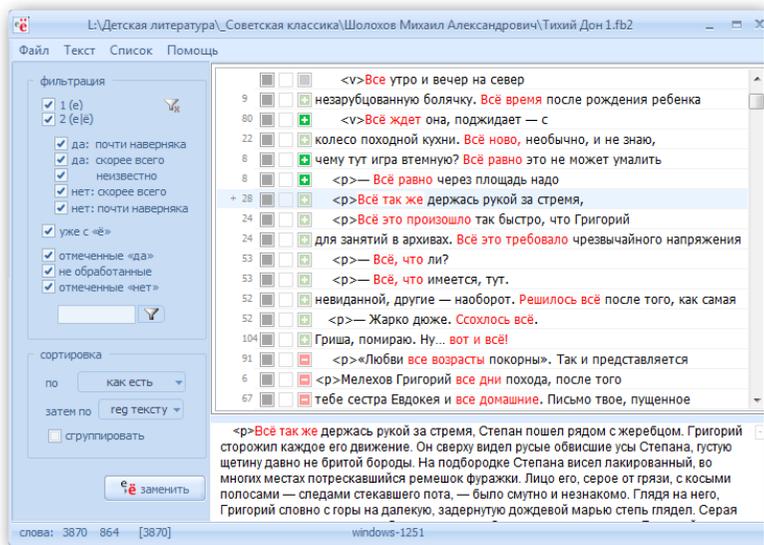


Введение

Программа **e2yo** предназначена для «ё-фикации» документов MS Word, Adobe InDesign (InCopy), rtf, fb2, html, xml, txt.

Слова, которые пишутся через «ё», разбиты на две группы. Однозначные — слова, которые без «ё» не употребляются (ещё, весёлый, днём, утёса и т.д.). Неоднозначные — слова, которые по тем или иным причинам (допустимость вариантов, разные слова, формы слова) могут писаться как через «е», так и через «ё» (всё — все, осознаёшь — осознаёшь).

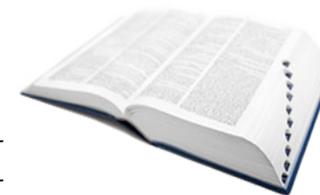
Основное отличие от остальных программ в том, что сначала создаётся список всех слов, в которых может стоять «ё», просматривается (в основном неоднозначные слова при помощи контекста), а потом осуществляется замена в документе отмеченных слов.



Документы в формате MS Word, Adobe InDesign, Adobe InCopy проверяются с помощью механизма «OLE Automation». Это значит, что документ должен быть открыт в соответствующей программе (сервер автоматизации), а **e2yo** запущена «поверх». Документы в формате rtf, fb2, html, xml, txt открываются программой непосредственно.

Словари

В основе — словарь русского языка для ispell Александра Лебедева.



словарь_однозначные.txt — словарь однозначных слов, представляющий собой просто список слов.

словарь_неоднозначные.txt — словарь неоднозначных слов, состоящий из строчек вида: слово, табулятор, «вероятность замены», где «вероятность замены» может принимать значения

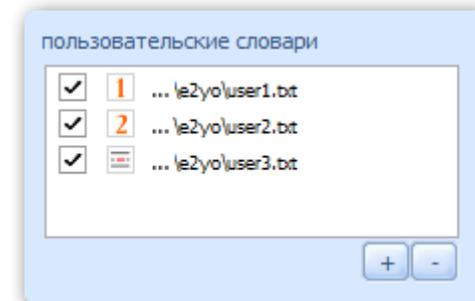
-2	нет, почти наверняка
-1	нет, скорее всего
0	непонятно
1	да, скорее всего
2	да, почти наверняка

словарь_уточнения.txt Словарь «контекстов». Использование регулярных выражений позволяет исключить значительную часть неоднозначностей, например, «в тёплом» — почти наверняка говорит о том, что пишется именно «тёплом». Словарь состоит из строчек вида: слово, табулятор, регулярное выражение, табулятор, «вероятность замены»

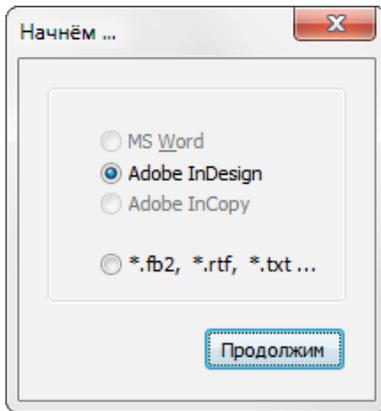
моём→(в|при|на|о) моем→1

тёплом→(в|о|на|при) теплом→2

В диалоге настроек можно добавить пользовательские словари всех трёх типов.



Порядок работы



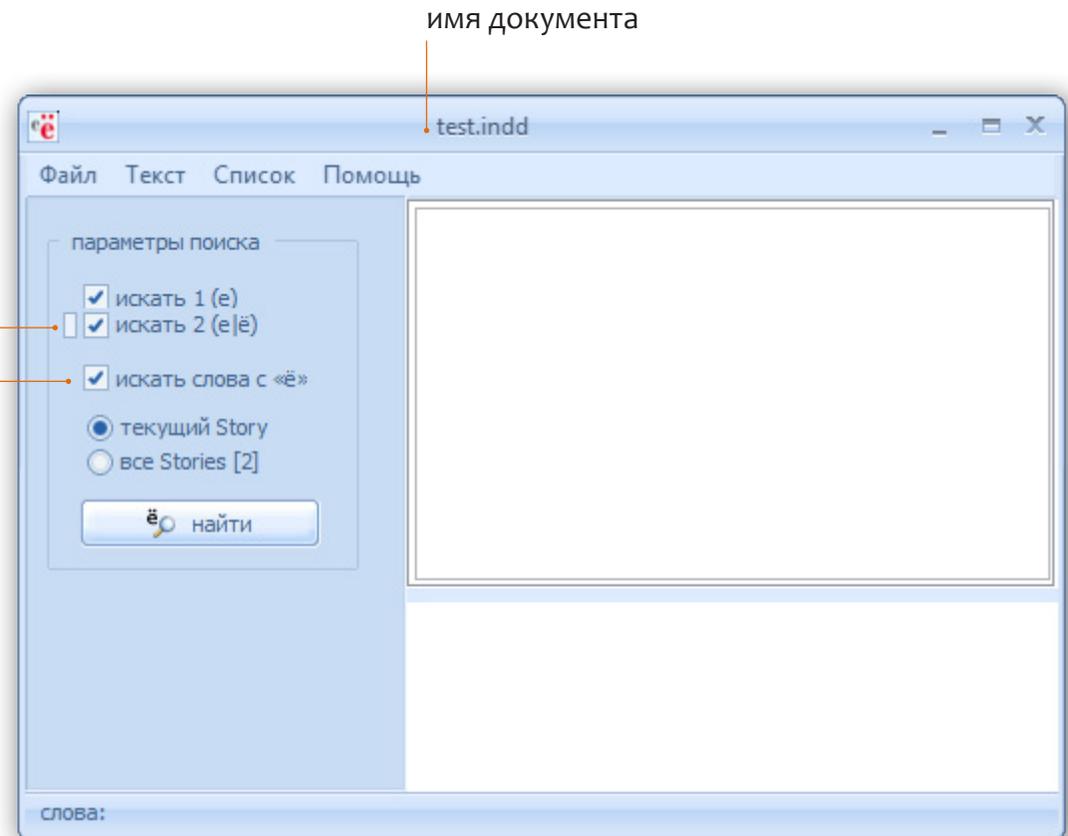
При запуске программа определяет, какие из программ (Word, InDesign, InCopy) запущены. Если таковые имеются, соответствующие элементы доступны для выбора. Для работы с файлами fb2, rtf, txt нужно отметить нижнюю строчку.

Далее, выбрав нужные опции, жмём кнопку «найти».

Если отмечено — контекстный поиск использует все правила

Например, для «все» определено несколько контекстных правил: всё шли и шли (все ([а-яё]{+} (и|да) |1), все смотрели (все +[а-яё]{3,}([и|ы]) и т. д. Так вот с отмеченной опцией будут обработаны все правила, а не первое совпавшее.

поиск слов в тексте с уже проставленной «ё»



Анализ

Хоть однозначные слова и не требуют пристальной проверки, но бывают случаи сокращений, приводящие к неоднозначностям. Замеченные мною и перенесённые в словарь неоднозначных слов слова: вел (вёл) — сокращение от великий, мед (мёд) — сокращение от медицин*. Не исключено, что в словаре однозначных ещё есть подобные или просто неоднозначные слова. Поэтому какой-то вариант просмотра нужен — либо в программе, либо последующая вычитка.

плюсик рядом с номером правила, говорит, что для данного вхождения их найдено несколько

индикатор замены

индикатор наличия «ё» в слове

индикатор вероятности замены

фильтрация

- 1 (e)
- 2 (e|ё)
- да: почти наверняка
- да: скорее всего
- неизвестно
- нет: скорее всего
- нет: почти наверняка
- уже с «ё»
- отмеченные «да»
- не обработанные
- отмеченные «нет»

сортировка

по

затем по

группировать

е ё заменить

слова: 3870 864 [3870] windows-1251

9 <v>Все утро и вечер на север

80 незарубцованную болячку. Все время после рождения ребенка

22 <v>Все ждет она, поджидает — с

8 колесо походной кухни. Все ново, необычно, и не знаю,

8 чему тут игра втемную? Все равно это не может умалить

8 <r>— Все равно через площадь надо

+ 28 <r>Все так же держась рукой за стремя,

24 <r>Все это произошло так быстро, что Григорий

24 для занятий в архивах. Все это требовало чрезвычайного напряжения

53 <r>— Все, что ли?

53 <r>— Все, что имеется, тут.

52 невиданной, другие — наоборот. Решилось всё после того, как самая

52 <r>— Жарко дюже. Ссохлось всё.

104 Гриша, помираю. Ну... вот и всё!

91 <r>«Любви все возрасты покорны». Так и представляется

6 <r>Мелехов Григорий все дни похода, после того

67 тебе сестра Евдокея и все домашние. Письмо твое, пущенное

<r>Все так же держась рукой за стремя, Степан пошел рядом с жеребцом. Григорий сторожил каждое его движение. Он сверху видел русые обвисшие усы Степана, густую щетину давно не бритой бороды. На подбородке Степана висел лакированный, во многих местах потрескавшийся ремешок фуражки. Лицо его, серое от грязи, с косыми полосами — следами стекавшего пота, — было смутно и незнакомо. Глядя на него, Григорий словно с горы на далекую, задернутую дождевой марью степь глядел. Серая