

Institut für Maschinelle Sprachverarbeitung  
Universität Stuttgart  
Pfaffenwaldring 5B  
D-70569 Stuttgart

Master thesis

# **Phonetic Representations of Speech for Human Pronunciation Feedback and Automatic Accent Transfer**

Isaac Riley

Studiengang: M.Sc. Computational Linguistics

Prüfer\*innen: Prof. Dr. Wolfgang Wokurek  
Prof. Dr. Antje Schweitzer

Betreuer: Prof. Dr. Wolfgang Wokurek

Beginn der Arbeit: 01.04.2021

Ende der Arbeit: 01.10.2021

## **Erklärung (Statement of Authorship)**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe und dabei keine andere als die angegebene Literatur verwendet habe. Alle Zitate und sinnngemäßen Entlehnungen sind als solche unter genauer Angabe der Quelle gekennzeichnet. Die eingereichte Arbeit ist weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen. Sie ist weder vollständig noch in Teilen bereits veröffentlicht. Die beigelegte elektronische Version stimmt mit dem Druckexemplar überein.<sup>1</sup>

(Isaac Riley)

---

<sup>1</sup>Non-binding translation for convenience: This thesis is the result of my own independent work, and any material from work of others which is used either verbatim or indirectly in the text is credited to the author including details about the exact source in the text. This work has not been part of any other previous examination, neither completely nor in parts. It has neither completely nor partially been published before. The submitted electronic version is identical to this print version.

## Acknowledgments

XXX

# Inhaltsverzeichnis

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Background</b>	<b>4</b>
<b>3</b>	<b>Related Work</b>	<b>4</b>
<b>4</b>	<b>Resources</b>	<b>4</b>
<b>5</b>	<b>Methods</b>	<b>5</b>
5.1	Phonetic Representations . . . . .	5
5.2	Automatic Pronunciation Feedback . . . . .	5
5.3	Accent Transfer in Phonetic Representation Space . . . . .	5
5.4	Speech Synthesis from Phonetic Representations . . . . .	5
<b>6</b>	<b>Results</b>	<b>5</b>
<b>7</b>	<b>Conclusions</b>	<b>5</b>

# **1 Introduction**

Accent

## **2 Background**

## **3 Related Work**

## **4 Resources**

## **5 Methods**

### **5.1 Phonetic Representations**

To learn representations of input speech samples, I use a neural network, or rather a family of similar neural networks all sharing the same sequence-to-sequence architecture. This architecture consists of an encoder, a decoder, and an attention mechanism, each of which will be discussed in greater detail below.

### **5.2 Automatic Pronunciation Feedback**

### **5.3 Accent Transfer in Phonetic Representation Space**

Text here.

### **5.4 Speech Synthesis from Phonetic Representations**

Text here.

## **6 Results**

## **7 Conclusions**

## **Literatur**