Orofacial somatosensory inputs enhance speech intelligibility in noisy environments

Rintaro Ogane, Jean-Luc Schwartz, Takayuki Ito

Orofacial somatosensory inputs enhance speech intelligibility in noisy environments

Rintaro Ogane[1], Jean-Luc Schwartz[1], Takayuki Ito[1,2]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, Grenoble, France, [2] Haskins Laboratories, New Haven, USA

* Institute of Engineering Univ. Grenoble Alpes

Contact : rintaro.ogane@gipsa-lab.grenoble-inp.fr

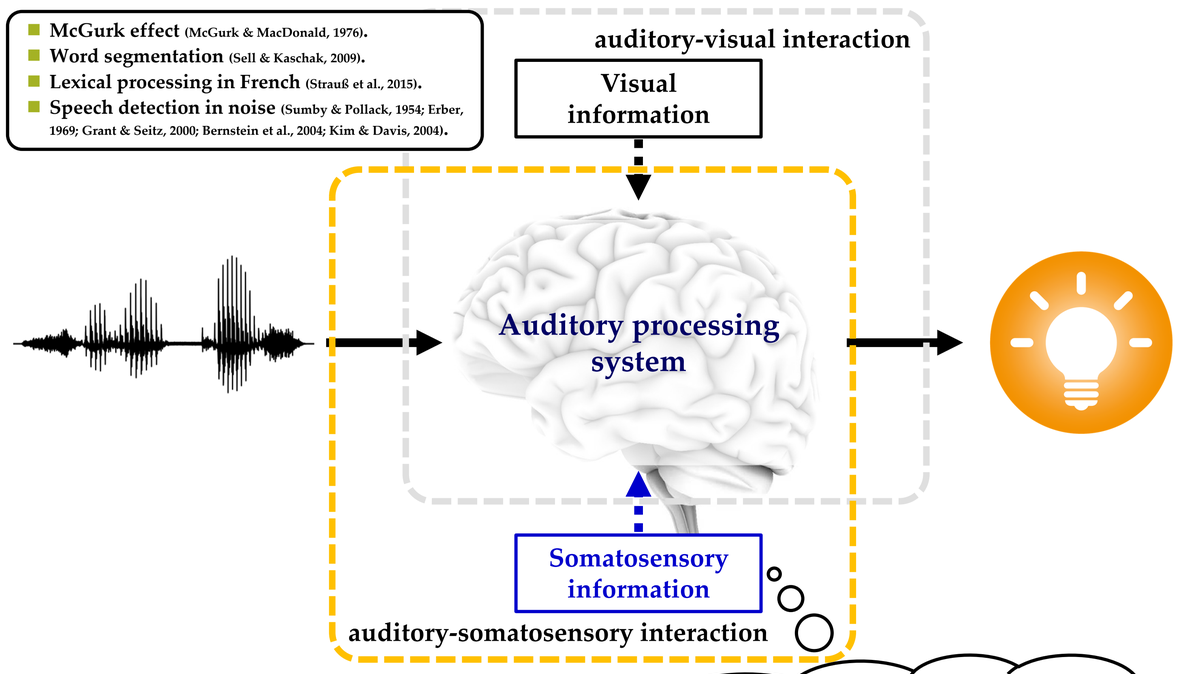
Summary

- **Somatosensory inputs** associated with **facial skin deformation** enhance **speech intelligibility in noise**, when the **somatosensory stimulation** is compatible with the **articulatory nature** of the corresponding speech sound.
- **The orofacial somatosensory system** may intervene in the process of **speech detection** in noisy environments.

Introduction

Speech perception is an **interactive process** with multiple modalities and some perceptuo(multisensory)-motor connections (Schwartz et al., 2012).

- McGurk effect (McGurk & MacDonald, 1976).
- Word segmentation (Sell & Kaschak, 2009).
- Lexical processing in French (Strauß et al., 2015).
- Speech detection in noise (Sumby & Pollack, 1954; Erber, 1969; Grant & Seitz, 2000; Bernstein et al., 2004; Kim & Davis, 2004).



How does **somatosensory input** affect the processing of speech ?

- Vowel perception (Ito et al., 2009, Trudeau-Fisette et al., 2017).
- Lexical perception (Ogane et al., 2019; 2020).

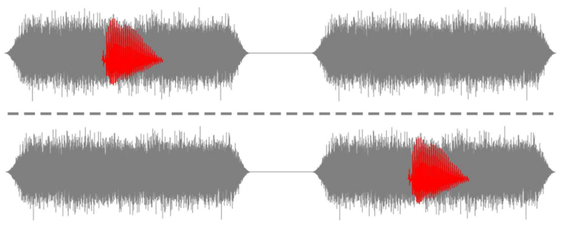
Do **somatosensory inputs** associated with **facial skin deformation** enhance **speech intelligibility in noise** ? — Q. 1

Do **somatosensory inputs** provide different effects in **different types of auditory stimulation** ? — Q. 2

Methods

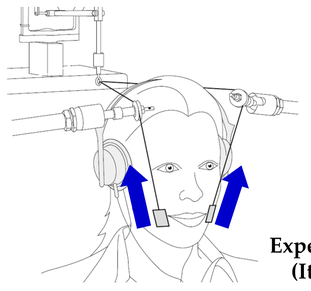
- Participants : 22 native French speakers.
  - 14 for Exp. 1 and 8 for Exp. 2.
- Speech materials.
  - /pa/ for Exp. 1 and /py/ for Exp. 2.
- Speech detection test.
  - Task : to identify which noise period includes the target speech sound ?

1st noise ? or 2nd noise ?



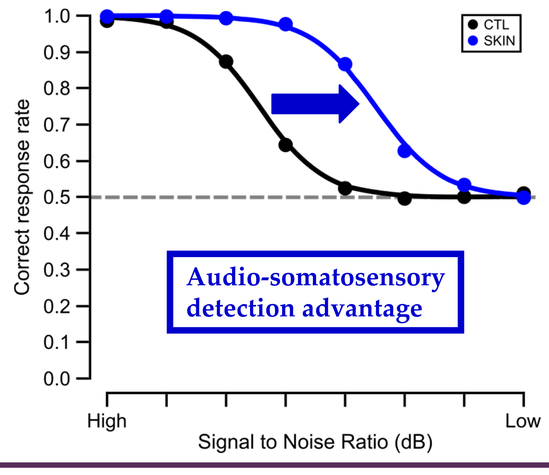
Note: two onsets of target speech sound were applied to avoid the participant's anticipation.

  - Speech stimulus was embedded in background noises (80 dB of SPL) with 8 SNR levels.
    - -8 dB to -15 dB for target /pa/.
    - -10 dB to -17 dB for target /py/.
  - Two experimental conditions were alternated every 8 trials.
    - SKIN : with somatosensory stimulation.
    - CTL : auditory-alone.
- Somatosensory stimulation on the face (SKIN).
  - Upward direction.
  - A half-wave 6 Hz sinusoidal pattern.
  - Applied in both noise periods.
  - The timing was adjusted to match the peak amplitude between somatosensory and auditory stimuli.



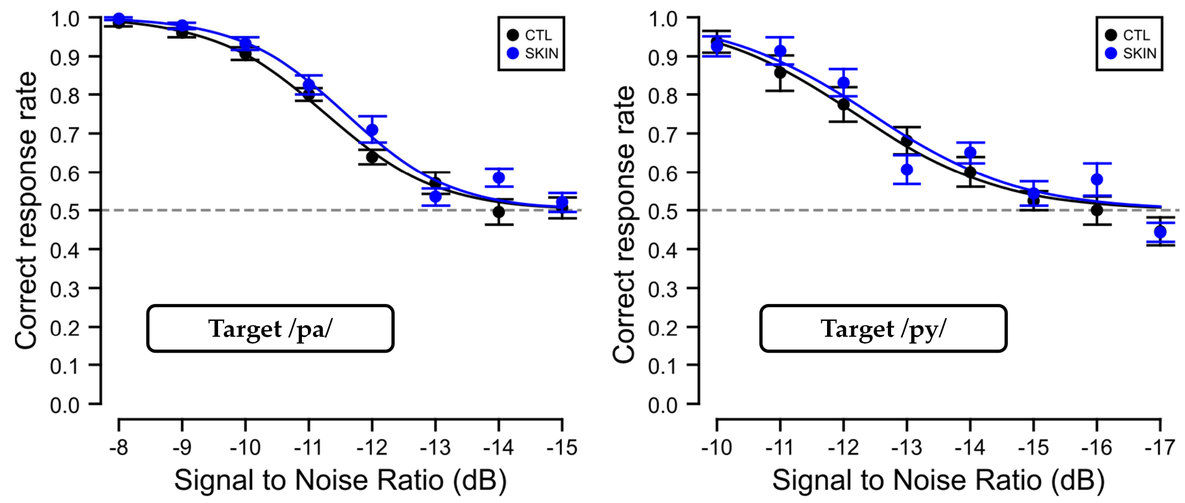
Experimental setup (Ito et al., 2009)

- Data analysis.
  - Mean probability of correct response rate across all SNR conditions.

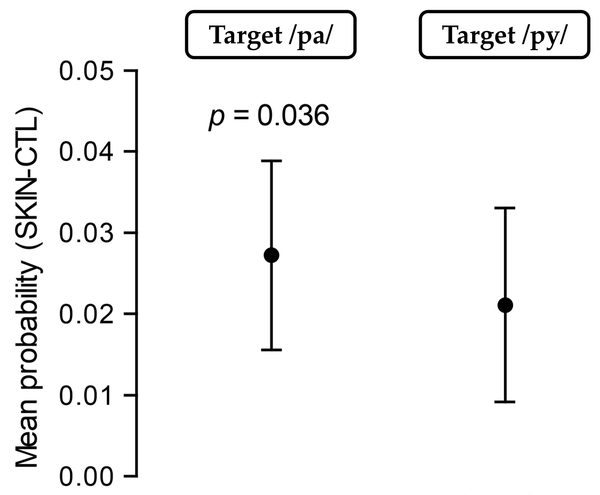


Results

- Correct response rate.



- Mean probability of correct response rate.

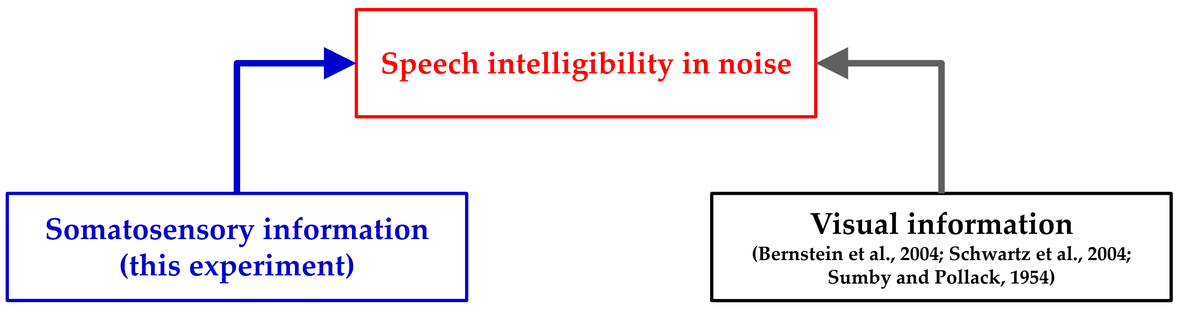


- Statistical results.
  - Mean probability for /pa/ was significantly different from zero ($t(13) = 2.33$, $p = 0.036$, one sample t-test).
  - No significant difference was found for /py/ (**) ($t(7) = 1.77$, $p > 0.12$, one sample t-test).

(**) to be confirmed with an increased number of participants for "py" (delayed by COVID).

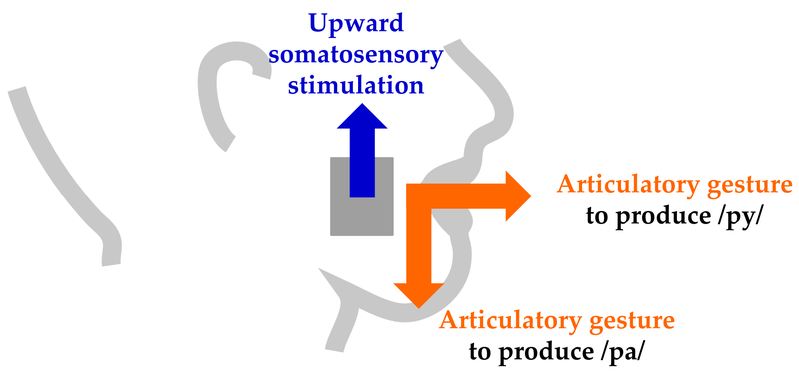
Discussion

- **Speech intelligibility in noise** was increased in **SKIN** compared to CTL.
  - ≈ 3% increased for speech target /pa/.
  - **Somatosensory effect** is consistent with audio-visual speech processing.



Additional sensory inputs associated with auditory information increase the intelligibility of speech sounds in noisy environment.

- A relationship between **somatosensory stimulation** & **articulatory gesture** in **auditory stimulation**.
  - **Speech intelligibility** increased only when the speech target was /pa/.



Somatosensory effect may appear when the **somatosensory stimulation** is matched with **articulatory gesture** in speech sound (Ogane et al., 2019; 2020).

Acknowledgments

- This work was supported by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 339152) and the National Institute on Deafness and Other Communication Disorders R01DC017439.
- We thank Gurvan Quiniou for data collection and analysis, and Silvain Gerber for statistical analysis. We also thank Coriandre Vilain for his technical support.