**FileProcessor**

Единый графический интерфейс, который позволяет производить:

1. Элайн размеченных интервалов границ слов, ЭДЕ, пауз и текстовой расшифровки. Для этой задачи существуют два варианта работы
   1. Кнопка «Выбрать файл». Выбираем любой файл, который относится к группе обрабатываемого файла (можно выбрать транскрипт, можно выбрать сег, можно выбрать речевой сигнал). Программа будет искать файл с расширением seg\_B1 и текстовую расшифровку (обязательно и сеги, и расшифровки должны быть в строго в кодировке UTF-8). На основе этих двух файлов будет произведен элайн. Будут созданы сеги: **seg\_B2** (туда будут вписаны слова), **seg\_R1** (границы ЭДЕ), **seg\_R2** (сюда будет записана информация об ошибках в разметке [несовпадение количества интервалов и слов для этой ЭДЕ в транскрипте + номер строчки в транскрипте, чтобы было удобно искать]; **если ошибок нет,** то на этот уровень записываются порядковые номера предложений. Граница предложения в расшифровке – пустая строка между двумя ЭДЕ), seg\_R3 – тут информация (либо количество символов “s” совпадает с количеством ЭДЕ в расшифровке, тогда там будет записано «ок», либо строчка с текстом «Проверьте количество меток границ ЭДЕ и что-то бла-бла-бла…».
   2. После того, как пользователь нажал «Выбрать файл» в поле сверху записывается путь к файлу, с которым происходит работа. У меня как правило возникает довольно много багов в разметке (где-то забыл интервал поставить, где-то не поставил границу начала эде на сигнале, где-то паузу забыл пометить), поэтому на кнопку «Выбрать файл» нужно было жать много раз. Меня это стало раздражать и я добавил кнопку «Обработать файл из окна». То есть по нажатию этой кнопки проводиться попытка элайна файла, путь к которому прописан в текстовом поле вверху программы. Эта штука очень сильно экономит время и позволяет не раздражаться от необходимости тысячу раз выбирать файл, который надо обработать.
   3. Наконец, есть возможность «Выбрать директорию». Тогда будет произведена попытка элайна всех файлов в директории, но на практике я это ни разу не использовал. Все-таки сделал один файл, потом начинаешь с ним плясать, чтобы сделать элайн.
   4. Кнопка «Выбрать TextGrid» нужна только в том случае, если разметка производилась в Praat, а не в Wave Assistant и вы хотите произвести элайн грида и транскрипта. Я пробовал размечать и там и тут. В принципе одно и тоже, поэтому в конце концов остановился на работе с Wave Assistant. Но если вдруг кому нужно сэлайнить грид и транскрипт, то пожалуйста
2. Кнопка «Добавить речевые сбои». Когда произведен элайн границ слов на речевом сигнале и транскрипт, когда создан conll для расшифровки, то за счет элайна conll и seg\_B2, мы добавляем уровни разметки seg\_G3 и seg\_Y2 информацию о речевых сбоях. Список всех речевых сбоев, которые может на данный момент найти скрипт и разметить на сигнале можно найти в файле **scripts\_for\_processing/speech\_repairs\_list.txt.**  Если нужно список дополнить, то открываем блокнотом и добавляем информацию либо в FRS, либо в SR. После сохранения можно запускать программу, информация от туда будет подхвачена.