

# К проблеме справедливости и стандартизации расчета объемов в письменном переводе: история, практика, автоматизация, перспективы



Максим Берендяев  
Операционный директор (1-й зам ГД)  
переводческая компания «АКМ-Вест» (Москва)  
группа компаний АКМ Translations

доцент базовой кафедры технологий и автоматизации перевода С(А)ФУ,  
член правления Союза переводчиков России, член Ассоциации преподавателей перевода

# План

Банальности и фон  
Мутная вода

Единицы измерения: слова, строки, страницы, символы, часы

Объект измерения: текст по входу или выходу, кабина или человек, иллюстрации

Осложнения: повторы, совпадения в CAT tools, цифры, NT, память, МП

Memsource vs Smartcat

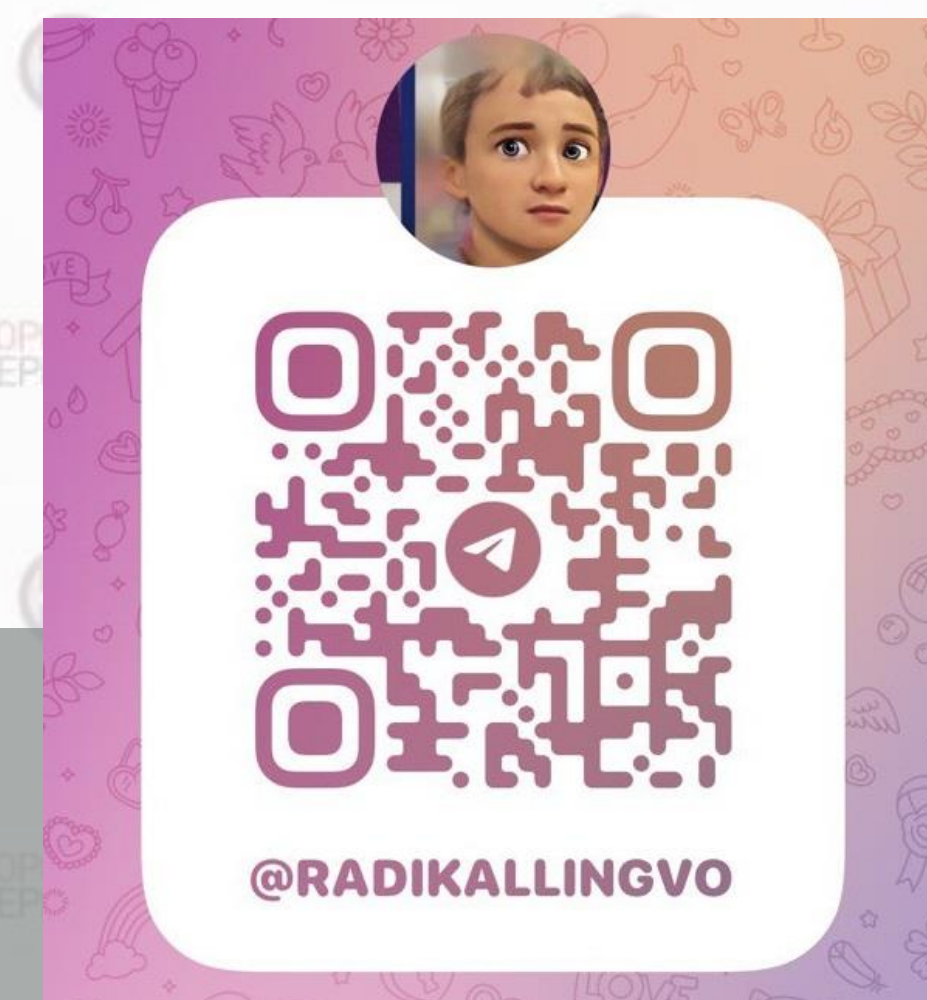
Арифметика (правила пересчета) единиц и систем

Разве надо переводить цифры???

Стандарты (Руководство для закупщиков) от АПК

Эксперименты АКМ-Вест

Школа дидактики перевода Н.Н. Гавриленко



# Банальности

Объем vs. сроки vs. качество\*

Объем и технологическая напряженность

Объем и усилия вендор-менеджера\*\*

Объем и усилия по проектному менеджменту\*\*\*

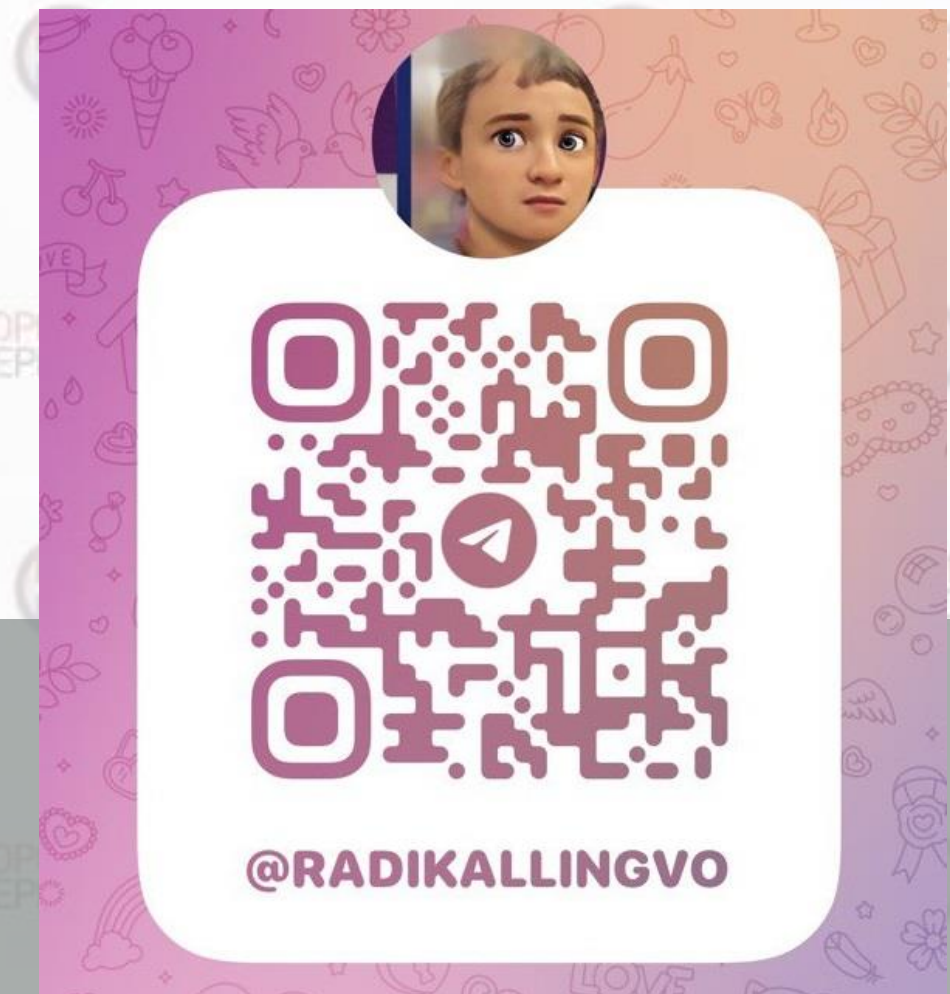
Объем работы – основная измеримая характеристика переводческого проекта. В отличие от качества его можно хоть как-то точно определить, но не тут-то было...

\* субъективно + «эффект наблюдателя»

\*\* это 1 из 3 core functions в LSP

\*\*\* это еще 1 из 3 core functions в LSP

Школа дидактики перевода Н.Н. Гавриленко



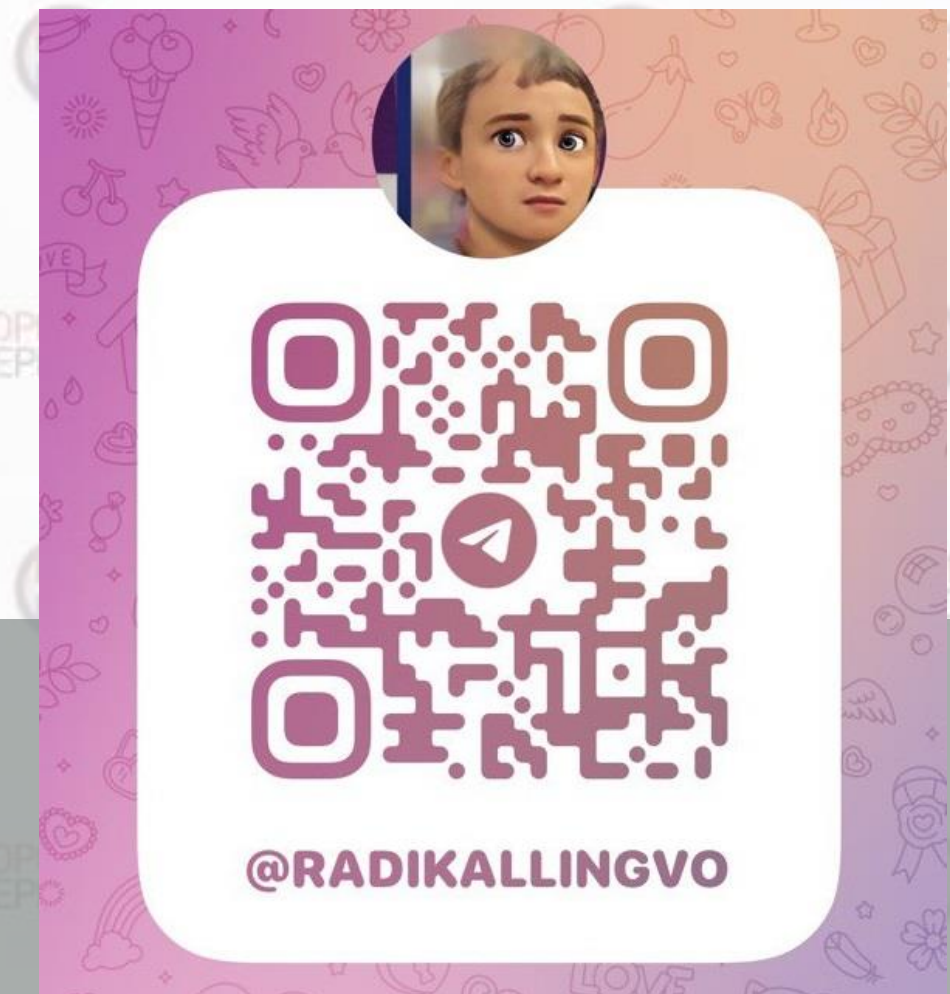
# Увертюра

В ходе работы инициативной группы СПР по созданию рекомендаций о минимальной стоимости перевода через несколько часов выяснилось,

что участники (даже представители компаний) спорили о конкретных цифрах, но предварительно не договорились о том,

- какой тематики текст, какой сложности текст (вопрос задали реальные переводчики, 10+ стр в день),
- какого размера страница: А4, 250 слов, 1800 зсп (бухгалтерия и продавники достали калькуляторы)
- как считать цифры, непереводимые элементы (попытались вставить «технологии» и СТО и СОО)
- не исключаем ли другие языки, что делать со смешанными текстами (чуть не обиделись «неангличане»)
- переводим ли мы с английского на русский или с русского на английский (а разве можно переводить не на родной, а на иностранный язык, подумли - в своем праве - теоретики)
- какой объем имеется в виду: тот **перевод**, который получит заказчик и покажет бухгалтерии, или тот объем текста, который отправят на изучение, осмысление и переработку переводчик...

Фанфары.  
ЗАНАВЕС



# Откровенно

Основные правила переводческого бизнеса в контексте расчета объемов:

Заказываешь услуги перевода, считай по входу. Иначе обманут.

Оказываете услуги перевода, считай по выходу. Иначе обманут.

Исключения - далее. Но все равно обманут.



Влада

@SamennicovaV

Два самых важных совета для бизнеса:

1. Оказываешь услуги? Всегда бери предоплату, иначе кинут.
2. Заказываешь услуги? Никогда не давай предоплату, иначе кинут.



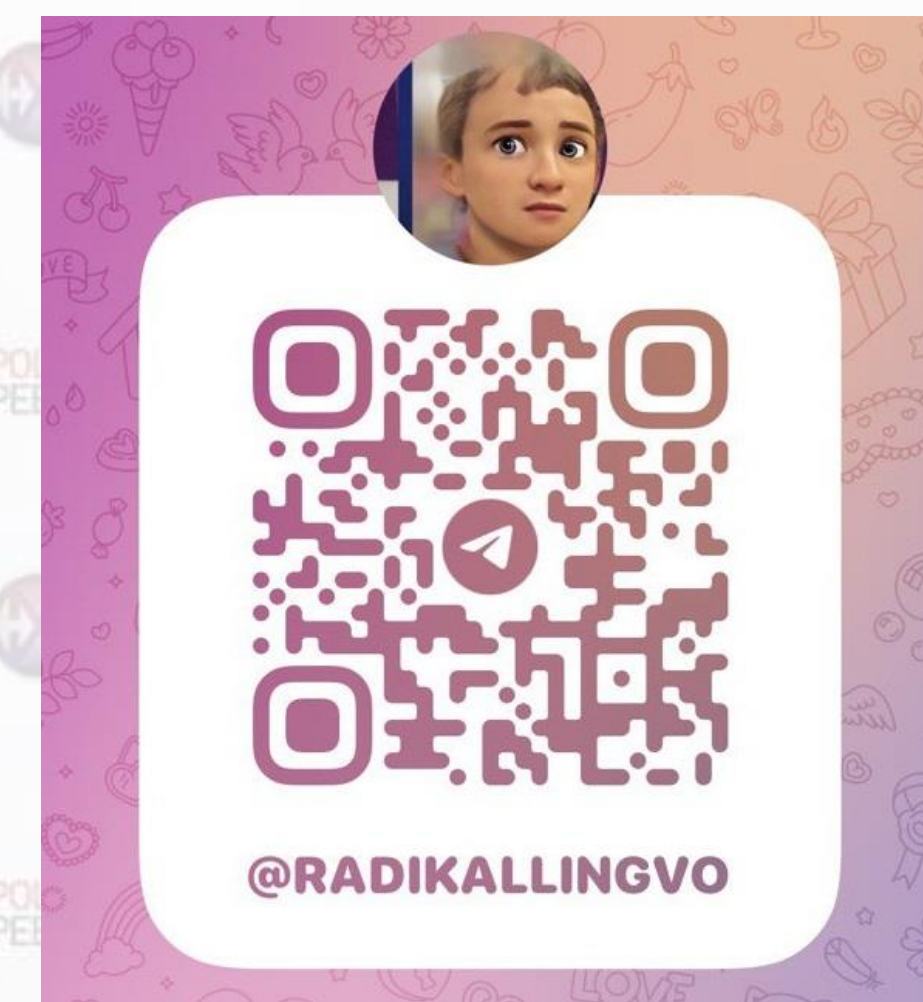
@RADIKALLINGVO

# Любой устный и письменный переводчик, редактор, консультант и психотерапевт умеет генерировать объем к оплате:

- генерировать у клиента потребность в работе (переводе)
- генерировать новую работу
- не спешить (в интересах заказчика и во имя качества коммуникации)... и тянуть время
- растягивать работу на все имеющееся в наличии время +10%
- «наливать» объем нечестными способами
- расшифровывать аббревиатуры и заниматься многословием и даже писать месяц буквами
- придумывать сотни причин, почему редакция (верстка) отняла так много времени...



Он переводит на почасовке!



# Каждое успешное переводческое предприятие умеет сокращать объем к оплате:

- справляться с большими объемами за счет команды параллельных исполнителей
- делать из 1 сверхсрочного проекта (наценка +100%???) 10 несрочных
- выполнять стажерами предперевод и «лочить» NT, цифры
- выполнять своими силами (в штате, без переменных расходов) и «лочить» предперевод сегментов с наиболее высокой повторяемостью внутри текста и совпадением с базами, чтобы не платить за них на outsource
- определять эшелоны работы переводчиков, чтобы лучший или самый дешевый выполнил самую важную часть текста и платить первому переводчику 100%, а остальным в зависимости от совпадений, но в идеале ничего (10%, если есть в договоре)
- (самые отчаянные, не мы) запускать часть проекта на машинный перевод с постредактированием и «подмешивать» его в базы — отмывание МП
- работать строго онлайн, чтобы объем к оплате каждому переводчику уменьшался за счет новых совпадений, которых не было
- использовать для оплаты редакторов или ПЕМТ-редакторов edit distance

В результате, при работе онлайн в CAT tool объем трудозатрат переводчика никогда нельзя определить заранее



## Традиции единиц измерения

- Цивилизованные Америка и Европа, английский язык, слова, слова, слова (ха-ха, 1 000 000 vs 1,000,000)
- Немцы, слова, артикли, сложные сказуемые, строки и Panzerabwehrlenkraketenstartfahrzeug (ха-ха, самоходная пусковая установка противотанковой управляемой ракеты (СПУ ПТУР)), туда же финны и венгры
- Консонантное письмо. Иероглифы. Хангиль и китайские иероглифы
- Отсталая Россия, СНГ, почти все внутрироссийские тендеры (!) и страницы Всесоюзного центра переводов по 1800 з. с. п.
- Мифология вокруг незаменимого эффективного слова ТРАДОС (см. тж. анализ Phrase в любых единицах измерения и даже в знаках без пробелов, по входу и по выходу, с учетом edit distance для глубокой правки любых, даже 100-процентных сегментов, или в зависимости от процента совпадения с базой на момент подтверждения сегмента независимо от правки)





## 5 футов эссе по переводоварению

# 465  
Имя 5 футов эссе по переводоварению  
Автор abakumova.v-narfu  
Создан 18.10. 19:30  
Статус Завершено  
Срок 01.11. 23:00

Задания Editing Новые Редактировать

#	Подтверждено	Файл
1	100%	The Present and Future of M

Анализ Редактировать Пересчитать Удалить

Анализировать

Тип По умолчанию

Имя Анализ #{innerId}

Включить совпадения с памятью переводов

Включить неточные совпадения в тексте

Разделить неточные совпадения в тексте

Включить непере译имые элементы

Включить МТ совпадения

Не учитывать подтвержденные сегменты

Не учитывать заблокированные сегменты

Не учитывать цифры

Анализ по поставщику

Анализ по языкам

**Единицы подсчета**

исходник

перевод

Анализировать

# Кошки

## 5 футов эссе по переводоварению

# 465  
Имя 5 футов эссе по переводоварению  
Автор abakumova.v-narfu  
Создан 18.10. 19:30  
Статус Завершено  
Срок 01.11. 23:00

Задания Editing Новые Редактировать

#	Подтверждено	Файл
1	100%	The Present and Future of M

Анализ Редактировать Пересчитать Удалить Скачать

Анализировать

Тип Постредактирован

Имя Анализ #{innerId}

Анализировать постредактирование памяти переводов

Анализировать постредактирование непере译имых элементов

Анализировать постредактирование машинного перевода

Не учитывать подтвержденные сегменты

Не учитывать заблокированные сегменты

Не учитывать цифры

Анализ по поставщику

Анализ по языкам

**Единицы подсчета**

исходник

перевод

Анализировать

## 5 футов эссе по переводоварению

Задания Editing Новые Редактировать

#	Подтверждено	Файл
1	100%	The Present and Future of M

Анализ Редактировать Пересчитать Удалить

Анализировать

Тип Сравнить

С Translation

Имя Анализ #{innerId}

Не учитывать цифры

Анализ по поставщику

Анализ по языкам

**Единицы подсчета**

исходник

перевод

Анализировать

# Умные кошки

## Переводим для студентов магистратуры англ-рус

СКАЧАТЬ

СВЕРНУТЬ ВСЕ РАЗВЕРНУТЬ ВСЕ

Общая статистика (без учета внутренних неполных совпадений)							
	Слов(а)	%	Сегментов	Страниц(ы)	Символов без пробелов	Символов с пробелами	Эффективных слов для клиента ?
Новый	17814	98.1	1092	71.26	92914	109783	17814
50%-74%	0	0	0	0	0	0	0
75%-84%	69	0.38	21	0.28	423	471	62.1
85%-94%	10	0.06	5	0.04	71	76	7
95%-99%	27	0.15	19	0.11	175	183	13.5
100%	43	0.24	37	0.17	193	199	4.3
101%	0	0	0	0	0	0	0
102%-103%	6	0.03	6	0.02	34	34	0
Повторы	190	1.05	73	0.76	1217	1339	0
Кросс-файловых повторов	0	0	0	0	0	0	0
Итого	18159	100	1253	72.64	95027	112085	17900.9
Только цифры или символы	0	0	45	0	70	70	0
Всего (включая цифры и символы)	18159	100	1298	72.64	95097	112155	17900.9

Jakub\_Absolon\_Human\_Translator\_4\_0\_The\_ultimate\_guide\_to\_Machine.pdf

### Настройки расчета

Статистика от 15.09.2023 11:15

Тарифная сетка: АКМ Трап

### Настройки

Чтобы обновить данные, перестройте статистику

- Память переводов: 3
- ПРАКТИКА изучение Mem
- ТЕХНОЛОГИИ ПЕРЕВОДА
- ТЕХНОЛОГИЯ ПЕРЕВОДА
- Файлы: 1

ОБЗОР ЗАДАЧИ НАСТРОЙКИ ФАЙЛЫ СТАТИСТИКА СМЕТЫ КОМАНДА ЛИНГВИСТИЧЕСКИЕ РЕСУРСЫ ...

620 00

Общая статистика (без учета внутренних неполных совпадений)							
	Слов(а)	%	Сегментов	Страниц(ы)	Символов без пробелов	Символов с пробелами	Эффективных слов для клиента ?
Новый	1	0.01	1	0	2	3	1
50%-74%	0	0	0	0	0	0	0
75%-84%	0	0	0	0	0	0	0
85%-94%	0	0	0	0	0	0	0
95%-99%	0	0	0	0	0	0	0
100%	0	0	0	0	0	0	0
101%	0	0	0	0	0	0	0
102%-103%	18158	99.99	1252	72.63	95025	112082	0
Повторы	0	0	0	0	0	0	0
Кросс-файловых повторов	0	0	0	0	0	0	0
Итого	18159	100	1253	72.64	95027	112085	1
Только цифры или символы	0	0	45	0	70	70	0
Всего (включая цифры и символы)	18159	100	1298	72.64	95097	112155	1

Jakub\_Absolon\_Human\_Translator\_4\_0\_The\_ultimate\_guide\_to\_Machine.pdf

### Настройки расчета

Статистика от 03.11.2023 13:51

Тарифная сетка: АКМ Трап

### Настройки

Чтобы обновить данные, перестройте статистику

- Память переводов: 3
- ПРАКТИКА изучение Memsour...
- ТЕХНОЛОГИИ ПЕРЕВОДА Уче...
- ТЕХНОЛОГИЯ ПЕРЕВОДА МА...
- Файлы: 1
- Jakub\_Absolon\_Human\_Transla...

ПЕРЕСТРОИТЬ СТАТИСТИКУ

## § 11 Honorar für Übersetzungen

(1) Das Honorar für eine Übersetzung beträgt 1,55 Euro für jeweils angefangene 55 Anschläge des schriftlichen Textes (Grundhonorar). Bei nicht elektronisch zur Verfügung gestellten editierbaren Texten erhöht sich das Honorar auf 1,75 Euro für jeweils angefangene 55 Anschläge (erhöhtes Honorar). Ist die Übersetzung wegen der besonderen Umstände des Einzelfalls, insbesondere wegen der häufigen Verwendung von Fachausdrücken, der schweren Lesbarkeit des Textes, einer besonderen Eilbedürftigkeit oder weil es sich um eine in Deutschland selten vorkommende Fremdsprache handelt, besonders erschwert, beträgt das Grundhonorar 1,85 Euro und das erhöhte Honorar 2,05 Euro. Maßgebend für die Anzahl der Anschläge ist der Text in der Zielsprache; werden jedoch nur in der Ausgangssprache lateinische Schriftzeichen verwendet, ist die Anzahl der Anschläge des Textes in der Ausgangssprache maßgebend. Wäre eine Zählung der Anschläge mit unverhältnismäßigem Aufwand verbunden, wird deren Anzahl unter Berücksichtigung der durchschnittlichen Anzahl der Anschläge je Zeile nach der Anzahl der Zeilen bestimmt.

(2) Für eine oder für mehrere Übersetzungen aufgrund desselben Auftrags beträgt das Honorar mindestens 15 Euro.

(3) Soweit die Leistung des Übersetzers in der Überprüfung von Schriftstücken oder Aufzeichnungen der Telekommunikation auf bestimmte Inhalte besteht, ohne dass er insoweit eine schriftliche Übersetzung anfertigen muss, erhält er ein Honorar wie ein Dolmetscher.



# 5000 за страницу

## ГОНОРАР ЗА ПЕРЕВОД

(1) Гонорар за перевод составляет 1,55 евро за каждые полные и неполные 55 знаков (ударов) письменного текста (базовый гонорар). Для текстов, которые не доступны в электронном редактируемом виде, гонорар увеличивается до 1,75 евро за каждые полные или неполные 55 знаков (ударов по клавишам) (повышенный гонорар). Если перевод особо затруднен из-за особых обстоятельств конкретного случая, в частности из-за частого использования технических терминов, трудной разборчивости текста, особой срочной необходимости или из-за того, что это иностранный язык, который редко встречается в Германии, базовый гонорар составляет 1,85 евро, а повышенный гонорар – 2,05 евро. Определяющим при определении числа знаков (ударов по клавишам) является текст на целевом языке; однако если в исходном языке используются только латинские символы, определяющим является число знаков в тексте на исходном языке. Если подсчет знаков требует непропорциональных усилий, их число определяется с учетом среднего количества знаков на строку в зависимости от количества строк.

(2) Гонорар за один или несколько переводов, выполненных по одному и тому же заказу, составляет не менее 15 евро.

(3) Если работа переводчика заключается в проверке письменных документов или записей телекоммуникационного характера на предмет определенного содержания без необходимости выполнения письменного перевода, он получает вознаграждение в качестве устного переводчика.

# Много букв: Почему в объем перевода входят цифры? А разве цифры надо переводить?

Во время перевода переводчик должен обязательно преобразовывать даты, время и числа в соответствии с правилами и принятым форматом целевого языка перевода, страны и региона, для которых предназначается перевод.

Нередко при переводе месяц, записанный цифрами, меняется на слова или наоборот, а указание времени дня переводится из 24-часового формата в 12-часовой, так как в разных странах принята разная запись дат и времени, и есть такие, в которых, например, цифры года записываются комбинацией иероглифов и цифр. И даже для записи номеров телефонов во многих странах существуют свои принципы группировки цифр и разделители между цифрами. Все эти особенности локалей надо обязательно учитывать при переводе и локализации.

Переводчику следует также обращать внимание на десятичный знак (десятичный разделитель), который отделяет целое число от десятичной дроби: в русском и немецком языках это запятая, а в английском – точка. Если работа ведется в системах автоматизации, особенно без функции предварительного просмотра, переводчик, например, должен убедиться в том, что перед ним число с десятичной точкой, а не элемент нумерации параграфов и пунктов в тексте, так как число 1.5 необходимо передавать на русском языке как 1,5, а номер параграфа 1.5 оставлять без изменений.

Отличаются между собой и разделители групп разрядов: в русском языке это, как правило, пробел или точка, в немецком – точка, а в английском – запятая.

Поэтому задачу по переводу цифр в тексте нельзя решить простым переносом оригинала в перевод – наоборот, надо быть особо внимательным, учитывать контекст, чтобы, например, при переводе с русского языка на английский передать число 25,768 как 25.768 и не ошибиться в 1000 раз.

Следует отметить, что Международный стандарт ISO 31-0 с 1992 года во избежание разночтений рекомендует не использовать для группировки цифр числа, точки и запятые, а лишь неразрывные пробелы. Однако до сих пор этим стандартом руководствуются не всегда. Поэтому нередко переводчик должен внимательней анализировать смысл оригинала, проверять информацию, учитывать контекст или запрашивать дополнительную информацию у заказчика.

Кроме того, в некоторых странах большие числа делятся на группы не из трех, а из двух или четырех цифр. А само деление чисел на группы разрядов не осуществляется, если перед десятичным знаком не больше четырех или, как вариант, пяти цифр.

Еще одной яркой иллюстрацией сложностей, возникающих при переводе цифр, является тот факт, что в арабском языке обычный текст пишут справа налево, а арабские цифры (которые по-арабски на самом деле называются «индийскими») – наоборот, слева направо, как и традиционно у большинства народов.

Кроме того, если перевод осуществляется не в редактируемом файле, а с листа, переводчик должен внимательно переносить цифры в текст перевода с учетом требований целевого языка, чтобы не допустить ошибок и опечаток, а редактор (или специалист по контролю качества) уделять повышенное внимание их проверке, так как любые цифры относятся к прецизионной информации и, как правило, обладают очень высокой смысловой значимостью.

## Арифметика

Винсент: — А знаешь, как в Париже называют четвертьфунтовый чизбургер?

Джулс: — Что, они не зовут его четвертьфунтовый чизбургер?

— У них там метрическая система. Они вообще там не понимают, что это за хрен четверть фунта.

— И как они его зовут?



— Они зовут его «Роял чизбургер».

— «Роял чизбургер»? А как же тогда они зовут «Биг-Мак»?

— «Биг-Мак» это «Биг-Мак», только они называют его «Лё Биг-Мак».



# Арифметика + эксперименты

Как увеличивается или уменьшается текст при переводе в разных языковых парах?

Коэффициенты увеличения или уменьшения размера текста перевода относительно текста оригинала зависят от **тематики** и языковой **пары**, а также от **характера** текста, наличия в нем **аббревиатур, сокращений, цифр, имен собственных** и т. д.

В маркетинговом переводе нередко перед переводчиком ставится задача **как можно лаконичней** передать мысли на языке перевода и постараться **удалить стилистически неуместное** на языке перевода **многословие**, даже если оно отмечается в тексте оригинала.

Однако, как правило, для европейских языков отмечается **тенденция к увеличению объема** текста перевода относительно объема текста оригинала, так как письменный переводчик в отсутствие особых указаний заказчика на сохранение объема (например, такие указания даются при переводе текстов для локализации сайтов, приложений и ПО) **склонен конкретизировать информацию** в интересах получателя перевода.

В рамках производственной практики в 2021-2022 годах студенты ВШЭ и САФУ поставили эксперимент с тремя разными текстами и замерили изменение объема в словах и в знаках с пробелами для большинства языковых пар, в которых мы работаем. Для тех языковых пар, по которым студенты-практиканты не смогли привлечь профессиональных переводчиков, в целях оценки изменения объема использовался машинный перевод с подключением разных сервисов.

Представленные результаты носят справочный характер, но могут быть использованы для предварительного расчета объемов. В любом случае порядок расчета объемов и единица расчетов определяются договором. Несмотря на то, что основной единицей измерения в АКМ является страница объемом 1800 знаков с пробелами, в ходе практики **также** рассчитывали изменение объема в словах.

Некоторые значения, например, для китайского языка кажутся нам неправдоподобными, соответственно, в ходе следующей практики в 2023 году мы проведем проверку этих значений, а также приглашаем всех коллег по отрасли к проведению исследований и экспериментов в данной области для составления общеотраслевого справочника.



# Стандарт по единицам измерения для закупщиков переводческих услуг от АПК

АПК (по сути, в целях противодействия нечестной конкуренции) разработала проект  
Руководства для заказчиков («стандарт»)  
по единицам измерения в переводе.

<https://t.me/apkrussia/68> или наводитеесь на QR

## Состав документа о единицах измерения в разработке:

1. Письменный перевод. Латиница, кириллица. Состав услуги.
2. Письменный перевод. Иероглифические языки и вязь. Состав услуги.
3. Устный последовательный перевод.
4. Устный синхронный перевод.
5. Верстка, особые форматы.
6. Возможные дополнительные требования (типа SLA – особые сроки, особое качество, носители, гостайна, доступность 24 часа, сроки реагирования, доставка).
7. Аудио, видео, субтитрование, скрипты, распознавание речи.
8. Организация удостоверения, нотариального удостоверения, легализации, сопутствующих нотариальных услуг.
9. Подбор кадров и обучение.



	С русского языка, в словах	... в знаках с проб. (страницах)	На русский язык, в словах	... в знаках с проб. (страницах)
<b>Английский</b>	<b>1,29</b>	<b>1,03</b>	<b>0,87</b>	<b>1,04</b>
Белорусский	1,01	0,99	1,02	1,02
Украинский	1,02	0,98	1,03	1,03
<b>Немецкий</b>	<b>1,17</b>	<b>1,11</b>	<b>0,93</b>	<b>0,97</b>
Французский	1,37	1,17	0,81	0,93
Испанский	1,35	1,08	0,84	0,94
Итальянский	1,17	1,08	0,87	0,98
Латышский	0,98	0,97	1,06	1,07
Литовский	0,97	0,95	1,06	1,01
Молдавский	1,21	1,09	0,84	0,91
Португальский	1,29	1,05	0,82	0,94
Румынский	1,21	1,11	0,8	0,91
Эстонский	0,81	0,88	1,21	1,1
Азербайджанский	0,93	0,96	1,06	1,03
Армянский	1	1,02	0,99	0,99
Болгарский	1,18	1,01	0,89	0,99
Грузинский	0,87	1	1,11	0,98
Казахский	0,91	0,94	1,07	0,98
Киргизский	0,83	0,89	1	0,90
Польский	1	1,01	1,01	1
Сербский	1,1	0,94	0,96	1,07
Словацкий	1,06	0,98	1,01	1,06
Словенский	1,09	0,96	0,95	1,07
Таджикский	1,09	1	0,89	0,9
Туркменский	0,9	0,96	1,03	0,97
Узбекский	0,92	0,95	1,07	0,94
Чешский	1,01	0,89	1,02	1,11
Хорватский	1	0,95	1,04	1,13
Албанский	1,28	1,01	0,82	1
Венгерский	0,97	0,98	1,07	0,99
Голландский	1,2	1,03	1,01	0,96
Греческий	1,22	1,08	0,8	0,87
Датский	1,21	0,99	0,92	1,03
Македонский	1,09	0,98	0,82	0,94
Норвежский	1,19	1,03	1	1,1
Финский	0,86	1,01	1,28	1
Шведский	1,18	1,05	1,05	1,14
Турецкий	0,94	0,94	1,13	1,12
Вьетнамский	1,78	1,03	0,73	1,1
Индонезийский	1,05	1,01	0,91	0,89
Монгольский	0,97	1,04	0,9	0,92

## «ЗАПАДНЫЕ»

### языки

Увеличение или уменьшение объема текстов при переводе с русского языка и на русский язык, в словах и в знаках с пробелами (в переводческих страницах по 1800 печатных знаков, включая пробелы, знаки препинания и цифры).



МЫ

Ну и на тот случай, если что-то забыл.

Я тут:

РАЗ

А если вы переводчик,  
то тут: [HR@akmw.ru](mailto:HR@akmw.ru)

ГЛЯД

Если в письме есть  
резюме и цены с  
единицами

ИМТВОЙ

Измерения, то отвечаю

ТАЛАНТ

лично всем. Каждому. Всегда.

