

# PROMT

Постредактирование машинного перевода



# Содержание

1. Место МП в переводческом процессе
2. Экспресс-оценка применимости МП
3. Улучшение качества МП
4. Стратегии постредактирования МП
5. Оценка эффективности интеграции МП

# Место МП в переводческом процессе

## Классическая схема:

Источник → переводчик → реципиент

## Схема с МП:

Источник → МП → переводчик → реципиент

## Эффекты от изменения схемы:

1. Фокус работы переводчика смещается с перевода на редактирование.
2. В отличие от типичного процесса с живым переводчиком и редактором, здесь имеем дело с механизированной письменной речью (робот).

# Место МП в переводческом процессе

## Особенности работы с роботом-переводчиком

### Преимущества

1. Возможность настройки
2. Предсказуемость поведения
3. Отсутствие ошибок из-за человеческого фактора

### Недостатки

1. Требуется навыков и умения в обращении
2. Требуется понимания алгоритмов работы
3. Вероятность трудноисправимых ошибок разработчиков

# Место МП в переводческом процессе

**Первые впечатления** от машинного перевода аналогичны реакции редактора на неидеальный перевод:

1. Все плохо, выкинуть и переписать
2. Вот здесь грубейшая ошибка и тут просто смешно
3. А вот тут сообразил, молодец...

Готовый перевод = сумма работы МП и переводчика, при этом работа переводчика не равна нулю.

Ключевой фактор для качественного и эффективного редактирования – степень доверия автору редактируемого текста.

# Содержание

1. Место МП в переводческом процессе
2. Экспресс-оценка применимости МП
3. Улучшение качества МП
4. Стратегии постредактирования МП
5. Оценка эффективности интеграции МП

# Экспресс-оценка применимости МП

## Подход к оценке:

- Акцент на удобство редактирования МП
- Базовые сведения о технологии
- Доступная функциональность с точки зрения организации процесса перевода
- Применимость МП к текстам определенного типа
- Возможности для повышения степени доверия к МП

# Экспресс-оценка применимости МП

Критерии для оценки МП:

- Эквивалентность по числу компонентов предложения и эквивалентность перевода по каждому компоненту

Исходный текст:

Being a leading group in the field, we are seeking for best methods

Машинный перевод:

Будучи одной из ведущих в этой области, мы ищем наиболее эффективные методы

Пропущенные составляющие: group

Неэквивалентный перевод: best

# Экспресс-оценка применимости МП

Критерии для оценки МП:

- Знание предметной области

Исходный текст:

Drilling mud is pumped down the pipe to cool the bit

Машинный перевод:

Буровой раствор закачивается вниз по трубе  
для охлаждения бита

Неэквивалентный перевод для предметной области: bit

# Экспресс-оценка применимости МП

## Критерии для оценки МП

- Устойчивость соблюдения терминологии

### Исходный текст:

The document contains a heading with the heading data inserted by the system.

### Машинный перевод:

Документ содержит заголовок с возглавляющими данными, вставленными системой.

Несоблюдение терминологии: heading

# Экспресс-оценка применимости МП

## Критерии для оценки МП

- **Связность:** слова-связки, служебные части речи, окончания

### Исходный текст:

The level set by the user is incorrect.

### Машинный перевод:

Уровня, установленного пользователем является неверным.

Ошибки в связности: падеж слова «уровень» и зависимых от него, пропущенная запятая после «пользователем»

# Экспресс-оценка применимости МП

## Выводы по оценке МП

Критерий	Оценка	Вывод
Эквивалентность	Нет	МП придется использовать с особой осторожностью, проверяя каждое слово.
	Да	Можно перейти к более серьезным вещам, чем подсчет значимых слов в исходном и переводном предложении.
Предметная область	Нет	Заготовить бумажный словарь по отрасли, найти консультантов.
	Да	Сосредоточиться на более точной передаче смысла второстепенных понятий, а не перепечатывании базовых слов.
Терминология	Нет	Выделить время на проверку каждого термина по глоссарию.
	Да	Найти более удачные решения по сочетаемости, тщательно выстроить логические связи между предложениями.
Связность	Нет	Запастись справочником по грамматике и тезаурусом.
	Да	Подумать о возможности разнообразить текст синонимами и/или тоньше проработать стиль.

# Содержание

1. Место МП в переводческом процессе
2. Экспресс-оценка применимости МП
3. **Улучшение качества МП**
4. Стратегии постредактирования МП
5. Оценка эффективности интеграции МП

# Улучшение качества МП

Настройка МП – отдельный этап перед началом процесса перевода, который имеет смысл при:

- 1) большом объеме однородных текстов,
- 2) жестких требованиях к терминологии и стилю,
- 3) работе в команде переводчиков,
- 4) крайне сжатых сроках.

# Улучшение качества МП

## Основные шаги

1. Определение тематики подлежащего переводу текста и подбор подходящих (специализированных) словарей.
2. Получение списка наиболее частых терминов-оборотов в переводимых текстах, определение их переводных эквивалентов и занесение в словарь.
3. Получение списка наиболее часто встречающихся слов и проверка их текущего перевода системой МП.
4. Установка требуемых правил перевода.

# Улучшение качества МП

## Результаты:

Повышение оценки применимости МП по вышеперечисленным критериям, в частности: выдерживание терминологии и предметной области.

= >

Снижение объема редактирования МП переводчиком:

- за счет настройки предметной области – с 80 % до 50 %,
- за счет настройки терминологии – с 50 % до 30 %.

= >

Повышение общего качества готового перевода.

# Содержание

1. Место МП в переводческом процессе
2. Экспресс-оценка применимости МП
3. Улучшение качества МП
4. Стратегии постредактирования МП
5. Оценка эффективности интеграции МП

# Стратегии постредактирования МП

Полное редактирование с проверкой каждого компонента текста по исходному фрагменту

## Исходный текст:

To this end, final analysis results are not quite obvious.

## Машинный перевод:

С этой целью окончательные результаты анализа не совсем очевидны.

## Перевод переводчика:

Заключение, полученное в результате анализа, не дало окончательного ответа на этот вопрос.

(При этом в глоссарии: analysis results = результаты анализа)

Объем постредактирования предложения: 100 %

# Стратегии постредактирования МП

Использование МП исключительно в качестве подсказки по терминологии

Исходный текст:

To this end, final analysis results are not quite obvious.

Машинный перевод:

С этой целью окончательные результаты анализа не совсем очевидны.

Перевод переводчика:

Окончательные результаты анализа, тем не менее, не дали ответа на этот вопрос.

(При этом в глоссарии: analysis results = результаты анализа)

Объем постредактирования предложения: 82 %

# Стратегии постредактирования МП

Использование МП в качестве грамматического подстрочника с подбором синонимичных выражений

Исходный текст:

SWIFT also brings the financial community together to work collaboratively to shape market practice and define standards

Машинный перевод:

SWIFT также примиряет финансовое сообщество для работы совместно, чтобы сформировать практику рынка и определить стандарты

Перевод переводчика:

SWIFT также объединяет финансовое сообщество для совместной работы в целях формирования рыночной практики и определения стандартов

Объем постредактирования предложения: 65 %

# Стратегии постредактирования МП

Чтение МП и исправление вызывающих сомнения мест с обращением к исходному тексту

Исходный текст:

Distributor may not take legal action against any infringer without SDA's prior written consent.

Машинный перевод:

Дистрибьютор может не подать в суд против никакого нарушителя без предварительного письменного согласия SDA.

Перевод переводчика:

Дистрибьютор не имеет права подавать в суд на какого-либо нарушителя без предварительного письменного согласия SDA

Объем постредактирования предложения: 34 %

# Стратегии постредактирования МП

Чтение МП и исправление очевидных погрешностей без обращения к исходному тексту

Исходный текст:

Our study focuses on emergent visual rhetorical practices that can inspire advertisers.

Машинный перевод:

Наше исследование сосредотачивается на визуальных риторических методах на стадии становления, которые могут вдохновить рекламодателей.

Перевод переводчика:

Наше исследование сосредотачивается на визуальных риторических приемах, находящихся на стадии становления, которые могут вдохновить рекламодателей.

Объем постредактирования предложения: 14 %

# Стратегии постредактирования МП

Сводная таблица по стратегиям редактирования

Стратегия	Объем постредактирования	Имеет ли смысл использовать МП?	Имеет ли смысл настраивать МП?
Полное редактирование	100 %	Нет	Нет
Подсказка по терминологии	80 %	Только при работе в команде	Да
Грамматический подстрочник	60 %	Да	Нет
Анализ сомнительных мест	35 %	Да	Да
Корректурa	20 %	Да	Да

# Стратегии постредактирования МП

Основные критерии для выбора стратегии редактирования:

- знание предметной области
- степень владения исходным и целевым языком
- сроки и предполагаемый объем постредактирования
- опыт работы с данной системой МП
- степень доверия машинному переводу
- наличие технических средств для анализа внутренней структуры МП

# Содержание

1. Место МП в переводческом процессе
2. Экспресс-оценка применимости МП
3. Улучшение качества МП
4. Стратегии постредактирования МП
5. Оценка эффективности интеграции МП

# Оценка эффективности интеграции МП

## Основные эффекты от интеграции:

- 1) Увеличение скорости работы переводчика
- 2) Снижение объема работы при редактировании переводов, полученных от команды переводчиков
- 3) Повышение качества готового перевода

## Дополнительные затраты на интеграцию:

- 1) Техническое и организационное внедрение, обучение
- 2) Настройка МП по проекту

## Влияющие факторы:

- 1) Квалификация переводчиков
- 2) Удобство технической инфраструктуры

# Оценка эффективности интеграции МП

Пример статистики по проекту:

Перевод:

- 66 страниц
- с английского на русский
- каталог продукции

Статистика ТМ: 19 % совпадений

Участники: 2 переводчика, редактор

Система МП: PROMT 9.5

# Оценка эффективности интеграции МП

## Пример статистики по проекту:

### Этап предварительной настройки МП:

- Количество строк в частотном списке – 305
- Количество терминов, занесенных в словарь, – 110
- Время настройки – 1:14

На 3 страницы текста приходится 5 терминов, заносимых в словарь.

Скорость настройки:

около 200 строк частотного списка в час, ИЛИ

около 78 терминов в час, ИЛИ

около 47 страниц в час

# Оценка эффективности интеграции МП

## Пример статистики по проекту:

Этап постредактирования МП переводчиками:

в режиме фриланс – 26 страниц/сутки

при почасовом контроле времени – 6 страниц/час

Этап редактирования текстов, полученных от переводчиков:

средний объем исправлений – 5 %

скорость редактирования – 16,5 страниц/час

# Оценка эффективности интеграции МП

Пример статистики по проекту:

Итого на 66 страниц:

Настройка МП = 1,4 ч.

Перевод = 11 ч.

Редактирование = 4 ч.

= 16,4 ч.



**Спасибо за внимание!**

**Ирина Малышева**  
**Руководитель проекта, PROMT**  
**[irina.malysheva@promt.ru](mailto:irina.malysheva@promt.ru)**