

# Лабораторная работа 2.1.



Исследовать текст

«Ромео и Джульетта» В. Шекспира

в переводе Б. Пастернака и Т. Щепкиной-Куперник

*(теоретико-множественный подход к моделированию*

*лингвистических процессов)*

# Подготовительная работа

1. Откройте перевод Б. Пастернака и скопируйте из Акта II сцену 2 (со слов «Сад Капулетти»)
2. Создайте **Текстовый документ** в своей папке с именем **Пастернак.txt** и вставьте в него скопированный текст.
3. Откройте перевод Т. Щепкиной-Куперник и скопируйте из Акта II сцену 2 (со слов «Сад Капулетти»)
4. Создайте **Текстовый документ** в своей папке с именем **Щепкина-Куперник.txt** и вставьте в него скопированный текст.
5. Запустите программу **WordTabulator**.

# Загрузка файлов в WordTabulator

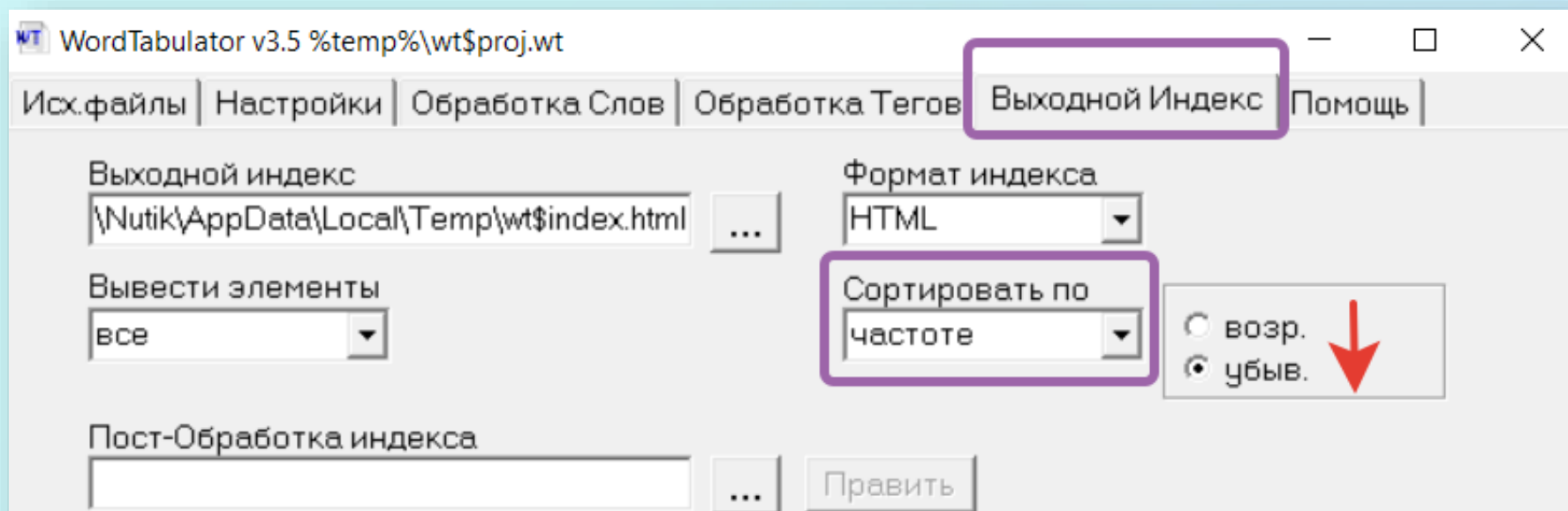


The screenshot shows the WordTabulator v3.5 interface with the following elements:

- Menu: Исх.файлы | Настройки | Обработка Слов | Обработка Тегов | Выходной Индекс | Помощь
- File Filter: \*.txt (circled in red, labeled 1)
- Buttons: A toolbar with icons for file operations. A red circle labeled 2 highlights the 'Load Files' icon (a document with a plus sign). Another red circle labeled 4 highlights the 'Start' icon (a green person). A third red circle labeled 5 highlights the 'View Results' icon (glasses).
- File List: A tree view showing a folder 'Множество Входных Текстов' (labeled 3) containing two files:
  - D:\YandexDisk\Преподавание\ОМОИ\Пастернак.txt
  - D:\YandexDisk\Преподавание\ОМОИ\Щепкина-Куперник.txt
- Status Bar: Shows processing progress: 'Обработано файлов: 2', 'Всего элементов: 729 (3508 байтов)', 'Уник. элементов: 402', 'Формирую выходной индекс...', 'Затрачено времени: 140 msec'. At the bottom, it says 'Затрачено времени 0:00:00'.

1. Чтобы видеть txt файлы, смените фильтр
2. Станьте на "Множество входных текстов" и нажмите кнопку загрузки файлов.
3. Убедитесь, что нужные файлы загружены, раскрыв (+) слева от папки.
4. Нажмите кнопку с зеленым человечком, чтобы начать построение частотного словаря.
5. Посмотрите на результат, нажав на "очки".

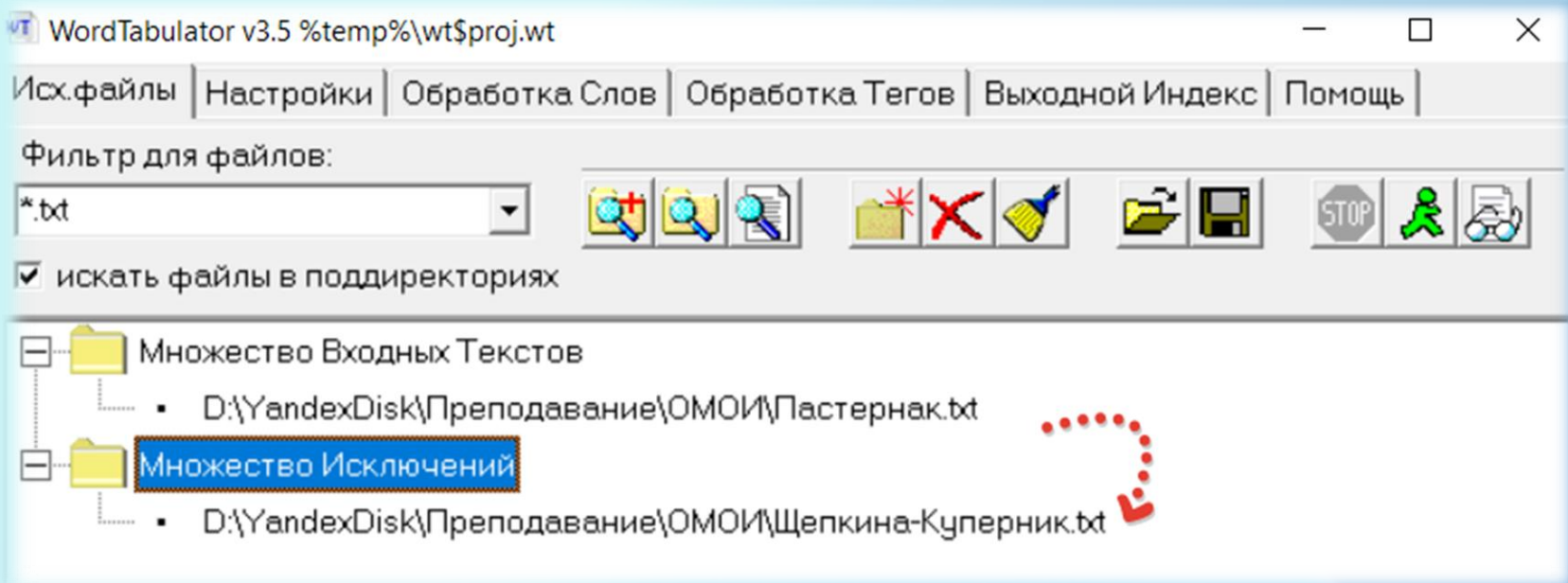
# Настройка словаря – Выходной индекс



На вкладке "Выходной индекс" можно настроить вид частотного словаря:

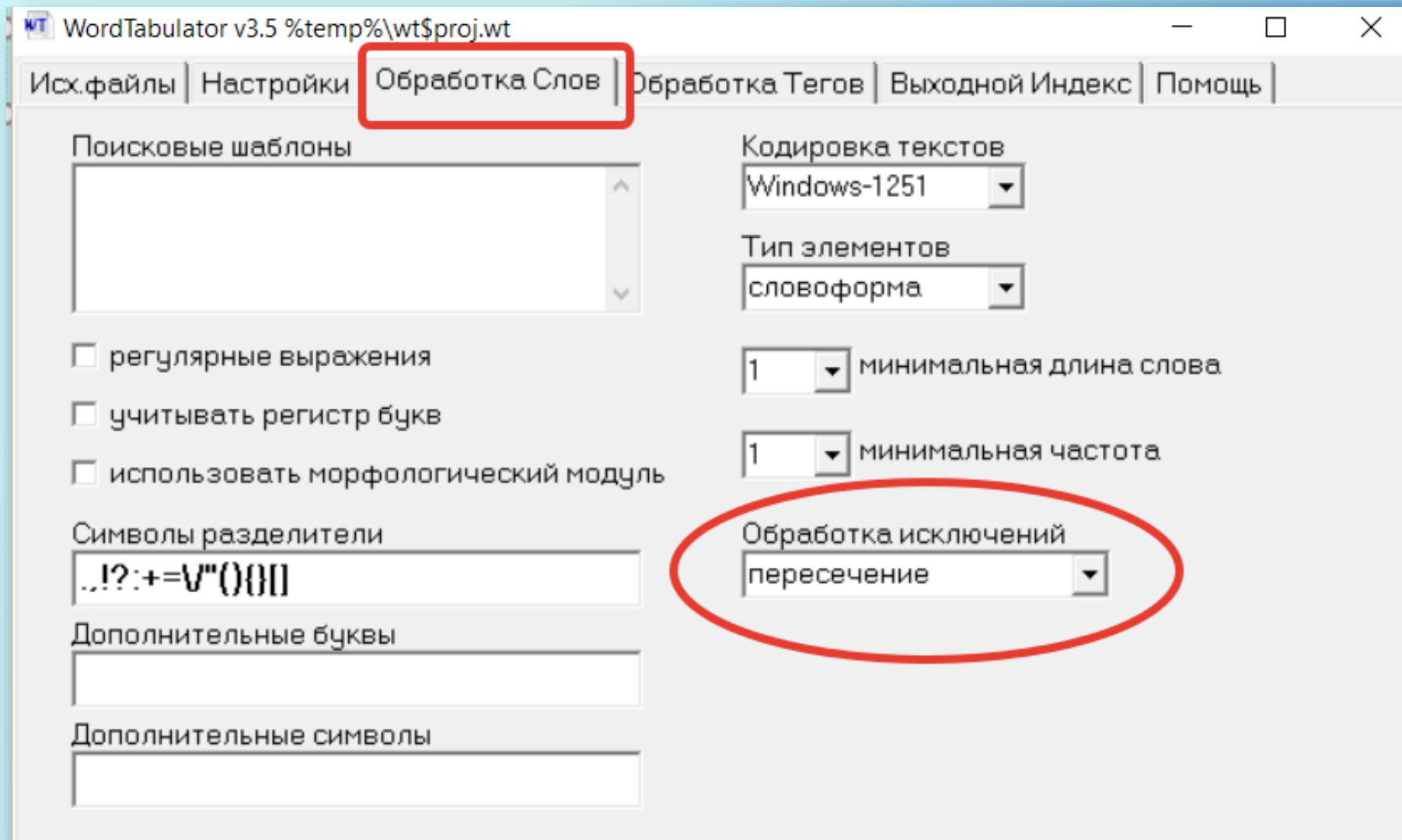
- 1) сортировать по "частоте".
- 2) по убыванию (от большего к меньшему)

# Множество исключений



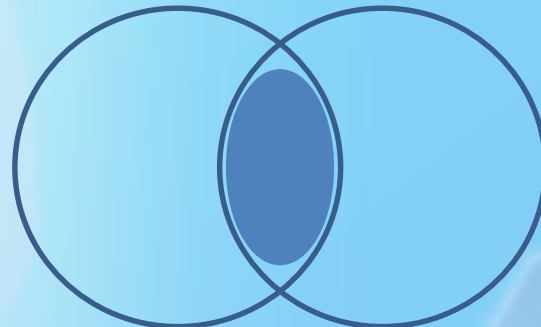
- Перетащите файл с текстом **Щепкиной-Куперник** из Множества Входных Текстов в папку **Множество Исключений**, в этом случае можно будет настроить взаимодействие между этими двумя папками

# Настройка – Обработка исключений

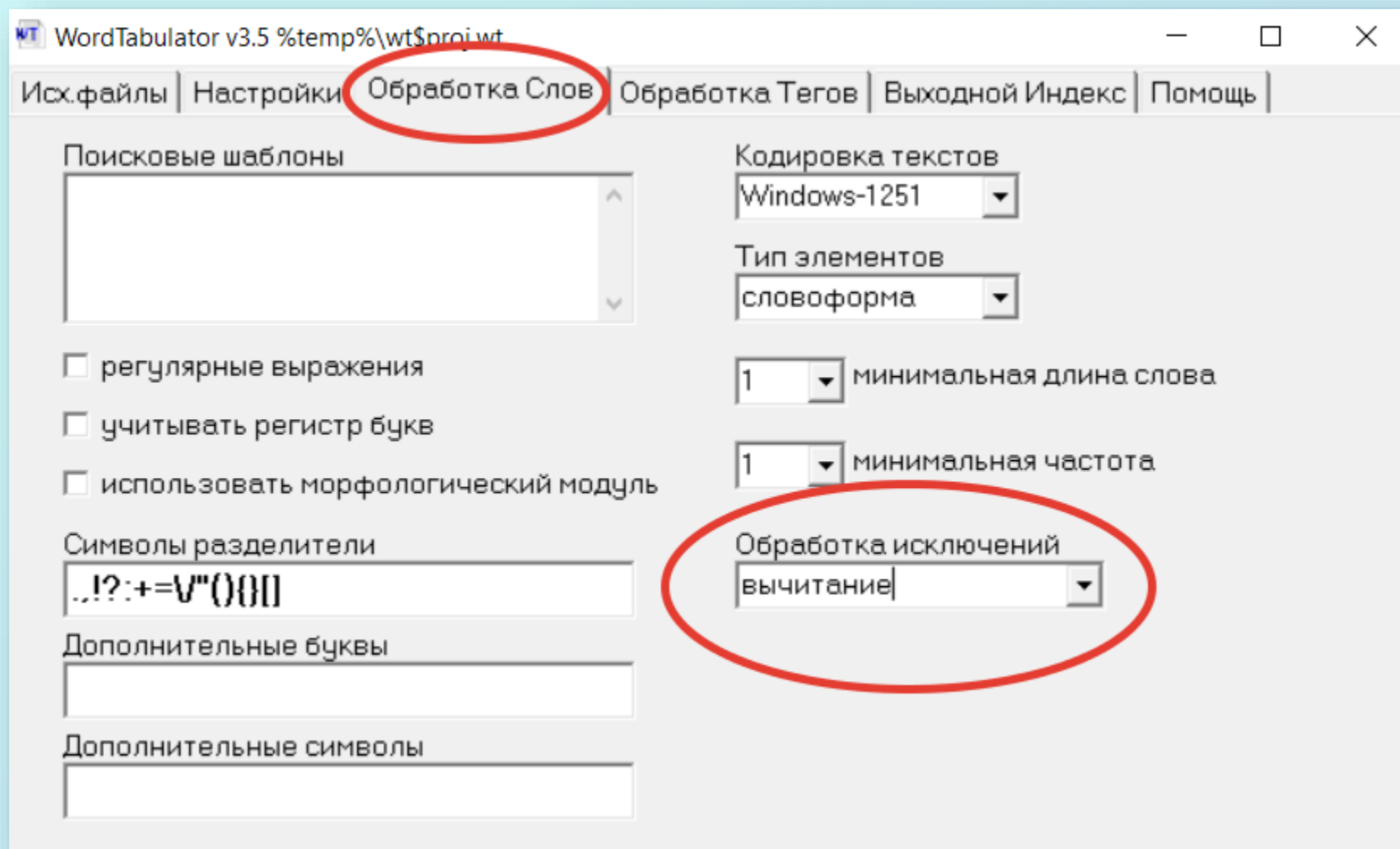


# Пересечение

1. После указанных настроек нажмите на обработку (зеленый человечек) и на **просмотр** (очки).
2. Выделите весь текст командой **Ctrl+A** и скопируйте его (Ctrl+C).
3. Создайте в папке файл RTF с именем **Лабораторная по Шекспиру.rtf** и вставьте скопированный текст на чистый лист (Ctrl+V).
4. Переименуйте строку «Тип элементов: *словоформа*» в «**Общая лексика (A  $\cap$  B)**»



# Настройка исключений – Вычитание

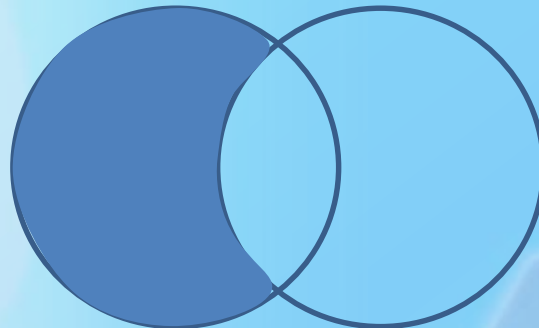


- Смените обработку исключений на «**Вычитание**»

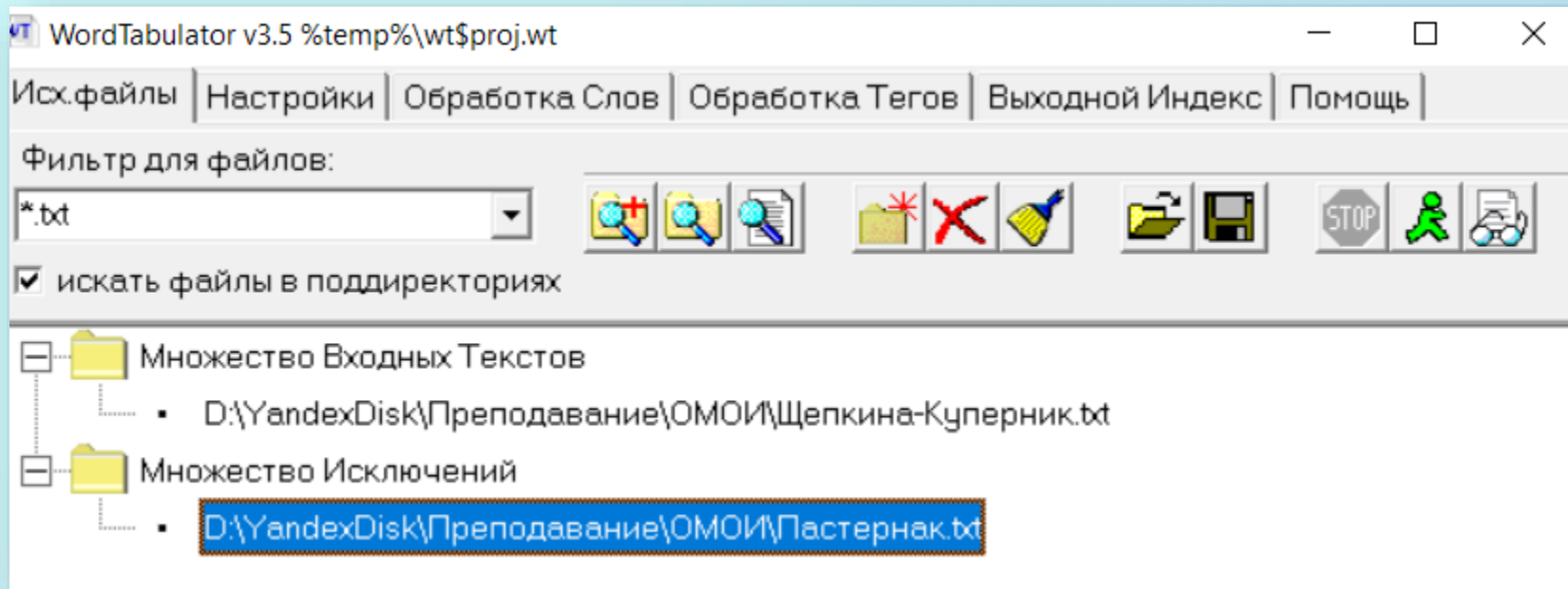


# Вычитание (уникальная лексика)

1. После указанных настроек нажмите на обработку (зеленый человечек) и на **просмотр** (очки).
2. Выделите весь текст командой **Ctrl+A** и скопируйте его (Ctrl+C).
3. Перейдите в конец лабораторной командой **Ctrl+End** и нажмите **Ctrl+Enter**, чтобы начать новую страницу, затем вставьте скопированный текст (Ctrl+V).
4. Переименуйте строку «Тип элементов: *словоформа*» в «**Уникальная лексика Б. Пастернака (A \ B)**»



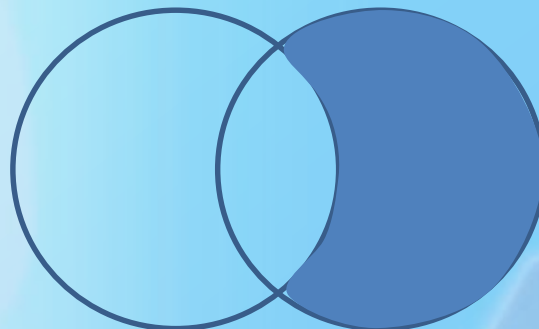
# Перестановка файлов



- Перетащите файл с текстом **Щепкиной-Куперник** назад в папку **Множество Входных Текстов**, а файл с текстом **Пастернака** переместите в папку **Множество Исключений**, обработку слов оставьте той же (Вычитание).

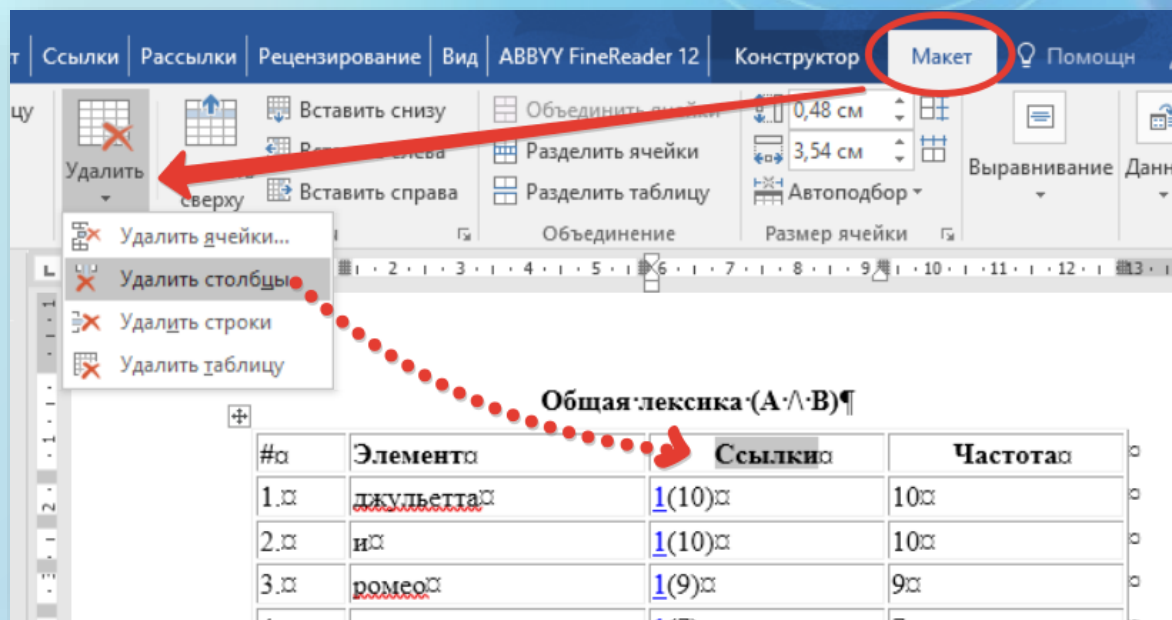
# Вычитание (уникальная лексика)

1. После указанных настроек нажмите на обработку (зеленый человечек) и на **просмотр** (очки).
2. Выделите весь текст командой **Ctrl+A** и скопируйте его (Ctrl+C).
3. Перейдите в конец лабораторной командой **Ctrl+End** и нажмите **Ctrl+Enter**, чтобы начать новую страницу, затем вставьте скопированный текст (Ctrl+V).
4. Переименуйте строку «Тип элементов: *словоформа*» в «**Уникальная лексика Т. Щепкиной-Куперник (B \ A)**»



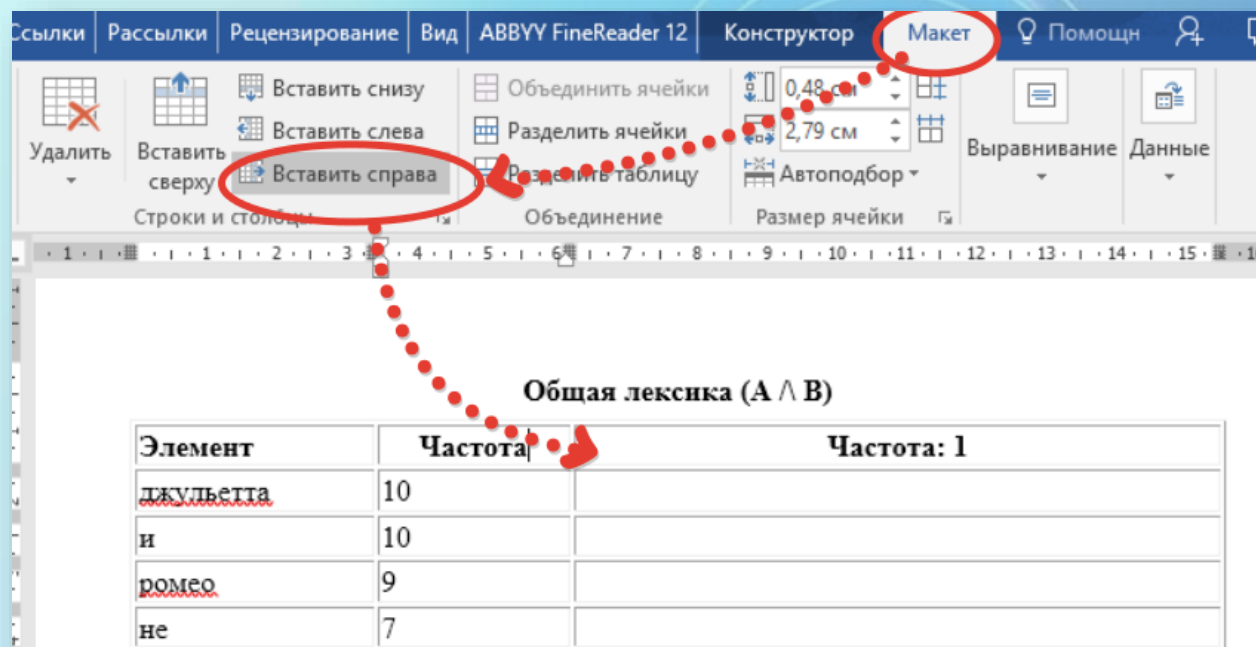
# Обработка таблиц


1. Программу WordTabulator можно закрыть.
2. Отредактируйте все три полученные таблицы по следующим шагам:
  - Удалите столбцы # и «Ссылки», которые дублируют информацию частотных столбцов и «раздувают» размер файла своими гиперссылками.



# Обработка таблиц - 2

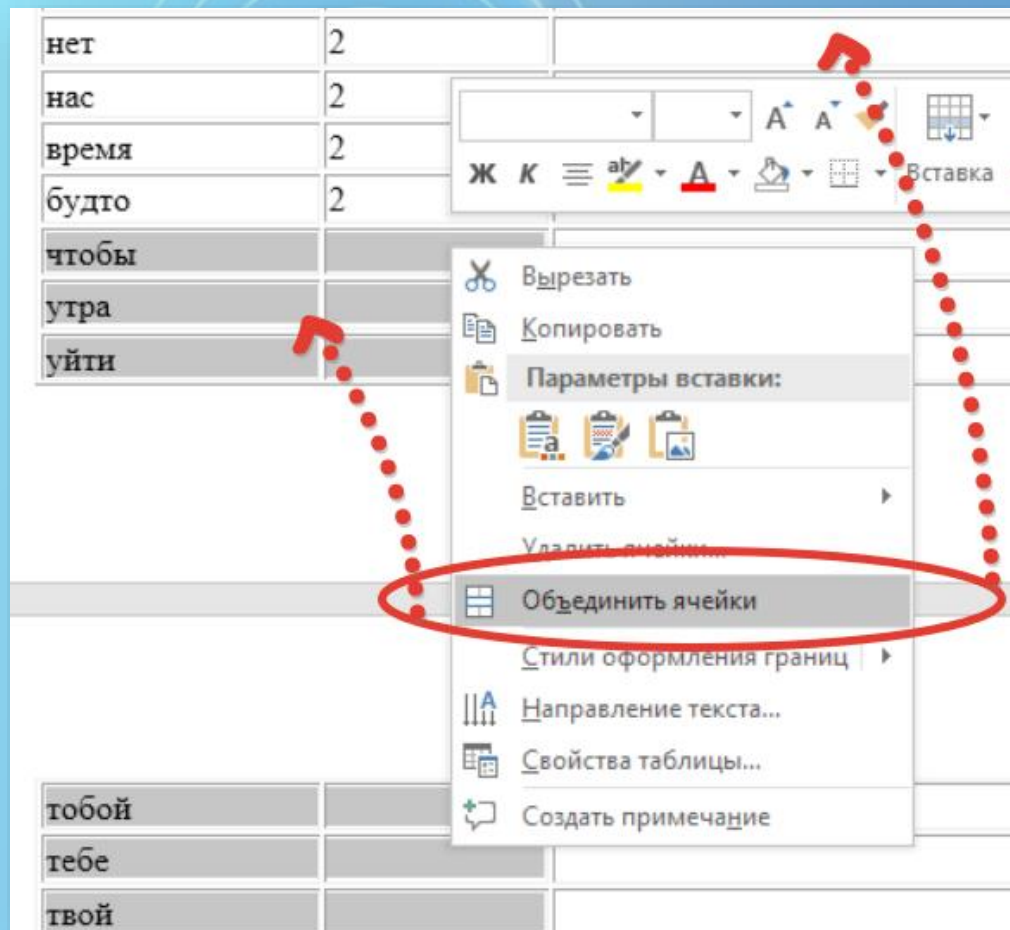
- Добавьте справа столбец, в который позже будут вставлены слова, имеющие частоту в 1 употребление.



- Очистите столбец «Частота» в тех ячейках, где указаны только единицы.
- Отсортируйте по алфавиту слова, употребленные по 1 разу (выделить ячейки с ними и нажать )

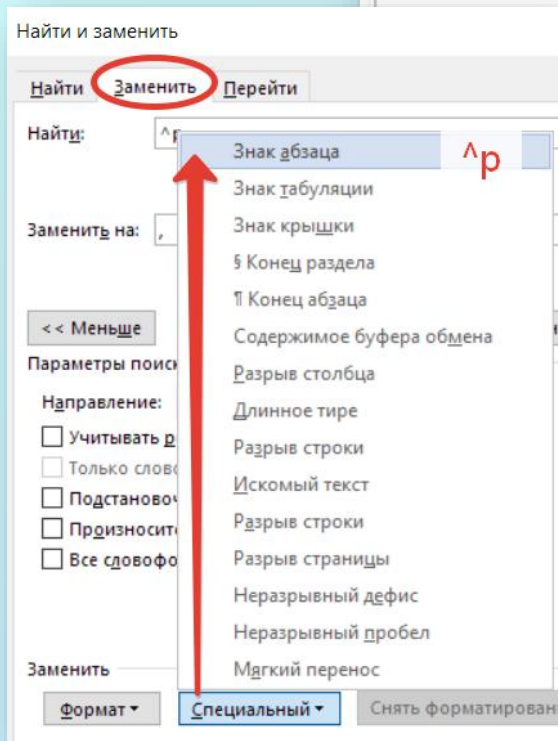
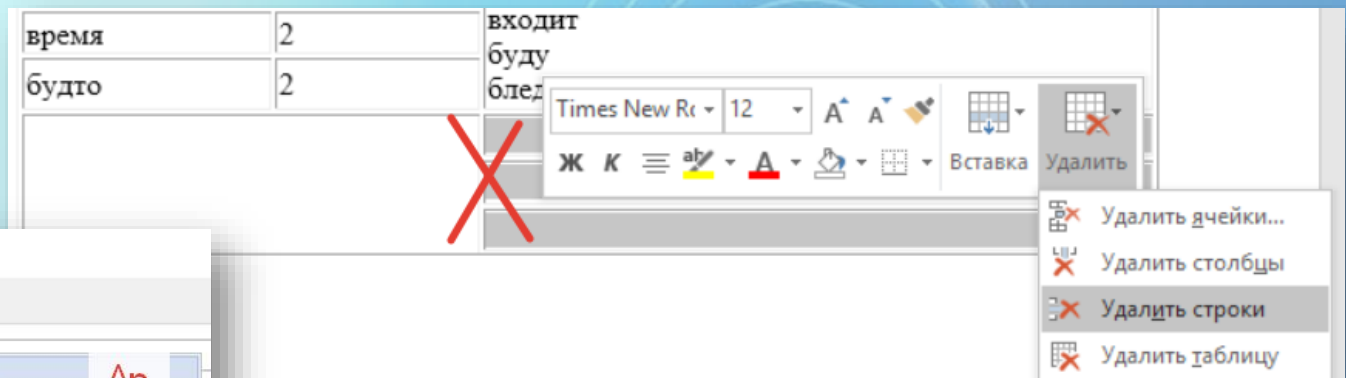
# Обработка таблиц - 3

- Объедините ячейки со словами, которые употреблены единожды в одну большую ячейку.
- Объедините ячейки столбца «Частота: 1» тоже в одну большую ячейку (но до начала нижней большой ячейки).
- Перенесите содержимое нижней большой ячейки в верхнюю правую.



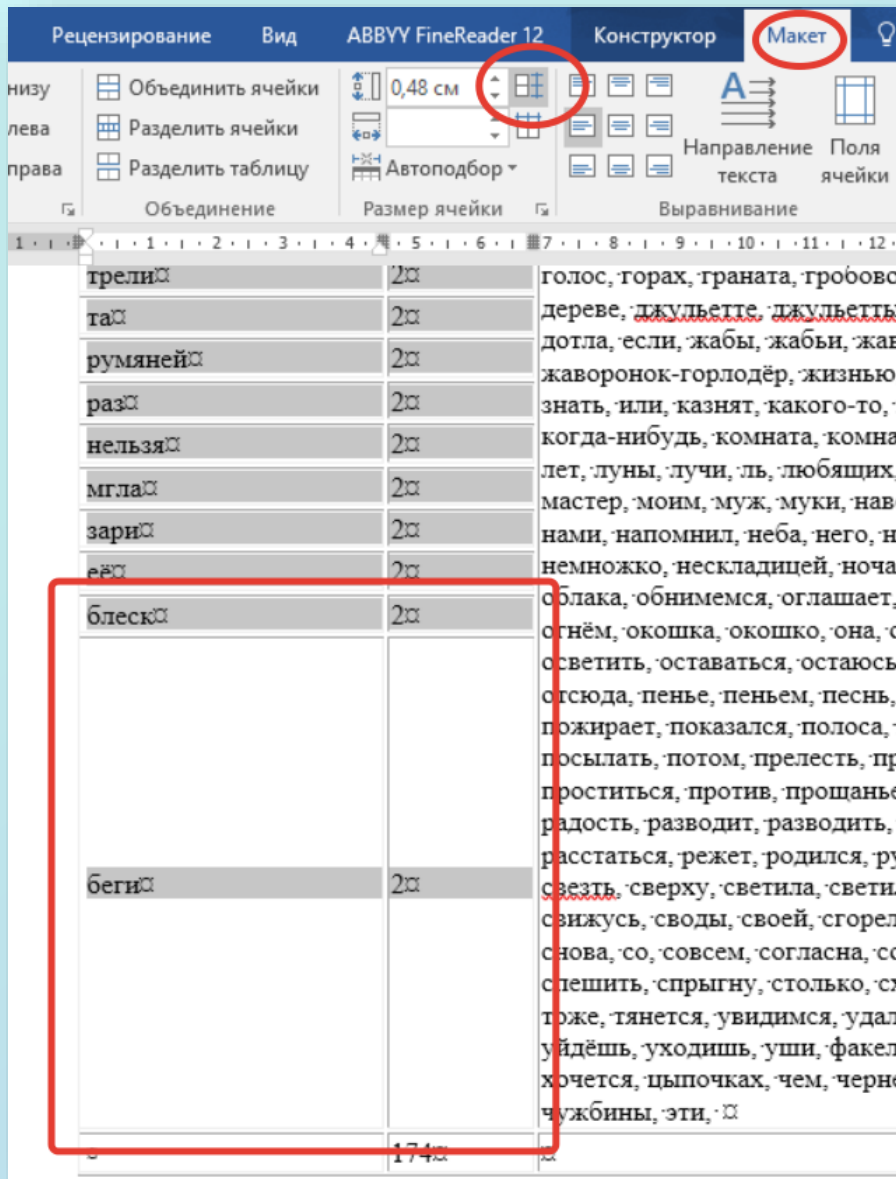
# Обработка таблиц - 4

- Удалите образовавшиеся пустые строки (только не удаляйте последнюю строку с общим количеством)



- Включите отображение непечатаемых символов (Ctrl+Shift+8)
- Замените Знак абзаца на запятую и пробел, чтобы сделать перечень одночастотных слов компактным.

# Обработка таблиц - 5



- При необходимости распределите высоту строк автоматически (выделить строки и выбрать команду «Выровнять высоту строк»).
- Перейдите в начало документа командой **Ctrl+Home** и **Ctrl+Enter**, чтобы отправить таблицы на отдельный лист.
- На первом листе напишите «Лабораторная работа 2.1» и свою фамилию.



# Анализ полученных данных

Анализируя полученные в таблицах данные, опишите концептуальные различия в текстах перевода.

## **Оформление в редакторе MS Word:**

- шрифт – Times New Roman;
- размер – 14 пт;
- межстрочный интервал – 1,5;
- все поля по 2–2,5 см.;
- страницы нумеруются;
- выравнивание абзацев – по ширине.

---

**Пришлите файл** с анализом и таблицами преподавателю на адрес **dopinfo@ya.ru**