

Natural Sounding Speech with Tacotron 2

Aiden Moosa

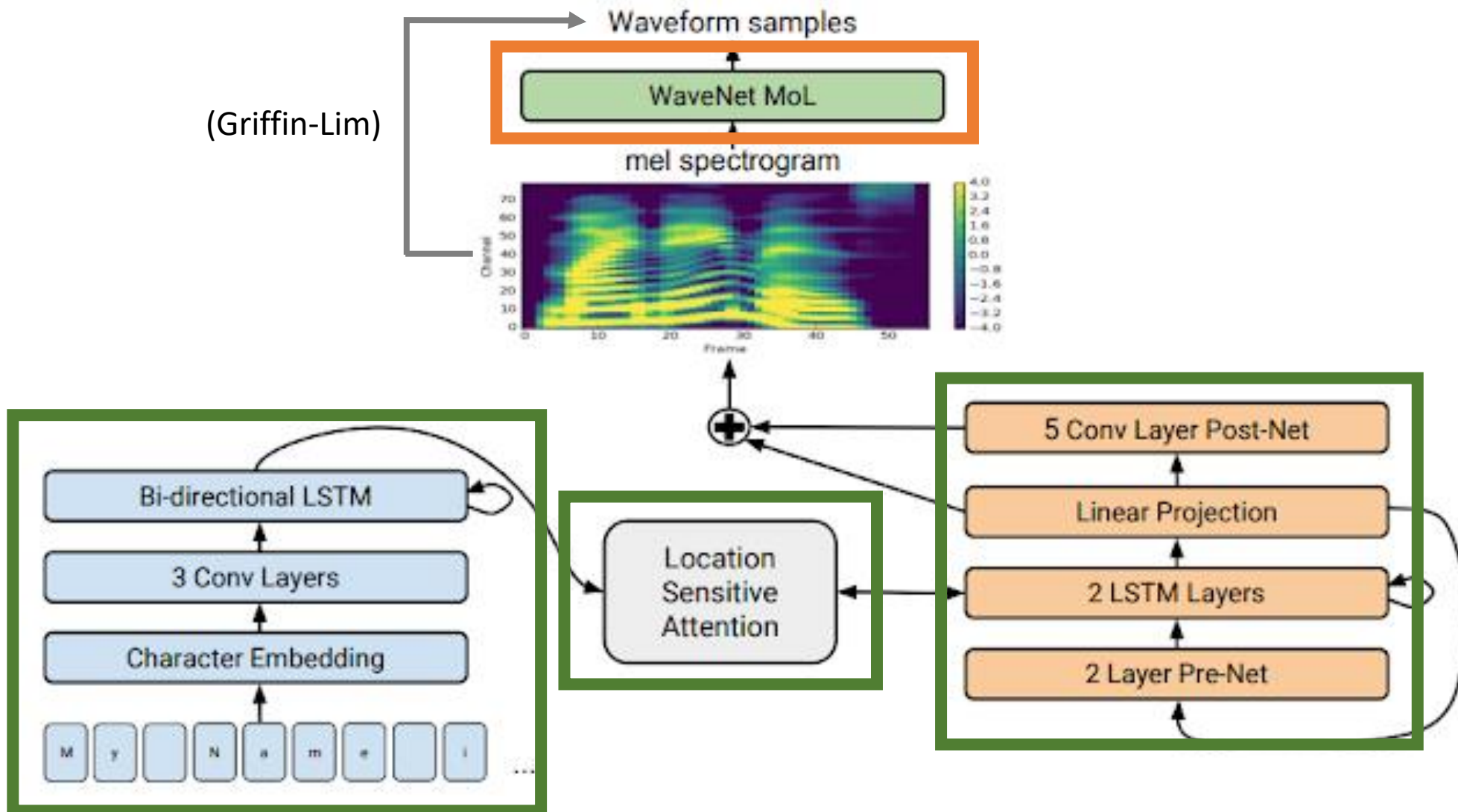
Supervisor: Benjamin James Day

Overseers: Andrew Pitts, Rafal Mantiuk

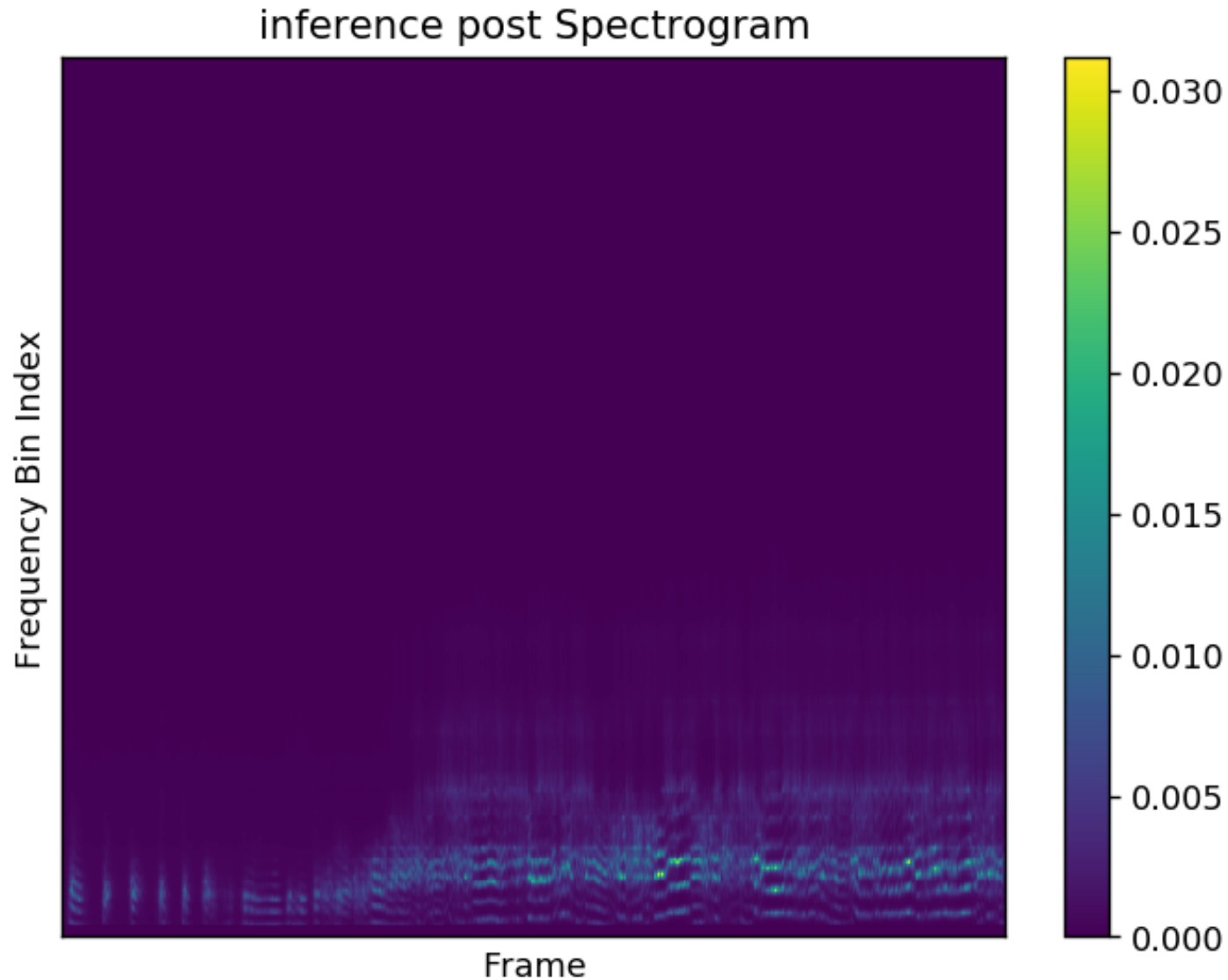
Introduction

- Newest text-to-speech system by Google
- Fully end-to-end (no linguistic/acoustic features)
- Improvements on Tacotron
- Recreation attempts somewhat successful
- Looking to more thoroughly document implementation

Progress - Implementation



Preliminary Results - Inference



“Hello, I am Tacotron 2.”



Preliminary Results - Training

Iteration 16 reference Spectrogram



Frame



Iteration 16 post Spectrogram



Frame



Iteration 2496 reference Spectrogram



Frame



Iteration 2496 post Spectrogram



Frame



Iteration 5360 reference Spectrogram



Frame



Iteration 5360 post Spectrogram



Frame

