MASTER Cognitive Sciences - 2021-2022

Impact of bone conduction on the perception of one's own speech and singing: experiments and modeling

Context

Speech production is constantly guided by the auditory feedback. It is obvious in the babbling toddlers, who explore repeatedly the relation between the articulatory commands and their acoustical result, and by trial and error end up mastering the sounds of their mother tongue. Feedback also guides speech in the adult: for instance, perturbing speech by introducing a sensory delay, or shifting frequency bands leads to a rapid adaptation of vocal production (see e.g. Rochet-Capellan et al. 2011). Speech is thus a plastic feedback system in which auditory signals are used in quasi real-time (modulo conduction and processing delays). The same can be said of singing, in which other dimensions (timbre, pitch…) of the auditory feedback are important. At present, it is unclear how this feedback system works.

Speech or singing sounds are not only conducted by the air. Part of the sound is transmitted from its sources in the vocal tract (vocal folds, radiated sound at the mouth or the nose…) through the skull and the different structures of the inner body to the inner ear — the cochlea. This *bone conduction* accounts for around half of the signal arriving on the cochlea (von Békésy 1949). The sound we hear when we speak is thus a mixture between the aerial- and the bone-conducted components of the speech sounds (this partially explains the odd feeling you can have when you first listen to your own voice on a recording. It can't be you! In fact, it is, but it's only the aerial part of your speech). Yet, most research has only considered the aerial feedback of the produced speech/singing sounds, because it is easy to measure, and corresponds to what the listener can hear. This study aims to reconsider bone conduction as a key parameter in the control of speech or song production.

To tackle this issue, we have designed an experimental device that can record separately aerial and an estimate of the bone-conducted vocal sounds. Our preliminary results tend to show interesting complementarities between aerial and bone conduction signals. Nasal sounds for instance tend to have more bone-conducted transmission than oral sounds. Internal resonances correlated to tongue position can be observed in the bone-conducted sound, and not in the aerial sound. This could allow feedback control strategies that were not thought possible. This project funded by the ANR aims at furthering this understanding of the relations between vocal production and acoustical feedback in order to evaluate the impact of bone conduction on the learning of speech and singing and on their plasticity in the adult.

Work program

a) Data acquisition

Part of the acoustical signal conducted by the bony structure is radiated by the tympanic membrane. We will make use of a specially developed experimental setup to record this tiny signal with a probe microphone inserted into the subject's ear. An « earbox » isolates the aerial (AC) and bone (BC) components of the acoustical transmission (see Figure 1). Pilot studies have allowed characterization of this device and the evaluation of AC/BC transmission (Bderi 2019). In the present internship, the student will have to refine an AC/BC sound equivalence procedure using a sound cancelation technique, compare the direct measure of skull vibration through a laser vibrometer to the in-ear recording method, and finally to acquire a set of joint AC/BC recordings of speech (phonemes, syllables and whole sentences) and singing in a large group of subjects.

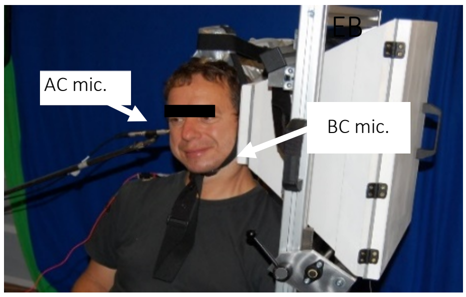
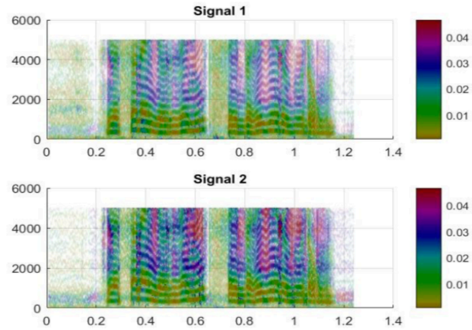


Figure 1 Experimental setup isolating the aerial component of speech (AC mic, close to right ear) from the bone conduction component (BC mic, inserted in the subject's left ear, acoustically isolated by the earbox EB).

b) Modeling

Once these data acquired, we will develop a statistical model of the relation between aerial- and bone-conducted voice in order to evaluate when these signals diverge. In technical terms, we will repurpose a technique called « voice conversion ». This will allow to evaluate the information that is unique to each signal, and to test to what extent one can predict the bone signal from the aerial signal (and conversely). Furthermore, we will explore the discriminability of phonemes and musical timbres in the bone-conducted signal compared to the aerial signal. Last, we will also probe the inter-subject variability of the model, i.e. to what extent it depends on the speaker's morphological features.



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

Figure 2: Example of voice-conversion. AC signal is converted into BC. The colors display the magnitude difference between original BC and AC to BC converted signal.

c) Experimental validation: perceptual experiment

If time allows, wee will test whether our reconstruction of the total auditory feedback, using both AC and BC sound, is credible. In a second session, subjects will be required to perceptually evaluate the closeness of a voice stimulus played in headphones to their own voice while they speak or sing. Auditory stimuli will be different mixes of the AC and BC sounds recorded in one of the previous speech/singing trials.

This internship starts an ANR project that comprises funding for a PhD (2022-2025) probing these issues deeper by analyzing the behavioral compensations elicited by a real-time modification of the bone-conducted voice.

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.