УДК 655.28.022.1

АНАЛИЗ ТРАДИЦИОННОЙ МЕТОДИКИ РАСЧЕТА ЕМКОСТИ АВТОРСКОГО ТЕКСТОВОГО ОРИГИНАЛА

Евсеев Илья Владимирович
старший преподаватель кафедры технологии полиграфического производства, кандидат технических наук
Московский государственный университет печати имени Ивана Федорова
127550 Россия, г. Москва, ул. Прянишникова, д. 2А
ilevseev@yandex.ru

***Аннотация**. В работе проведено исследование влияния некоторых параметров набора на точность расчета емкости авторского текстового оригинала.*

Ключевые слова: кегль, длина строки, среднее число знаков в строке, емкость авторской информации, авторский текстовый оригинал.

Одним из базовых параметров расчета объема издания с помощью традиционной методики [1, 2] служит объем авторской информации в авторских листах. Расчет количества авторских листов проводится в соответствии с «Инструкцией по исчислению объема литературного произведения в авторских листах (утв. приказом ОГИЗ РСФСР от 22.02.1940 № 51)» (далее Инструкции). При таком подсчете авторских листов неполные строки считают за полные. Тем не менее, некоторые издательства проводят расчет авторских листов, используя данные из вкладки «Статистика» в текстовом редакторе MS Word. А он не учитывает неполные строки (т.е. он не учитывает часть концевой строки не заполненную текстом, а также абзацные отступы). Естественно издательствам выгоднее рассчитывать емкость авторской информации с помощью статистики MS Word, а авторам, наоборот, выгоден расчет с помощью Инструкции.

Следует уточнить, что в Инструкции не указываются кегль и формат набора, при которых следует проводить расчет. Там говорится, что число строк подсчитывается отдельно по каждому шрифту [3]. Из этого следует, что рассчитывать объем издания в авторских листах можно при любом шрифте и формате. То есть автор может принести рукопись, набранную любым текстовым кеглем на любой формат строки (как правило, менее 210 мм), в соответствии с требованиями к подаче материалов конкретного издательства [4].

В соответствии с существующим стандартом, авторский оригинал (рукопись) должен соответствовать следующим техническим требованиям: изготавливаться на одной стороне формата А4 (210 х 297 мм) или близкого к нему формата; абзацный отступ должен быть одинаковым и равен трем ударам на пишущей машинке, допускается отступ в пять ударов по всему оригиналу; печать основного текста через 2 интервала; строчные буквы шрифта высотой не менее 2 мм; число знаков в строке 60±2 для пишущих машинок с крупным очком литер; на одной странице сплошного текста должно быть (29 ± 1) строк (для словарей — 20 строк); поля: верхнее 20 мм, правое 10 мм [5]. Как видно из текста, эти требования предъявляются к тексту, отпечатанному на пишущей машинке моноширинным шрифтом. Емкость полной текстовой полосы при этом будет варьироваться от 1682 зн. до 1860 зн.

Чтобы достичь необходимой емкости набора в текстовых редакторах необходимо подобрать соответствующий кегль и шрифт. Например, есть мнение, что наиболее близкой к машинописной странице может быть страница, созданная в MS Word с параметрами:

- формат бумаги: А4;
- поля: левое — 3,5 см; правое — 1,5 см; верхнее — 2,0 см; нижнее — 1,9 см;
- шрифт: моноширный типа Courier New или Lucida Console;
- кегль: 12 пт;
- междустрочный интервал: двойной (24 пт).

При этом электронная страница имеет 30 строк, в каждой по 62 знака, что суммарно составляет стандарт машинописной страницы в 1860 знаков [6]. Непонятно, должен ли быть включен или отключен в таком случае алгоритм переносов слов.

На практике оказывается, что каждое издательство имеет свои требования к оформлению текстовых материалов. Так, при подготовке текста данной статьи использовались такие требования: формат бумаги А4; поля: по 3 см с каждой стороны листа (длина строки 204 см); шрифт: Times New Roman; кегль: 14 пт; межстрочный интервал двойной; абзацный отступ: 1 см.

В соответствии с этими требованиями в строке получается приблизительно 63 знака (с учетом пробелов), число строк 21, что в результате образует емкость наборной полосы в 1323 знака (при условии включенного алгоритма разбивки слов на переносимые части).

В связи с вышеизложенным, интересным является рассмотрение условий расчета емкости авторского текстового оригинала по Инструкции, при которых будет наименьшее отличие емкости авторской информации от статистики MS Word.

На этапе предварительного эксперимента была обнаружена зависимость между кеглем шрифта основного текста и объемом авторской информации в авторских листах. Для экспериментального подтверждения этой зависимости были созданы файлы в программе MS Word со следующими параметрами:

- кегль 8–26 пт. с шагом 2 пт. (10 файлов). Такой широкий диапазон кеглей был использован, чтобы посмотреть влияние среднего числа знаков в строке на изменение емкости авторской информации;
- гарнитура Times New Roman (выбрана как наиболее часто используемая);
- формат полосы А4;
- тексты разбиты на абзацы, но без абзацного отступа;
- тексты (табл. 1): текст № 1, текст № 2, текст № 3;
- длина строки 163 мм (сумма левого и правого полей 4,7 см).

Таблица 1. Характеристика используемых текстов

Текст	Средняя длина слова, зн.	Процент пробелов, %	Количество знаков и пробелов по статистике MS Word, зн.	Количество абзацев
Текст № 1	5,248	16	201544	878
Текст № 2	6,14	14	99676	216
Текст № 3	7	12,54	447771	1385

После создания файлов в каждом из них выбиралось (в соответствии с Инструкцией) 10 полноформатных строк из разных абзацев (исключая строки со знаком переноса и иностранными словами). В каждой строке считалось количество символов, включая межсловные пробелы. После чего рассчитывалось среднее число знаков в строке, как среднее арифметическое по 10 строкам. Затем подсчитывалось количество строк в тексте с помощью инструмента «Статистика» в MS Word.

Произведением среднего числа знаков в строке на количество строк, занимаемых текстом рукописи, определялся объем литературного произведения в знаках при каждом кегле и для каждого текста.

Так как расчет объема издания по традиционной методике [1, 2] использует объем авторской информации в авторских листах, рассчитанные емкости были поделены на емкость одного авторского листа (40000 зн.) (рис. 1).

Кроме того, с помощью инструмента «Статистика» MS Word было получено общее число знаков и пробелов рукописи для каждого текста отдельно и произведено сравнение с результатами расчетов емкости авторской информации (рис. 2).

Также было рассмотрено влияние средней длины слова (относительного количества межсловных пробелов в тексте) на среднее число знаков в строке (рис. 3).

Анализ проведенных экспериментов позволил сделать следующие выводы.

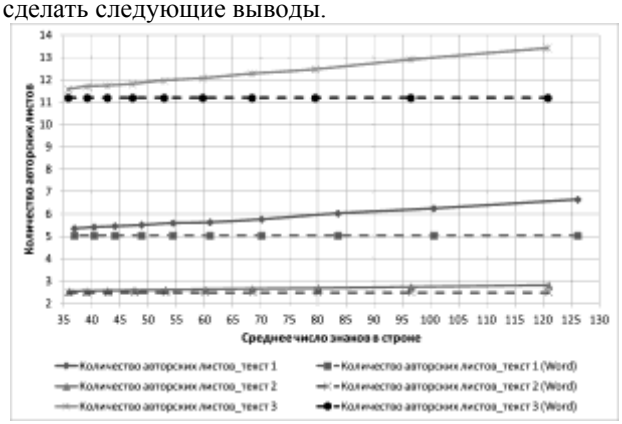


Рис. 1. Зависимость количества авторских листов от среднего количества знаков в строке для трех текстов

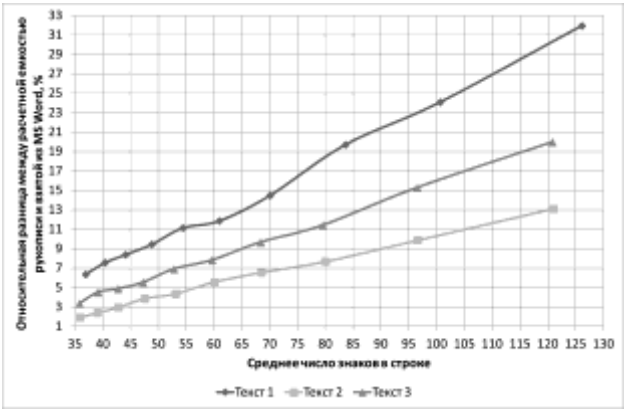


Рис. 2. Относительная разница между расчетной емкостью рукописи и взятой из MS Word, в зависимости от среднего количества знаков в строке для трех текстов

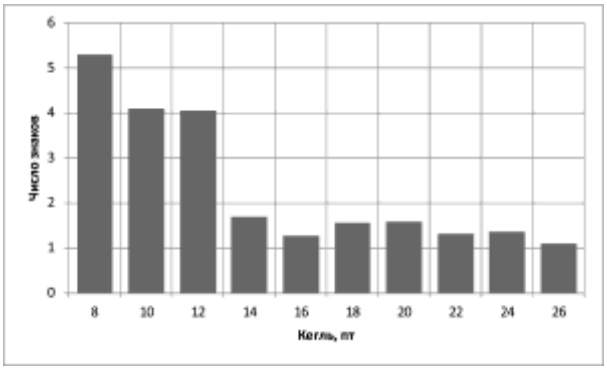


Рис. 3. Максимальная разница между средним числом знаков в строке для трех текстов в зависимости от кегля набора

Уменьшение кегля набора при постоянстве формата набора (т.е. увеличение среднего числа знаков в строке) для одного и того же текста приводит

к увеличению объема рукописи в авторских листах и, наоборот. Причем, чем больше реальная емкость рукописи в MS Word, тем больше разница между ней и расчетной емкостью рукописи. Емкость рукописи, рассчитанная по Инструкции всегда больше емкости в реальных печатных знаках из «Статистики» MS Word. Это связано с тем, что «Статистика» не учитывает «непечатаемые знаки». Под «непечатаемыми знаками» будем понимать часть концевой строки, незаполненной реальными напечатанными знаками и межсловными пробелами. Для исследуемых текстов эти «непечатаемые знаки» могут составлять до 32% от реального количества знаков. С увеличением среднего числа знаков в строке количество «непечатаемых знаков» возрастает. Поэтому если концевая строка не заполнена, например, на 80% от полной строки при кегле 8 пт количество «непечатаемых знаков» в строке будет больше, чем при той же пропорции для большего кегля.

С увеличением кегля набора уменьшается его влияние на изменение объема авторской информации. Так при смене кегля с 8 пт на 10 пт для текста № 3 число авторских листов уменьшится на 0,5, а при смене кегля с 24 пт на 26 пт только на 0,1 авт. лист. Для рассмотренных текстов и параметров форматирования наибольшая точность расчета емкости рукописи достигается при кегле набора 26 пт. Однако есть один нюанс: первоначальные результаты расчета среднего числа знаков в строке дали ощутимую погрешность при кегле свыше 14 пт, что можно объяснить увеличением разброса величины межсловного пробела. Поэтому для повышения точности расчета в больших кеглях использовалось 200 строк, вместо 10.

Влияние средней дины слова (количества межсловных пробелов в тексте) на среднее число знаков в строке наиболее ощутимо в мелких кеглях (в том числе и применяемых для основного текста издания при верстке), а при кегле свыше 14 пт практически сходит на нет (при этом для рассмотренных текстов разница составит 1,5 знака). Чтобы при формате бумаги А4, кегле 14 пт, гарнитуре Times New Roman и включенном алгоритме переносов слов получить среднее число знаков в строке приблизительно 60, необходимо, чтобы сумма левого и правого полей составляла около 6,7–7 см.

В заключение хочется отметить, что для повышения точности расчета объема издания по классической методике необходимо разработать стандартные условия для расчета объема рукописи в авторских листах, отвечающие современным возможностям компьютерных издательских систем, либо полностью отказаться от такого способа получения исходных данных для расчета объема издания.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Волкова Л.А.* Технология обработки текстовой информации. Технологический дизайн. Ч II: компьютерная обработка текста: учебное пособие — М.: Изд-во МГУП, 2007. — 344 с.
2. *Ревякова О.Н.* Моделирование технологии подготовки изданий на донаборной стадии: дис. канд. техн. Наук — М.: Изд-во МГУП, 2002.
3. Инструкция по исчислению объема литературного произведения в авторских листах [Электронный ресурс] — Режим доступа: https://ru.wikisource.org/wiki/Инструкция_по_исчислению_объема_литературного_произведения_в_авторских_листах (дата обращения: 20.11.2015).
4. Определение объема текста для подсчета размера вознаграждения при выполнении письменных переводов [Электронный ресурс] — Режим доступа: http://www.akmw.ru/info-price.html (дата обращения: 20.11.2015).
5. ГОСТ 7.89-2005 СИБИД. Оригиналы текстовые авторские и издательские. Общие требования — М.: Стандартинформ, 2006. — 18 с.
6. Машинописная страница [Электронный ресурс] — Режим доступа: https://ru.wikipedia.org/wiki/Машинописная_страница (дата обращения: 20.11.2015).
7. *Демидов Д.Г., Федоренко Н.М.* Введение в электронные издания: конспект лекций. М.: МГУП. — 2009. — 150 с.

ANALYSIS OF THE TRADITIONAL METHOD OF CALCULATING THE NUMBER OF CHARACTERS OF THE AUTHOR'S TYPESCRIPT

Ilya Vladimirovich Evseev
Moscow State University of Printing Arts
127550 Russia, Moscow, Pryanishnikova st., 2A

Annotation. *In article studied the influence of some parameters set on the accuracy of calculation of the number of characters in the author's typescript.*

Keywords: type size, line measure, the average number of characters per line, the capacity of the author's information, the author's original text.