**Уровень 1**

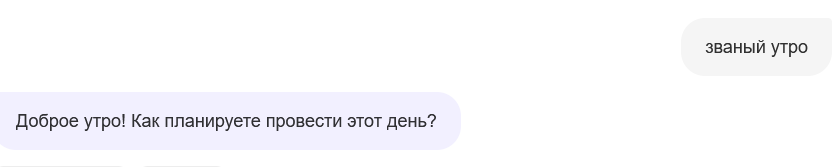
**Стр 87**

Вопросы:

1. К наиболее важным задачам речевой обработки и анализа можно отнести распознавание говорящего по голосу и преобразование речи в текст.
2. Решение этих задач позволяет реализовать:автоматическое создание транскрипции, синхронное преобразование живой речи в текстовый формат, управление голосовыми помощниками, выявление особенностей речи говорящего, выделение наиболее значимых слов в речи, идентификация и верификация личности, создание и разработка технологий для людей с ограниченными возможностями, создание диалоговых систем и чат-ботов.
3. Верификация работает по системе “свой/чужой”: в ходе этого процесса мы сравниваем речевой сигнал говорящего с хранящимися.
4. Идентификация - система понимает какой из заранее известных пользователей говорит с ней.
5. Решение данной проблемы позволяет развивать искусственный интеллект(например, у голосовых помощников), а также призвано помочь людям с ограниченными возможностями.
6. Первые образцы основывались на эталонных моделях и дробили речь на составляющие, что приводило ко множеству ошибок. Сейчас же в основе современных систем распознавания речи лежат базы данных коллокаций и используются N-grammы.
7. За счёт использования баз данных коллокаций, N-gramm, машинного обучения.
8. Анализ сигнала - речевой сигнал очищается от лишних шумов:  
   Распознавание сигнала - определяются звуки через модель, объединяем в значимые единицы  
   Преобразование в текст - с помощью модели языка и баз данных распознаем нераспознанные, переобразуем в текст

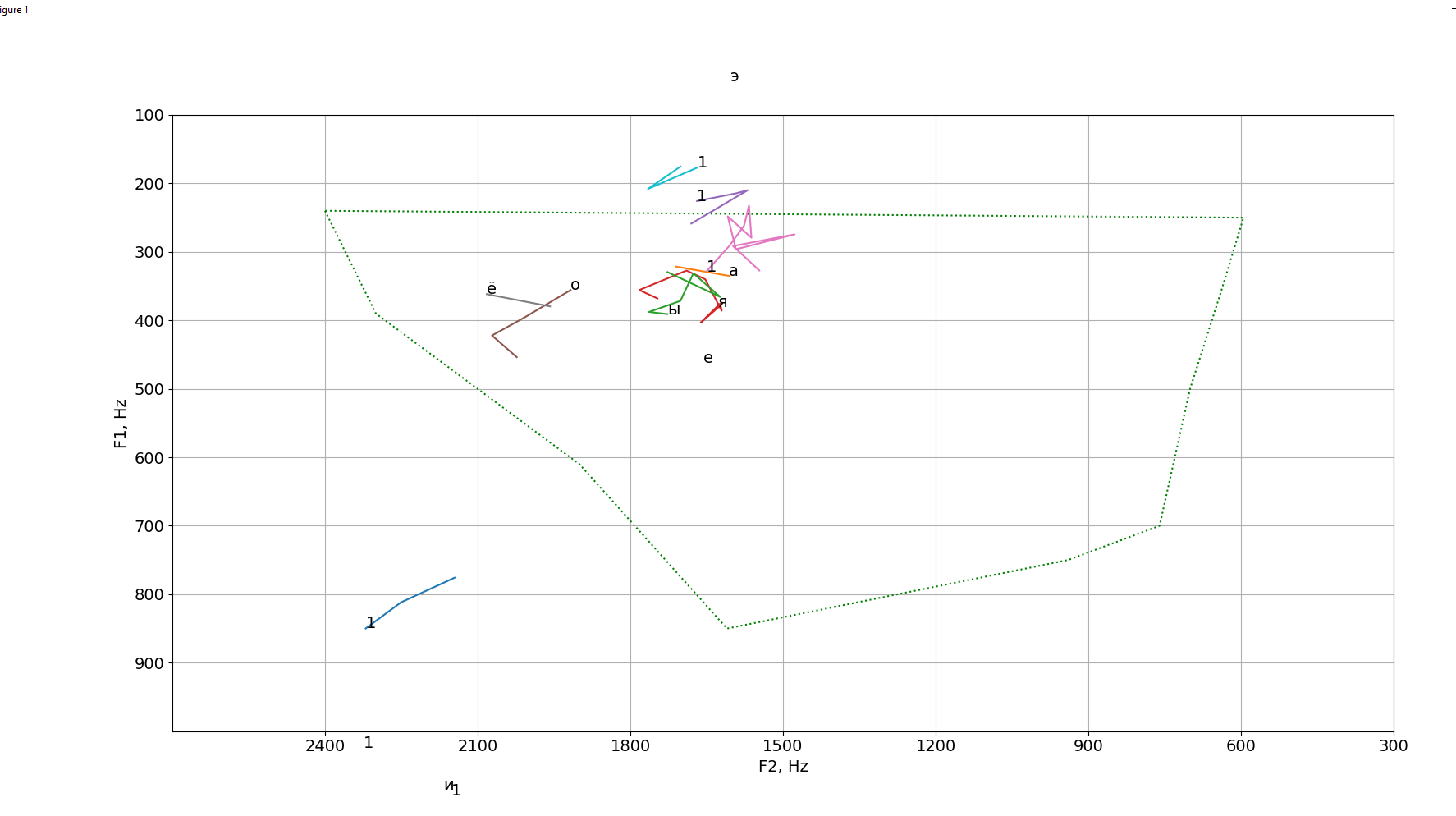
Лабораторная работа:

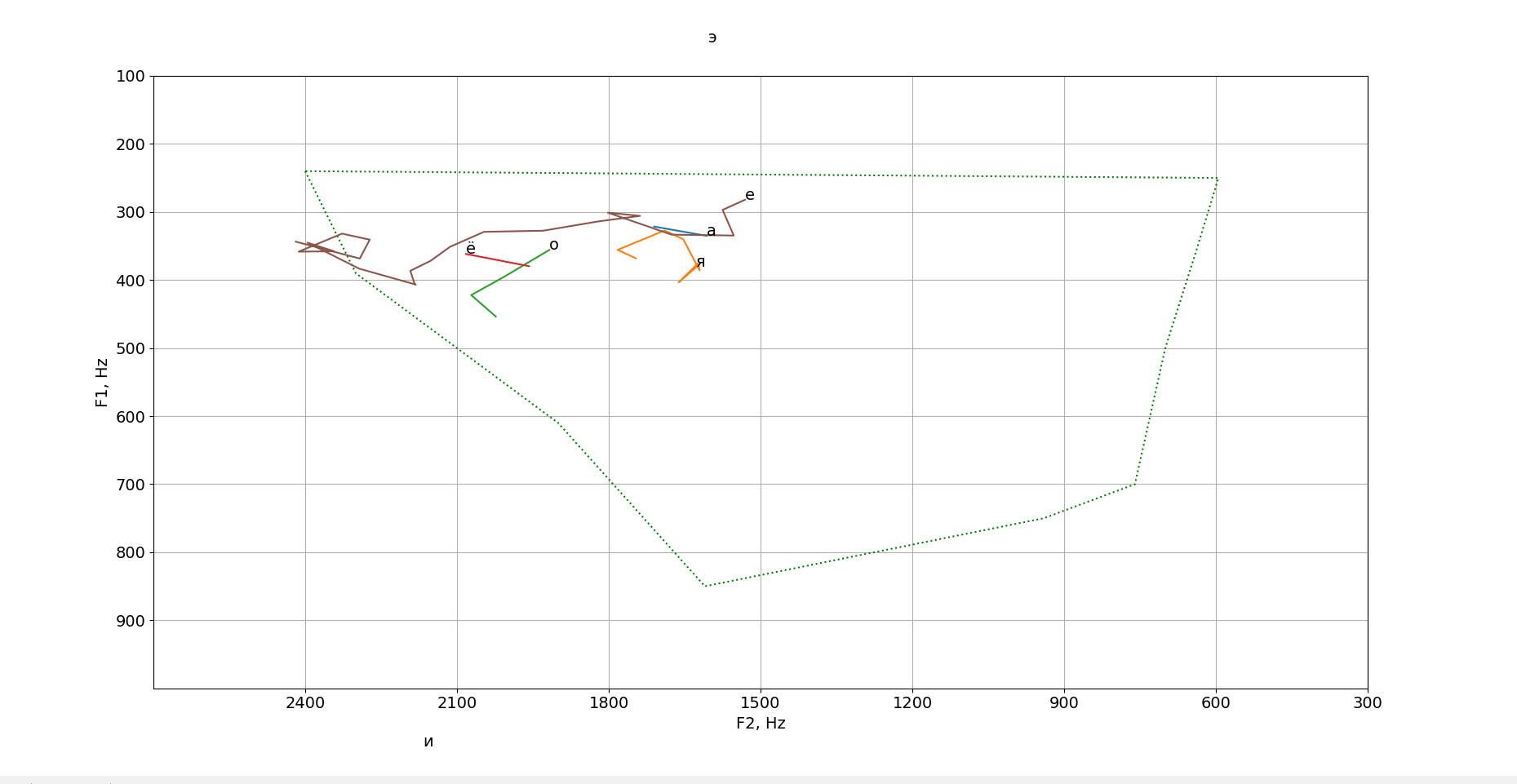
Для проверки я взял Алису и Яндекс переводчик. Обе программы хорошо распознают русский и английский, посторонние шумы в виде побочного диалога - не мешают. При если произносить слова не согласованные между собой, то обе программы распознают также с нарушенным распознаванием. При несвязном наборе речи(использовался текст песни "Продается х5 93 года"): Яндекс транслейтер смог уловить почти все слова и смысловые связки верно, в то время как Алиса не смогла их даже обработать.



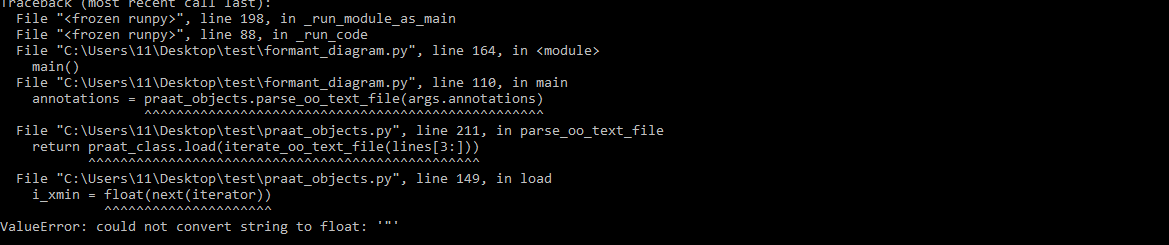
**стр 100**

Лабораторная работа:

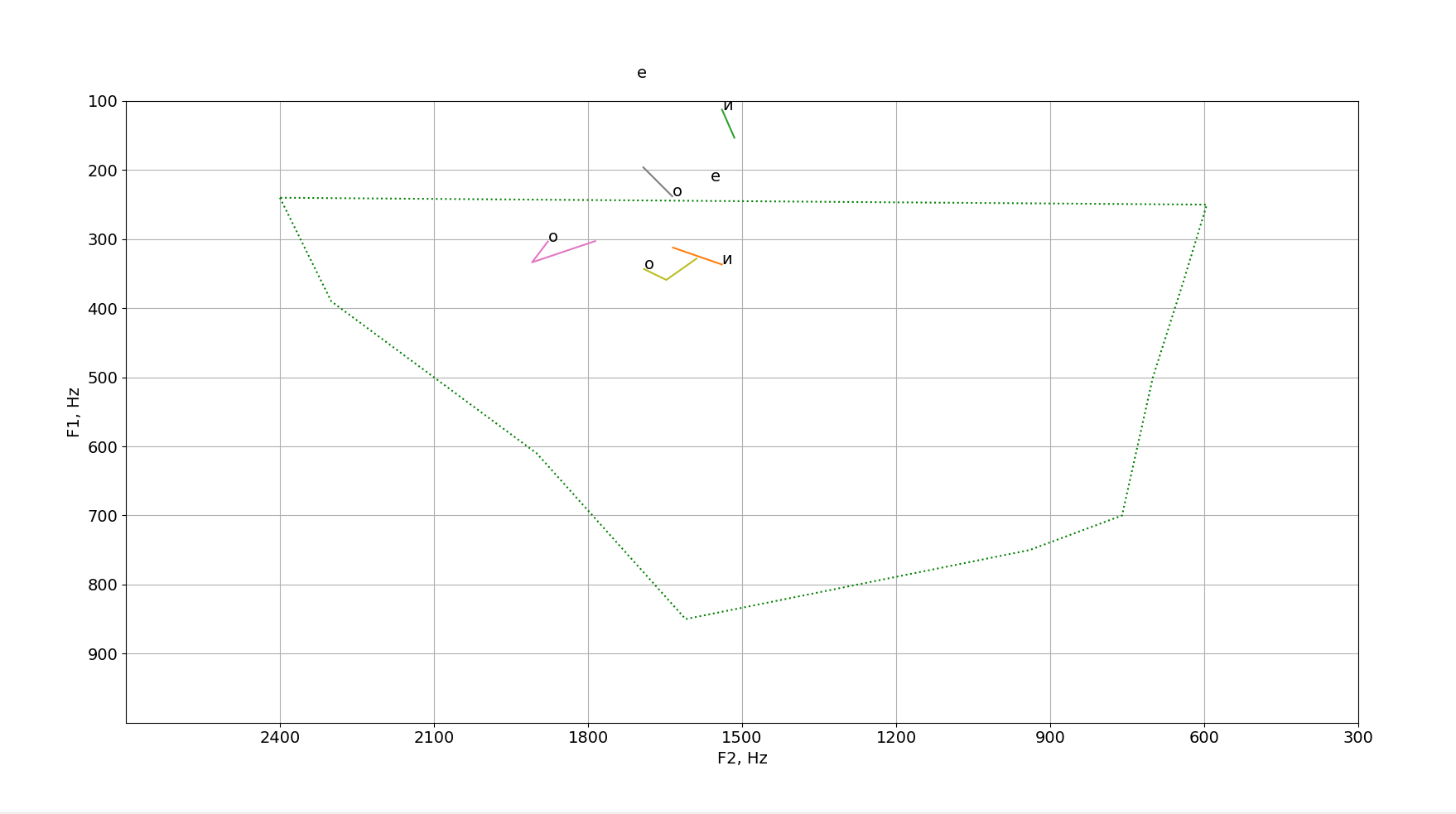




Проблема с тем, что если у нас будут “промежутки” между гласными, то код падает с ValueError. Я нашёл два костыля: Либо заполнять промежутки чем-то, что сказывается на точности, либо добавить в код условие, что если строка пустая - пропускать её.



результат вышел отличный от представленного в методичке, но всё же можно проследить связь, что гласные в положении между твёрдыми согласными не сильно отличаются от гласных между мягкими. Сопоставить с диаграммой МФА возможно.

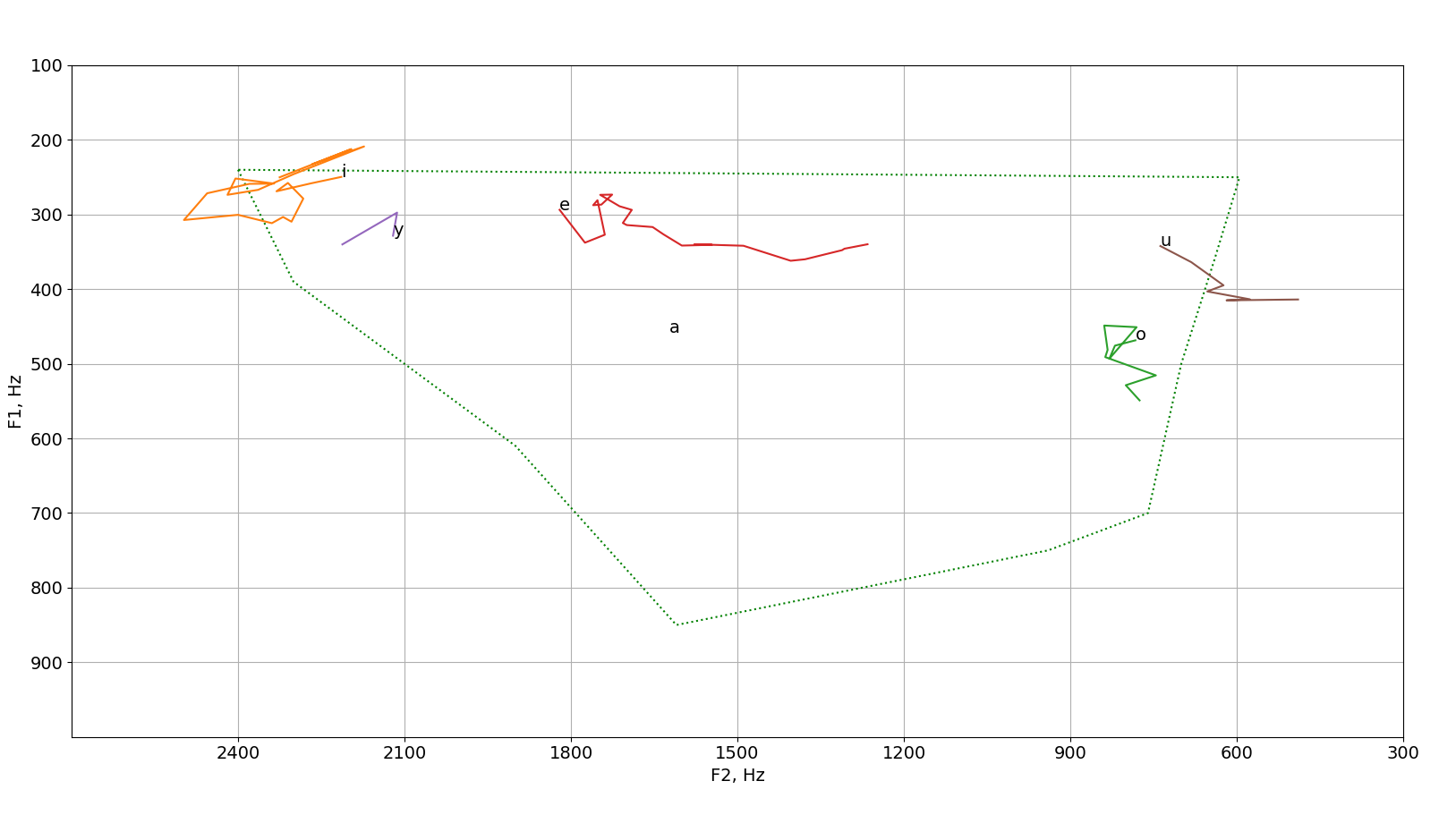


В ударной позиции гласные произносятся четко и ясно, сохраняя основное звучание.

В предударной позиции гласные сохраняют основное звучание, но произносятся с меньшей силой и длительностью.

В «прочей» позиции гласные подвергаются качественной редукции.

английские слова:



**b**

**er**

**b**

**er**

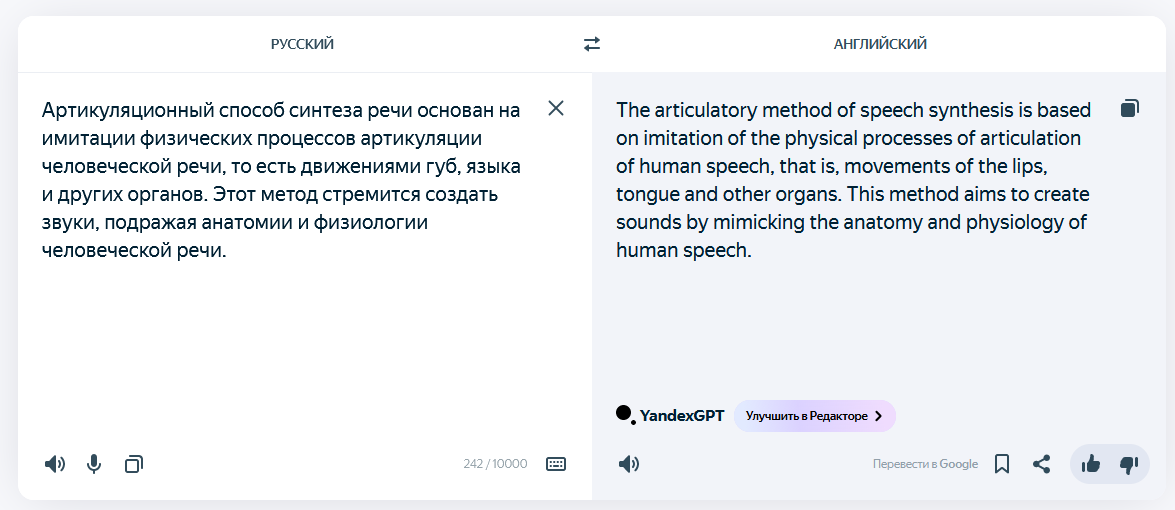
**стр 117-118**

Вопросы:

1. Системы Text-to-Speech (TTS) преобразуют письменный текст в синтезированную речь.Основной задачей систем TTS является генерация естественно звучащей синтезированной речи из письменного текста. Это включает в спользуя техники синтеза речи, такие как конкатенативный синтез или синтез на основе формирования волны.
2. Основные критерии для оценки качества синтеза речи: разборчивость, т.е. возможность человеком декодировать речевой сигнал, и естественность - автоматически созданное высказывание не должно значительно отличаться от речи человека.
3. В основе компилятивного синтеза речи заложены заранее записанные сегментные единицы (от звуков речи до целых высказываний), т.е. на вход системы подаются подготовленные человеком образцы речевых сигналов, которые система объединяет и озвучивает в требуемой последовательности.
4. Два способа синтеза речи, основанные на акустических свойствах речи, это синтез на основе формантов и синтез на основе скрытых марковских моделей (HMM). Основное отличие между ними заключается в подходе к моделированию и генерации звуков.
5. Артикуляционный способ синтеза речи основан на имитации физических процессов артикуляции человеческой речи, то есть движениями губ, языка и других органов. Этот метод стремится создать звуки, подражая анатомии и физиологии человеческой речи.
6. Основные сложности для синтеза речи каждым способом включают в себя естественность звучания, высокую степень детализации, управление просодией (интонация, темп, паузы) и др. Пути их решения включают в себя использование глубоких нейронных сетей, улучшение баз данных для синтеза, разработку более точных моделей артикуляции и другие технические усовершенствования.

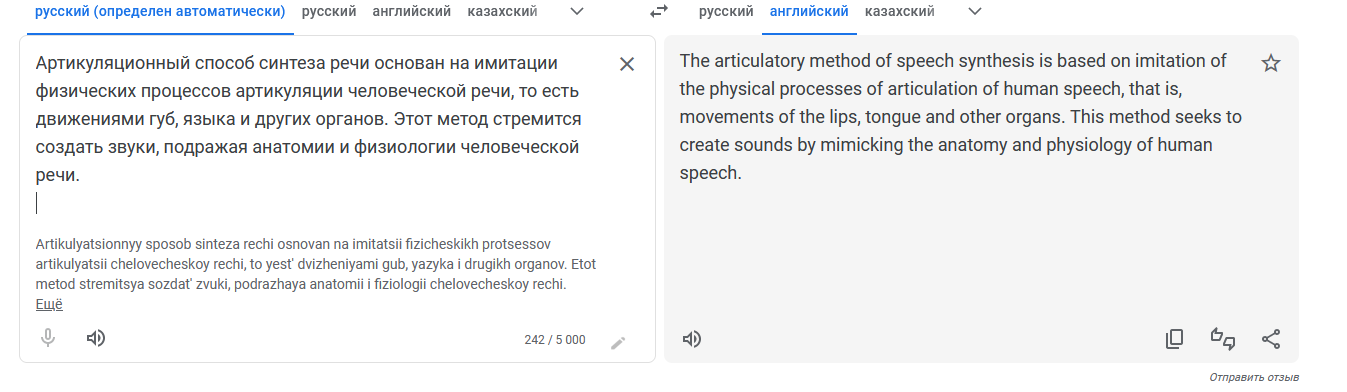
Лабораторная работа: oddcast - так себя показал, ударения поставлены плохо.

яндекс переводчик:



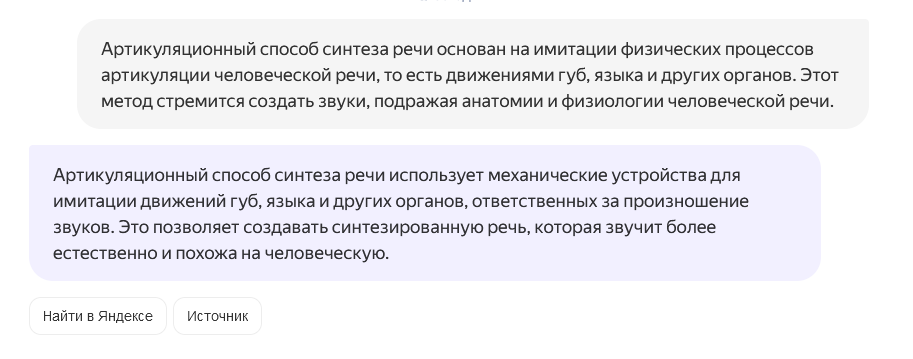
нормально показывает себя

гугл переводчик:



тоже нормально

Яндекс Алиса:



хорошо работает