## Motivation

## weaknesses

## 





Mismatch.

### Cannot

### $\prod$ Γ

### real-world









Difficult





Limited

professional

## performance

# model.

super

# distortion



# handled

## knowledge



# **VoiceFixer**

## real-world

### ı



### T

Hiah











# dimensionality

### resolution

# **Spectrogram**













### percept

### )ار

### environment?



Interference from other sound sources love eating donuts!

Sound Source

Acoustic Transfer function

Sensory Input

What did I heard?



|aɪ lʌ... 'iːtɪ... 'd...ʊ...s|

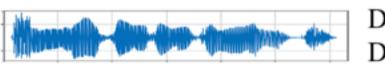
What did he say?

- Multimodality information.
- Linguistic knowledge.
- Common senses.

|aɪ lʌv ˈiːtɪŋ ˈdəʊnʌts|

What was it sound like? What does he imply?

- Speaker age, emotion, sex.
- Acoustic environment.
- Phonetic and other prior knowledges.



Donut is a kind of food. Donut is delicious.

Bilateral Anterior STG

Primary Auditory Cortex → Planum Temporale (PT) (PAC)

**Initial Processing** 

Identification, segregation and matching onto previously learnt spectemporal representations

**Auditory Spatial Analysis** 

Left Posterior STG Left Inferior Frontal Gyrus

Imaginery and Comprehension

**High Level Perception** 





# knowledge











