DHIVEHI AUTOMATIC SPEECH RECOGNITION SYSTEM

CPT311 Applied IT Project

Date of Submission: 18th November 2018

Supervisor: Adam Khalid

Ibrahim Hassaan - S047603
Mohamed Ifham - S048740
Adam Raaif Nasheed - S054659
Yameen Mohamed - S055680
Abstract

This report details the work done to create a speech recognition solution for Dhivehi language. The system was developed using CMUSphinx speech recognition toolkit, which requires the development of a text corpus to use as output, a phonetic dictionary (list of phonemes), a language model (probabilistic representation of word occurrences in language) and an acoustic model (mapping voice features to text). The latter two can be trained using provided audio and text data.

The development of our ASR system was carried out in two phases. The first phase dealt only with numbers (covering real numbers from 0 (inclusive), up to but not including 1 trillion). The second phase dealt with the entire Dhivehi language (with exceptions: it does not support thikijehi thaana and can only pick up the common Malé dialect).

The system developed during the first phase had an accuracy rate of 75% (which barely passed our set minimum acceptable rate), while the system developed during the second phase had an accuracy rate of 42.5% (which failed our set minimum acceptable rate).
Acknowledgements

We would like to express our sincere gratitude to the lecturers at the Faculty of Engineering, Science and Technology, without whose support this work would not have been possible. We are especially thankful to our supervisor Mr. Adam Khalid, for his encouragement and guidance throughout the completion of this project. We would also like to thank our advisor Mr. Basim Abdulla for being the inspiration behind us starting this project, Ms. Mariyam Nadhira for her assistance in providing rooms to set up our recording equipment, and Ms. Asma Rushdy and her students for spending their valuable time to contribute to our voice data collection.

We would also like to extend our appreciation to the lecturers at the Faculty of Arts: namely Ms. Azeela Afeef for providing us her Dhivehi language corpus for use in our project and giving helpful information on Dhivehi linguistics, Mr. Mohamed Riyaz for his expertise on Dhivehi phonetics, and Messrs. Anwar Ibrahim and Abdul Rasheed Ali for their helpful advice and inspiration.

We wish to acknowledge the assistance given by the Academy of Dhivehi Language and its president Mr. Ashraf Ali, for providing valuable time, information and resources to help us achieve our goals.

We would also like to thank the staff of Public Service Media, for providing audio and text data with which to train and test the system.

Finally, we wish to offer our thanks to all our family and friends, for their invaluable support and confidence in us, and their time spent to help make this concept a reality.
# Table of contents

Abstract .................................................................................................................................................. i

Acknowledgements .................................................................................................................................. ii

List of figures .......................................................................................................................................... vi

1. Context and preliminary investigation ......................................................................................... 1

1.1. Introduction ..................................................................................................................................... 1

1.2. Reasons for such a system ............................................................................................................. 1

1.3. Literature review ........................................................................................................................... 2

1.3.1. Current work done .................................................................................................................... 2

1.3.1.1. CMUSphinx requirements ........................................................................................................ 3

1.3.2. Acoustic feature extraction ....................................................................................................... 4

1.3.2.1. The mel-scale .......................................................................................................................... 5

1.3.2.2. Filterbanks ................................................................................................................................ 6

1.3.2.3. Cepstra .................................................................................................................................... 7

1.3.2.4. MFCC calculation .................................................................................................................... 7

1.3.3. Language model .......................................................................................................................... 8

1.3.4. Acoustic model ........................................................................................................................... 9

1.3.4.1. Dhivehi phonetics .................................................................................................................... 10

1.4. General methodology .................................................................................................................... 11

1.4.1. The ASR system ...................................................................................................................... 11

1.4.2. Recording .................................................................................................................................. 13

1.4.3. Training ..................................................................................................................................... 13

1.5. Tools used ....................................................................................................................................... 14

1.5.1. Speech recognition ................................................................................................................... 14

1.5.2. API, website ............................................................................................................................. 14

1.5.3. Telegram bot ............................................................................................................................ 14

1.5.4. Smartphone app ....................................................................................................................... 14
Dhivehi Automatic Speech Recognition System

1.5.5. Miscellaneous ................................................................. 14

2. Phase One: Number Recognition ............................................ 15

2.1. Methodology ........................................................................ 15
   2.1.1. Recording process ...................................................... 15
   2.1.2. Review process ......................................................... 16

2.2. Testing and evaluation ............................................................ 17
   2.2.1. Choosing test phrases ................................................ 17
   2.2.2. Test success criteria and results .................................... 17

3. Phase Two: General Speech Recognition ..................................... 18

3.1. Methodology ........................................................................ 18
   3.1.1. Text data collection ..................................................... 18
   3.1.2. Crowdsourcing mechanisms ......................................... 20
      3.1.2.1. API for collecting and reviewing voice data ............ 20
      3.1.2.2. API for demo ............................................................ 21
      3.1.2.3. Website for collecting and reviewing voice data ...... 23
      3.1.2.4. Website 1 for demo (Funa) ..................................... 25
      3.1.2.5. Website 2 for demo (Koi) ....................................... 27
      3.1.2.6. Smartphone app for collecting voice data (Kanvaaru) 29
      3.1.2.7. Telegram bot for collecting voice data .................... 30
      3.1.2.8. Telegram bot for demo ........................................... 33
   3.1.3. Recording process ....................................................... 34
   3.1.4. Review process ............................................................ 36

3.2. Analysis ............................................................................... 37
   3.2.1. Noises and noise removal ........................................... 37
   3.2.2. Phonetics and linguistics ............................................. 39

3.3. Testing and evaluation ........................................................... 40
   3.3.1. Effect of language model ............................................. 41
Dhivehi Automatic Speech Recognition System

3.3.2. Effect of voice pitch.................................................................42
3.3.3. Effect of data.............................................................................42

4. Critical evaluation..............................................................................43
   4.1. Success of project........................................................................43
   4.2. Suggestions for improvement if repeated, and future enhancements.........44
   4.3. Experiences gained......................................................................45

References.................................................................................................46

Appendices.................................................................................................47
   Appendix A: Phrase list used for recording voices for Phase One (number) model ........47
   Appendix B: Phrase list A used for recording voices for Phase Two (main) model 52
   Appendix C: Phrase list B used for recording voices for Phase Two (main) model........56
   Appendix D: List of countries in Dhivehi................................................65
   Appendix E: List of team meetings..........................................................74
   Appendix F: Recording sessions timetable for Phase Two..........................77
List of figures

Figure 1: The relationship between the hertz- and mel-scales .................................................. 5
Figure 2: Mel filterbanks .................................................................................................................. 6
Figure 3: Process of acoustic modeling ............................................................................................ 9
Figure 4: Methodology used to create Dhivehi ASR system ............................................................ 12
Figure 5: Example list of phrases for number model ........................................................................ 15
Figure 6: Accuracy of Phase One model .......................................................................................... 17
Figure 7: API help endpoint ............................................................................................................. 21
Figure 8: API response if wrong file type is supplied ..................................................................... 21
Figure 9: Successful conversion and output ...................................................................................... 22
Figure 10: Website recording page .................................................................................................. 23
Figure 11: Website main review page, showing list of recordings ................................................... 24
Figure 12: Listening to a recording on the website .......................................................................... 24
Figure 13: Decision screen of the website ....................................................................................... 24
Figure 14: Web UI for Funa .............................................................................................................. 25
Figure 15: Web UI for Koi ................................................................................................................. 28
Figure 16: Smartphone app screens ................................................................................................ 29
Figure 17: Telegram bot gender selector .......................................................................................... 30
Figure 18: Gender confirmation message ......................................................................................... 30
Figure 19: Welcome back message .................................................................................................. 31
Figure 20: Dictation prompt ............................................................................................................ 31
Figure 21: Successful submission .................................................................................................... 32
Figure 22: If the user sends a picture in reply instead ....................................................................... 32
Figure 23: User statistics .................................................................................................................. 33
Dhivehi Automatic Speech Recognition System

Figure 24: Telegram bot for demo .................................................................33
Figure 25: Example list of phrases for main model ........................................34
Figure 26: Example waveform showing noise ...............................................37
Figure 27: Example waveform without noise ...............................................38
Figure 28: Sudden noise removal .................................................................38
Figure 29: Audio data used in training the base model ..................................40
Figure 30: Audio data used in testing the hypothesis ......................................40
Figure 31: Test results ..................................................................................40
Figure 32: WER comparison results ............................................................41
Figure 33: Effect of using more data to train the model ..................................42
1. Context and preliminary investigation

1.1. Introduction

Dhivehi, the official language of the Maldives, has little been explored by the outside world until the late 20th century. Since 1980s with its introduction by Mahal Unit Press, Dhivehi typography has moved into the tech world with abundant fonts. However, little has been explored in Dhivehi language in other technological aspects such as natural language processing, automatic speech recognition, etc.

Hence, to evolve Dhivehi into the new technological era, Dhivehi has to be recognized by computer systems when spoken. Thus, an Automatic Speech Recognition (ASR) system for Dhivehi language is crucial.

1.2. Reasons for such a system

- Technologically advanced facilities such as ASR are lacking for Dhivehi. There have been attempts to address the shortcoming, with little to show for it.
- Such a system can improve access to technology, especially for people with limited or no knowledge of English, the de facto lingua franca of the modern age.
- Such a system can improve accessibility for people with special needs.
- Such a system can be expanded upon to allow further conveniences (eg: domotic appliance control).
- Such a system can contribute to the betterment of Dhivehi language. It can help create a market for Dhivehi based gadgets.
1.3. Literature review

1.3.1. Current work done

(by Yameen Mohamed)

Research into speech recognition has been going on for several decades, with several different approaches to the problem over the years. However, there have been no developments with regards to Dhivehi language in particular.

The following ASR software was researched to find the most ideal tool for the job of Dhivehi ASR:

- CMUSphinx
- Kaldi
- IBM Watson
- Google Speech API

CMUSphinx was chosen as it was free, open-source, customizable to suit the needs of Dhivehi ASR, and could output results in real-time. Kaldi could have better performance but does not output in real-time. Both Google’s and IBM’s solutions, despite having better accuracy, were not free for use.
1.3.1.1. **CMUSphinx requirements**

CMUSphinx is a speaker-independent, continuous recognition system which is designed to handle large vocabularies. It consists of hidden Markov acoustic models and an n-gram statistical language model which recognize speech using mel-frequency cepstral coefficients (MFCCs) extracted from split audio streams or .wav files (Kai-Fu, Hsiao-Wuen, & Reddy, 1990). Creating an ASR system using CMUSphinx requires the following components:

- **Text corpus:** One of the major reasons that Dhivehi currently lags behind in technological advancements is the lack of a comprehensive text corpus. There have been talks of work being done by the Academy of Dhivehi Language to create a corpus (SunOnline, 2018), but so far only small-scale corpora exist for personal projects ([Sofwath], 2018). Projects such as ASR require a much larger corpus, however.

- **Phonetic dictionary:** There exists a list of around 32000 Dhivehi words online ([kudanai], 2012), which could be converted to their phonemic representation. This list contains just root words, however (e.g. the list contains the word *aharen* but not the related word *aharenge*).

- **Language model:** Language models are derived from corpora, but as no suitable corpus for Dhivehi exists, no language model for Dhivehi was found.

- **Acoustic model:** Unlike other common languages, Dhivehi language does not have many audio data sources publicly available that could be used to generate a good acoustic model, nor was it feasible to adapt an existing acoustic model from another language.
1.3.2. **Acoustic feature extraction**  
(by Adam Raaf Nasheed)

One of the first steps in recognizing speech is to clean up the input audio of any component that is unnecessary (i.e., components that do not have any linguistic property, such as background noise). This process is known as feature extraction.

The general idea behind feature extraction is to generate a digital representation that closely mimics human speech and auditory systems, which can then be used in statistical analysis. There have been several methods or processes developed that help solve this problem. One of the most commonly used methods with current ASR systems is mel-frequency cepstral coefficients.
1.3.2.1. The mel-scale

The mel-scale (named after the word “melody”) is a scale of measurement of frequency or pitch. It is based upon the perception of sound. The perceived distance between equidistant frequencies on the mel-scale are equal, unlike the hertz-scale. This scale is useful because humans do not perceive pitch linearly: the difference between two low-frequency hertz values is more apparent to a listener than two similarly spaced high-frequency values.

As the mel-scale is based upon perception and is not a standardized unit, there is no universally agreed upon formula for conversion between the mel-scale and the hertz-scale (the standard unit of measurement of frequency). An approximated formula for the conversion of hertz to mel is

\[ m = 2595 \log_{10} \left( 1 + \frac{h}{700} \right) \]

where \( m \) is the frequency in mels, and \( h \) is the frequency in hertz (O'Shaughnessy, 1987; Xu, et al., 2004). This logarithmic relationship can be seen in Figure 1.

![Figure 1: The relationship between the hertz- and mel-scales](image)
1.3.2.2. Filterbanks

Another feature of the human ear (or more specifically, the cochlea) is how it acts as a filter for certain frequencies, as it concentrates on those rather than the entire spectrum. These filters are spaced further apart the higher the frequency is (Lopez-Poveda & Meddis, 2001). This feature can also be mimicked by the use of filterbanks based on the mel-scale.

![Figure 2: Mel filterbanks](image)
1.3.2.3. **Cepstra**

A cepstrum is the result of applying the following formula to a spectrum:

$$\mathcal{F}^{-1}(\log(|\mathcal{F}(f(t))|^{2}))$$

where \( f(t) \) is the spectrum based on time \( t \), \( \mathcal{F}(x) \) is the discrete Fourier transform (DFT) of \( x \), and \( \mathcal{F}^{-1}(x) \) is the inverse discrete Fourier transform of \( x \). The name “cepstrum” is a play on the word “spectrum” where the first few characters are written in reverse order.

1.3.2.4. **MFCC calculation**

Mel-frequency cepstral coefficients (MFCCs) are coefficients that make up a mel-frequency cepstrum (MFC). These coefficients can be calculated by applying the following steps (Xu, et al., 2004):

- Divide the input signal into short frames or windows (about 20ms each).
- Take DFT of each window to get power spectra.
- Use the mel filterbanks on power spectra to discard noises that typically won’t be heard by the human ear, and get sum of amplitudes for each filterbank region to produce a mel spectrum.
- Take logarithm of the mel spectrum, then get discrete cosine transform of the log values. MFCCs are the amplitudes of the resulting spectrum.

Variances in some of the steps (e.g. length of each frame, size or spacing between filterbanks, number of filterbanks used) can yield different results (Fang, Guoliang, & Zhanjiang, 2001). These coefficients are a good feature for speech recognition systems, for use in further analysis or interpretation.
1.3.3. Language model
(by Mohamed Ifham)

Language model is an art which helps to operate in major areas of interaction between human and machine to deliver good performance, such as spelling corrections, handwriting recognition, machine translation, optical character recognition (OCR), speech recognition, etc (Goodman, 2001). The purpose of a language model is to provide probable sequence of words, which is derived from corpora.

In automatic speech recognition (ASR), acoustic model alone cannot deliver expected output, without language model. There are several kinds of issues which need to be solved, such as word ambiguity and word boundaries. Especially in continuous speech, language model is needed to minimize the search area and get better, more accurate results. Therefore, it is an important component of ASR. There are different types of language models. However, each of these models have its own pros and cons. Some of these models perform better alone, while others work better when combined. The most popular two models are statistical language models and neural language model.

In statistical language model, it is not only dealing with allocation of possible sentences in language, it can do much more. It assigns probable arrangement of all words and predict next word in line (Goldberg, 2017). Neural language models (NLM) has grown its popularity due to its outstanding performance. NLM deliver better solution for challengers that identify traditional language models when it deals with larger model (Goldberg, 2017).
1.3.4. Acoustic model
(by Ibrahim Hassaan)

Acoustic model is a statistical representation of the relationship between feature vector of audio signals and the linguistic properties or phonemes of speech. The model is achieved by training the speech recordings with corresponding transcripts in a specialized software. Some of the most common types of acoustic models are neural networks, hidden Markov model (HMM), maximum entropy models and segmental models (Droppo, 2004). CMUSphinx uses HMM acoustic models.

In a typical HMM acoustic model, system uses a list of phonemes which are trained (eg: A, IH, U, E) and a phonetic dictionary or lexicon which contains a list of words with corresponding phoneme set.

Example phonetic dictionary:

```
EH K EH QQ
EH G AA R AH
```

The decoder of ASR uses feature vector of the speech to find a matching pattern in HMM. When a matching phoneme is found in HMM, the decoder keeps track of all the phonemes until a pause or word break. Then the decoder uses the phoneme set to identify the spoken word in phonetic dictionary.

![Figure 3: Process of acoustic modeling](image)

With a large vocabulary, the accuracy of the word predicted by the acoustic model can be very low as there are more possibilities. Thus, a language model is used to give a sequence of words that are more probable.
1.3.4.1. Dhivehi phonetics

The phonemic inventory of Maldivian consists of 29 consonants and 10 vowels. Like other modern Indo-Aryan languages, the Maldivian phonemic inventory shows an opposition of long and short vowels, of dental and retroflex consonants as well as single and geminate consonants (Maumoon, 2002).

Dhivehi is a very phonetic language. This makes it easy to make acoustic model than with English. However, no there is no proper research on phonetics of Dhivehi language.

Even with high phonetic pattern in Dhivehi language, there are special cases where written Dhivehi and spoken Dhivehi are different. There are three main cases where written Dhivehi and spoken Dhivehi differs (Mohamed, 1958):

1. When specific consonants (thaa, shaviyani) comes together with a sukun (eg: baiy ބަތްް).

2. When a lone noonu comes in a word (eg: an’bu އަނބ ް).

3. When two vowels sound comes together specially in a linguistic pattern called aibaifili (eg: jahaifi ޖަހައިފި).
1.4. General methodology

1.4.1. The ASR system

The Dhivehi ASR System will be able to take input from an audio source (whether spoken directly or through an audio file), detect any Dhivehi speech within that audio, and then output what was spoken in Thaana. The system does this by applying probabilistic analysis based on a large set of text data previously collected (the corpus). Hence, it will not have any intrinsic knowledge of Dhivehi grammar or spelling rules.

It does this by first splitting the input audio into small chunks, filtering out the noise and normalizing it to make it ideal for identification of phonemes (distinct sounds that make up spoken words). The system then listens for cues in the audio (such as audio that matches a known set of phonemes that the system was previously trained on, or silences that help distinguish between words), and then uses statistical analysis to deduce the words (and by extension, the phrase or sentence) being said. Because of the method used, the accuracy of the outputs can improve over time as it receives more voice and dictionary data to train on.

The system will use an iterative model of development, because the model is continually improved based on the data that is gained, and it can be expanded upon later. This report will detail two such phases of development: the first phase will be restricted to just number recognition, and the next phase will deal with the entire language (with some caveats, discussed later in this report).

One of the earliest development steps is to gather text data from multiple sources. This can be from news articles, books, journals, etc. This data is preprocessed and used to develop a text corpus for Dhivehi (a large unstructured set of texts used for linguistic analysis), a phonetic dictionary (a map of vocabulary words to their phonemic equivalents), and a language model.

The aforementioned text data is then split into phrases for use with voice data collection methods. There are three collection methods planned for development: a Telegram bot, a website and a smartphone app. Since voice collection is crowdsourced, collected voice data needs to undergo manual reviewing so that we can ensure that the system is not being fed garbage data.

Training and developing an acoustic model is the next stage. As more and more voice data is gathered, the system can train using that data and continually improve its speech recognition.
The final stage of each iteration of the project is testing the trained models. A section of the voice data can be set apart for this purpose, while the rest is used to train the system. The system can also be tested by speaking to it in real-time. If the accuracy of the recognition by the system is greater than 75%, the project can be concluded as successful.
1.4.2. Recording
Each speaker was provided with the set of 151 phrases. As a general guideline during recording, the speakers were asked to maintain a set of parameters.

- **Sampling rate:** 16 kHz
- **Bitrate:** 16 bits per sample
- **Channel:** Mono

For consistency and ease of training, Audacity was used as the recording tool with a preset setting as shown below. For speakers who opted to record from mobile phones or any other device where controlling the said parameters was not possible, the recordings were later post-processed with Audacity with these controls and parameters in place.

1.4.3. Training
The following files are needed to train the system using CMUSphinx:

- **Fileids file:** A text file where all the recorded .wav file IDs (paths) are listed. Each line contains the names of the recordings (utterance IDs) and the path in a file system.
- **Transcription file:** A text file containing the corresponding text for each audio file and the file name of audio file.
- **Phonetic dictionary file:** A text file containing every word of training data, and its corresponding phonetic transcription.
- **Phoneset file:** A text file containing the list of all unique phones used in the phonetic dictionary file.
- **Filler dictionary file:** A text file containing the filler phones (eg: breaths, “hmm”, etc.)
1.5. Tools used

1.5.1. Speech recognition
- CMUSphinx 4
- PocketSphinx
- Visual Studio (for compiling on Windows)

1.5.2. API, website
- Laravel

1.5.3. Telegram bot
- Netbeans
- Telegram for iOS/Android
- Telegram bot API via telegrambots Java library

1.5.4. Smartphone app
- Ionic
- React Native
- Visual Studio Code

1.5.5. Miscellaneous
- Python for text data cleaning, transcript creation, phonetic dictionary creation
- PHP, jQuery for web recording and review
- Brackets for text editing
- PhpStorm for web development
- Adobe Audition for visualizing formants and other audio features
- Audacity for editing audio files during reviewing process
- ffmpeg for converting audio files to required format
2. Phase One: Number Recognition

2.1. Methodology

2.1.1. Recording process

For the numbers model, 151 phrases were generated such that permutations of each distinctly pronounced numbers (1 to 29) and the place values (30, 40 and so on for tens; 100, 200 for hundreds, etc. up to the billions) were included. Each phrase is stored on a separate text file. The files were named incrementally, with each prefixed with “n Xxx” where x is the chronological number.

Sample phrases:

n_068.txt - އެއްްސަތޭކަްސްދޮޅަސްރަބިރިލިޔަންްތިސްްސަތޭކަްސޯޅަްމިރަޔަންްތަރ ްސަތޭކަްސޯޅަް

n_079.txt - ސެއްްސަތޭކަްފަސްދޮޅަސްްހަްބިރިލިޔަންްރަސްްސަތޭކަްސޯޅަވރަސްްމިރަޔަންްތަރ ްސަތޭކަްސަރ

(See Appendix A for the full list of phrases used.)
2.1.2. Review process

After the recording, the recorded audio files are reviewed. During this process, each audio file was listened by the reviewer for the lapses and noises and overall speech disposition, with respect to the duration between the words and syllables. Any Freudian slips will be addressed as follows:

- **Re-recording:** For significant variances, the speaker is re-invited to record the mismatching phrases. Often the recordings with considerable background noise are more prone to errors. Hence those recordings are replaced with new recordings. Also some of the recordings were partially done where possible. In these cases the tone, tempo, amplitudes, distortions etc. are often difficult to match. In the entire process, the partial adoption was only done and the manual touches were minimal to maintain the organic sound profile for better results.
2.2. Testing and evaluation

2.2.1. Choosing test phrases

Random phrases were chosen from the number speech corpus. The phrases chosen comprises less than 10% of the corpus. Due to the lack of data, similar transcript phrases were used between test subjects.

2.2.2. Test success criteria and results

As numbers corpus is a restricted domain, the success of the model is expected to have more than 70% accuracy.

The model was 75% accurate: of 1244 words, 928 was accurate.

Figure 6: Accuracy of Phase One model
3. Phase Two: General Speech Recognition

3.1. Methodology

3.1.1. Text data collection

Data for this model was accumulated from various online sources including local online news websites, religious transcripts, legislative publications and other miscellaneous sources.

For the preliminary data collection, more than 600Mb of data was collected by web-scraping. However, it hardly had any corresponding audio or video data available online. Neither the government organizations nor anybody known working in the field happened to have any such data that could be used, except for the speeches delivered by the president at various occasions which were made available on the website presidency.gov.mv. Nevertheless, this dataset alone was not sufficient as the model demands for multiple voices to be fed for the same text so that the model can better understand the vocal properties involved in each sound unit in the text. Having those speeches recorded by other people was impossible given the time constraint and the limitation of resources.

Consequently, the only option left was to collect audio data for the textual data payload that was already collected. The large text file was then tokenized into phrases such that one phrase does not have more than 18 words. The number of resultant phrases exceeded a million, and it was surely good enough data to develop the model given that we collect the respective set of audio.

Upon random sample reviewing, it proved to need cleaning. Texts collected from local news (which was the major chunk of the collected large set) had logical and semantical errors. Also, since there were no context-aware filters applied at any phase in the data collection or processing, the phrases contained biased, confidential, sensitive contents as well. On a separate note the texts contained borrowed words, foreign words (primarily English and Arabic) and words that do not exist in the prescribed dictionary which are going to be nothing but noise for our model since the dictionary, phonetic model and language model do not account for anything that does not exist in the vocabulary published by the authorities.

Furthermore, delimiting the sentences to a fixed length also yielded to logically invalid phrases. Hence the collected data further need an intense cleaning process which required a considerable amount of time and human resources.
For the cleaning, a heuristic approach was adopted. The foreign elements were removed and adjusted wherever possible, eg:

- *Khaa* as in *khabaru* -> *Haa* (*habaru*)
- *Qaafu* as in *qaumu* -> *Gaafu* (*gaumu*)

Some of the sounds were included in the phonetic dictionary and also some of the words were added to the dictionary considering their frequency of occurrence in the general language usage, eg: *Allah, assalaam alaikum*, etc.
Crowdsourcing mechanisms

3.1.2.1. API for collecting and reviewing voice data

(by Yameen Mohamed)

Audio recording even by a single person for the kind of huge dataset that resulted was again impossible given the time and resources. Thus a mass contribution mechanism was planned. As such an API-based web application was developed which catered to the following:

1. Contributor
   a. Registration
   b. Authorization
   c. Contribution
   d. Statics

2. Moderator
   a. Authorization
   b. Contribution analysis
   c. Moderation

3. Administrator
   a. User creation
   b. Data import (phrases)
   c. Data revision

Upon creating a phrase, the API does the necessary pre-checks and calculates the statistics for the phrase. The statistics include:

- number of words
- number of syllables
- the set of individual syllables
- a rough estimate of the duration of the audio file
- a recommended tempo rate at which the speaker should speak

Also upon submission of audio file, the API cross-checks the audio file against the metadata calculated for the phrase to automate a limited but valid checks to avoid manual validations. Also a tempo rate is calculated from the audio so that it could be used during training to improve the process.
Dhivehi Automatic Speech Recognition System

3.1.2.2. API for demo
(by Yameen Mohamed)

This API takes voice file as input, and outputs the speech-to-text result in JSON.

The API has two endpoints:

- Help (see Figure 7)
- Audio submission and conversion (see Figure 8 and Figure 9)

![Figure 7: API help endpoint](image)

![Figure 8: API response if wrong file type is supplied](image)
Prerequisites:

The API requires the user to have basic knowledge in executing Python scripts. The API is developed with Python3 with the following dependencies:

- flask (Flask, request)
- werkzeug.utils (secure_filename)
- os, shlex, subprocess, codecs, json

The CMUSphinx toolkit needs to be installed and the environment variables must be set as mentioned in the documentation.

To run the API service execute the following command in the terminal with in the root directory of the project:

```
>> python3 app.py
```
3.1.2.3. **Website for collecting and reviewing voice data**  
(by Ibrahim Hassaan)

Along with the API, a website was made at https://dhivehi.info to appeal to the general public for contributions. The contribution panel and administration panel was also hosted as part of the website.

When a regular user logs in to the web page, Figure 10 shows the first screen they see. The screen will be loaded with a unique phrase for the user. Upon clicking the record button, the syllables get highlighted at a given rate. Upon finishing the dictation, the user has to stop the recording. Then, the user can review the recording and upload the recording if they choose to do so.

![Website recording page](image)

*Figure 10: Website recording page*

Review page consists of 3 major screens:

- Screen listing the phrases pending review (see Figure 11)
- Screen where the reviewer can listen to a selected recording (see Figure 12)
- Screen where the reviewer can accept or reject a recording, and optionally add a remark (see Figure 13)
Dhivehi Automatic Speech Recognition System

Figure 11: Website main review page, showing list of recordings

Figure 12: Listening to a recording on the website

Figure 13: Decision screen of the website
3.1.2.4. Website 1 for demo (Funa)
(by Mohamed Ifham)

![Figure 14: Web UI for Funa](image)

**Prerequisites**
- Java
- Gradle
- Terminal
- Internet connection (for dependencies while building)

**Dependencies**

```groovy
dependencies {
    compile group: 'org.springframework.boot', name: 'spring-boot-starter-websocket', version:'2.0.5.RELEASE'
    compile group: 'org.springframework.boot', name: 'spring-boot-starter-tomcat', version:'2.0.5.RELEASE'
    compile group: 'javax.servlet', name: 'jstl', version:'1.2'
    compile group: 'org.apache.tomcat.embed', name: 'tomcat-
```

Faculty of Engineering, Science and Technology
To build and run the solution, execute the following commands on the terminal with in the project root folder:

```bash
>> ./gradlew build
```
```
>> gradle bootRun
```
3.1.2.5. Website 2 for demo (Koi)
(by Ibrahim Hassaan)

This is a simplified way to run Dhivehi ASR in a webserver.

Requirements

- Python 3+
- Webserver configured to run Python
- Internet connection (for script to download required files)

Setup

1. Extract koi-asr.zip to host folder
2. Run script.bat for DOS systems, script.sh for UNIX systems
3. Allow webserver to run Python scripts

Add “AddHandler cgi-script .py” to httpd config file or to virtual host file:

```<VirtualHost *:80>
   AddHandler cgi-script .py
   DocumentRoot "E:/www/speech"
   ServerName beta.test
</VirtualHost>```

4. If you are running the app on an unsecure host, add the domain to secure zone of the browser. For example, in Chrome:
   a. Go to chrome://flags/
   b. Search "Insecure origins treated as secure"
   c. Add the domain in the text area, enable, and re-launch Chrome
Dhivehi Automatic Speech Recognition System

Usage

Figure 15 shows the web UI of Koi. Press record button and speak. When done, stop recording and submit. The response will be displayed on-screen, as shown in the figure.

Figure 15: Web UI for Koi
3.1.2.6. **Smartphone app for collecting voice data (Kanvaaru)**

*(by Mohamed Ifham)*

A mobile application was developed for both iOS and Android to ease user contribution. The application supported all the aforementioned API features.

The first screen of the app (shown in Figure 16) shows a description of what the app is about, and an optional form for collecting some useful data.

Once the user is registered, they are taken to the recording screen (shown in Figure 16). Here, the contributor can see their ID, and the currently displayed phrase number. Sound control buttons for recording and playback are at the bottom of the screen. There are also two larger buttons in the middle of the screen: one to change the phrase, and the other to upload the recording.

*Figure 16: Smartphone app screens*
3.1.2.7. **Telegram bot for collecting voice data**

*(by Adam Raaif Nasheed)*

A Telegram-based bot was also developed to quickly propagate the outreach. The bot was similar to the personal assistants and it also supported all the features that API provided besides the user experience that comes with Telegram platform.

When a user first contacts the bot (the first message sent to any Telegram bot is always “/start”), they are given the choice of letting the system know their gender (see Figure 17). This data is useful as male and female voices have very different pitches.

Upon tapping one of these buttons, their selection is saved on the server, and the message is edited to show confirmation (see Figure 18).

If the user sends the “/start” command to the bot again after they have selected a gender (i.e. pressed either Male or Female on the selection above), they are greeted normally without the option to re-select gender (see Figure 19).

In either case (i.e. whether the user is newly registered or returning), they are then given the prompt to use the command “/dictate” or “/stats” (see Figure 18 and Figure 19).

![Figure 17: Telegram bot gender selector](image1)

![Figure 18: Gender confirmation message](image2)
When the user sends the “/dictate” command, they are given detailed instructions on how to achieve best results, then given a phrase to dictate (see Figure 20).

Once the user records their voice and sends their dictation, it is stored in the server, and the user is sent a confirmation message and the option to keep going (see Figure 21).

If the user sends an erroneous message in reply to the prompt, they are shown an error message with instructions on how to try again (see Figure 22).
Dhivehi Automatic Speech Recognition System

Figure 21: Successful submission

Figure 22: If the user sends a picture in reply instead
Dhivehi Automatic Speech Recognition System

When the user sends the “/stats” command, they are given a few statistics with regards to their submissions (see Figure 23). They can view their gender selection, and get a sense of their approval ratio here.

![Figure 23: User statistics](image1)

### 3.1.2.8. Telegram bot for demo

*(by Adam Raaif Nasheed)*

A second Telegram bot was also created to act as a demo for the intermediate (Phase One) model. It simply accepted voice input (as .wav file upload, or through Telegram voice recording feature), and processed it using the Phase One model to produce and display an output (see Figure 24).

![Figure 24: Telegram bot for demo](image2)
3.1.3. Recording process

The dataset was trimmed to accommodate the absolute minimal to train the model. As such a set of 83 phrases were made up from the data collected. During the phrase selection, it was primarily kept in focus to include all the phonemes in the phoneme set. More importance was given to the words with more likelihood of occurrence. During this process, grammar or language was not considered hence the readers might find the phrases a bit difficult to narrate.

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th>Last Modified</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>m_01.txt</td>
<td>AL</td>
<td>Nov 15, 2018</td>
<td>155 bytes</td>
</tr>
<tr>
<td>m_02.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>210 bytes</td>
</tr>
<tr>
<td>m_03.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>315 bytes</td>
</tr>
<tr>
<td>m_04.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>295 bytes</td>
</tr>
<tr>
<td>m_05.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>305 bytes</td>
</tr>
<tr>
<td>m_06.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>310 bytes</td>
</tr>
<tr>
<td>m_07.txt</td>
<td>AL</td>
<td>Nov 14, 2018</td>
<td>270 bytes</td>
</tr>
</tbody>
</table>

(See Appendix B for the full list of phrases used)

For the purpose of collecting audio for the phrases, 3 classrooms at Faculty of Engineering, Science and Technology were booked during a weekend (see Appendix F). About 25 people were invited to record the 83 phrases. A sophisticated microphone with attached equalizer and morpher was set in one of the classrooms. Mobile phones were used in the other room as the device for recording. The maximum number of speakers accommodated in parallel were 3.

A set of general instructions were given to each speaker prior to the recording session. The rules were however applied to the speakers recorded with the computer aided with the microphone. The mobile based recordings are always accompanied with one of the members of the project team.
The instructions were:

1. Keep reading from the first phrase till the end
2. Give a pause of about 3 seconds between each two phrases
3. In case there happens a lapse read it from the immediate last word
4. Do not have silences in between that are more than 3 seconds.

For the mobile-based recordings, each phrase was recorded to a separate audio file.
3.1.4. **Review process**

Reviewing the audio file was done with Audacity. Some of the recordings had one file for each phrase, and others had a single file for all the phrases. During this process, the parameters mentioned earlier were enforced.

- **Phrase splitting:** For those audio files with all the phrases comprised in the file, it needed to be split for each phrase separately. For this, Audacity’s labels and multiple export was used. The three-second gap was used as a delimiter for Audacity to automatically split the audio file into phrases. This task did not quite take a smooth road. Some of the audio files had much shorter gaps left, despite the clear instructions. Also some had longer pauses in between. Hence, the automatic split resulted in more than 83 files. To fix this, those audio files were manually processed. Each file was listened to and the separators labelled manually, and then were split to extract the 83 phrases.

- **Reviewing:** Each phrase per user was listened to for errors, and wherever applicable, the respective transcription file was updated. Since it was not possible for the re-recording at this stage, audio files were edited to remove the areas that could not be fixed. Silences and noisy patches found on the audio file were also removed. Corrections made by the person during recording were also either included in the transcript file or removed from the audio file.
3.2. Analysis

3.2.1. Noises and noise removal
To get better accuracy, the audio files were reviewed carefully and thoroughly. It was noticed that there was presence of traffic and other noises in the captured audio files. Therefore, Adobe Audition and Audacity was used to reduce the noises noticed in the audio files. However, another issue encountered when trying to reduce noise was the audio becoming more robotic, and the speaker’s voice losing its natural flavour. Some audio files had to be discarded because the noise could not be corrected effectively.

As seen in Figure 26, both waveform and spectral frequency display provide a clear picture of the audios. This audio was recorded from a small speech in the Faculty of Engineering, Science and Technology classroom Q3-09. The environment was good enough for human ears, yet in the recordings, a constant sound could be heard. This was not identified at the time of the recording. In waveform patterns of Figure 26, one can identify low amplitude waves at both ends of the graphical representation, which turned out to be noises generated from the air conditioning unit in the classroom.

![Figure 26: Example waveform showing noise](image)

To maintain the audio’s originality, the noise was sampled before reduction. Figure 27 shows a processed noise-reduced graphical representation of the same audio clip.
If there is any sudden extra fabricated sounds such as phone ringtones, it will easily be identified through frequency spectrum display. It will look like a sharp horizontal line at a specific frequency. Intensity or brightness indicates strength of the signal in specific frequency.

In the above diagram, the speaker is saying “theyvees milian”. In Figure 28A, there is a loud wind noise in between the two words. This kind of noise can be removed without affecting the speech as a whole; thus, the end product you can see in the waveform in Figure 28B.
3.2.2. Phonetics and linguistics

There are words that are written in the same way but spoken differently by different people of same dialects. Two spoken words with same meaning could be written in the same way (eg: hastheyka and ha satheyka). Attempts were made to recognize and included both spoken ways in the system.

In Dhivehi numerals, certain numerals can be pronounced in two or more ways. For example, 42 can be spoken as bayaalhees or saalhees dheyh. The phonation is also similar to the Sinhalese phonation of numerals. In our first model, we only used the more commonly used phonation of these, which is the latter type.

Another issue that was faced was the difference in the sound obtained with regard to gender and age. The first model included only male sounds, hence it was not able to recognize any female voices. Only once it was trained with a female voice was it able to recognize female voices. The issue was less prominent with different age categories such as those of a child and an adult, after the introduction of female voices. This may be because children have a voice pattern closer to that of females.

According to lecturers Messrs. Anwar Ibrahim and Abdul Rasheed Ali, it would not be acceptable to use most of the news articles published online, as they do not follow the correct guidelines or grammatical terms of the written language. Lecturer Ms. Azeeza Afeef adds weight to this. She adds that most of the online news do not recognize word segmentation as per new guidelines established in 2013. The young generation is still not aware of such guidelines in practice. She further points out that there is no proper systematic way in the writings of earlier days as well.

Regarding the dotted use of Dhivehi alphabets (also known as thikijehi thaana), the president of the Academy of Dhivehi Language Mr. Ashraf Ali mentions that it is not a necessity to use thikijehi thaana in the Dhivehi language. He points out that the dotted Thaana scripts were initially used to write Arabic words in its correct phonation, but even without the dotted alphabets, we can still use proper Dhivehi words. There was already an established model to use proper Dhivehi alphabets without using dotted Thaana scripts.
3.3. Testing and evaluation

The analysis is carried out using a base model which was incremented to test different hypotheses. The base model was created using an open domain Dhivehi language model which has a word count of 109638 words and with 48 minutes and 33 seconds of audio data. The hypothesis was tested using two sets of data (male and female). The hypothesis was tested using word error rate (WER).

<table>
<thead>
<tr>
<th>Type</th>
<th>Variations</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>4 (249 phrases)</td>
<td>00:36:15</td>
</tr>
<tr>
<td>Female</td>
<td>1 (83 phrases)</td>
<td>00:12:18</td>
</tr>
</tbody>
</table>

*Figure 29: Audio data used in training the base model*

<table>
<thead>
<tr>
<th>Sets</th>
<th>Variations</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1 (male)</td>
<td>3 (42 phrases)</td>
<td>00:04:51</td>
</tr>
<tr>
<td>Set 2 (female)</td>
<td>2 (28 phrases)</td>
<td>00:02:55</td>
</tr>
</tbody>
</table>

*Figure 30: Audio data used in testing the hypothesis*

<table>
<thead>
<tr>
<th>Test sets</th>
<th>Word error rate (in percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base model</td>
</tr>
<tr>
<td>Set 1</td>
<td>24.6</td>
</tr>
<tr>
<td>Set 2</td>
<td>54.57</td>
</tr>
</tbody>
</table>

*Figure 31: Test results*
3.3.1. **Effect of language model**

The base model and the test set 1 were used to analyze the relation between the WER of the model and the language model.

First, the WER of the model was analysed without including the phrases used for the test set 1 in language model. Later on the phrases used in test Set 1 was also added to language model.

When the phrase spoken is in the corpus, the analysis showed an average of 5.8% increase in accuracy of the model. Hence, accuracy depends heavily on language model.
3.3.2. **Effect of voice pitch**

To test the effect of pitch, the base model is incremented with 2 female voices and the WER is calculated. Later on the base model is increased with 2 male voices and the WER is calculated. The addition of female voices to base model increased the accuracy of female set while not much change was seen in male set. Also the addition of male voices to base model showed a significant increase in male voice test set than the female voice test set (see Figure 31).

Hence, depending on the user group of model, the model can be trained with the pitch range of the group. Therefore, if the system is used by more female users, training the model with female voice will perform better.

3.3.3. **Effect of data**

The system was trained with increasingly higher amount of voice data, and then tested for accuracy (100% - WER). The addition of data to train the model in each stage showed some improvement in accuracy of the model (see Figure 33).

![Figure 33: Effect of using more data to train the model](image-url)
4. Critical evaluation

4.1. Success of project

The system generated in Phase One of this project has an accuracy rate of 75%, which barely passes the set minimum acceptable accuracy rate in the proposal. The second, larger system created in Phase Two has an accuracy rate of 42.5%, which fails the minimum acceptable rate. This failure wasn’t altogether unexpected, as there are several known causes that could have contributed to this, including, but not limited to:

- Large amount of noise present in recording environment
- Misalignment of audio file and transcript
- Data collected was too little for the scope chosen. Only about 5 hours of voice data was collected in total.
- There were not enough variations in the data collected (too few female voices, no children’s voices, much of the phrases were based on news)
- Text data collected was not gotten from credible sources, and had many typos/mistakes
- Lack of resources (recording equipment, human resources for recording and review)
- Dictionary used does not contain every variation of every word. According to lecturer Mr. Anwar Ibrahim, there are 26 possible variations of every verb. It is unrealistic to expect each one to be covered in a randomly collected set of text, especially if not that large.
- Noise reduction proved to be ineffective in some cases, leading to use of or discarding of unusable data, of which the system already had very little
- Ambiguity in how phrases should be pronounced (eg: use of lone raa in borrowed English words (obzaav vs. obzarv))
4.2. Suggestions for improvement if repeated, and future enhancements

The ASR system was developed so that it has the ability to be incrementally updated. Hence, more text and voice data being available for training would naturally be an improvement.

The phoneme list used for this model was generated by simply assuming that each consonant and vowel is a new phoneme (with some exceptions, such as shaviyani-sukun and thaa-sukun). Unlike English, this works for Dhivehi because of how well the Thaana representation of a word can depict how it is pronounced, but is not always the case in practice (Mohamed, 1958). This method of phoneme generation also fails to capture some nuances that exist in spoken language. A better acoustic model can be produced once an official dictionary of the Dhivehi language, complete with the standard IPA pronunciations for words, is produced.

The voice data collected for creating the acoustic model was taken almost exclusively from news sources. Though attempts were made to get phrases from other sources (namely Faihthoora and drama scripts), these efforts failed to produce any results. Therefore, the acoustic model that was generated lacked variations in tone of voice. This system can be improved upon with the use of more varied phrases for the contributors to read.

The voice data collected were also predominantly from male youths. The system can be improved by using voices from many more people, from different age groups and an approximately equal distribution of genders.

Similar to how some English words are pronounced differently based on whether the speaker is British, Australian or Indian, there are multiple Dhivehi dialects spoken across different parts of the Maldives. These dialects may have different spellings for their words, and/or different pronunciations. An improvement on this system can take those other dialects into account, in the creation of the dictionary, acoustic and language models.

The dictionary for use with this system was created using existing corpora, and court documents available online. Therefore, the dictionary is not guaranteed to contain every Dhivehi word or their variations (although the larger the amount of text, the higher the number of unique words covered). Moreover, spaces between words in Dhivehi are inconsistent. For example, eotheeboalha and eothee boalha can both be understood and used perfectly fine, although only the latter is grammatically correct. Because the sources were blocks of text typed in by someone, there are many instances of the former in the dictionary, but simply specifying space as the delimiter between words produced a good enough dictionary for our current purposes. Therefore, a better dictionary for use with this system can be produced once a part-
of-speech corpus is available for Dhivehi. As the part-of-speech type for words will be known in such a corpus, generating word variations using existing rules would be a lot easier.

4.3. Experiences gained

Heading into this project with fresh pairs of eyes, we attained a lot of knowledge in various fields, such as Dhivehi phonetic makeup, Dhivehi grammar and spelling rules, noise reduction methods and acoustic feature extraction methods.

We also picked up multiple software development tools and techniques, particularly in Java, Python, PHP and JavaScript programming languages.

We also learned the benefits of collaborative work in large-scale projects, and the importance of teamwork.
References


## Appendices

### Appendix A: Phrase list used for recording voices for Phase One (number) model

<table>
<thead>
<tr>
<th>#</th>
<th>Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>2</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>3</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>4</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>5</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>6</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>7</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>8</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>9</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>10</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>11</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>12</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>13</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>14</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>15</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>16</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>17</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>18</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>19</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>20</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>21</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>22</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>23</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>24</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>25</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>26</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>27</td>
<td>ساپچ را ها دنکار یا برای استفاده در اینترنت و لینک های شبکه، من و دیگر استفاده کنندگان استفاده می‌کنند.</td>
</tr>
<tr>
<td>Page No.</td>
<td>Dhivehi Text</td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>28</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>
Dhivehi Automatic Speech Recognition System
<table>
<thead>
<tr>
<th>Page</th>
<th>Dhivehi Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>91</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>92</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>93</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>94</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>95</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>96</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>97</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>98</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>99</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>100</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>101</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>102</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>103</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>104</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>105</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>106</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>107</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>108</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>109</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>110</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>111</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>112</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>113</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>114</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>115</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>116</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>117</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>118</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>119</td>
<td>dhivehi text</td>
</tr>
<tr>
<td>120</td>
<td>dhivehi text</td>
</tr>
</tbody>
</table>
Dhivehi Automatic Speech Recognition System

<table>
<thead>
<tr>
<th>Page</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
<td>Dhivehi automatic speech recognition system.</td>
</tr>
<tr>
<td>122</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>123</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>124</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>125</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>126</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>127</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>128</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>129</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>130</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>131</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>132</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>133</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>134</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>135</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>136</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>137</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>138</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>139</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>140</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>141</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>142</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>143</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>144</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>145</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>146</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>147</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>148</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>149</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>150</td>
<td>ރަބްނިރު ވަރެނުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
<tr>
<td>151</td>
<td>ގިލަޤްޤުދްޤަޕުނ ވަގިލަޤްޤުދްޤަޕުނ ވަދިހަދައާވަދިހަދައ</td>
</tr>
</tbody>
</table>

Faculty of Engineering, Science and Technology
### Appendix B: Phrase list A used for recording voices for Phase Two (main) model

<table>
<thead>
<tr>
<th>#</th>
<th>Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>40</td>
<td>41</td>
</tr>
<tr>
<td>42</td>
<td>43</td>
</tr>
<tr>
<td>44</td>
<td>53</td>
</tr>
</tbody>
</table>

Dhivehi Automatic Speech Recognition System

Faculty of Engineering, Science and Technology
Dhivehi Automatic Speech Recognition System

| 65 | ބަތްްފިލްމ ންްދޭއްްޕޮއިންޓްްތިނެއްްބިލިއަންްޑޮރަރ |
| 66 | ްމިއީްޔ ނިވާސަރްސްޓޫހިރ ގެްމިފަދަރެބ |
| 67 | ރިވިމްްސޫޓ ތަކ ގައިްކ ޑަކ ދިވްއެކިްސ ވިމްްސޫޓ |
| 68 | ވެރިންްގޮނޑ ދޮށ ގައިްސިށީންދެރައިގެން |
| 69 | ބޭރ ންްޑައިގްނޯޒްވިއެރ ރެވޭނެްކަމަކީރ |
| 70 | ބައީމަްފަރ ވާރްކާފައިރޮބްޒާވ ރ ން،ްމިހާރ |
| 71 | އަދިްރްސަތްތޮޅ ްމަދްރަސާގައިރާސައްކަސަރ |
| 72 | ސާރ ދ ްއެއްގޮތަރިސްހިރީވިއ އެހަވުދައިދަރ ރާތި |
| 73 | އާރިއެއްްހޯދައިްއޭބ ވެރިދެވިފައިވާްފަރާތް |
| 74 | ސާރ ދ ްއެއްވިރާސ ގައިރެވިރ ސްރްސްކ |
| 75 | ސާރ ދ ްއެއްގޮސްސްރ ގަސްސްކ ރައްސާރ ގައިރެވިރ |
| 76 | ވިދ އަރެވިދ އަރެވިދ |
| 77 | ހަބީބަށް،ްބަ�ތްޔިތ ންގެ،ްހަދ ޖެހ |
| 78 | އަދިްރްސަތޮސްސްސައްކ |
| 79 | އަދިްރްސަތޮސްސްސައްކ |
| 80 | އަދިްރްސަތޮސްސްސައްކ |
| 81 | އަދިްރްސަތޮސްސްސައްކ |
| 82 | އަދިްރްސަތޮސްސްސައްކ |
| 83 | އަދިްރްސަތޮސްސްސައްސެވ |

Faculty of Engineering, Science and Technology
### Appendix C: Phrase list B used for recording voices for Phase Two (main) model

<table>
<thead>
<tr>
<th>#</th>
<th>Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>
Dhivehi Automatic Speech Recognition System

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td>Text</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
</tr>
</tbody>
</table>
| 51   | دھیؤھی اٹھورٹی نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر نکر N

Faculty of Engineering, Science and Technology
Dhivehi Automatic Speech Recognition System

<p>| 78 | ހައްދާލަމ ންގެންދިއ މެވެ،ްހައްދ ންމަތީ،ްހައްދ ފަހަނަޅައިގެން |  |
| 79 | ހަވާދ ގެ، ްހަވާސާ، ްހަވާސާގައިރަމ |  |
| 80 | ހަވީރަ، ްހަވީރަވާދ ވިއިރ نز |  |
| 81 | ްދަވައިދެއްވާނަމޭ، ްހަ |  |
| 82 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 83 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 84 | ހަމަހިލާ، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 85 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 86 | ހަމަމިއާއެކ، ްހަމަތަކެކެވެމައ ވެދާނެ، ްހައްގ |  |
| 87 | ްހައްލ، ްހައްލ ހްވޭނޭ |  |
| 88 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 89 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 90 | ހަސި ރ ބ ައ ރ ބ ަލ، ްހަބ ، |  |
| 91 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 92 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 93 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 94 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 95 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 96 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 97 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 98 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 99 | ްދިއ މެވެ، ްހަދަރ ދީއެއް |  |
| 100 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 101 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 102 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 103 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 104 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 105 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 106 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 107 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |
| 108 | ހަބަރެއް، ްހަބީބ، ްހަބ، ްހަބ ރ |  |</p>
<table>
<thead>
<tr>
<th>Page</th>
<th>Dhivehi Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>ނަބިއްޔާ،ްނަވާވީސްޓީޗަރ ންނަށް</td>
</tr>
<tr>
<td>110</td>
<td>ހޯދަނީ،ްހޯދަންބާއެވެ،ްހޯދައިގެންރާއްޖެް</td>
</tr>
<tr>
<td>111</td>
<td>ހޮޓެލް，ްހޮވ ނީމައެއްނ ް</td>
</tr>
<tr>
<td>112</td>
<td>ނިއ ޔޯކް，ްނިވްޓަންް</td>
</tr>
<tr>
<td>113</td>
<td>ނިޒާމަކަށްވީއިރ ،ްނިޒާމ ،ްނިޔާވިް</td>
</tr>
<tr>
<td>114</td>
<td>ހޯލްލައ ންޖްމި，ްހޯލްޑިންގްސްް</td>
</tr>
<tr>
<td>115</td>
<td>ހިތްވަރަކީވެސް，ްހިތްވަރ ނެތ ން،ްހިތްވަރ ދެއްވީ，ްހިތްވަރ ދޭ</td>
</tr>
<tr>
<td>116</td>
<td>ހާތަހ ޒީބ ،ްނާސަރީއަށް</td>
</tr>
<tr>
<td>117</td>
<td>ހީވަނީބާއެވެ،ްހީވާް</td>
</tr>
<tr>
<td>118</td>
<td>ހ ރަސްނާރ ބިހި،ްހ ރަސްނޭޅޭނެ،ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ،</td>
</tr>
<tr>
<td>119</td>
<td>ރާތި，ްހިރޯޝީމާ،ްހިބިހި،ްހ ރަނީ，ްހ ވަރެއް</td>
</tr>
<tr>
<td>120</td>
<td>ހިތްހެޔޮ，ްހިތްހެޔޮކަމާއި</td>
</tr>
<tr>
<td>121</td>
<td>ހީރޯޝީމާ،ްހ ރަސްނޭޅޭނެ،ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ،</td>
</tr>
<tr>
<td>122</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>123</td>
<td>ްރަސްނާރ ބިހި，ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން，ްހ ރިމިންވަރެއް،ްހ ރީކީ،</td>
</tr>
<tr>
<td>124</td>
<td>ރާތި，ްހިރޯޝީމާ،ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ，</td>
</tr>
<tr>
<td>125</td>
<td>ކރީނީ，ްހ ވަރެއް</td>
</tr>
<tr>
<td>126</td>
<td>ްރިކާއެކ ،ްހެޔޮހިތ ން，ްހެޔޮނަތީދާ</td>
</tr>
<tr>
<td>127</td>
<td>ރަސްނާރ ބިހި،ްހ ރަސްނޭޅޭނެ،ްހ ރިހެން，ްހ ރިމިންވަރެއް،ްހ ރީކީ،</td>
</tr>
<tr>
<td>128</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>129</td>
<td>ްރަސްނާރ ބިހި，ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ،</td>
</tr>
<tr>
<td>130</td>
<td>ރާތި，ްހިރޯޝީމާ،ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން،ްހ ރިމިންވަރެއް，ްހ ރީކީ،</td>
</tr>
<tr>
<td>131</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>132</td>
<td>ރާތި，ްހިރޯޝީމާ،ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން،ްހ ރިމިންވަރެއް،ްހ ރީކީ،</td>
</tr>
<tr>
<td>133</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>134</td>
<td>ރާތި，ްހިރޯޝީމާ،ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން،ްހ ރިމިންވަރެއް，ްހ ރީކީ，</td>
</tr>
<tr>
<td>135</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>136</td>
<td>ރާތި，ްހިރޯޝީމާ，ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ，</td>
</tr>
<tr>
<td>137</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
<tr>
<td>138</td>
<td>ރާތި，ްހިރޯޝީމާ，ްހ ރަސްނޭޅޭނެ，ްހ ރިހެން，ްހ ރިމިންވަރެއް，ްހ ރީކީ，</td>
</tr>
<tr>
<td>139</td>
<td>ދީވަނީބާއެވެ，ްހ ވަފެނަކަށް</td>
</tr>
</tbody>
</table>

Faculty of Engineering, Science and Technology
<table>
<thead>
<tr>
<th>Dhivehi Automatic Speech Recognition System</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
</tr>
<tr>
<td>141</td>
</tr>
<tr>
<td>142</td>
</tr>
<tr>
<td>143</td>
</tr>
<tr>
<td>144</td>
</tr>
<tr>
<td>145</td>
</tr>
<tr>
<td>146</td>
</tr>
<tr>
<td>147</td>
</tr>
<tr>
<td>148</td>
</tr>
<tr>
<td>149</td>
</tr>
<tr>
<td>150</td>
</tr>
<tr>
<td>151</td>
</tr>
<tr>
<td>152</td>
</tr>
<tr>
<td>153</td>
</tr>
<tr>
<td>154</td>
</tr>
<tr>
<td>155</td>
</tr>
<tr>
<td>156</td>
</tr>
<tr>
<td>157</td>
</tr>
<tr>
<td>158</td>
</tr>
<tr>
<td>159</td>
</tr>
<tr>
<td>160</td>
</tr>
<tr>
<td>161</td>
</tr>
<tr>
<td>162</td>
</tr>
<tr>
<td>163</td>
</tr>
<tr>
<td>164</td>
</tr>
<tr>
<td>165</td>
</tr>
<tr>
<td>166</td>
</tr>
<tr>
<td>167</td>
</tr>
<tr>
<td>168</td>
</tr>
<tr>
<td>169</td>
</tr>
<tr>
<td>170</td>
</tr>
</tbody>
</table>
### Dhivehi Automatic Speech Recognition System

| 171 | އާފޯލިނގްޓުނައިޓް، އާލިއަންޒް |
| 172 | ބޯހަރ ،ރޯހިއްލާ،ބޯނަސް |
| 173 | ބޮޑާ，ބޮޑާހޭކ މާއިް |
| 174 | ކަލޭފާނ ،ބަވަރޭޖާް |
| 175 | އަވޯންޓެޑް，ބަރޫންް |
| 176 | ބޭއިހްތިޔާރ ގައިް |
| 177 | ބެންކޮކ ގަ،ބެންޗ ން |
| 178 | ބްލަޑް，ބްރިޓިޝްް |
| 179 | އިދޫކޮށް，ބަހ މަދ ސޯޅަ |
| 180 | ރެވޭބާވައެވެ，ބީއެމްއެލް |
| 181 | ރިރާވޭބާވައެވެ，ބީއެމްއެލް |
| 182 | ބަހްމަދ ޑޮކްޓަރ，ބަރޫދ |
| 183 | އެހެރް，ބެހޮޅި，ބެހްރެންބާރގް |
| 184 | އިސްކޫލެއް，ބަނޫ،ބަންހެނ ންގެގާތ |
| 185 | އިބްރަހިމޮވިޗަށް，ބަލީދިހައެއް |
| 186 | ރަށް，ބަބޫބަކ ރަށް |
| 187 | އިދޫކޮށް，އިރީރޭވެސް |
| 188 | ރީރޭވެސް |
| 189 | ރީރޭވެސް |
| 190 | ރީރޭވެސް |
| 191 | ރީރޭވެސް |
| 192 | އިސްކޫލެއް |
| 193 | ރީރޭވެސް |
| 194 | ރީރޭވެސް |
| 195 | ރީރޭވެސް |
| 196 | ރީރޭވެސް |
| 197 | ރީރޭވެސް |
| 198 | ރީރޭވެސް |
| 199 | ރީރޭވެސް |
| 200 | ރީރޭވެސް |
| 201 | ރީރޭވެސް |
Dhivehi Automatic Speech Recognition System

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

Facility of Engineering, Science and Technology
Dhivehi Automatic Speech Recognition System

| 233 | ીިރިހިރީރަންޓްކ ރ ން،ްޕްރޮސެސާ،ްޕްރޮޓ ން،ްޖ ޑީޝަލް،ްޖޫނޯ،ްޖޭޕީން،ްޖޮއޭން،ްޖޯސެއިން،ްޖޯޒެފް،ްޖޯޖް،ޗާބޫކަކ ންވެސް،ްޗާޕ ،ްޗާޕްކޮށް |
| 234 | އޮފްގެް |
| 235 | އަލީޔ ނިވަރސީޓީގަ |
| 236 | އި，ްގަރީޒާތަކ ގެް |
| 237 | އަލީޔ ނިވަރސީޓީގަ |
### Appendix D: List of countries in Dhivehi

<table>
<thead>
<tr>
<th>English name</th>
<th>Dhivehi name (without thikijehi thaana)</th>
<th>Dhivehi alternative name (some with thikijehi thaana)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somaliland</td>
<td>₪ގިކުލްސްސަރު</td>
<td></td>
</tr>
<tr>
<td>South Georgia and South Sandwich Islands</td>
<td>₪ރުފިމުހިރުފިމުހިރުރު</td>
<td></td>
</tr>
<tr>
<td>French Southern and Antarctic Lands</td>
<td>₪ހީތިސީސީސީރުރު</td>
<td></td>
</tr>
<tr>
<td>Palestine</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Aland Islands</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Nauru</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Saint Martin</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Tokelau</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Western Sahara</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Afghanistan</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Albania</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>American Samoa</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Andorra</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Angola</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Anguilla</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Antigua and Barbuda</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Armenia</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Aruba</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Bahamas</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Bahrain</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Barbados</td>
<td>₪ރާސާރުސާރު</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Belarus</td>
<td>ދެއްސެވިނިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Belize</td>
<td>ބަސްގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Benin</td>
<td>ދެއްސެވިނިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Bermuda</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Bhutan</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Bolivia</td>
<td>ދެއްސެވިނިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Botswana</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>British Virgin Islands</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Brunei Darussalam</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Myanmar</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Burundi</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Cambodia</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Cameroon</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Cape Verde</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Cayman Islands</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Central African Republic</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Chad</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Christmas Island</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Cocos Islands</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Comoros</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Democratic Republic of the Congo</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Republic of Congo</td>
<td>ދިވުނީރިފިގިހިބަސްގި</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Cook Islands</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Cote d'Ivoire</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Croatia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Cuba</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Curaçao</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Cyprus</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Djibouti</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Dominica</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>El Salvador</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Equatorial Guinea</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Eritrea</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Estonia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Falkland Islands</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Faroe Islands</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Fiji</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>French Polynesia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Gabon</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>The Gambia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Ghana</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Gibraltar</td>
<td>Миллиномаа</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>గ్రీస్</td>
<td></td>
</tr>
<tr>
<td>Greenland</td>
<td>గ్రీన్‌లాండ్</td>
<td></td>
</tr>
<tr>
<td>Grenada</td>
<td>గ్రెనాడా</td>
<td></td>
</tr>
<tr>
<td>Guam</td>
<td>్‌గ్యం</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>గ్రైమాటామా</td>
<td></td>
</tr>
<tr>
<td>Guernsey</td>
<td>గ్ర్యూర్న్సీ</td>
<td></td>
</tr>
<tr>
<td>Guinea</td>
<td>గ్ర్యునీ</td>
<td></td>
</tr>
<tr>
<td>Guinea-Bissau</td>
<td>గ్ర్యునీబిస్సాయు</td>
<td></td>
</tr>
<tr>
<td>Guyana</td>
<td>గ్యుయానా</td>
<td></td>
</tr>
<tr>
<td>Haiti</td>
<td>హేటి</td>
<td></td>
</tr>
<tr>
<td>Vatican City</td>
<td>వాటికన్సిటీ</td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>హోండ్యారా</td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>హంగరియా</td>
<td></td>
</tr>
<tr>
<td>Iceland</td>
<td>ఈస్కండిల్</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>ఇండియా</td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>ఇండోనిసియా</td>
<td></td>
</tr>
<tr>
<td>Iran</td>
<td>ఇరాన్</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>ఇరాక్</td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>ఈరల్యాద్యా ఇరేల్డ్</td>
<td></td>
</tr>
<tr>
<td>Isle of Man</td>
<td>ఈస్లోఫమాన్</td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td>ఇస్రేల్</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>ఇటిల్యాద్యా</td>
<td></td>
</tr>
<tr>
<td>Jamaica</td>
<td>జామాయికా</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>జాపాన్</td>
<td></td>
</tr>
<tr>
<td>Jersey</td>
<td>జరేసి</td>
<td></td>
</tr>
<tr>
<td>Jordan</td>
<td>జరోడన్</td>
<td></td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>కాజాకసాట్స్</td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td>కెనేయా</td>
<td></td>
</tr>
<tr>
<td>Kiribati</td>
<td>కిరిబాతి</td>
<td></td>
</tr>
<tr>
<td>North Korea</td>
<td>నథ్‌ఫర్థ్‌కొరెయ</td>
<td></td>
</tr>
<tr>
<td>South Korea</td>
<td>స్థ్‌ఫర్థ్‌కొరెయ</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Kosovo</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Kuwait</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Laos</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Latvia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Lebanon</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Lesotho</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Liberia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Libya</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Liechtenstein</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Lithuania</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Macedonia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Madagascar</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Malawi</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Malaysia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Maldives</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mali</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Malta</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Marshall Islands</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mauritania</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mauritius</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mexico</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Federated States of Micronesia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Moldova</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Monaco</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mongolia</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Montenegro</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Montserrat</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Morocco</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Mozambique</td>
<td>ދެގިހިރު</td>
<td>ބްކަރި ބިރަ ބްއެގިރު</td>
</tr>
<tr>
<td>Country</td>
<td>Name</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Namibia</td>
<td>ދިށުވިއިޏުބިއ</td>
<td></td>
</tr>
<tr>
<td>Nepal</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>New Caledonia</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Niger</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Niue</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Norfolk Island</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Northern Mariana Islands</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Oman</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Palau</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Panama</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Papua New Guinea</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Peru</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Philippines</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Pitcairn Islands</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Russia</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Rwanda</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Saint Barthelemy</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Saint Helena</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Saint Kitts and Nevis</td>
<td>ބަދީމުރ</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi Name</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>Saint Lucia</td>
<td>سنت لويس</td>
<td></td>
</tr>
<tr>
<td>Saint Pierre and Miquelon</td>
<td>سانت بيرن إند ميكلون</td>
<td></td>
</tr>
<tr>
<td>Saint Vincent and the Grenadines</td>
<td>سانت واينس ود الجراداين</td>
<td></td>
</tr>
<tr>
<td>Samoa</td>
<td>ساموا</td>
<td></td>
</tr>
<tr>
<td>San Marino</td>
<td>سان مارينو</td>
<td></td>
</tr>
<tr>
<td>Sao Tome and Principe</td>
<td>ساو تومه ود برينسيه</td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>سعودية</td>
<td></td>
</tr>
<tr>
<td>Senegal</td>
<td>سينغال</td>
<td></td>
</tr>
<tr>
<td>Serbia</td>
<td>سيربيا</td>
<td></td>
</tr>
<tr>
<td>Seychelles</td>
<td>سيشيليس</td>
<td></td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>سيراليون</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>سنغافورة</td>
<td></td>
</tr>
<tr>
<td>Sint Maarten</td>
<td>سينت مارتن</td>
<td></td>
</tr>
<tr>
<td>Slovakia</td>
<td>سلوسكيا</td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td>سلونيا</td>
<td></td>
</tr>
<tr>
<td>Solomon Islands</td>
<td>سولومن ايلاند</td>
<td></td>
</tr>
<tr>
<td>Somalia</td>
<td>سوماليا</td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>زومبي</td>
<td></td>
</tr>
<tr>
<td>South Sudan</td>
<td>زومبي نورث</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>سپانیا</td>
<td></td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>سری لنکا</td>
<td></td>
</tr>
<tr>
<td>Sudan</td>
<td>سودان</td>
<td></td>
</tr>
<tr>
<td>Suriname</td>
<td>سورینا</td>
<td></td>
</tr>
<tr>
<td>Svalbard</td>
<td>سولاپبارد</td>
<td></td>
</tr>
<tr>
<td>Swaziland</td>
<td>سوازی‌یلند</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>سوئدیکا</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>زویترلند</td>
<td></td>
</tr>
<tr>
<td>Syria</td>
<td>سیریا</td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>تائوان</td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td>تاجیکستان</td>
<td></td>
</tr>
<tr>
<td>Tanzania</td>
<td>تانزانیا</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Dhivehi</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>ދިވެހިވުދ</td>
<td></td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>ދިވެހިއަރުމުއިބިޤ</td>
<td></td>
</tr>
<tr>
<td>Togo</td>
<td>އެހަފިދ</td>
<td></td>
</tr>
<tr>
<td>Tonga</td>
<td>ދިވެހިއަނިކްތަޤ</td>
<td></td>
</tr>
<tr>
<td>Trinidad and Tobago</td>
<td>ދިވެހިއާރުބީށުރުދ</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>ދިވެހިހުބިސ</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>ދިވެހިބަމުދ</td>
<td></td>
</tr>
<tr>
<td>Turkmenistan</td>
<td>ދިވެހިރުކްނަރަމިލ</td>
<td></td>
</tr>
<tr>
<td>Turks and Caicos Islands</td>
<td>ދިވެހިނުދުމުހިދ</td>
<td></td>
</tr>
<tr>
<td>Tuvalu</td>
<td>ދިވެހިއަދީނުމ</td>
<td></td>
</tr>
<tr>
<td>Uganda</td>
<td>ދިވެހިއަދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Ukraine</td>
<td>ދިވެހިއަބީރުކ</td>
<td></td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>ދިވެހިއުދްމުރުބ</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>ދިވެހިރުބުމުދުރ</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>ދިވެހިއުދުފުދުރ</td>
<td></td>
</tr>
<tr>
<td>Uruguay</td>
<td>ދިވެހިބަމުރ</td>
<td></td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>ދިވެހިއަބީރުކ</td>
<td></td>
</tr>
<tr>
<td>Vanuatu</td>
<td>ދިވެހިރުބުމުދުރ</td>
<td></td>
</tr>
<tr>
<td>Venezuela</td>
<td>ދިވެހިރުބުމުދުރ</td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>ދިވެހިރުބުމުދުރ</td>
<td></td>
</tr>
<tr>
<td>US Virgin Islands</td>
<td>ދިވެހިއަދީނުމ</td>
<td></td>
</tr>
<tr>
<td>Wallis and Futuna</td>
<td>ދިވެހިރުބުމުދުރ</td>
<td></td>
</tr>
<tr>
<td>Yemen</td>
<td>ދިވެހިއަދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td>ދިވެހިއަދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>ދިވެހިއަދުމުރ</td>
<td></td>
</tr>
<tr>
<td>US Minor Outlying Islands</td>
<td>ދިވެހިއުދްމުރުބ</td>
<td></td>
</tr>
<tr>
<td>Antarctica</td>
<td>ދިވެހިއުދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Northern Cyprus</td>
<td>ދިވެހިއުދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>ދިވެހިއުދުމުރ</td>
<td></td>
</tr>
<tr>
<td>Heard Island and McDonald Islands</td>
<td>ދިވެހިއުދުމުރ</td>
<td></td>
</tr>
</tbody>
</table>

Faculty of Engineering, Science and Technology
<table>
<thead>
<tr>
<th>British Indian Ocean Territory</th>
<th>مالاقا</th>
<th>Macau</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>مالاقا</td>
<td></td>
</tr>
</tbody>
</table>
Appendix E: List of team meetings

This list contains short descriptions of all meetings attended as a team, including supervisory meetings and interview sessions. This is not an exhaustive list of all time spent on the project, as much of it was also worked on individually.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Venue</th>
<th>Present</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jul 24</td>
<td>08:30 am - 10:00 am</td>
<td>Royal Garden</td>
<td>4</td>
<td>Discussion about splitting work.</td>
</tr>
<tr>
<td>Jul 24</td>
<td>09:30 pm - 10:30 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Discussion about splitting work.</td>
</tr>
<tr>
<td>Jul 28</td>
<td>02:00 pm - 03:30 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Proposal/PSF discussion.</td>
</tr>
<tr>
<td>Jul 28</td>
<td>05:30 pm - 08:30 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Proposal/PSF discussion.</td>
</tr>
<tr>
<td>Jul 29</td>
<td>03:00 pm - 11:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Proposal/PSF draft finalization.</td>
</tr>
<tr>
<td>Aug 9</td>
<td>05:00 pm - 05:30 pm</td>
<td>FEST</td>
<td>4</td>
<td>Khalid sir Supervisory meeting 1</td>
</tr>
<tr>
<td>Aug 11 - 12</td>
<td>04:00 pm - 12:00 am</td>
<td>MNU Central</td>
<td>4</td>
<td>Proposal finalization.</td>
</tr>
<tr>
<td>Aug 17</td>
<td>02:30 pm - 05:30 pm</td>
<td>Harbour Food Court</td>
<td>4</td>
<td>Reporting progress on individual work. Discussing and resolving any issues.</td>
</tr>
<tr>
<td>Aug 29</td>
<td>02:00 pm - 06:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Reporting progress on individual work. Working together.</td>
</tr>
<tr>
<td>Aug 30</td>
<td>11:00 am - 11:30 am</td>
<td>MNU Central</td>
<td>3 (Y busy)</td>
<td>4 FA lecturers Meeting with FA lecturers for linguistic/phonemic help &amp; research.</td>
</tr>
<tr>
<td>Sep 2</td>
<td>01:00 pm - 01:30 pm</td>
<td>MNU Central</td>
<td>2 (Y busy, R busy)</td>
<td>Riyaz sir Meeting for help with understanding Dhivehi phonetics, with respect to ASR</td>
</tr>
<tr>
<td>Sep 2</td>
<td>02:00 pm - 04:30 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Reporting progress on individual work. Working together.</td>
</tr>
<tr>
<td>Sep 3</td>
<td>10:00 am - 10:30 am</td>
<td>Royal Garden</td>
<td>4</td>
<td>Preparation for supervisory meeting</td>
</tr>
<tr>
<td>Date</td>
<td>Time</td>
<td>Location</td>
<td>Name</td>
<td>Event Details</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Sep 3</td>
<td>10:30 am - 11:00 am</td>
<td>FEST</td>
<td>Khalid sir</td>
<td>Supervisory meeting 2</td>
</tr>
<tr>
<td>Sep 4</td>
<td>01:00 pm - 02:00 pm</td>
<td>MNU Central</td>
<td>3 (Y busy)</td>
<td>Meeting for linguistic help (for preparation of phonetic dictionary) and corpus</td>
</tr>
<tr>
<td>Sep 4</td>
<td>02:00 pm - 05:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Discussion post-meeting about how to proceed with phonetic dictionary - Include only one variation for words that could be spelled in more than one way (unless spelling is vastly different) - Use right phoneme for words that are pronounced differently than they are spelled</td>
</tr>
<tr>
<td>Sep 10</td>
<td>05:00 pm - 08:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Training using numbers data (for midterm submission) - data preprocessing</td>
</tr>
<tr>
<td>Sep 11</td>
<td>08:30 pm - 11:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>- Training using numbers data (for midterm submission) - testing - Discussion about midterm submission documentation</td>
</tr>
<tr>
<td>Sep 17</td>
<td>02:00 pm - 03:00 pm</td>
<td>PSM building</td>
<td>3 (Y busy)</td>
<td>PSM Staff - Asking for training data (recordings and transcripts) from news/drama</td>
</tr>
<tr>
<td>Sep 19 - 20</td>
<td>07:00 pm - 12:00 am</td>
<td>MNU Central</td>
<td>4</td>
<td>Mid-term submission finalization.</td>
</tr>
<tr>
<td>Sep 24</td>
<td>06:30 pm - 08:00 pm</td>
<td>Harbour Food Court</td>
<td>4</td>
<td>Getting ready for mid-term presentation</td>
</tr>
<tr>
<td>Sep 25</td>
<td>04:30 pm - 06:30 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Getting ready for mid-term presentation</td>
</tr>
<tr>
<td>Sep 26</td>
<td>12:00 pm - 03:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Finalizing mid-term presentation</td>
</tr>
<tr>
<td>Sep 26</td>
<td>07:00 pm - 10:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Finalizing + rehearsing mid-term presentation</td>
</tr>
<tr>
<td>Sep 27</td>
<td>09:00 am - 11:00 am</td>
<td>FEST</td>
<td>4</td>
<td>Judges - Mid-term presentation</td>
</tr>
<tr>
<td>Sep 30</td>
<td>10:00 am - 12:00 pm</td>
<td>Academy of Dhivehi Language</td>
<td>4</td>
<td>Judges - Ask for help with corpus + reviewing process</td>
</tr>
<tr>
<td>Oct 10</td>
<td>09:30 am - 10:30 am</td>
<td>Royal Garden</td>
<td>4</td>
<td>Discussing further actions + preparation for supervisory meeting</td>
</tr>
<tr>
<td>Oct 10</td>
<td>04:30 pm - 07:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>System presentation method(s) development</td>
</tr>
<tr>
<td>Oct 16</td>
<td>09:00 pm - 11:00 pm</td>
<td>MNU Central</td>
<td>4</td>
<td>Coming up with phrase list + discussion about what do with thikijehi thaana - Remove thikijehi thaana from system entirely (This does not include paviyani, shaa.)</td>
</tr>
</tbody>
</table>
Dhivehi Automatic Speech Recognition System

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Location</th>
<th>Group</th>
<th>Description</th>
</tr>
</thead>
</table>
| Oct 18 | 01:30 pm - 02:30 pm | FEST              | 4     | Khalid sir Supervisory meeting 3  
- Keep non-\textit{aaseenuthashi sukun} in loan words  
- Instruct contributors to read the phrases as-is (whenever possible)  
- Displayed output will also not have \textit{thikijehi thaana} |
| Oct 19 | 02:00 pm - 09:30 pm | FEST              | 4     | Contributors Recording voice data from contributors (see Appendix F) + reviewing voice data |
| Oct 20 | 10:00 am - 04:00 pm | FEST              | 4     | Contributors Recording voice data from contributors (see Appendix F) |
| Oct 23 | 01:00 pm - 02:00 pm | MNU Central       | 4     | Discussing review methods + other actions going forward |
| Nov 8  | 03:00 pm - 04:00 pm | PSM building      | 2     | PSM Staff  
Getting audio data from selected news segments  
- Signed a contract detailing how the data gathered cannot be used for other purposes |
| Nov 12 | 10:30 am - 01:00 pm | MNU Central       | 4     | Preparation for supervisory meeting |
| Nov 12 | 01:00 pm - 01:45 pm | FEST              | 4     | Khalid sir Supervisory meeting 4  
Reviewing audio recordings + discussion about reviewing process and documentation  
- \textit{eve}, \textit{sheve} must be corrected in transcript for wherever speaker does not pronounce the \textit{vaavu}  
- \textit{gai}, \textit{fai} is given a pass (i.e. not corrected); same goes for \textit{hus noonu} |
| Nov 14 | 04:30 pm - 08:00 pm | MNU Central       | 4     | Reviewing audio recordings + discussion about reviewing process and documentation  
- \textit{eve}, \textit{sheve} must be corrected in transcript for wherever speaker does not pronounce the \textit{vaavu}  
- \textit{gai}, \textit{fai} is given a pass (i.e. not corrected); same goes for \textit{hus noonu} |
| Nov 16 | 04:30 pm - 08:00 pm | MNU Central       | 4     | Reviewing audio recordings + splitting work for final documentation |
| Nov 18 | 08:00 pm - 10:00 pm | MNU Central       | 4     | Compiling individual work + putting finishing touches on final documentation |

* H : Hassaan, I : Ifham, R : Raaf, Y : Yameen
Appendix F: Recording sessions timetable for Phase Two

Three classrooms were booked over the course of two days for recording purposes. The plan was to reach out to whoever had the time to contribute, to schedule a slot with us. This table shows how many bookings were made prior to that weekend. Names are not included for privacy reasons.

<table>
<thead>
<tr>
<th>Time</th>
<th>Room 1 (FEST Q3)</th>
<th>Room 2 (FEST Q3)</th>
<th>Room 3 (FEST Q4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Friday, 19th October</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2pm – 3pm</td>
<td>Occupied</td>
<td>Occupied</td>
<td>Occupied</td>
</tr>
<tr>
<td>3pm – 4pm</td>
<td>Occupied</td>
<td>Occupied</td>
<td></td>
</tr>
<tr>
<td>4pm – 5pm</td>
<td>Occupied</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5pm – 6pm</td>
<td>Occupied</td>
<td>Occupied</td>
<td>Occupied</td>
</tr>
<tr>
<td><strong>Saturday, 20th October</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10am – 11am</td>
<td>Occupied</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11am – 12pm</td>
<td>Occupied</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12pm – 1pm</td>
<td></td>
<td></td>
<td><strong>BREAK</strong></td>
</tr>
<tr>
<td>1pm – 2pm</td>
<td>Occupied</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2pm – 3pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pm – 4pm</td>
<td>Occupied</td>
<td>Occupied</td>
<td></td>
</tr>
</tbody>
</table>

The extra time was spent reviewing recorded audio clips, and inviting other students who were in the building at the time to contribute if they wished to.