



HAL
open science

Mesures de directivité de sources non répétables avec un réseau de 960 microphones MEMS: application à la voix humaine

Paul Luizard, Matthieu Hartenstein, Hélène Moingeon, Cédric Pinhède, Marc Pachebat, Christian Ollivon, Francois Ollivier, Fabrice Silva

► **To cite this version:**

Paul Luizard, Matthieu Hartenstein, Hélène Moingeon, Cédric Pinhède, Marc Pachebat, et al.. Mesures de directivité de sources non répétables avec un réseau de 960 microphones MEMS: application à la voix humaine. CFA 2025 - 17e Congrès Français d'Acoustique, Société Française d'Acoustique (SFA), Apr 2025, Paris, France. <hal-05365400>

HAL Id: hal-05365400

<https://inria.hal.science/hal-05365400v1>

Submitted on 11 Feb 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



Mesures de directivité de sources non répétables avec un réseau de 960 microphones MEMS: application à la voix humaine

P. Luizard^{1,3,4}, M. Hartenstein², H. Moingeon², C. Pinhède¹,
M. Pachebat¹, C. Ollivon², F. Ollivier², F. Silva¹

¹Aix Marseille Univ, CNRS, Centrale Marseille, LMA, Marseille, France

²Sorbonne Université, CNRS, UMR 7190, Institut Jean Le Rond d'Alembert, Paris, France

³Audio Communication Group, Technische Universität Berlin, Germany

⁴Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France

Projet ANR **RayoVox** (2021-2026, collab. LMA – IJLRDA)

Caractérisation du rayonnement de la voix humaine : du conduit vocal au champ lointain



Introduction et motivation

Mesures de directivité de source sonore dans l'espace: deux approches possibles

- Approche séquentielle : nombre réduit de microphones mais nécessite répétitions [1,2,3]
- Approche unique : grand nombre de microphones, sans répétition [4,5,6]

Microphones MEMS abordables et suffisamment sensibles pour utilisation efficace en grand nombre

→ Dans cette étude, réalisation unique du champ acoustique capturé par 960 microphones

1. Description du système (architecture, capteurs)
2. Étalonnage acoustique (niveau sonore, phase)
3. Étalonnage spatial (position des capteurs)
4. Exemple de résultats pour un locuteur

[1] Trendelenburg, F. (1929). "Contribution to the question of voice directivity," Z. Techn. Phys. 11, 558–563.

[2] Katz, B. F., Prezati, F., & d'Alessandro, C. (2006). Human voice phoneme directivity pattern measurements. In 4th Joint Meeting of the ASA and the ASJ (p. 3359).

[3] Chu, W. T., & Warnock, A. C. C. (2002). Detailed directivity of sound fields around human talkers. Report B3144, 6.

[4] Monson, B. B., Hunter, E. J., & Story, B. H. (2012). Horizontal directivity of low-and high-frequency energy in speech and singing. JASA, 132(1), 433-441.

[5] Pörschmann, C., & Arend, J. M. (2021). Investigating phoneme-dependencies of spherical voice directivity patterns. JASA, 149(6), 4553-4564.

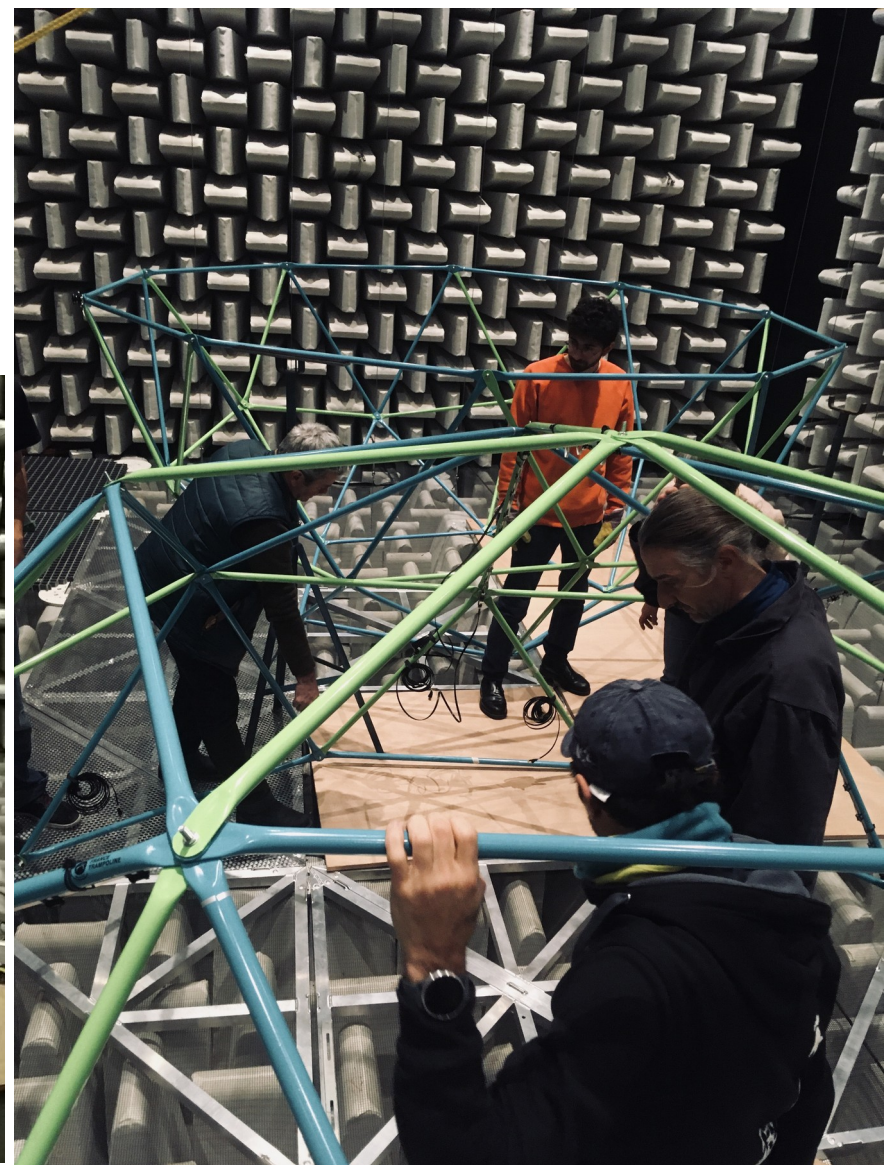
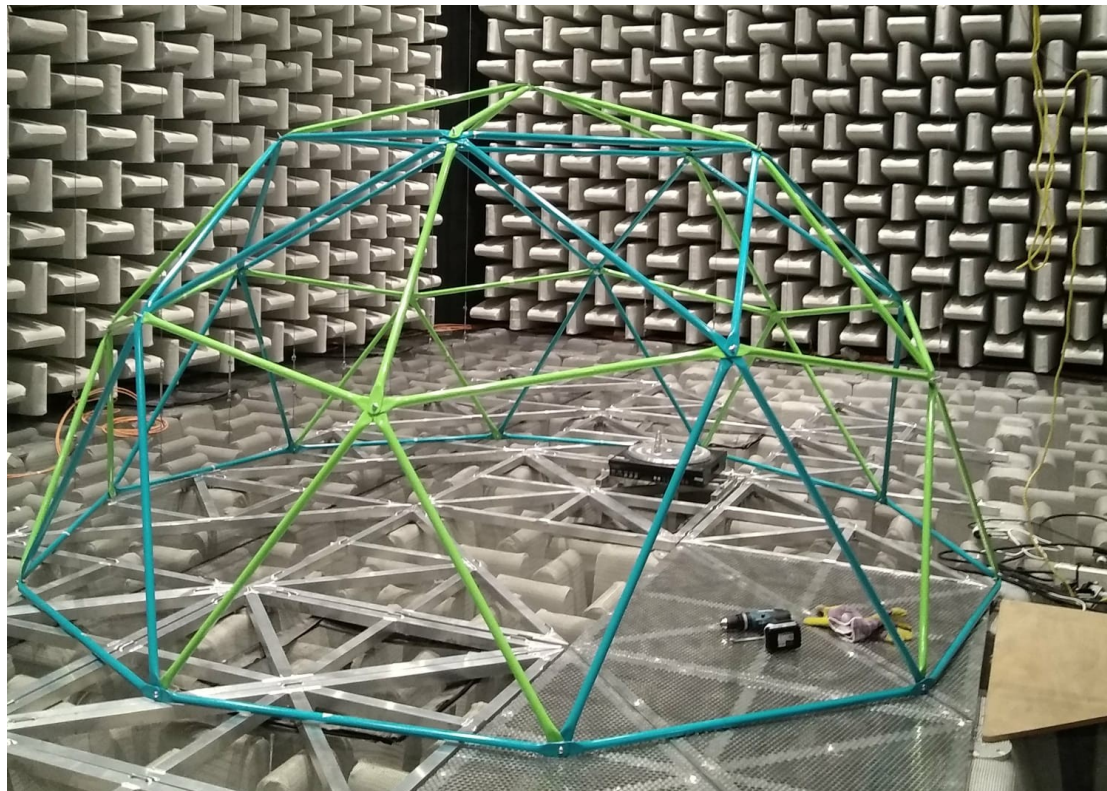
[6] Brandner, M., Frank, M., & Sontacchi, A. (2022). Horizontal and vertical voice directivity characteristics of sung vowels in classical singing. In Acoustics (Vol. 4, No. 4, pp. 849-866). MDPI.

Structure sphérique

Rayon : 1.80 m

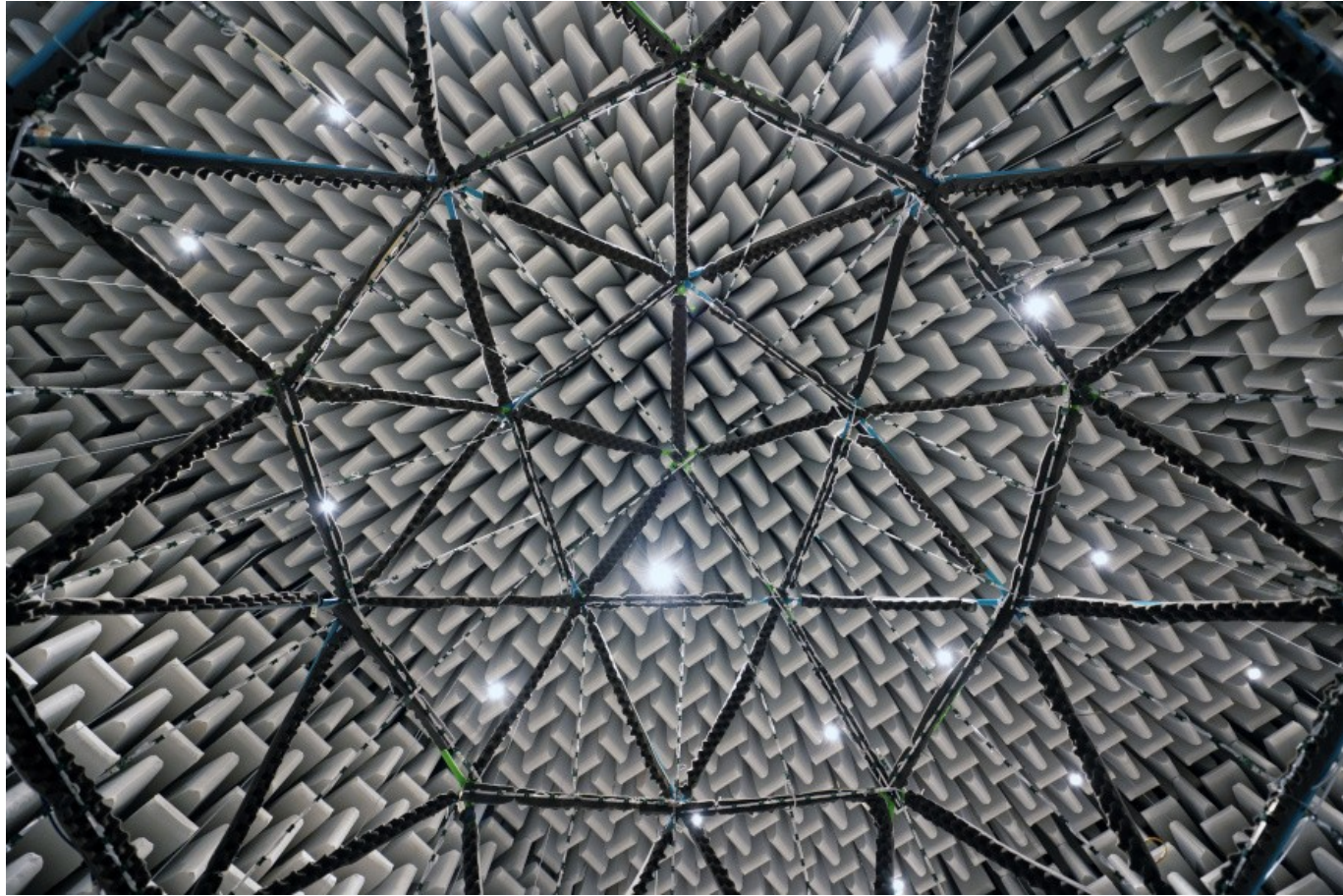
Barres en acier

- Longueur : 1.168 m (bleu) et 1.024 m (vert)
- Diamètre : 25 mm



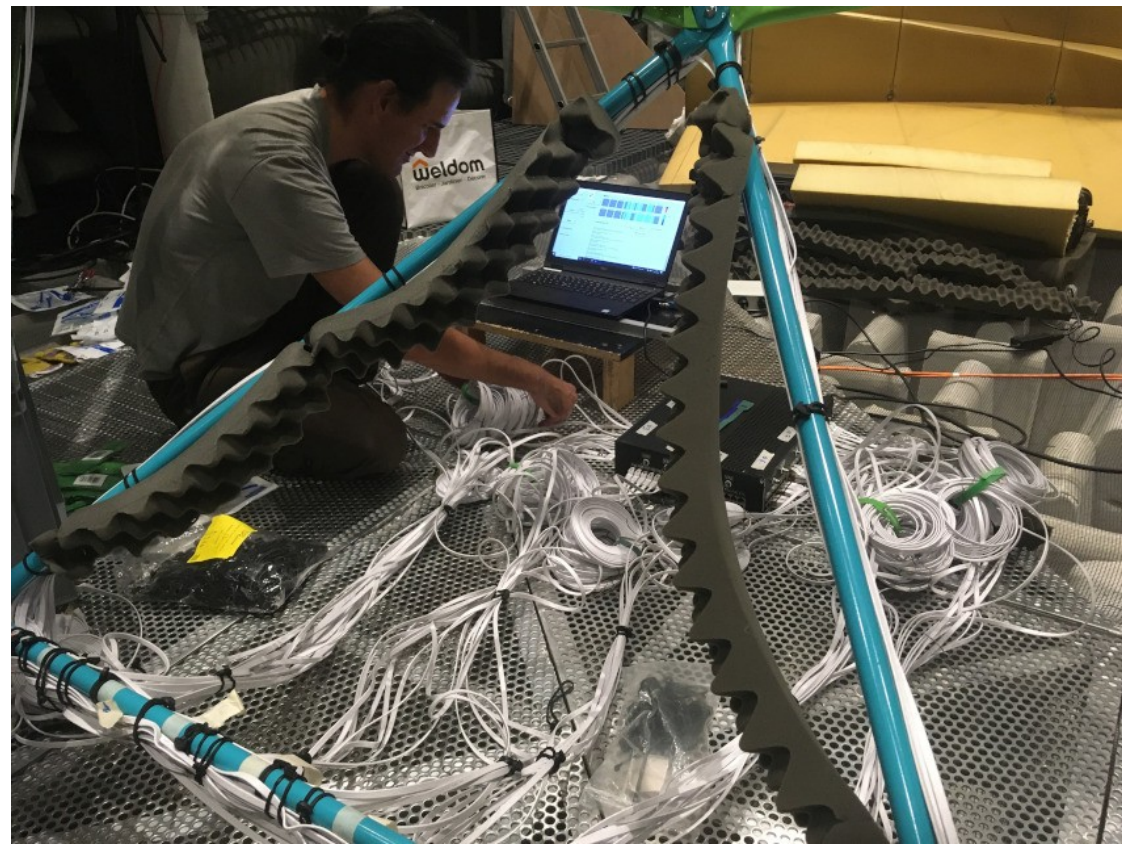
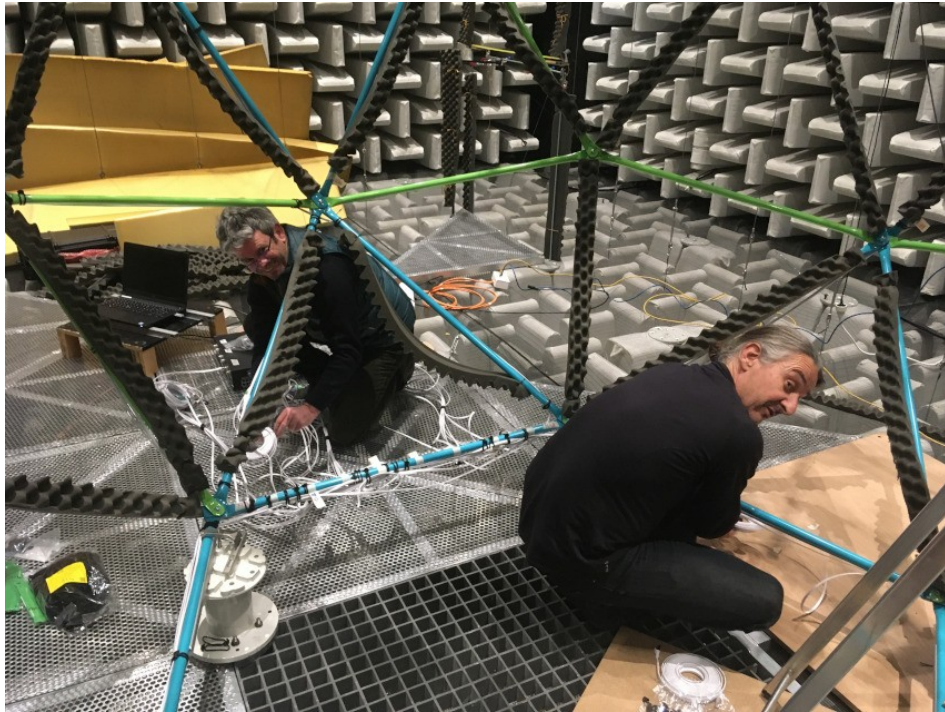
Structure sphérique

Placement des capteurs...



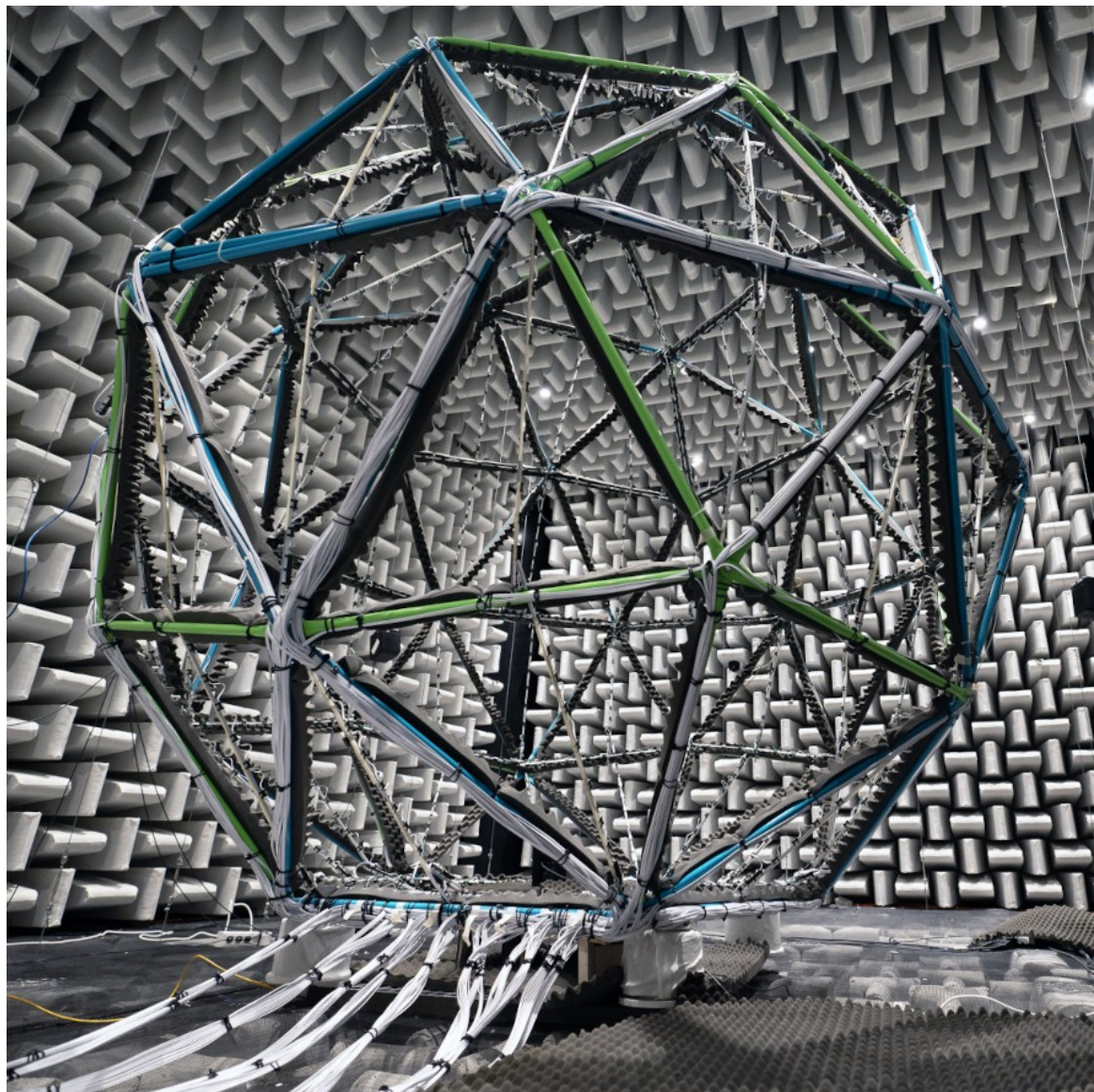
Structure sphérique

Câblage des capteurs...



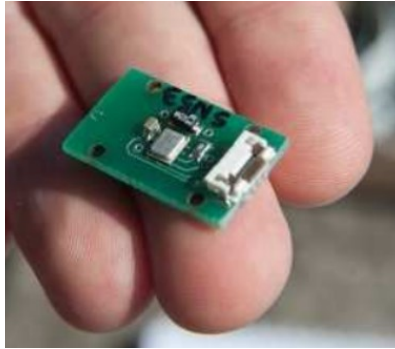
Structure sphérique

Le résultat final

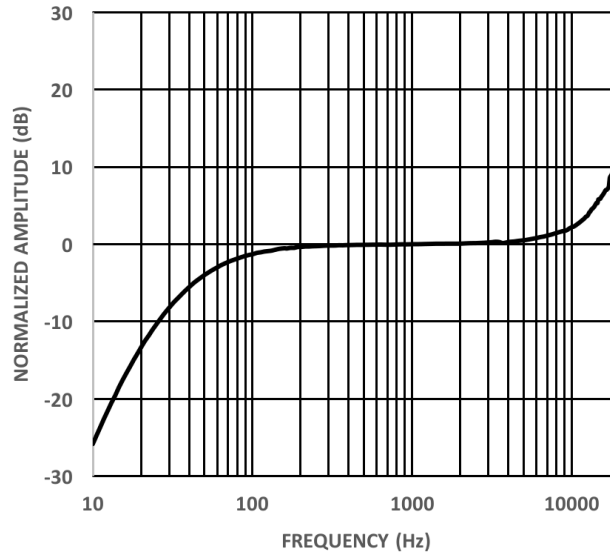


Microphones MEMS

Capteur MEMS :
Constructeur: InvenSense
Modèle: ICS-43434

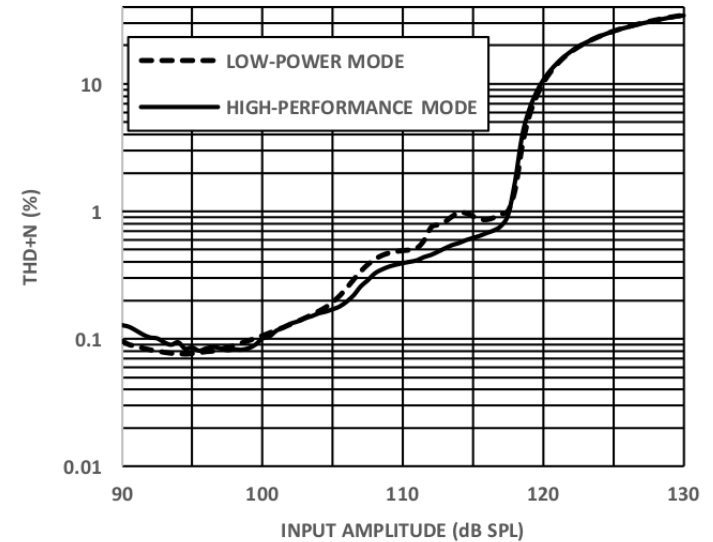


Réponse en fréquence
(fournie par le constructeur)



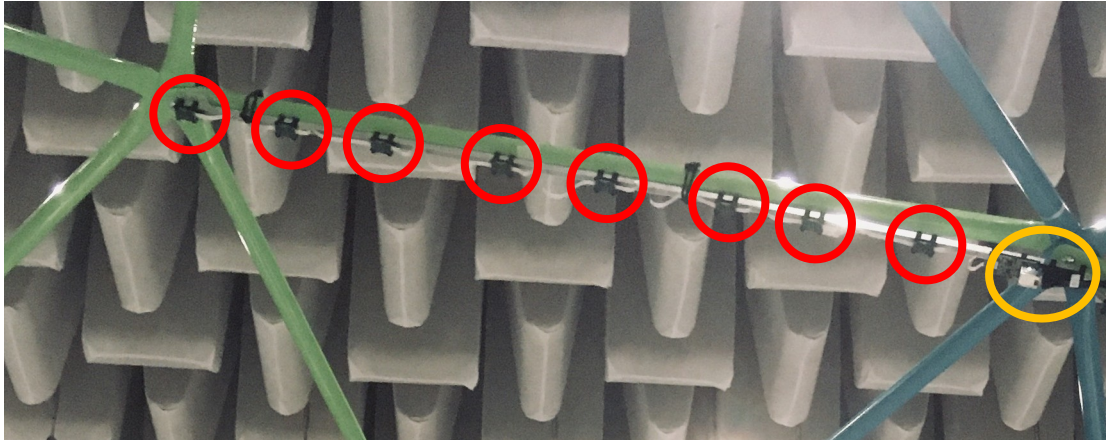
SPEC	HIGH PERFORMANCE MODE
Sensitivity	-26 dB FS \pm 1 dB
SNR	65 dBA
Current	490 μ A
AOP	120 dB SPL
Sample Rate	23 – 51.6 kHz

Distorsion harmonique (THD)
(fournie par le constructeur)



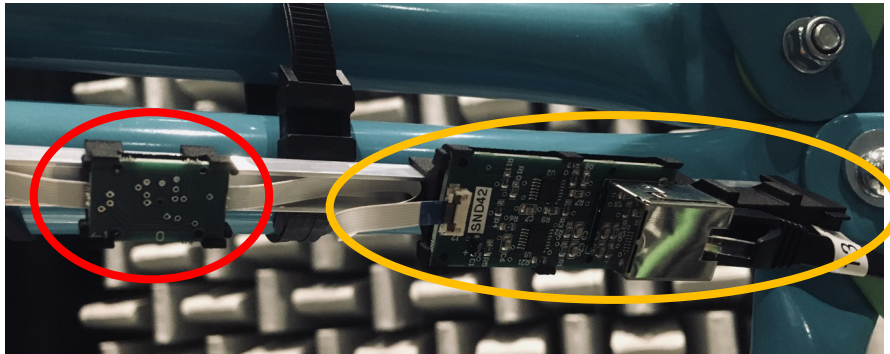
Microphones MEMS

8 MEMS sur bar rigide



MEMS individuels

Prise câble réseau (RJ45)



Concentrateur "MegaMicro"
(développé à Sorbonne Université)
Entrée : 120 câbles réseau (RJ45)
Sortie: 1 single câble USB3



120 barres de 8 MEMS = 960 MEMS

Étalonnage acoustique (niveau et phase)

Objectif : estimer la variabilité des microphones MEMS

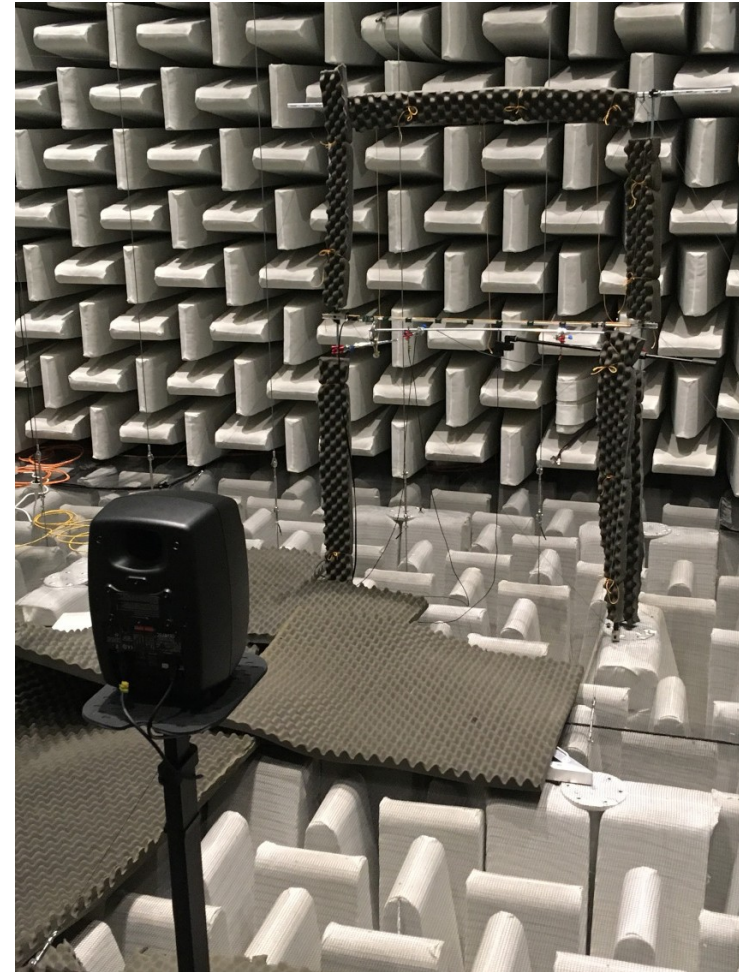
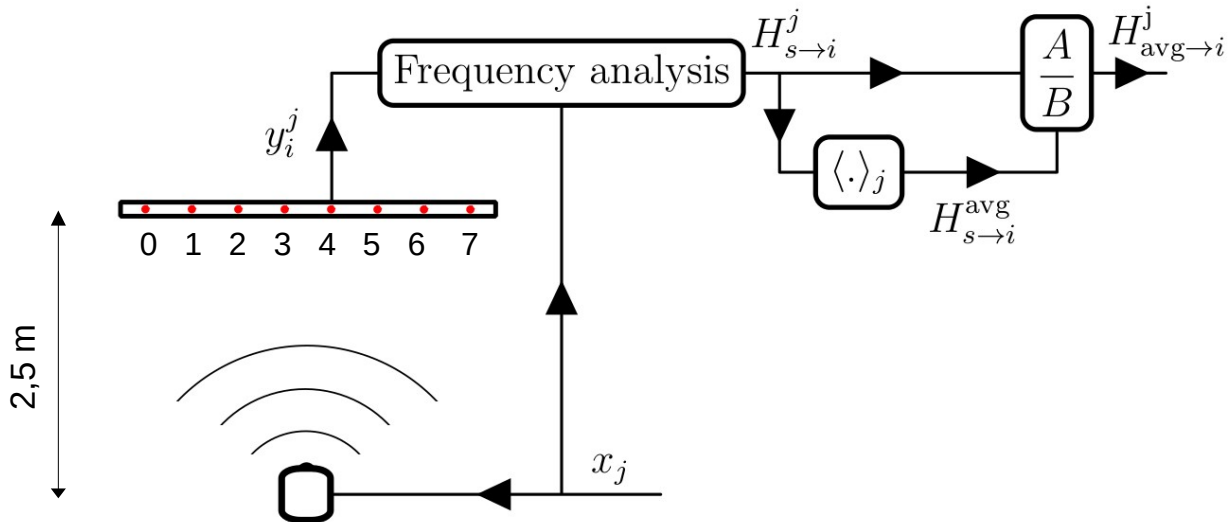
Source de référence : Genelec 8240A, sweep (30 s, 100 Hz-10 kHz)

Fonctions de transfert

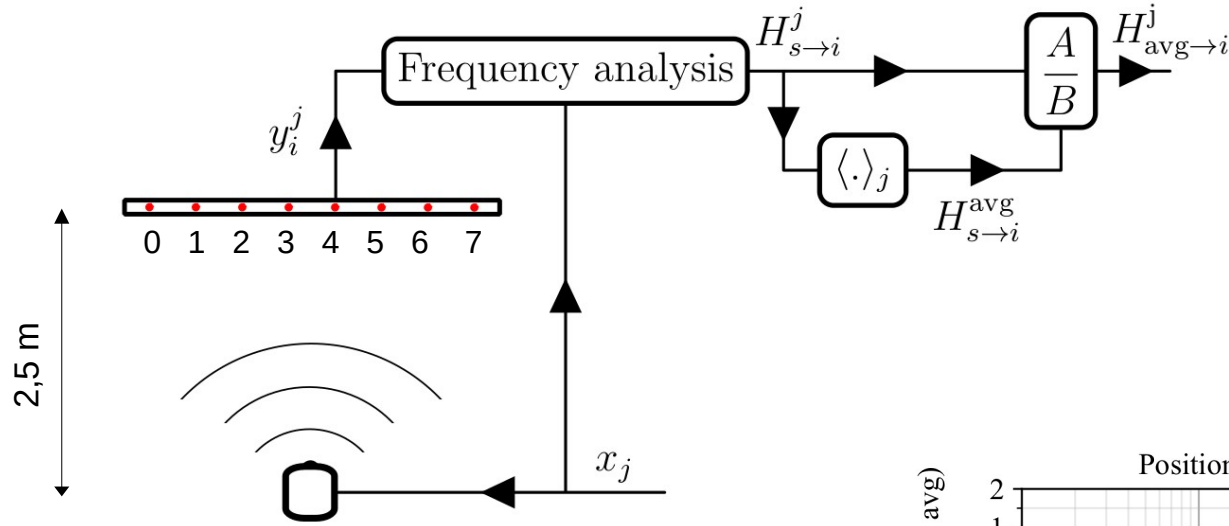
$$H = P_{xy} / P_x$$

P_{xy} : interspectre

P_x : autospectre

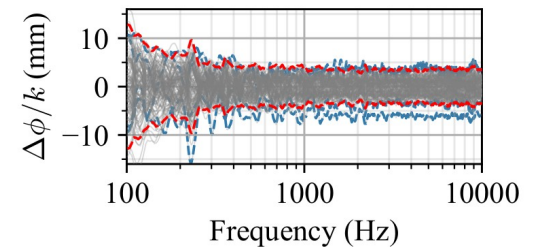
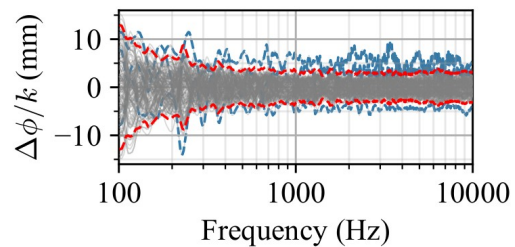
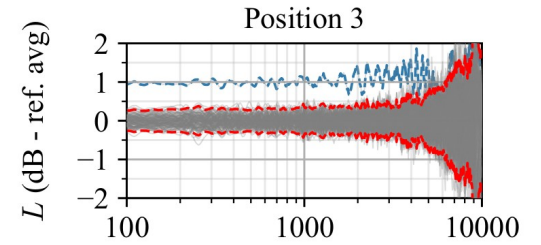
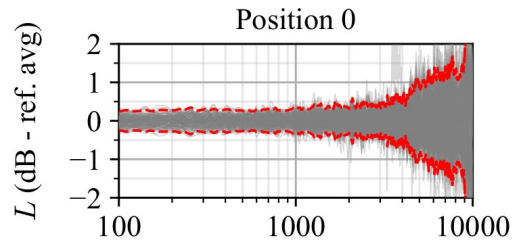


Étalonnage acoustique (niveau et phase)



Fonctions de réponse en fréquence des MEMS

- Rouge : std
- Bleu : MEMS écarté (réponse > 2 std)



Étalonnage spatial

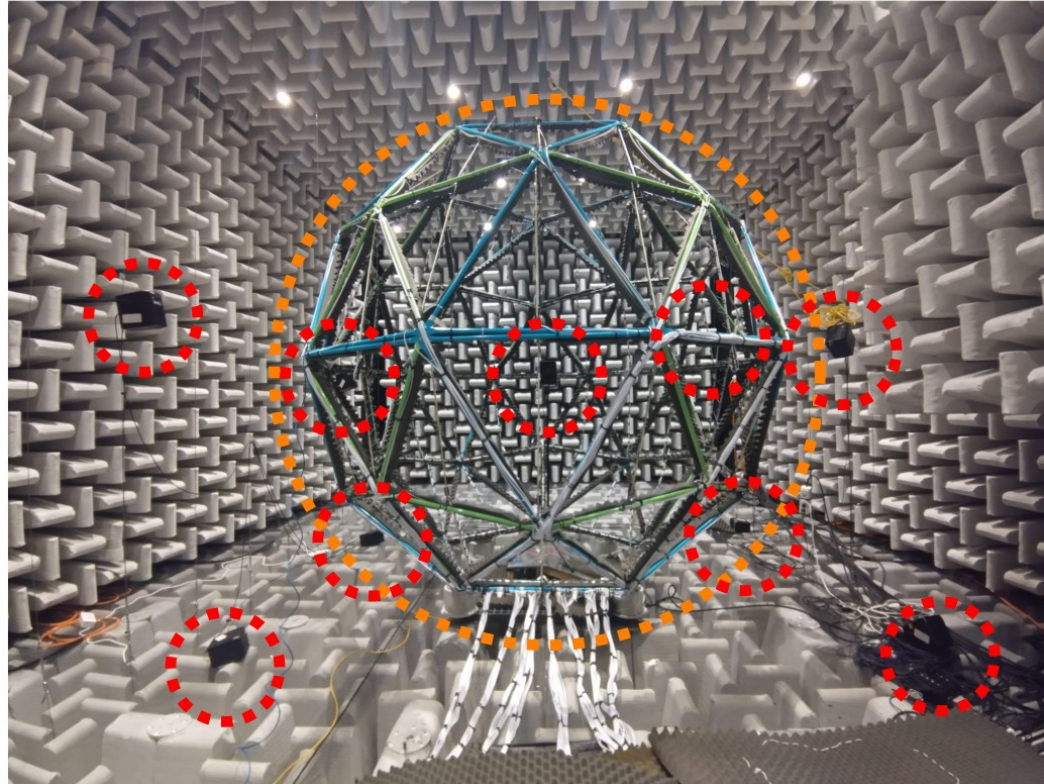
Objectif : estimer la position de chaque capteur

Procédure : Sweep (30 s, 100 Hz-10 kHz) joué par chaque source

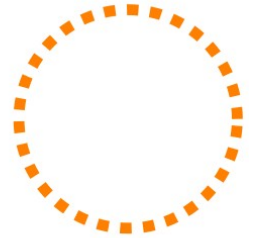
8 sources externes
1 source au centre

Algorithme RMDU [7]

- Basé sur différence de temps de vol entre chaque source et chaque microphone
- Mapping entre distances et positions



Réseau de
microphones

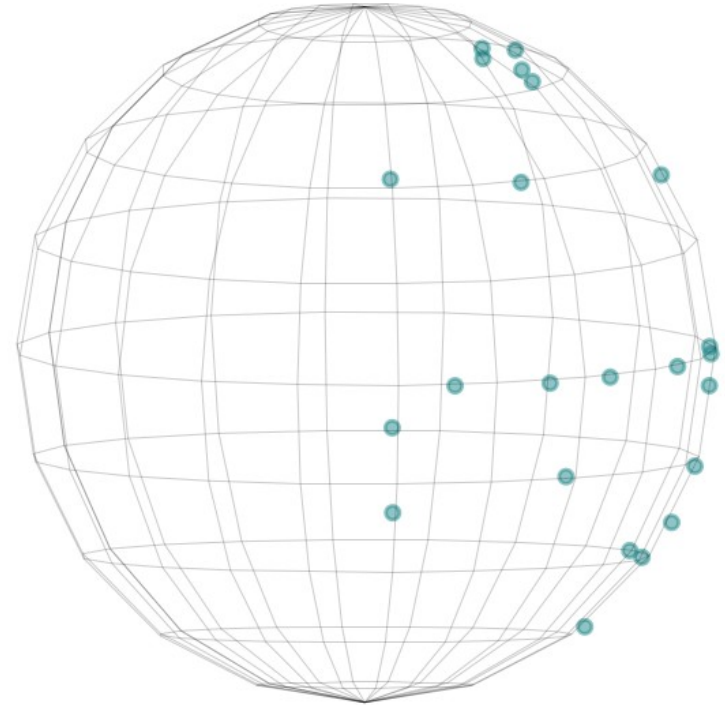
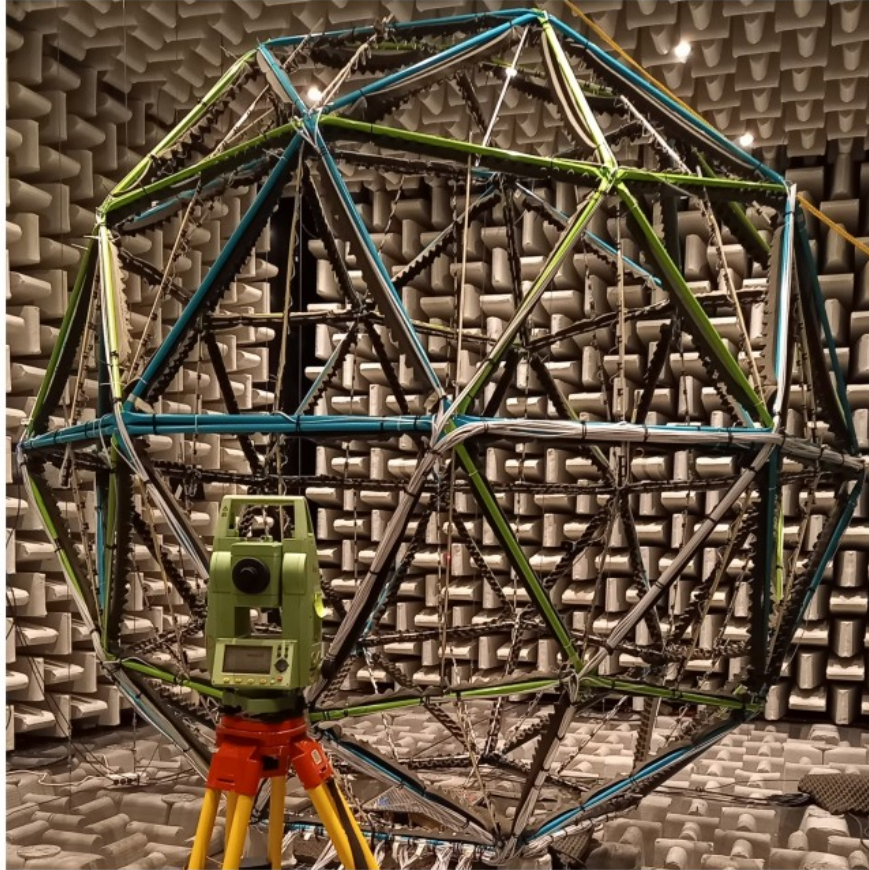


Sources
sonores



Étalonnage spatial

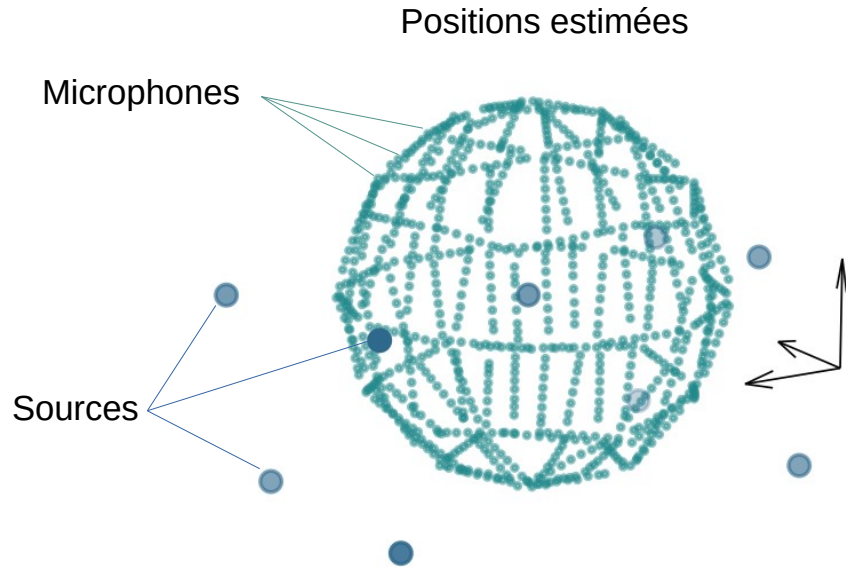
Comparaison à des mesures de référence au théodolite



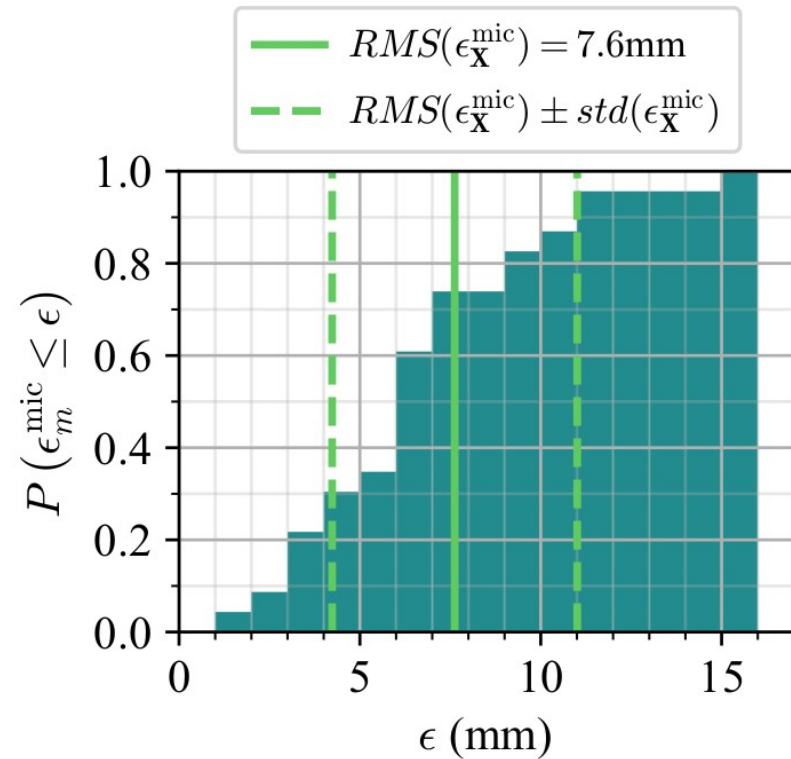
Étalonnage spatial

Résultats et erreur associée

Précision spatiale 7.6 mm ($\sim 0,2^\circ$ vu depuis 1,8 m)



Densité cumulée de l'erreur entre estimation et mesures sur 23 points



Application à la voix humaine

Tâche demandée au locuteur :

- Glissando sur environ 2 octaves (plus grave au plus aigu)
- Phonème tenu (5 voyelles, 5 consonnes) pendant 3 secondes
- Série de phrases phonétiquement équilibrées
- Petite chanson

[Mesures préliminaires, exploratoires]



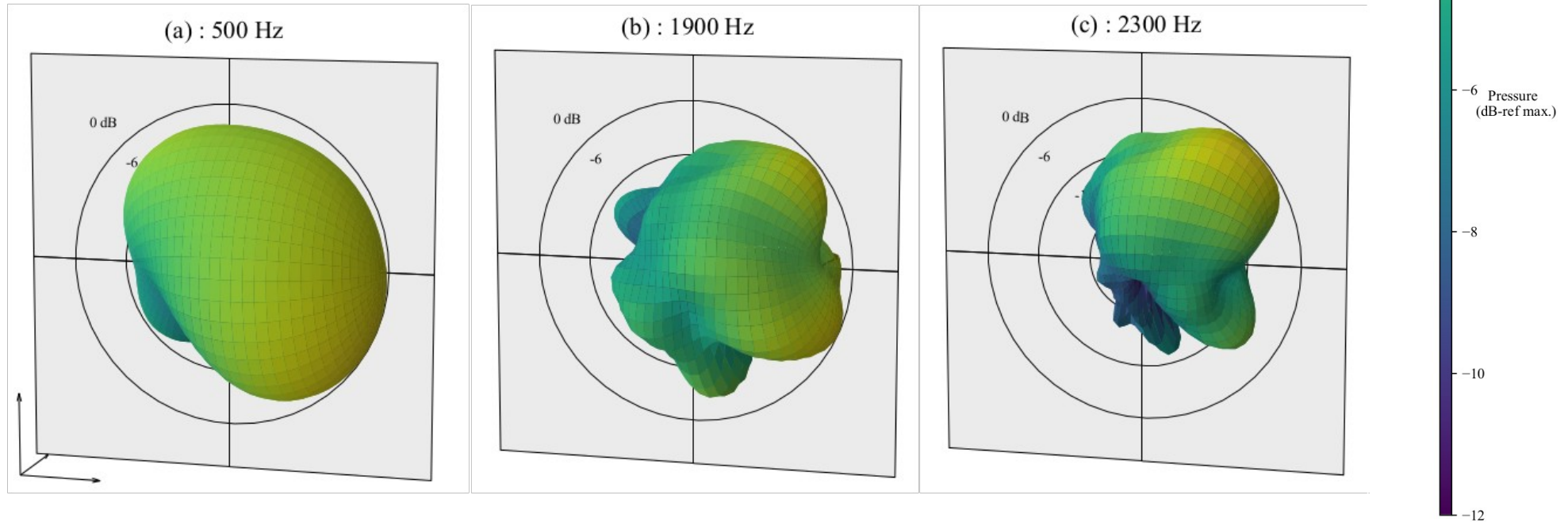
Application à la voix humaine

Ballons de directivité

1 locuteur

Voyelle /a/, glissando sur environ 2 octaves

(Long term average spectrum – LTAS, pour chaque microphone)



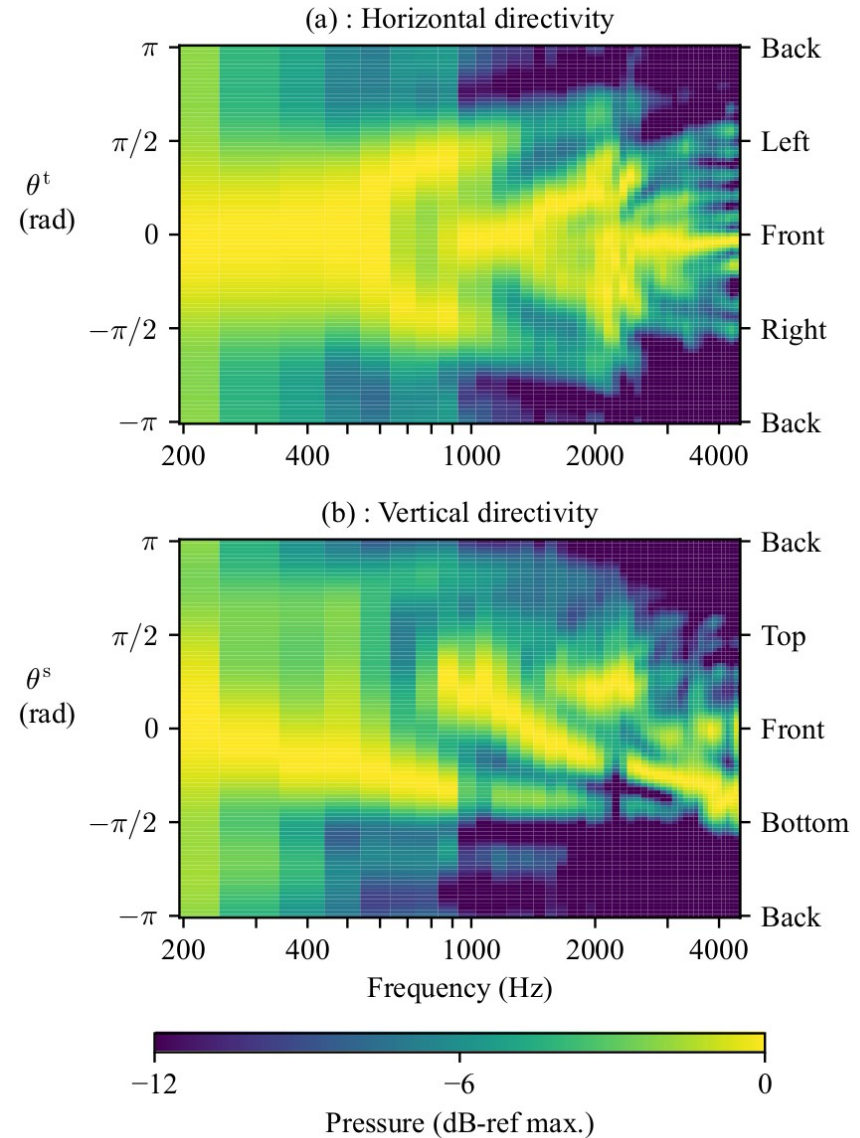
Application à la voix humaine

Représentation dans le plan horizontal et vertical

Figures d'interférences visibles

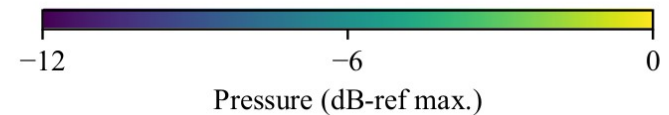
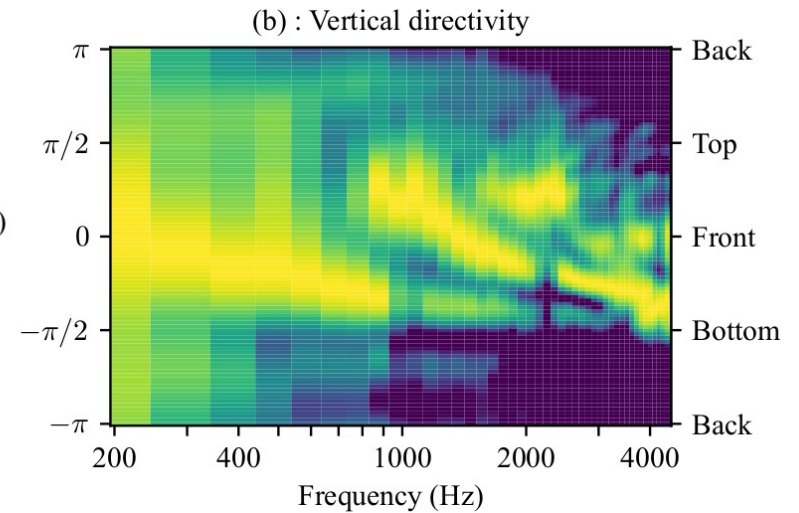
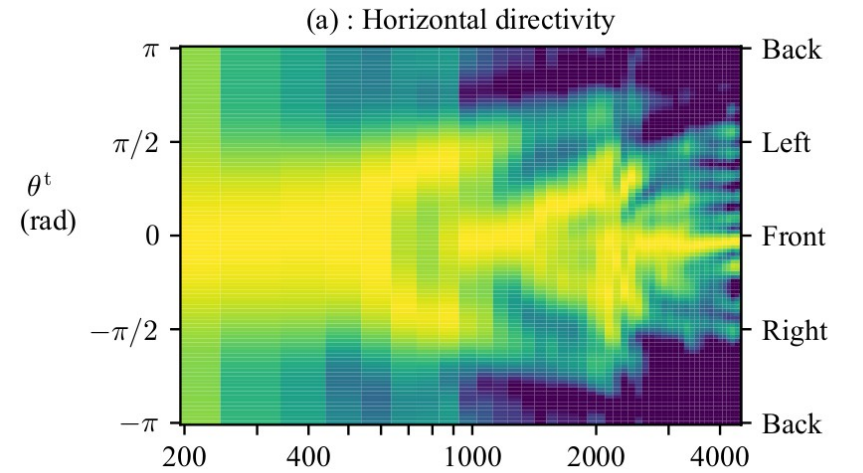
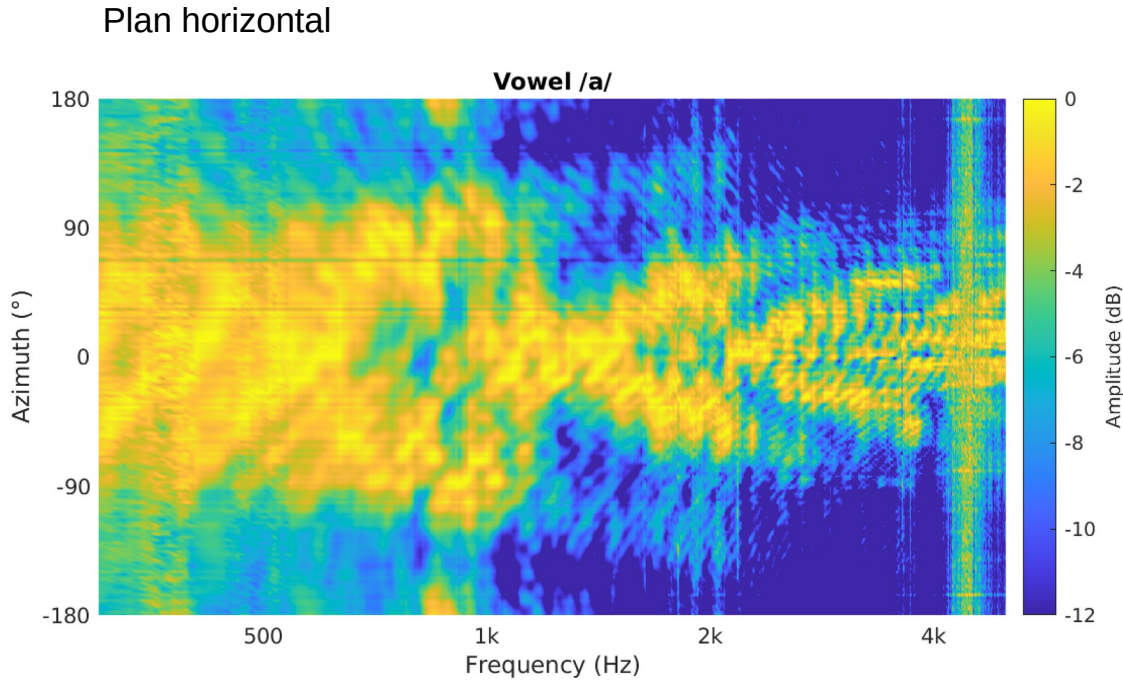
Horizontal : Symétrie droite-gauche pour $f < 2$ kHz

Vertical : Énergie plutôt vers le bas (sauf $1 < f < 1,5$ kHz)



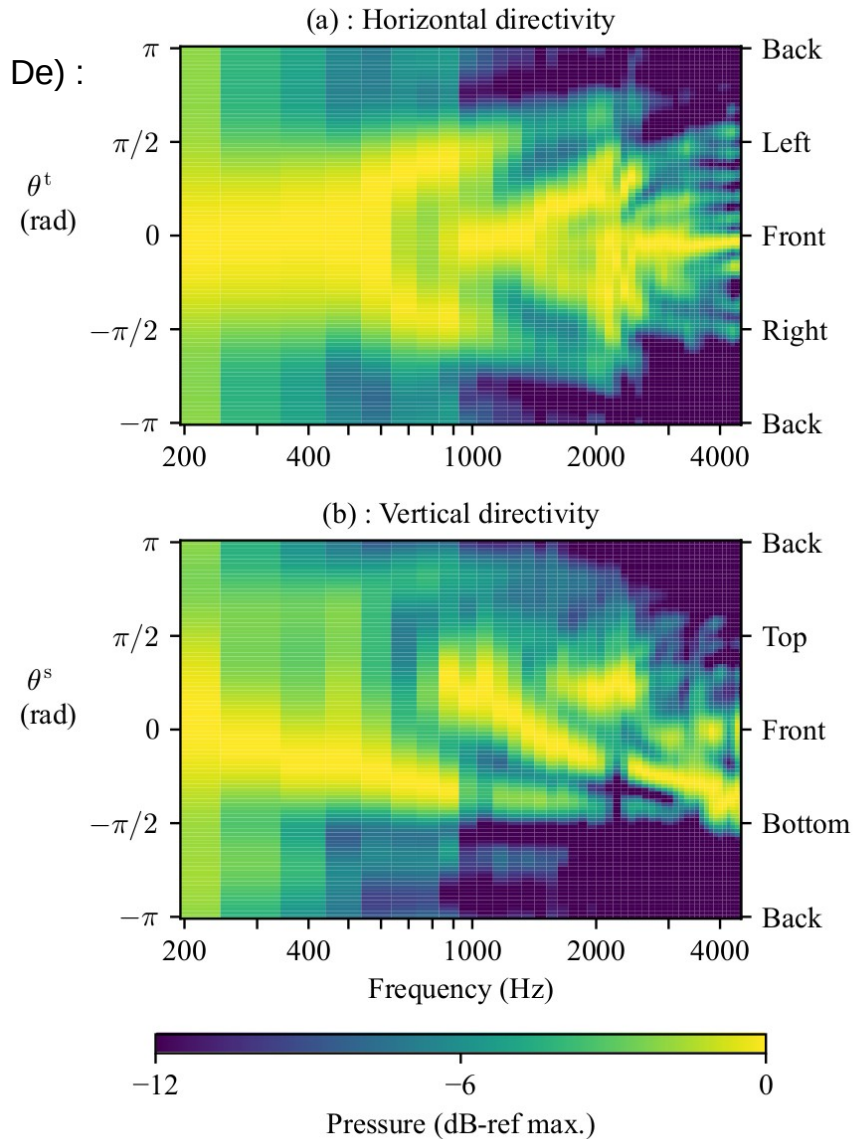
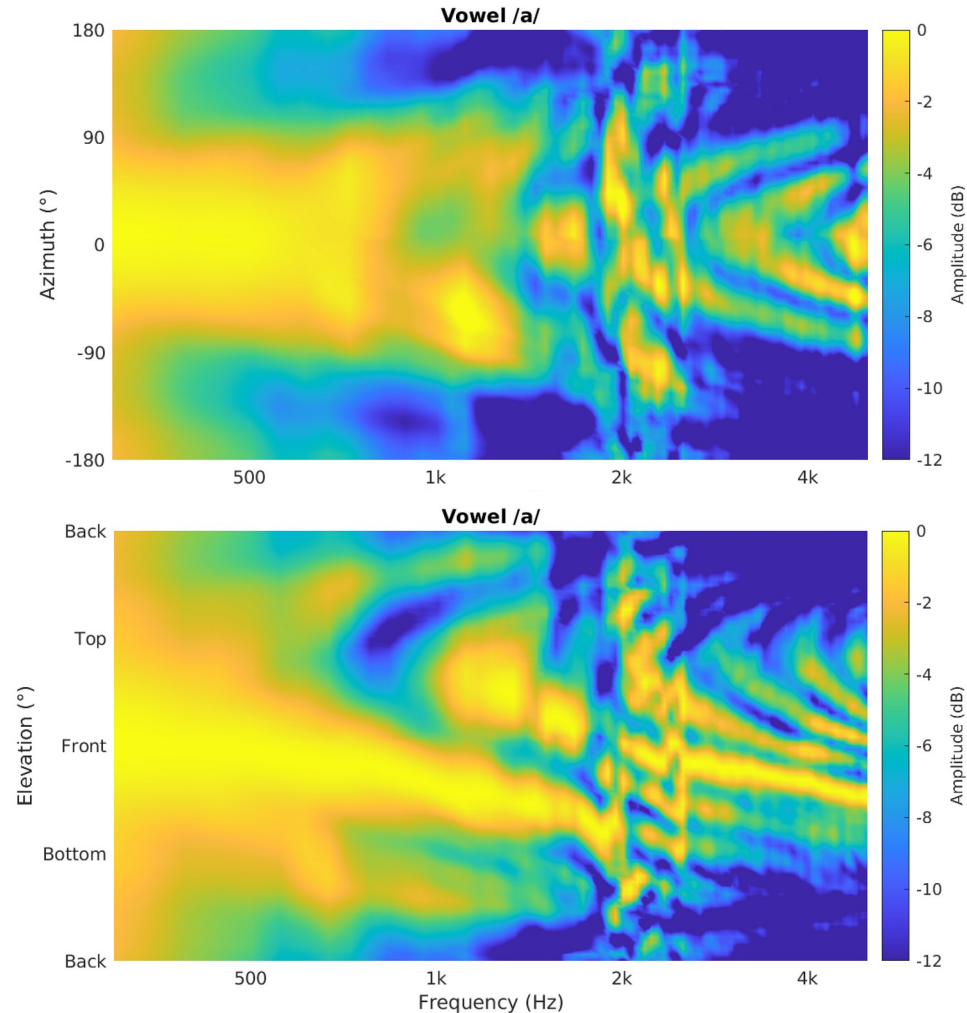
Application à la voix humaine

Comparaison avec locuteur mesuré sur autre système (Paris) :
Figures d'interférences similaires



Application à la voix humaine

Comparaison avec tête artificielle mesurée sur autre système (Aachen, De) :
Figures d'interférences similaires



Conclusion

- À notre connaissance, système de mesure de cette dimension unique pour captation en une seule prise (alternatives à 64 microphones max, ou séquentielles existent ailleurs)
- Méthodes d'étalonnage acoustique et spatial validées
- Système opérationnel jusqu'à l'ordre 9 (SH) avec 600 MEMS : résultats au moins comparables à ceux d'autres études utilisant
 - Sphère de 32 microphones [8]
 - Cercles de 64 microphones plan horizontal et vertical (résolution angulaire 11°) [9]
 - Mesures répétées avec tête artificielle (résolution angulaire 15°) [10]

Prochaines étapes

- Enregistrer plusieurs locuteurs et chanteurs, instruments de musique : base de données prévue en accès libre
- Analyser divers phonèmes et quantifier leurs différences
- Aspects perceptifs : quelle résolution spatiale est nécessaire pour rendu sonore réaliste ?

[8] Pörschmann, C., & Arend, J. M. (2024). On the impact of downward-directed human voice radiation on ground reflections. *Acta Acustica*, 8, 12.

[9] Brandner, M., Blandin, R., Frank, M., & Sontacchi, A. (2020). A pilot study on the influence of mouth configuration and torso on singing voice directivity. *The Journal of the Acoustical Society of America*, 148(3), 1169-1180.

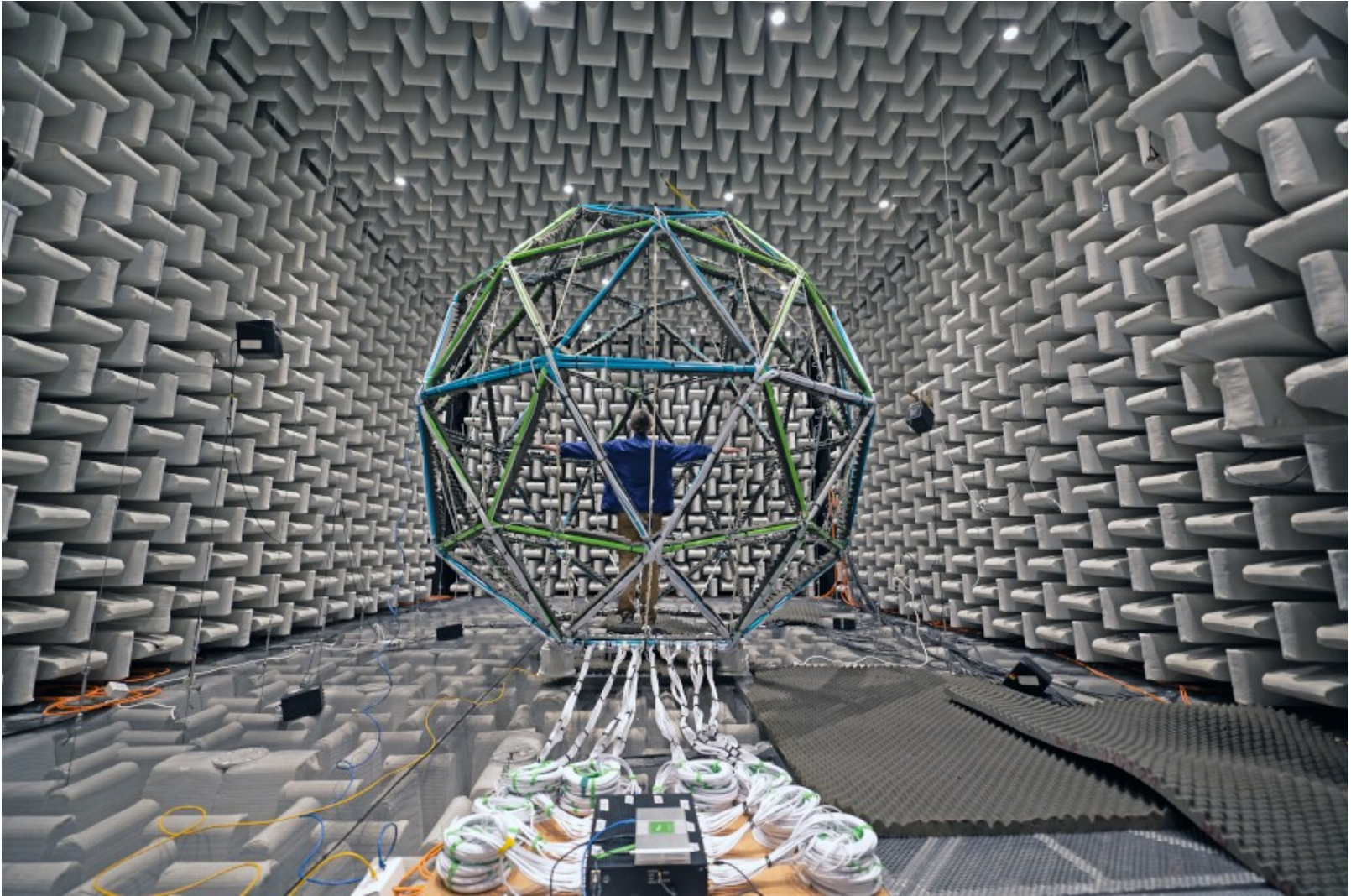
[10] Blandin, R., Geng, J., & Birkholz, P. (2023). Investigation of the influence of the torso, lips and vocal tract configuration on speech directivity using measurements from a custom head and torso simulator. *Acta Acustica*, 7, 39.

Funding sources

- ANR, projet “**RAYOVOX**”, 2021-2025, ANR-21-CE42-0017-01
- DFG project “**Spatial characteristics of the human voice**”, 2021-2024



Prix concours photo CNRS 2024



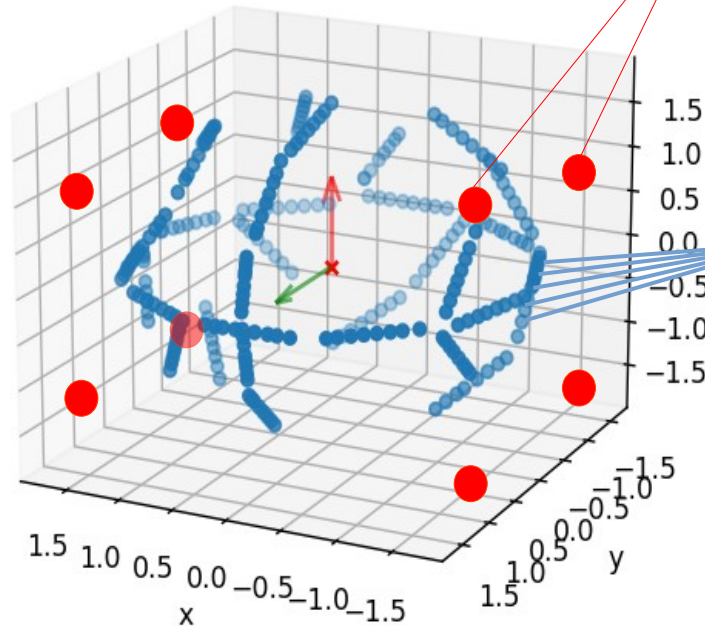
Annex - Geometric calibration

Using sources outside the sphere

Based on time of arrival from each source to each receivers
→ The more sources, the higher the spatial precision

- 1) Estimation of Time Difference of Arrival (TDOA) between nodes (MEMS or sources)
- 2) Building an Euclidean Distance Matrix (EDM)
- 3) Solving an optimization problem to map the nodes (MEMS and sources)
- 4) Robustness of the method increases with number of nodes (here, 256 mic positions and 8 source positions)

Source : Vanwynsberghe et al 2018, Geometric calibration of very large microphone arrays in mismatched free field (Jasa)



External sources

MEMS microphones

