MySungRussian Final Report

Programmer: He (Stephen Dai) 998177565

Cheng Liu 997563890

Specialist: Dann Mitton

Date: 7 April 2016

Word count: 1446

Introduction

Russian songs are written out using the Cyrillic alphabet. This poses a barrier for singers hoping to explore this rich musical repertoire, because most North American singers can't read the Cyrillic characters in which the Russian language is written. Today's singers do receive extensive training in the International Phonetic Alphabet (IPA). MySungRussian instantly converts Cyrillic text into familiar IPA symbols and optimizes them for classical singing using the rules of Russian lyric diction. These conversions, or *transcriptions*, can enable singers to engage immediately in Russian vocal repertoire without needing to read or understand Cyrillic.

In addition to the core Transcribe function, MySungRussian will include modules on instruction in the rules of Russian lyric diction, and a pronunciation coach comparing the singer's attempts with model audio examples. MySungRussian will suggest strategies to fine-tune nuanced pronunciation for singing.

Overall Design

The Specialist created the User Interface mockup. The team identified three main features: "Saved", "Transcribe" and "Learn". For primary navigation we used the Android Toolbar instead of the dropdown menu from the original design.

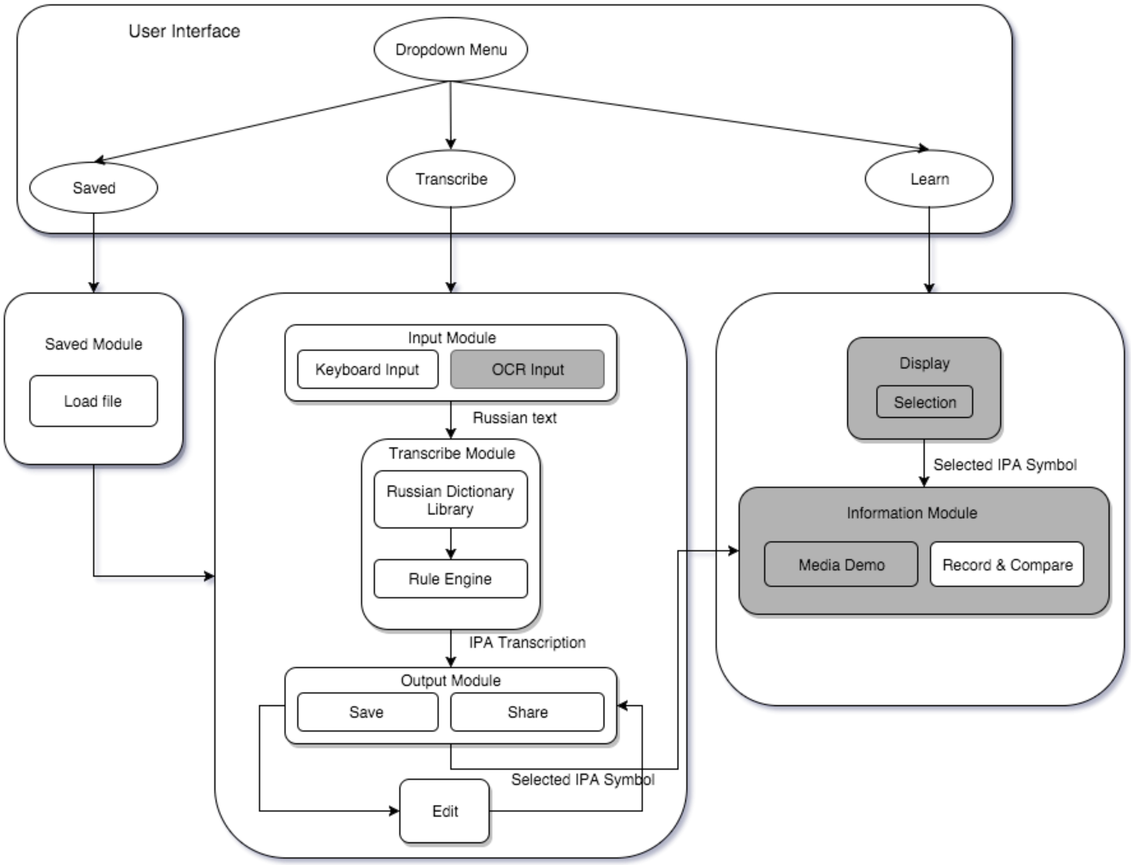


Figure 1: Block diagram of MySungRussian. Greyed-out modules were not implemented in Spiral 6.

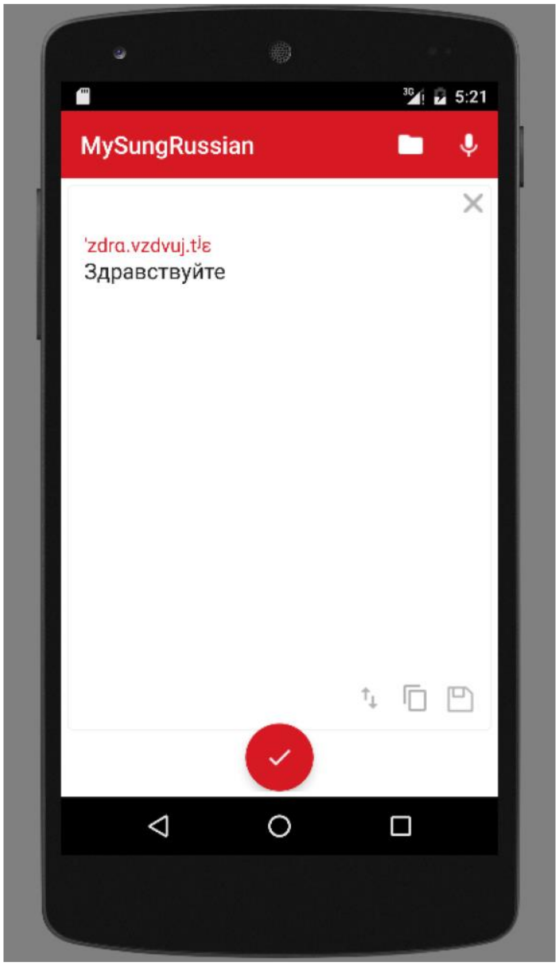
The core of MySungRussian is the "Transcribe" feature. The singer inputs Cyrillic text by directly typing or pasting it from the clipboard. The text is passed into the Rules Engine, which converts the Cyrillic into the symbols of the IPA and applies linguistic rules to simplify the output for optimal singing. Assigning the right syllable stress is crucial in Russian. Unfortunately, no rule governs syllable stress, so MySungRussian curates an online Russian Dictionary to determine the correct stressed syllable for each word. The IPA transcription is finally displayed, and the user can copy or edit the output as desired.

The "Saved" feature allows the user to store input and output on the phone. Users can browse previous transcriptions and edit transcriptions over time.

The “Learn” feature is designed help users familiarize themselves with pronouncing Russian text using the rules of Russian lyric diction. Accurate pronunciation of some IPA symbols can be a challenge. We explored the possibility analyzing the user’s pronunciation efforts and providing constructive feedback such as “try lowering your tongue”, or “rounding your lips more will help”. This feature would need to be able to perform digital signal processing on a capture of the user’s voice. Our app currently records and displays a 2-second spectrogram recording. We did not achieve any qualitative advice capability yet.

Functionality

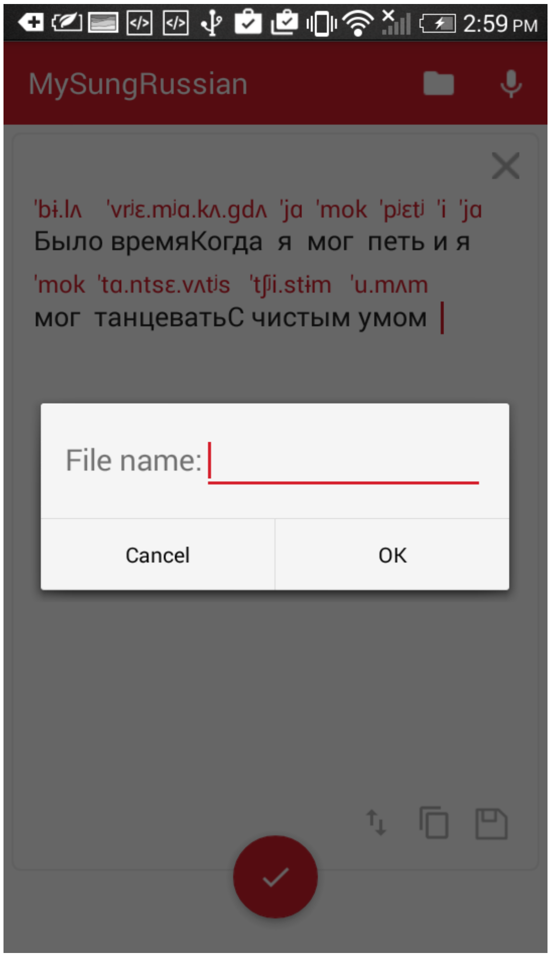
There are three main functions: Transcribe, Saved and Learn. The Transcribe is the main function placed on the initial page of the app. (Figure 2)



'zdrɑ.vzdvuj.tʲɛ
Здравствуйте

Figure 2: Transcribe function, MySungRussian's opening screen

The white space takes user inputs of Russian words and displays corresponding IPA symbols on top of the words in red. The interpretation is processed by a special engine we designed called Rule Engine. Currently, the Rule Engine is capable of handling basic cases. However, there remain many rules and corner cases to be implemented. On the bottom right corner, there are three assisting functions. The first one on the left is used to edit the IPA symbols manually. The Russian IPA symbols are not unique under some situations and thus we need this IPA editing function. The second one is Copy/Paste. The last one is Save. The app pops up a window ask user for file name on save button clicked. (Figure 3)



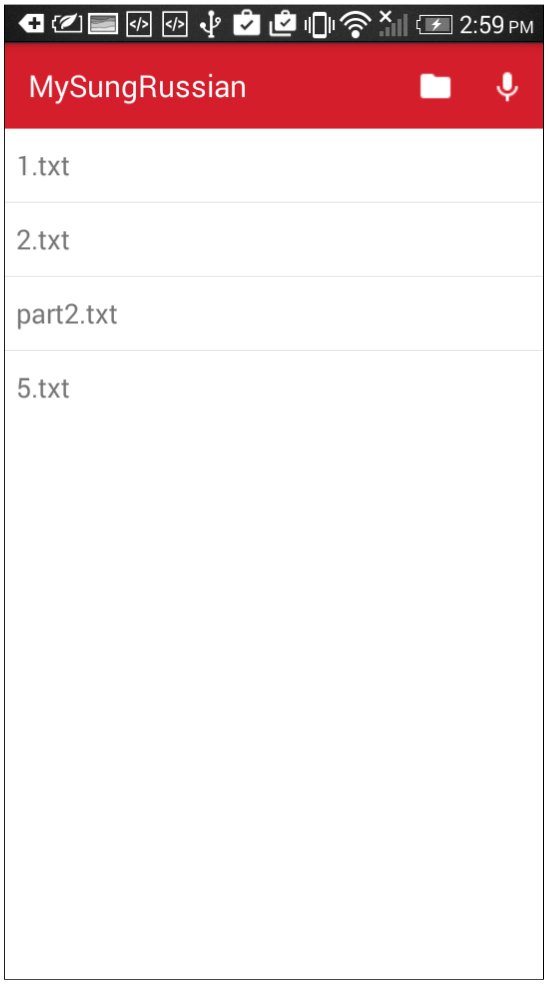
ˈbɨ.lʌ ˈvrʲɛ.mʲɑ.kʌ.gdʌ ˈjɑ ˈmok ˈpʲɛtʲ ˈi ˈjɑ
Было времяКогда я мог петь и я
ˈmok ˈtɑ.ntsɛ.vʌtʲs ˈtʃʲi.stɨm ˈu.mʌm
мог танцеватьС чистым умом

File name:

Figure 3: Promote user input for file name

The file name should not be empty; otherwise file will not be saved. On the top right corner the X denotes a button that clears the text field.

The Saved function lists the entire user-saved files. (Figure 4) The list items are clickable for user to review the content. The content are displayed in the transcribe layout so that user can edit. Listed files can be deleted by long clicking. The app shall pop up a delete confirmation window on long click. (Figure 5)



MySungRussian

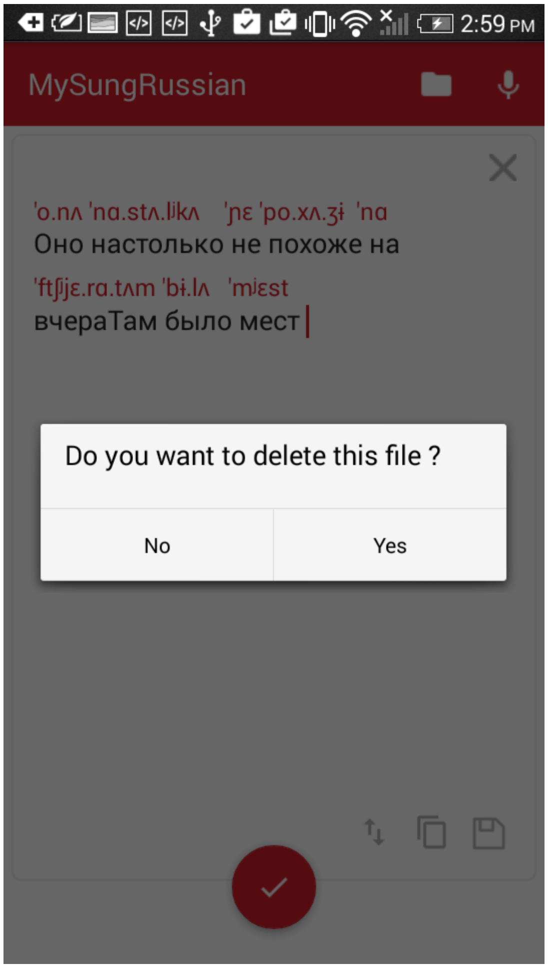
1.txt

2.txt

part2.txt

5.txt

Figure 4: Saved file list



ˈo.nʌ ˈnɑ.stʌ.lʲkʌ ˈɲɛ ˈpo.xʌ.ʒɨ ˈnɑ
Оно настолько не похоже на

ˈftʃʲjɛ.rɑ.tʌm ˈbɨ.lʌ ˈmʲɛst
вчераТам было мест

Do you want to delete this file ?

Figure 5: Confirm delete file

The Learn function is designed to encourage users to learn the accurate pronunciation by comparing the model audio to their own efforts. The app shall give advice based on frequency domain analysis. As for now, the app is able to record the voice and display its spectrogram using Fast Fourier Transform (Figure 6).

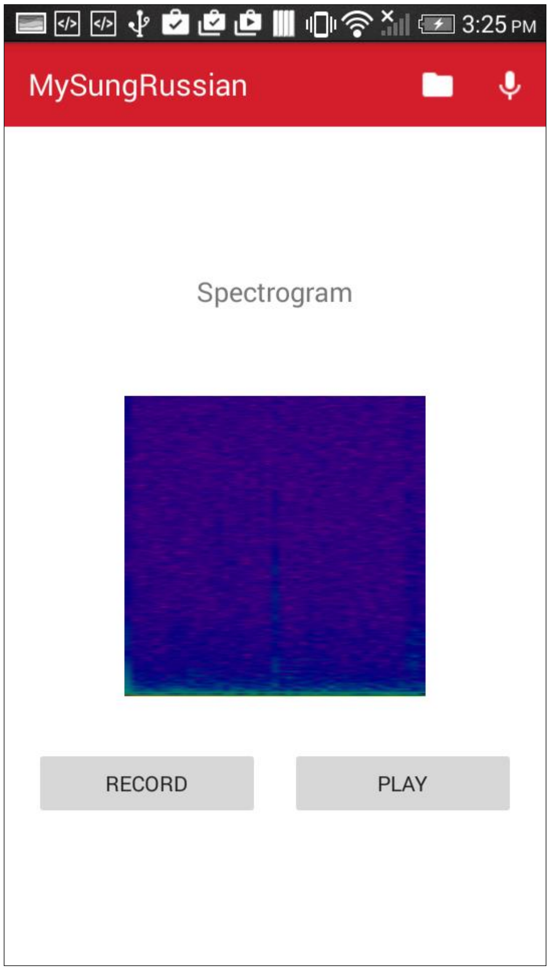


Figure 6: Learn layout

The spectrogram shows the energy intensity at each frequency (y-axis) at a give time (x-axis) in different colour, red being strong while blue being weak. The team has encountered problems trying to achieve the goal of providing constructive feedback. We need to do more research to implement an algorithm for analyzing frequencies of different IPA pronunciations. Therefore, this function is incomplete.

Key Learning

We have all learned and gained value throughout this project. Our experiences can be categorized into two broad categories: Communication and Technical Research. For communication, we started well and held many meetings (both online and fact-to- face) to share our thoughts. However, we could have improved our efficiency by summarize our conversations by taking minutes at our meetings.

We could have been more effective with our Technical Research. The spectrogram function is new to every team member. We underestimated this challenge. If we were to start the project over, we would begin research on this topic sooner, and solicit help from experienced specialists.

Contributions

Dann is the Specialist in our team. He conceived the app and gave direction at the project outset. He designed a prototype UI. Dann gave instruction on transcribing Russian words to IPA symbols, including in-person explanations and summarizing his research in a document that formed the basis of our Rule Engine. He also edited all of our text for assignment submissions.

Cheng was responsible for building both the UI and conceiving the sound analysis module. She set up and refined our UI. Cheng learned basic transcription rules, and wrote a helper class for determining syllable stress. Near the end of our project, Cheng focused on voice recording and researching spectrogram analysis.

He (Stephen) created the Rule Engine and completed functions like Save and file management. Stephen learned the Russian IPA rules from Dann's document and organized these rules into programmer view. He implemented these rules in the app later on.

Specialist Statement

Specialist Dann Mitton is a professional opera singer who will emerge from his Doctor of Musical Arts (DMA) in Performance degree with a specialization in voice pedagogy and a minor in Russian. Dann's dissertation thesis is an inquiry into the possibility that singing and performing in Russian can catalyze the technical development of low male voice singers.

While similar apps already provide reliable transcriptions of text optimized for classical singers exist for French, German and Italian text (six of these were reviewed in Assignment S1, Part II), MySungRussian is currently the sole app in the world to do so for Russian. No one else has taken on the sophisticated rules governing Russian lyric diction in an app format. Creating this app is a modern, important step in Dann's emergence as a contemporary expert in the field of Russian Lyric Diction for classical singers. There is a demand for this tool evidenced by the enthusiasm and anticipation arising from each discussion about the app with voice teachers, singers and coaches. Its completion and dissemination will effectively extend the body of performance-based singers' resources in Russian.

Future Work

We need to continue to improve our Rule Engine's sophistication and reliability. At the moment it is incomplete. There are many more rules governing contextual nuances and special cases in the linguistic literature that have not been encoded in the current Rule Engine. These are essential for a reliable, authoritative transcription tool. As the level of complexity increases, the Rule Engine is more likely to make mistakes or face conflicts. The algorithm needs to be designed thoughtfully for this expanding scalability, and tested against many different cases for robustness.

MySungRussian should be equipped with more input/output and file sharing methods. We originally proposed to capture Cyrillic text from photos using OCR technique, and allowing users to share their transcription results via email or uploading to the cloud.

Finally, we barely addressed the "Learn" feature. Only a skilled expert can infer how sounds are being pronounced from reading the spectrogram as it stands. More research needs to be done to design an algorithm for handling such a complex task. We had further hoped to integrate a means of advising the user on how to perfect their pronunciation. This kind of analysis paired with constructive feedback currently lies beyond our reach.

Source code link: https://github.com/StephenWo/MySungRussian