Stage 2024, Master 2

Correlation of mouth size to dynamic voice directivities

Place: LAM – d'Alembert, Sorbonne Université, 4, place Jussieu 75005, Paris
Duration: 5 months
Internship Director: Samuel D. Bellows (LAM)
Codirector: Brian F.G. Katz (LAM)
Renumeration: The monthly payment is fixed by CNRS at 550€.

Context:
Sound radiation from the human voice has applications in telecommunications, architectural acoustics, and auralizations. One of the primary difficulties of voice directivity analysis is that speech is inherently a time-varying process. As a result, most analysis of voice radiation are grounded in analysis of time-averaged, phonetically-balanced speech or steady-state phonemes. These conditions, however, ignore important dynamic speech characteristics such as co-articulation.

Recent measurements of voice directivity have employed both a 24-channel rotating microphone and a 180-channel MEMS microphone to capture spoken and sung voice radiation. Additionally, synchronized video recordings provide a means to evaluate the effect of dynamically varying mouth shape on the resultant voice directivity.

Subject of Internship:
The aim of the present internship is to create an audio-visual analysis of dynamic voice directivity to better understand voice radiation in realistic conditions. During the internship, the intern is expected to either make use of open-source computer vision tools to extract time-varying mouth sizes or to develop a simple machine-learning-based algorithm. The data will then be combined with dynamic voice directivites to study correlations between mouth opening size and sound source directivity.

Required Skills:
MATLAB or Python programming skills. Some experience in machine learning, particularly in image recognition, would be beneficial.

Contact:
Samuel D. Bellows: samuel.bellows@dalembert.upmc.fr

References:
[1] Brian F. G. Katz, Fabien Prezat, and Christophe d'Alessandro. Human voice phoneme directivity pattern measurements. J. Acoust. Soc. Am., 120(5 Supplement):3359–3359, 11 2006.
[2] Brian FG Katz and Christophe D'Alessandro. Directivity measurements of the singing voice. In Proceedings of the 19th International Congress on Acoustics, Madrid, 2007.
[3] Timothy W. Leishman, Samuel D. Bellows, Claire M. Pincock, and Jennifer K. Whiting. High-resolution spherical directivity of live speech from a multiple-capture transfer function method. J. Acoust. Soc. Am., 149(3):1507–1523, Mar. 2021.