Orofacial somatosensory inputs enhance speech intelligibility in noisy environments
Rintaro Ogane, Jean-Luc Schwartz, Takayuki Ito

To cite this version:
Rintaro Ogane, Jean-Luc Schwartz, Takayuki Ito. Orofacial somatosensory inputs enhance speech intelligibility in noisy environments. ISSP 2020 - 12th International Seminar on Speech Production, Dec 2020, Providence (virtual), United States. hal-03083564

HAL Id: hal-03083564
https://hal.science/hal-03083564
Submitted on 19 Dec 2020

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.
Orofacial somatosensory inputs enhance speech intelligibility in noisy environments

Rintaro Ogane¹, Jean-Luc Schwartz¹, Takayuki Ito¹,²
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).

Methods

- Participants : 22 native French speakers.
- 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.

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

Results

- Correct response rate.

  - Target /pa/ vs. /py/.

  - Mean probability for /pa/ was significantly higher than for /py/.

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.

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

Acknowledgments

- This work was supported by the European Research Council under the European Community’s Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 295312) and by National Institute on Deafness and Other Communication Disorders BR00870/00.
- We thank Corina Quince for data collection and analysis, and Shilpa Carbone for statistical analysis. We also thank Corinna Vidal for his technical support.

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