



# Reconstructing voice identity from fMRI of the temporal voice areas

Charly Lamothe (INT, LIS)

Supervisors: **Prof. P. Belin** (INT, Dept of Psych, Univ of Montréal) & **Prof. T. Artières** (LIS, ECM)

Collaborators: *Dr. BL. Giordano* (INT, INP), *Dr. S. Ayache* (LIS), *Dr. E. Thoret* (ILCB, PRISM), *Dr. R. Trapeau* (INT), *Dr. S. Takerkart* (INT), *PhD M. Obliger* (INT), *Dr. R. Marxer* (LIS)





1/17

#### Background

#### ERC-2017-ADG-788240: A Comparative Study of Voice Perception in Primates







\*\*\*\*

pca

vls

50

vls

50

pca

4/17

\*\*\*\*

pca

vls

0

## **Voice latent space properties**



# Auditory fMRI responses to voice identity



(Pernet et al., 2015)

Stimuli were balanced

gender, age, language (#8) and speaker's identity (~140).

٠

in

6/17

# Question



We first asked whether a linear relationship could be established between the voice latent space and fMRI activity patterns.

## Encoding of auditory fMRI responses to voice identity with voice latent space



# Question



We then asked whether the voice latent space represent dissimilarities like the brain responses do.

# Representational geometries of auditory fMRI responses to voice identity with voice latent space



# Question



We asked whether this linear mapping could reconstruct voice.

## **Reconstructed voices and spectrograms**



#### **Objective evaluation of the reconstructed voices**





- Gender best decoded from A1;
- Age and identity best decoded from TVAs;
- VLS > PCA;
- TVAs > A1.

- Behavioral experiment. subjects on Prolific listening to the original and reconstructed voices:
  - Task 1 Loudness change detection ~3 min
  - Task 2 Speaker discrimination ~2h
  - Task 3 Naturalness judgment ~30 min
  - Task 4 Gender categorization ~30 min
  - Task 5 Age categorization ~30 min
- #9 subjects

#### Subjective evaluation of the reconstructed voices



- We proposed a simple encoding-decoding pipeline that allows to reconstruct voice from non-invasive measures;
- Our 'voice latent space' seems to share encoding similarities with the voice processing system, to a greater extend than the baseline PCA;
- We showed that the 'voice quality' of our model's reconstructions is better preserved than for the PCA;
- ➢ In general, the VLS is closer to the TVAs encoding than to A1.

- Finalizing the paper ;
- Investigate the temporal dimension with MEG acquisition. 2 subjects listened to 7h of short voice samples.

Thank you!

# Preprocess auditory fMRI responses to voice identity

