# Seeing our own voice: an electrophysiological study of audiovisual speech integration during self perception

Avril Treille, Coriandre Vilain, Sonia Kandel, Marc Sato

# SEEING OUR OWN VOICE: AN ELECTROPHYSIOLOGICAL STUDY OF AUDIOVISUAL SPEECH INTEGRATION DURING SELF PERCEPTION

*Avril Treille[1], Coriandre Vilain[1], Sonia Kandel[1] & Marc Sato[2]*

[1] *GIPSA-lab, Département Parole & Cognition, CNRS & Grenoble Université, Grenoble, France*
[2] *Laboratoire Parole & Langage, CNRS & Aix Marseille Université, France*

## Introduction

To recognize one's own face and voice is key for our self-awareness and for our ability to communicate effectively with others. Interestingly, recent studies suggest that better recognition of one's actions may result from the integration of sensory inputs with our own sensory-motor knowledge. However, whether hearing our voice and seeing our articulatory gestures facilitate audiovisual speech integration is still debated.

## Methods

❖ Participants

18 healthy adults, right-handed native French speakers.

❖ Stimuli

- Syllables : /pa/, /ta/, /ka/
- Modalities : auditory (**A**), visual (**V**), audio-visual (**AV**) and incongruent audio-visual (**AVi**, self auditory signal, other visual signal)
- Half of the stimuli were related to the participant (self condition), the other half to an unknown speaker (other condition).
- A total of 1176 stimuli were created

❖ Tasks

1) Before experiment, a short training was performed.

3) EEG session : a three-alternative forced-choice identification task, with participants instructed to categorize each perceived syllable with their right hand, after an audio "beep".

❖ Data acquisition

- EEG data were continuously recorded from 64 scalp electrodes (international 10–20 system) using the Biosemi ActiveTwo AD-box EEG system operating at a sampling rate of 256 Hz.
- Two additional electrodes served as reference [CMS] & [DRL]
- One other external reference electrode was at the top of the nose. The electrooculogram controlling for horizontal (HEOG) and vertical (VEOG) eye movements were recorded using electrodes at the outer canthus of each eye as well as above and below the right eye. Before the experiment, the impedance of each electrode was adjusted to get low offset voltage and stable DC.

❖ Analysis

- *Behavioral analyses:*

| | Behavioral |
|---|---|
| **%** | ANOVA : Modality (**A, AV, V, AVi**), speaker (**Self/other**), syllables (**/pa/, /ta/, /ka/**) |

- *EEG analyses on fronto-central electrodes (F3/F4/C3/C4/Fz/Cz):*

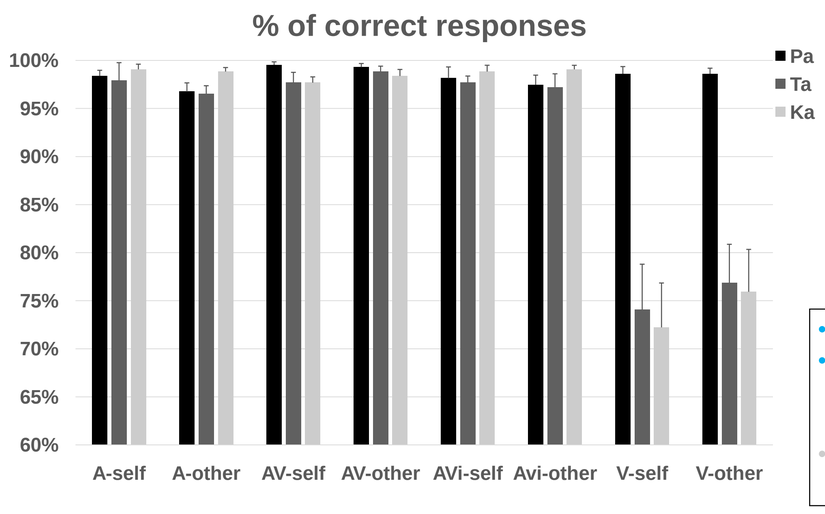
| | | | |
|---|---|---|---|
| **Pre-processing** | - Re-referenced off-line to the nose<br>- Epochs : 1000ms (baseline from -500 to -400ms to the acoustic syllable onset) | | - Filtering : 2-30 Hz<br>- Rejection : ±60 µV |
| **NI & P2** | | **Amplitude** | **Latency** |
| | **Speaker's effect** | ANOVA : Auditory modality (**Self/other**), Visual modality (**Self/other/None**) | ANOVA : : Auditory modality (**Self/other**), Visual modality (**Self/other/None**) |
| | **Audio-visual integration** | ANOVA : Signal type (**Bimodal/Sum**), Auditory modality (**Self/Other**), Visual modality (**Self/Other**) | ANOVA : Signal type (**Bimodal/Sum**), Auditory modality (**Self/Other**), Visual modality (**Self/Other**) |

- *Correlation between EEG and behavioral data*

Correlations between Integration (EEG; AV- A+V) & Visual identification (%)

## Results

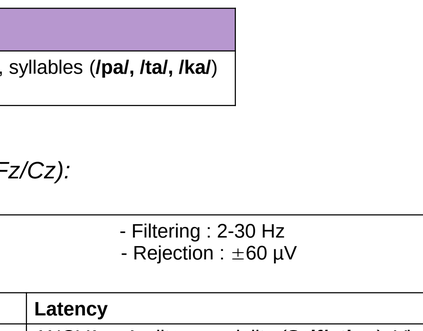
### 1) Behavioral - % correct responses (p<.001)

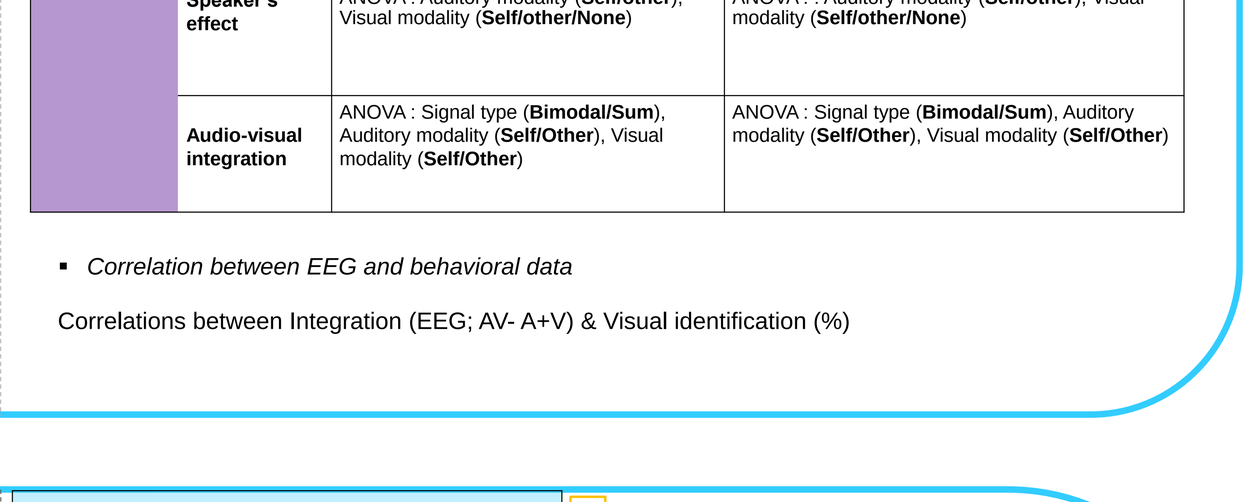



- A = AV = Avi > V
- /pa/ = /ta/ = /ka/ except for V-self and V-other
- No effect of the speaker

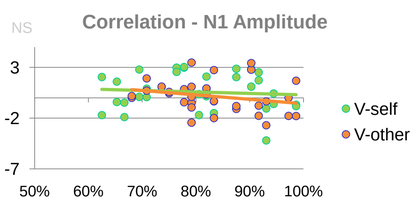
### 2) EEG - Integration (AV <> A+V)

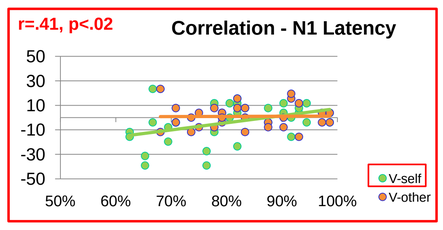
- **P2 amplitude : AV < A+V** (p<. 02)
- **=> integration**

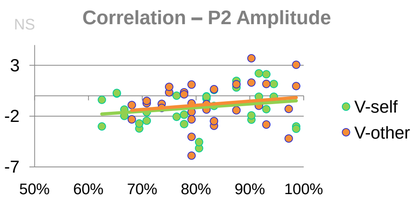


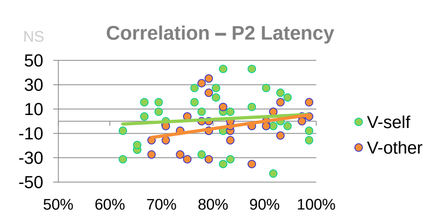




### 4) Correlations between Integration (EEG; AV- A+V) & Visual identification (%)



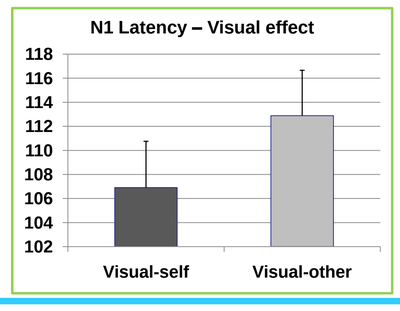




- Amplitude N1 & P2 : No efffect
- Latence P2 : No effetc
- **Latence N1 : Self : r=.41, p<.02;** other : r=.01, p<.94

- **Self : N1 latency is negatively correlated with visual saliency**




### 3) EEG – Self effect on integration

- **Visual-Self : reduced N1 latency** (p<.02)

## Discussion

1) Behavioral results : All modality was perfectly perceived except in the visual modality. /pa/ was better perceived than /ta/ and /ka/ during visual presentation.
2) Integration for both self and other signals : Early integration processing (on P2) during AV-self and AV-other speech perception compared to A+V.
3) Speaker effect on N1 latency : Compared to Visual-other, Visual-self stimuli induced a temporal facilitation on N1 during integration mechanisms.
4) Correlations on N1 latency for self visually ambiguous syllables : A negative correlation was oberved between visual-self identificcation and integration results on N1.

=> In line with previous EEG studies on multimodal speech perception, our results point to the existence of early integration mechanisms of auditory and visual speech information. Crucially, they also provide evidence for a processing advantage when the perceptual situation involves our own speech productions mostly for visually ambiguous syllables. Viewing our own utterances leads to a temporal facilitation of the integration of auditory and visual speech signals.

Part of this research was supported by a grant from the European Research Council (FP7/2007-2013 Grant Agreement no. 339152, "Speech Unit(e)s")
Correspondence: avril.treille@gipsa-lab.inpg.fr