

MASTER Sciences Cognitives - 2021-2022

Impact de la conduction osseuse sur la perception de la parole et du chant: expériences et modélisation

Contexte

La production de la parole est guidée en permanence par le retour auditif. C'est évident chez le jeune enfant qui babille, explorant ainsi de manière répétée les relations entre action articuloire et perception auditive. Mais une plasticité de ces relations est maintenue chez l'adulte : perturber la parole d'un sujet par l'introduction d'un délai, ou le déplacement de certaines bandes de fréquences, entraîne une adaptation rapide de sa production vocale (p.ex. Rochet-Capellan et al. 2011). La parole est donc un système bouclé, où les retours auditifs sont utilisés en temps réel (modulo des délais de conduction et traitement du signal). On peut en dire autant du chant, où d'autres dimensions du retour auditif sont pertinentes (timbre, hauteur,...). Actuellement, le système audio-moteur est un des rares systèmes sensorimoteurs où les retours sensoriels sont mal connus.

Les sons de la parole ou du chant ne sont en effet pas seulement transmis par vibration aérienne. Une partie du son voyage depuis ses sources dans le conduit vocal jusqu'à l'oreille interne (la cochlée) à travers le crâne et les tissus mous de la tête. On sait que cette conduction acoustique interne (ou **conduction osseuse**) permet la transmission d'à peu près la moitié du signal acoustique arrivant sur la cochlée (von Békésy 1949). Le son que le cerveau perçoit est un mélange entre cette partie osseuse et la partie aérienne (ce qui explique qu'on soit généralement surpris en entendant pour la première fois un enregistrement de sa propre voix, qui ne correspond pas avec notre propre perception interne). Toutefois, l'écrasante majorité des recherches sur la voix humaine n'ont considéré que la partie aérienne du son produit, parce qu'il est facile à mesurer et qu'il correspond à ce que l'auditeur perçoit.

Ce stage cherche à évaluer l'impact de la conduction osseuse sur le contrôle de la production de la parole et du chant. Pour cela, nous avons mis au point un dispositif expérimental qui permet d'enregistrer indépendamment les sons aériens et une estimation des sons transmis par conduction osseuse. Nos résultats préliminaires suggèrent des complémentarités intéressantes entre signaux aériens et osseux. Les sons nasaux, par exemple, sont plus transmis par conduction osseuse. Certaines résonances internes du conduit vocal corrélées à la position de la langue ne s'observent que dans la partie osseuse. Ceci pourrait permettre des stratégies de contrôle audio-moteur qui n'ont pas été envisagées jusqu'à présent. Ce projet financé par l'ANR vise à améliorer notre compréhension des relations entre production vocale et retour acoustique afin d'évaluer l'impact de la conduction osseuse sur l'apprentissage de la parole et du chant, et leur plasticité chez l'adulte.

Programme de travail

a) Acquisition de données

Une partie du signal acoustique transmis par conduction osseuse est rayonné par le tympan. On mettra en œuvre un dispositif qui permet d'enregistrer ce faible signal à l'aide d'un microphone « capillaire » inséré dans le conduit auditif du sujet. Une « boîte à oreille » permet d'isoler les composantes aériennes (CA) et osseuses (CO) de la transmission acoustique (cf. Figure 1). Des expériences pilotes ont permis de caractériser ce dispositif et l'évaluation des différences CA/CO (Bderi 2019). Le/la stagiaire devra tester une amélioration de la méthode d'équivalence subjective entre sons CA/CO (utilisant l'annulation de signal), comparer la mesure directe par vibromètre laser à notre méthode intra-

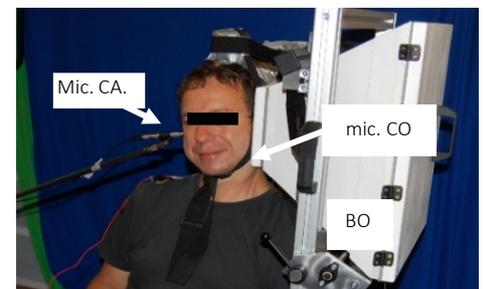
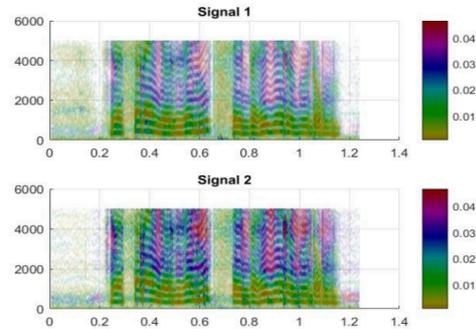


Figure 1. Dispositif expérimental permettant d'isoler la composante aérienne des vocalisations (micro CA) de la composante osseuse (micro intra-auriculaire CO, isolé acoustiquement par la boîte à oreille BO).

auriculaire, et enfin réaliser un ensemble d'acquisitions audio CA/CO de parole (phonèmes, syllabes, phrases) et de chant sur un ensemble de locuteurs.

b) Phase de modélisation

Une fois ces acquisitions réalisées, on développera un modèle statistique de la relation entre les composantes aériennes et osseuses du signal, afin d'évaluer quand et comment ces signaux divergent. Pour cela, on tirera parti d'une technique maintenant bien rodée appelée « conversion de voix » afin d'évaluer l'information propre à chaque signal, et tester dans quelles limites on peut prédire la partie osseuse du signal à partir de la partie aérienne (et réciproquement). Grâce à ce modèle, on s'intéressera à la discriminabilité des phonèmes et des timbres musicaux dans la partie osseuse du signal et on la comparera à la discriminabilité aérienne. Enfin, on évaluera la variabilité inter-sujet d'un tel modèle (dans quelle mesure dépend-il des caractéristiques morphologiques du locuteur ?).



(Signal 1: original bone signal, Signal 2: aerial signal transformed to bone signal, y-axis: frequency (Hz), x-axis: time (s))

Figure 2: Exemple de conversion de voix. Le signal aérien CA est converti en signal osseux CO. Les couleurs dénotent la différence d'amplitude entre le signal CO réel et celui prédit à partir de la composante CA.

c) Validation expérimentale

Si le temps le permet, on testera enfin si notre reconstruction du retour acoustique total, contenant le son CA et CO, est crédible. Les sujets seront conviés à une seconde session dans laquelle ils évalueront perceptivement la proximité entre 1) un stimulus de voix diffusé dans un casque et 2) leur propre voix pendant qu'ils parlent ou chantent. Les stimuli auditifs seront différents mélanges de sons CA et CO enregistrés dans les essais précédents.

Ce stage amorce un projet ANR comprenant un financement de thèse (2022-2025), dans laquelle on poursuivra ces investigations en analysant les compensations comportementales consécutives à une modification en temps réel de la voix osseuse.

Contact

Pierre Baraduc, CR CNRS, Pierre.Baraduc@grenoble-inp.fr, 04 76 82 71 50

Coriandre Vilain, IGR UGA, Coriandre.Vilain@grenoble-inp.fr, 04 76 82 77 80

Nathalie Henrich-Bernardoni, DR CNRS, Nathalie.Henrich@grenoble-inp.fr, 04 76 57 45 36

Funding

INCEPTION-CONTROL project (ANR 2021, Neurosciences intégratives et cognitives).

References

- Bderi M. (2019). Impact of bone conduction on speech perception. Internship report, M2 Movement-Ergonomics-Disability
- Békésy, von, G. (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *J Acoust Soc Am* 21:217-232.
- Pörschmann, C. (2000). Influences of bone conduction and air conduction on the sound of one's own voice. *Acustica - Acta Acustica* 86:1038-1045.
- Reinfeldt, S., Östli, P., Håkansson, B., Stenfelt, S. (2010). Hearing one's own voice during phoneme vocalization — transmission by air and bone conduction. *J Acoust Soc Am* 128:751-762.
- Rochet-Capellan, A., Ostry, D. J. (2011). Simultaneous acquisition of multiple auditory-motor transformations in speech. *J Neurosci*, 31(7):2657-2662.