gipsa-lab

UNIVERSITÉ Grenoble Alpes

cnrs

MASTER Cognitive Sciences - 2020-2021

Impact of bone conduction on the perception of its own speech: 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). Nevertheless, it is unclear how this feedback system works.

Speech 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 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 production. It is a delicate task. Bone conduction is difficult to characterize, among others, because one cannot easily put microphones inside your head, or record from your cochlear nerve!

To tackle this issue, we have designed an experimental device that can record separately aerial and an estimate of the bone-conducted speech 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 aims at furthering this understanding between speech production and acoustical feedback in order to evaluate the impact of bone conduction on speech plasticity.

Work program

a) *Data acquisition*

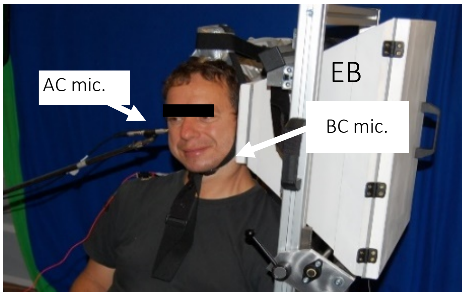
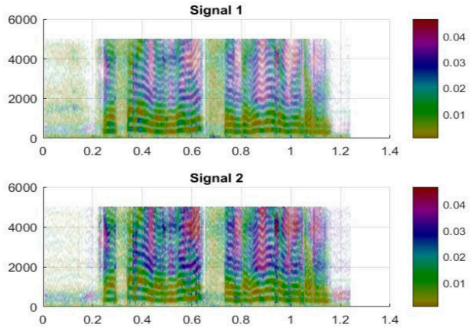


Figure 1 Experimental setup isolating the aerial component of speech (AC mic, close to right ear) from the bone conduction component (BC mic, inserted in the subject's left ear, acoustically isolated by the earbox EB).

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). A first internship has focused on the characterization of this device and the evaluation of AC/BC transmission in 2 pilot studies (Bderi 2019). In the present stage, the intern will have to master this setup in order to acquire a set of joint AC/BC recordings of phonemes, syllables and of a speech corpus on a large group of speakers.

b) Modeling

Once these data acquired, we will strive to understand the transformations between aerial- and bone-conducted voice by developing a statistical model of their relationships. To this aim, we will repurpose a well-known technique called « voice conversion » in order 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, in other words, guess from the outside the full auditory feedback of a speaker. Furthermore, we will explore the discriminability of phonemes 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

We will test whether the adaptation to an auditory perturbation depends on the information comprised in the bone-conducted part of the signal. We expect that adaptation to a transformation between phonemes P1 and P2 introduced with headphones will be maximal if P1 and P2 are weakly distinguishable in the bone-conducted signal (in other terms, if bone and aerial signals are not markedly incongruent), and conversely.

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

Funding

CondOss project.

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.