@yandex.ru

ПРОЕКЦИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА КОГНИТИВНЫЕ МОДЕЛИ СТИЛИСТИКИ В СМЕШАННОМ ПЕРЕВОДЕ НАУЧНО-ИННОВАЦИОННОГО ДИСКУРСА[©]

Исследование выполнено при финансовой поддержке Министерства образования и науки Российской Федерации в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы (тема: «Проведение научных лингво-педагогических исследований применения инновационных технологий в области языкового образования высшей технической школы»; ГК № 14.740.11.1399).

Многообразие и сложность задач межкультурной коммуникации (МКК) вызывает необходимость включения в число специфических объектов функционирующих в виртуальной среде - искусственного интеллекта (ИИ), в частности, машинного перевода (МП). В современных системах МП применение языка понимается как чисто лингвистический процесс конструирования высказываний, «совершаемый по языковым нормам, заложенным в системе языка, а не в системе его применения» [2, с. 184-188].

Для обеспечения высокого уровня переводимости программы автоматического перевода за рубежом широко практикуется использование контролируемого языка, подразумевающее применение разной степени строгости правил для конкретных пар языков и систем МП, что значительно повышает качество перевода, сокращая потребность в постредактировании [5, р. 50-54]. К наиболее известным контролируемым языкам относится упрощенный технический английский (Simplified Technical English, STE) [7].

Во многих случаях разносторонней коммутативной практики, например, реализации научно-инновационного дискурса (НИД) [3, с. 95-99], такой подход в качестве общего неприемлем. Наиболее оптимальным способом возможного системного применения языка становится использование выработанных на основе знаний принципов, правил, норм релевантных для различных дискурсивных практик с учетом усиления антропоцентрического фактора [2, с. 184-188]. Однако и в этих условиях когнитивную лингвистическую модель следует рассматривать в проекции ИИ, создающего дополнительные возможности обеспечения идентичности языковых знаний субъектов социума на семантико-прагматическом уровне общения. В связи с этим возникает проблема оформления текстов НИД в процессе смешанного перевода (СП) [3, с. 95-99] в стилевых жанрах и идиостилях. Стиль представляет собой связь, как внутреннего содержания, так и внешней формы, служащей оформлением, образной «оболочкой» текста. Для «присутствия» человека и ИИ в тексте когнитивный подход предполагает наличие кодов: функциональных стилей, речевых жанров, композиционно-речевых форм и др. Стилиевые жанры обладают набором неоднородных признаков, имеют подвижные границы, строятся по полевому принципу, образуя многочисленные промежуточные формы, а тексты различаются в зависимости от теории и практики употребления языка, стилистической структуры текстов, функциональных стилей и др. [1; 2, с. 184-188].

Лексическое наполнение и композиционная форма любого произведения зависят от индивидуально-типологических характеристик личности. Под идиостилем понимается система логико-семантических способов репрезентации доминантных личностных смыслов концептуальной системы автора текста, объективированная в эстетической деятельности и предполагающая индивидуальную трансформацию языковых выражений [4]. В ряде случаев идиостиль может рассматриваться как отклонение от конвенционального стиля.

Возникает вопрос: каково место ИИ в этой «системе систем», и как когнитивные аспекты проявляются и применяются в этой специфической речевой деятельности? Рассмотрим сложившиеся когнитивные представления и практические стереотипы перевода с использованием ИИ.

Лексическая база МП обычно состоит из фиксированных общего (генерального) и специализированного словарей, а также из изменяемого (пользовательского) и вариативного (электронного) словарей [7]. Гене-

ральные словари - основные словари каждого из направлений перевода, постоянно участвуют в процессе перевода и содержат общеупотребительную лексику. Специализированные словари предназначены для расширения словарной базы системы в определенной предметной области с целью перевода предметно-ориентированных текстов. Пользовательские словари, количество которых не ограничивается, создаются и пополняются самостоятельно. В них происходит активная работа со словарной базой: вводятся новые вокабулы, сохраняются как новые и измененные словарные статьи из других словарей, так и внутренняя информация (уникальная терминология, принятые обозначения, аббревиатуры). Электронные словари содержат множество вариантов уточнения ПТ, ориентированные на различные предметные области и известные как *PROMT Electronic Dictionary*, *Lingvo*, *Context*, *MultiLex* и др. Значение многовариантных слов (словосочетаний), зависит от контекста ИТ, не выделяется МП самостоятельно и находится в компетенции переводчика.

Подключение словарей означает выбор из списка словарей, словарей необходимых для данного направления перевода текущего текста определенного шаблона тематики. Программа просматривает словари по порядку номеров списка подключенных словарей (определяющего их приоритет), извлекая перевод слова из первого же словаря, его содержащего. Создание и настройка пользовательского словаря с максимальным приоритетом (первым в списке подключенных словарей) рекомендуются даже при наличии соответствующего специализированного словаря. Наиболее эффективно подключение к системе перевода одного-двух специализированных словарей и одного пользовательского словаря.

С целью многократного оперирования созданными лингвистическими настройками для перевода текстов схожей тематики в определенной предметной области введено понятие шаблона с модифицируемой тематикой, наследующей все его компоненты для каждого направления перевода. Шаблон тематики автоматически определяется списком ключевых слов и словосочетаний, исходя из наибольшего количества совпадений слов из списка со словами документа. Для новых ключевых слов отдельно вводятся все необходимые словоформы. Существует модифицируемый базовый шаблон тематики («Общий»), основанный на общем словаре.

Варианты перевода многовариантных слов отображаются в форме: главный перевод, далее все остальные варианты перевода, либо как словарная статья, содержащая исходное слово первым в списке переводов. Выбор варианта перевода доступен для абзацев, не отредактированных вручную после перевода. Многозначность терминов требует учета контекста при выборе значения даже в одной предметной области.

Для непереводаемого в словарях термина подбирается переводной эквивалент из справочников, специальной литературы, или создается новый эквивалент на основе моделей образования терминов. В других случаях термин переводится описательным путем с дополнительным сохранением его денотата на языке исходного текста (ИТ). С целью упрощения анализа структуры предложения, указывается принадлежность данного слова к семантическому классу ИТ.

Резервированные слова на языке ИТ (имена собственные, аббревиатуры, математические формулы, специальные наименования, совпадающие со значимыми словами (например, *Bill Gates*)) помечаются как не требующие перевода и могут транслитерироваться или выводиться в переводном тексте символами исходного языка. Списки резервированных слов входят в состав шаблонов тематик или хранятся в отдельном файле. Резервирование слова и словосочетания в программе основывается на правиле сегментации, и может быть использовано для адаптации текстов с учетом конкретных требований.

Подключение стандартных препроцессоров, позволяет исключать некоторые конструкции (имена файлов или адреса электронной почты, служебные символы, смайлы, дата и время и т.д.) из процесса перевода, оставляя их без изменений, либо трансформировать согласно правилам языка переводного текста (ПТ). Препроцессоры можно подключать, отключать, настраивать и изменять их приоритет.

Настройка правил перевода позволяет выбирать модель перевода некоторой лингвистической конструкции, не управляемой одной модификацией словарных данных. Различные значения правил перевода применимы при переводе отдельных абзацев или всего текста. Для каждого направления перевода существует собственный список правил перевода, настраивающих программу на перевод конкретного текста (информационных сообщений, документов, инструкций, контрактов, личной переписки и др.). Для каждого правила перевода одно значение принимается программой по умолчанию (как правило, наиболее общее), но может быть изменено для шаблона тематики, для перевода конкретного документа или для перевода отдельных абзацев документа.

Таким образом, ориентирование ИИ на заданную предметную область осуществляется тематикой текста, определяемой по ключевым словам: базисным концептам, терминам, категориям. Если в параллельной картине мира существует знак с соответствующим значением, который может быть распознан ИИ, используются генеральный и пользовательский словари. Если знак существует, но значение его не может быть распознано применяются простые и регулярные выражения. Если же не существует соответствующего знака, передается его форма (изображение).

В основе технологии *Translation Memory* (ТМ) лежит принцип сбора и хранения фрагментов (сегментов) ИТ и ПТ (оригинал и его перевод), которые находятся в базе данных (БД) переводов и при совпадении не менее чем на некоторое процентное соотношение могут использоваться при последующем выполнении переводов той же тематики. МП при помощи системы ТМ возможен при наличии заранее заполненных баз данных ТМ, предназначенных для сохранения сегментов ИТ вместе с ПТ для их последующего использования в процессе перевода. В процессе МП сравнивается сегмент ИТ с сегментами из подключенных баз ТМ до совпадения с минимальным заданным процентом совпадения при переводе (указывается в скобках после фрагмента). Использование баз ТМ требует правильного разбиения ИТ на сегменты и профессионального

выполнения перевода сегментов ИТ при создании баз ТМ. Сегмент понимается как единица текста одновременно подаваемая на перевод. Правило сегментации задает границу сегмента текста оригинала и определяется простым или регулярным выражением (шаблоном для искомой последовательности символов).

На данном этапе развития технологии МП в ряде случаев перевода отдельных частей текста постредактирование ПТ требует больших когнитивных усилий, чем выполняемый переводчиком традиционный перевод. Постредактирование в значительной мере может быть ускорено тщательным редактированием ИТ, настройкой параметров программы МП на определенную предметную область, установкой специализированных словарей, ориентированных на содержание и набор языковых средств переводимых текстов, использованием баз ТМ, в значительной степени обеспечивающих лексическое и семантическое соответствие ИТ и ПТ.

В отличие от МП в СП межкультурной коммуникации возникает активная деятельность автора, переводчика, ИИ и реципиента, что позволяет осуществлять перевод с учетом стиля и идиостиля ИТ. Стилистика, изучающая закономерности функционирования языковых средств в задачах сообщения, объединяет определенные данные грамматики, лексикологии. Для облегчения коммуникации используется не полный язык во всех его связях и возможностях, а язык с определенными жесткими ограничениями основных параметров и характеристик, что проявляется в стилистике, которая таким образом близко соприкасается с проблематикой языковой коммуникации, теорией речевых актов, прагматикой языка [2, с. 184-188].

Для выявления стиля и индивидуальных особенностей автора, необходим формально-стилистика анализ ИТ, включающий определение значимых черт стиля на основе лексико-синтаксического исследования, и характеристику их частностей, особенности перехода от формы к содержанию, особенностей формы представления категории, ее признаки, частоту ее представления в текстах и жанрах текстов, ее дополнительные функции. Анализ языковых средств ведет к выявлению системы стилистических категорий плана содержания, которые вместе с типичными средствами их выражения образуют функционально-стилистика поля [Там же].

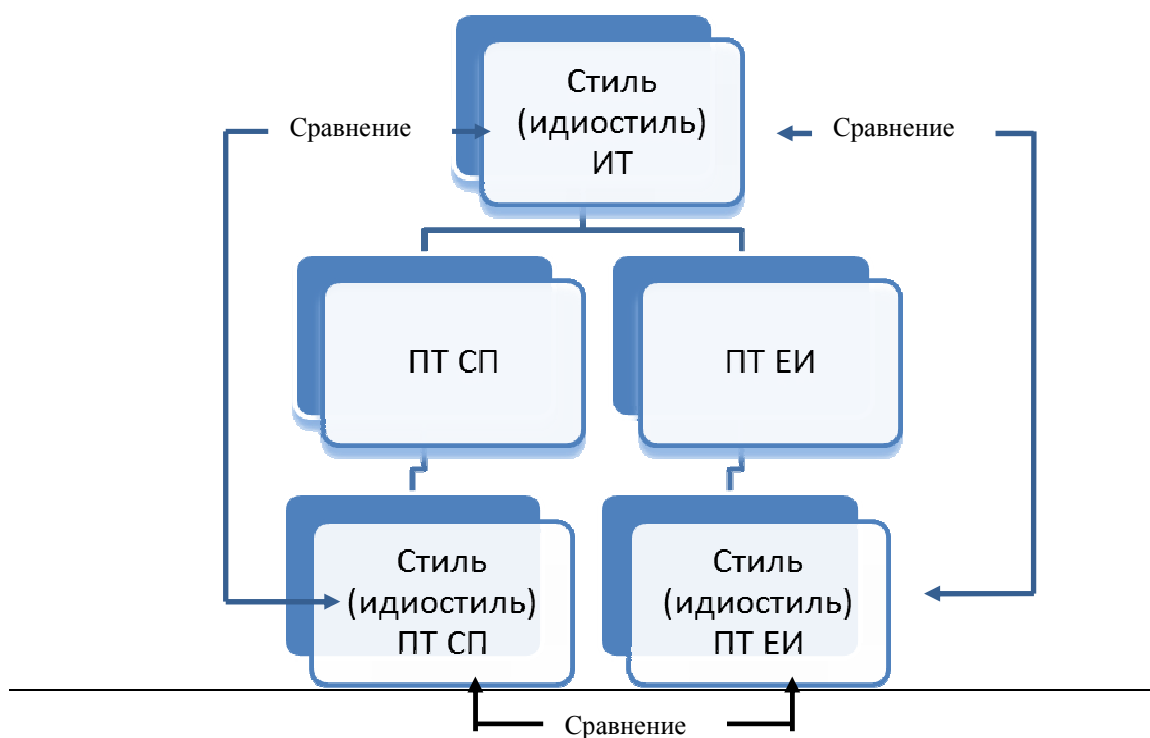


Рис. Репрезентации стиля и идиостиля ИТ в ПТ

При репрезентации стиля и идиостиля ИТ в ПТ стилистика текста в настройках МП учитывается опосредствованно и частично. Только на этапе постредактирования СП стиль рассматривается как принцип построения целостной формы ИТ и реализуется обратный процесс - от содержания к форме, стилю ПТ, на основе отбора средств выражения среди возможных синонимических средств языка; выявляется, какими конкретными средствами в языке данного ПТ должна быть выражена та или иная коммуникативная интенция стилистического плана, и тем самым определяются в целом способы организации аутентичного высказывания на данном языке в определенных ситуациях при переводе [1]. Этот процесс состоящий из комплекса компаративных взаимодействий стилей и идиостилей СП, субъекта перевода и ИТ (Рис.) обеспечивает репрезентацию языковых средств, использованных в ИТ, обусловленных основной идеей, которая на этапе постредактирования становится вполне определенной, а ПТ по содержанию - целостным.

Таким образом, для более полного обеспечения соответствия ИТ и ПТ в СП необходимыми условиями является активное и корректное использование различных интеракций с ИИ (опции, правил, настроек и др.),

к которым относятся: создание и редактирование списков ключевых слов настройка тематик текста, формирование упорядоченных списков подключённых словарей, зарезервированных слов, подключённых препроцессоров; подключённых баз ТМ, настройка правил сегментации и параметров работы с базами ТМ, задание значений правил перевода, установка опций транслитерации и показа вариантов перевода слов, а также полное привлечение когнитивного стилистического потенциала субъекта перевода.

Список литературы

1. **Арнольд И. В.** Стилистика. Современный английский язык. М.: Флинта; Наука, 2002. 384 с.
2. **Брандес М. П.** Стилистика текста: теоретический курс. М.: Прогресс-Традиция; ИНФРА-М, 2004. С. 184-188.
3. **Котельникова Е. В.** Когнитивная модель смешанного дискурсивного перевода // Модернизация современного общества: проблемы, пути решения: материалы I международной научно-практической конференции. Ставрополь: Центр научного знания «Логос», 2011. Ч. 2. С. 95-99.
4. **Пищальникова В. А.** Проблема идиостиля: психолингвистический аспект. Барнаул, 1992. 73 с.
5. **Phillips A. P.** Is Translation Memory Worth the Investment?: Web Methods' Experience with Two TM Tool Suites // Multi-Lingual Computing & Technology. 2005. Vol. 16. № 2 (70). P. 50-54.
6. **PROMT Standard 8.0** [Электронный ресурс]. URL: <http://standart.promt-translator.ru> (дата обращения: 20.01.2009).
7. **www.simplifiedenglish-aecma.org** (дата обращения: 20.10.2011).

**ARTIFICIAL INTELLECT PROJECTIONS ON STYLISTICS COGNITIVE MODELS
IN SCIENTIFIC-INNOVATIVE DISCOURSE MIXED TRANSLATION**

Evgeniya Vladimirovna Kotel'nikova

Department of English Language for Classical Specialities

Faculty of Linguistics and Journalism

Rostov State Economic University

jenni.cot@yandex.ru

In the processes of the mixed translation of intercultural communication while representing the style and idio-style of the original text, in the translated text the stylistics of text is taken into account in machine translation settings mediately and partly as opposed to vocabulary and semantics. At the stage of mixed translation post-editing style is considered as the principle of the original text integral form constructing and the inverse process is realized - from the translated text content to its form and style on the basis of expression means choice among possible synonymic language means.

Key words and phrases: mixed translation; idio-style; machine translation; original text; translated text; scientific-innovative discourse.

УДК 81'42

Статья посвящена описанию специфических особенностей универсального концепта «внимание» и его языковой репрезентации в русском языке в сравнительном анализе с языковыми единицами английского языка. Гипотетически предполагается, что актуализация рассматриваемого концепта имеет свою специфику с точки зрения описания, в том числе и с точки зрения аксиологического модуса, в зависимости от ментальной культуры нации, сформированной на определенном этапе исторического развития. По результатам лингвистического анализа видится возможной реконструкция некоторого фрагмента языковой картины мира российского сообщества.

Ключевые слова и фразы: концепт «внимание»; языковая репрезентация; аксиологический модус; ментальная культура; языковая картина мира; концептосфера.

Ирина Анатольевна Куприева, к. филол. н.

Кафедра английского языка и методики преподавания

Белгородский государственный университет

kuprieva@yandex.ru

**ЛИНГВОКУЛЬТУРОЛОГИЧЕСКИЕ АСПЕКТЫ
ЯЗЫКОВОЙ РЕПРЕЗЕНТАЦИИ МЕНТАЛЬНЫХ КУЛЬТУР®**

Исследование выполнено при финансовой поддержке Министерства образования и науки Российской Федерации в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы (тема: «Проведение научных лингво-педагогических исследований применения инновационных технологий в области языкового образования высшей технической школы»; ГК № 14.740.11.1399).

Сфера изучения культуры и ментальности - удел довольно широкого пласта теоретических исследований на стыке разных гуманитарных дисциплин. В частности, для филологии данная область интересна, прежде всего, с точки зрения рассмотрения варианта объяснения межфеноменального взаимодействия - ментальности и культуры (ментальной культуры) - как возможности уточнения понятия ментальности и обнаружения