

UNIVERSITÉ DE LIÈGE

FACULTÉ DES SCIENCES APPLIQUÉES

---

# Automatic Voice Cloning Across Languages

---

Author:

Corentin JEMINE

Supervisor:

Prof. Gilles LOUPPE

Academic year 2018 - 2019



*Graduation studies conducted for obtaining the Master's degree  
in Data Science by Corentin Jemine*

# 1 Abstract

To do when I'll have a good overview of the project. Try to answer:

- What is the goal of the application? What are its requirements, what is the setting, what kind of data are we going to use it on?
- What is zero-shot voice cloning? How does it fit in here (difference between an online and offline approach)?
- What are the particularities of our implementation (both model and datasets), what are its upsides and downsides (for example: requires huge datasets but fast inference)?
- What did we ultimately achieve? How good are our results?

# 2 Introduction

Concise presentation of the problem

Methods before deep learning

Wavenet

Tacotron

Deep voice (1, 2, 3 + few samples) Tacotron 2

SV2TTS

Extensions?