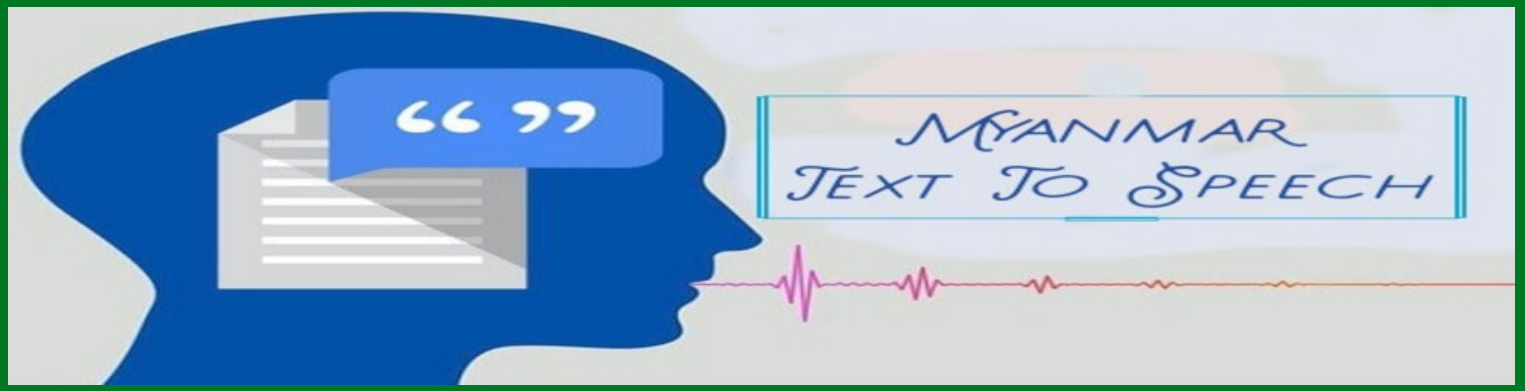


# NLP Project Presentation

## Group Members


| | | | |
|---|---|---|---|
| Badounmar | Ph.D(IT)-18 | Nan Yu Hlaing | Ph.D(IT)-16 |
| Hnin Yu Hlaing | Ph.D(IT)-10 | Thida San | Ph.D(IT)-12 |
| Hlaing May Tin | Ph.D(IT)-6 | Zun Hlaing Moe | Ph.D(IT)-17 |

# Outline

- Motivation
- Introduction
- Tacotron2
- Dataset Preparation
- Experiment
- Implementation
- Demonstration
- Conclusion

# Motivation

1. Myanmar TTS has been developed by statistical parametric speech synthesis method: HMM, DNN.

2. We focus Deep learning based end to end speech synthesis(Tactron2) with small corpus in our project.

3. To investigate end-to-end generated text-to-speech (TTS) model with syllable and word-level.

# Introduction

1. The modern TTS trend is more complex.
2. End-to-End speech synthesis is a new research direction in the deep learning area in recent years.
3. In this project, we experimented TTS by Tacotron2 which is to synthesize speech directly from the characters.

# Tacotron2

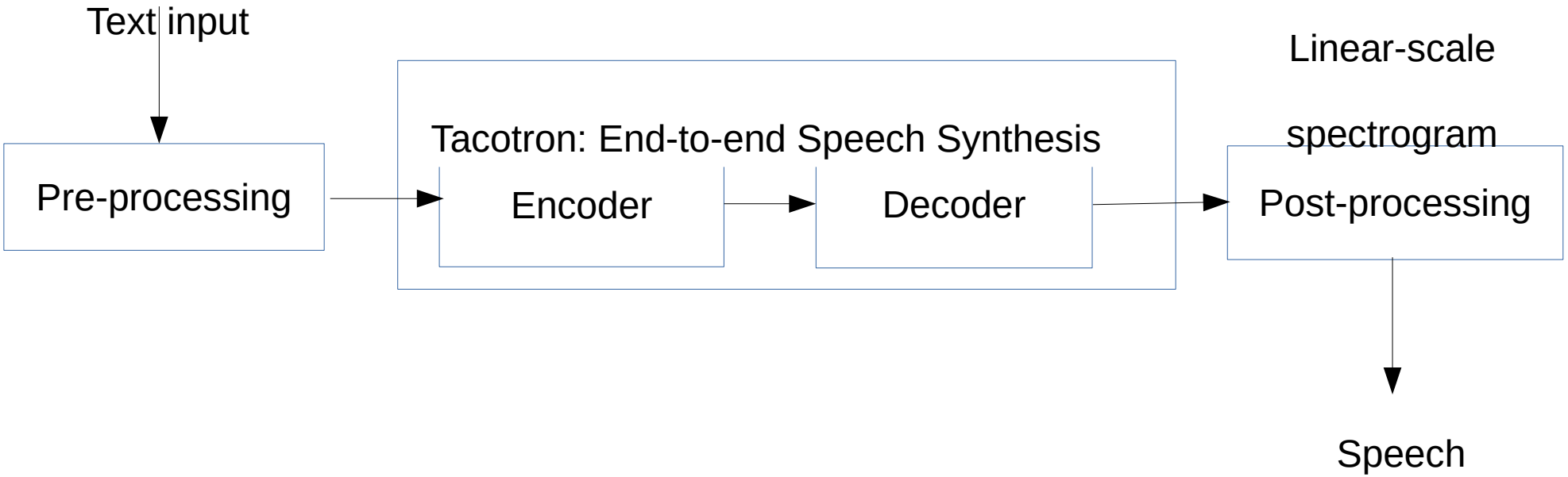


Fig1: Block Diagram of Tacotron2

# Data set preparation

1. Collected data from Grade 2 Myanmar textbook.
2. Prepared Myanmar sentences and audio files.
3. Applied input text as syllable and word-level.
4. Used a regular expression perl script developed by Saya Ye Kyaw Thu for the syllable segmentation.
5. Utilized speech data that has been recorded for Myanmar Braille TTS.

# Implementation

1. We used Tensorflow framework to implement tacotron.
2. Other requirements: numpy, scikit-learn, librosa, falcon, tqdm, matplotlib.
3. Training time is a week for ten sentences.
4. The maximum input text length is 117.

# Experiment

1. Syllable level
2. Word level

# Syllable Level

Input text sentence ဒုတိယတန်း မြန်မာ ဖတ်စာ ။

Segmented sentence ဒု တိ ယ တန်း မြန် မာ ဖတ် စာ ။

Text and audio pair for Training

ဒု တိ ယ တန်း မြန် မာ ဖတ် စာ ။,g2_0001.wav,
အ ခန်း ၁ ။ ,g2_0002.wav,

Train data with tacotron2 model.

Linear spectrogram , Mel spectrogram and alignment of encoder and decoder

# Syllable Level Test Output

Test sentence အ ခန်း တစ်

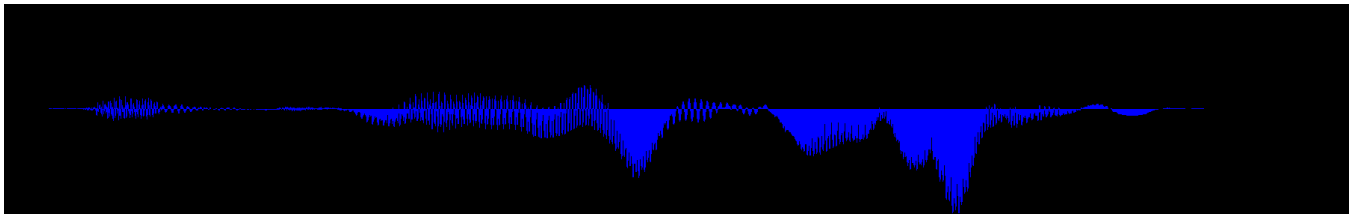

Fig2: Test output Wave form

Test sentence ဘယ် ကို သွား ခဲ့ သ လဲ ။

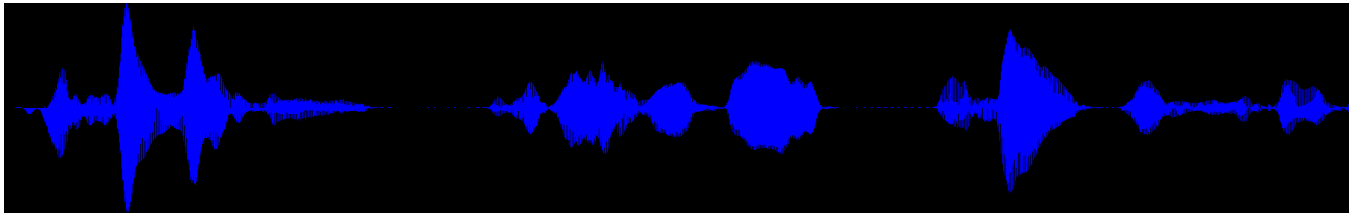

Fig3: Test output Wave form

# Experiment

1. Syllable level
2. Word level

# Word Level Test Output

Test sentence

အ ခန်း တစ်

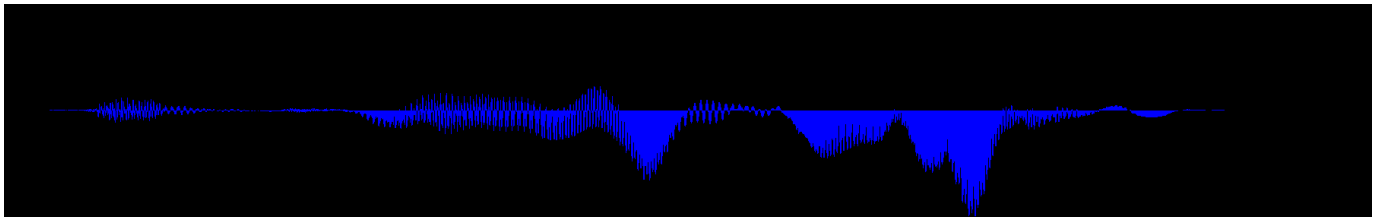

Fig4: Test output Wave form

Test sentence

ဘယ် ကို သွား ခဲ့ သ လဲ ။

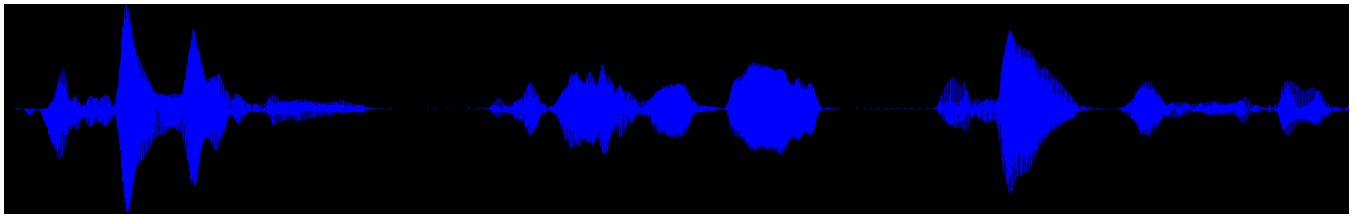

Fig5: Test output Wave form

# Evaluation

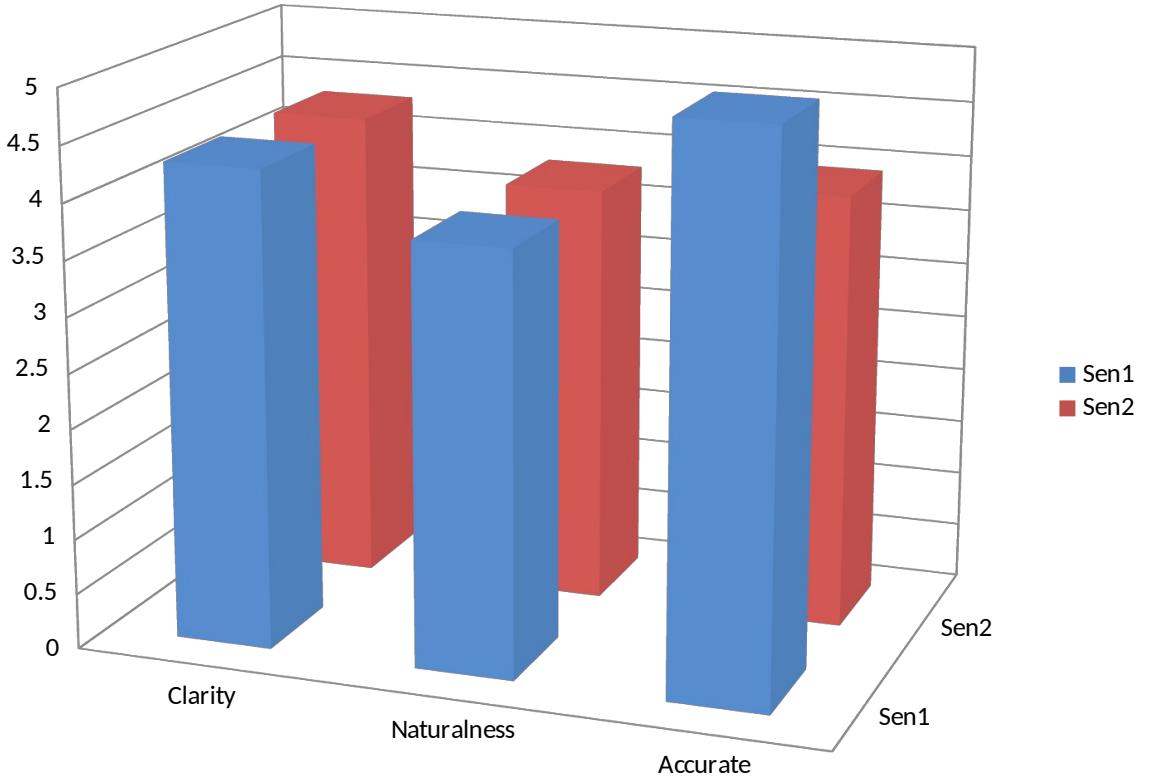


Fig6: MOS results for test sentences

# Demonstration

We created Myanmar Text to Speech User Interface for this project.


LET'S GO UI

Fig7: User Interface of this project

# Conclusion

1. This mini project was contributed the evaluation of the end-to-end speech synthesis with tacorton model.
2. We evaluated end-to-end TTS with the small corpus by using ten sentences from Grade 2 Myanmar Basic Education Textbook.
3. Although the syllable-level achieved 3.8 MOS score, the word-level achieved 3.4 on closed tests with listeners.
4. syllable-level also obtained more clearance result than word-level.
5. Good speech output depends not only on recording condition but also on tone signature.


Thank You
Thank You
Thank You!!!!