



HAL
open science

To integrate the unknown: touching your lips, hearing your tongue, seeing my voice

Avril Treille, Coriandre Emmanuel Vilain, Jean-Luc Schwartz, Marc Sato

► To cite this version:

Avril Treille, Coriandre Emmanuel Vilain, Jean-Luc Schwartz, Marc Sato. To integrate the unknown: touching your lips, hearing your tongue, seeing my voice. ICAC 2014 - 5th International Conference on Auditory Cortex, Sep 2014, Magdeburg, Germany. hal-01298240

HAL Id: hal-01298240

<https://hal.science/hal-01298240>

Submitted on 5 Apr 2016

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.

To integrate the unknown:

touching your lips, hearing your tongue, seeing my voice

Avril Treille¹, Coriandre Vilain¹, Jean-Luc Schwartz¹

& Marc Sato²

¹ GIPSA-lab, Département Parole & Cognition, CNRS & Grenoble Université, Grenoble, France

² Laboratoire Parole & Langage, CNRS & Aix-Marseille Université, Aix-en-Provence, France

ERC Speech Units

Introduction

Seeing the speaker's articulatory gestures significantly enhances auditory speech perception. A key issue is whether cross-modal speech interactions only depend on well-known auditory and visual modalities or, rather, might also be triggered by other sensory sources less common in speech communication but likely to exist in the listener's motor knowledge.

The present electro-encephalographic (EEG) and functional magnetic resonance imaging (fMRI) studies investigate cross-modal interactions between auditory, haptic, visuo-facial and visuo-lingual speech signals during the perception of our own productions or/and those of other speakers.

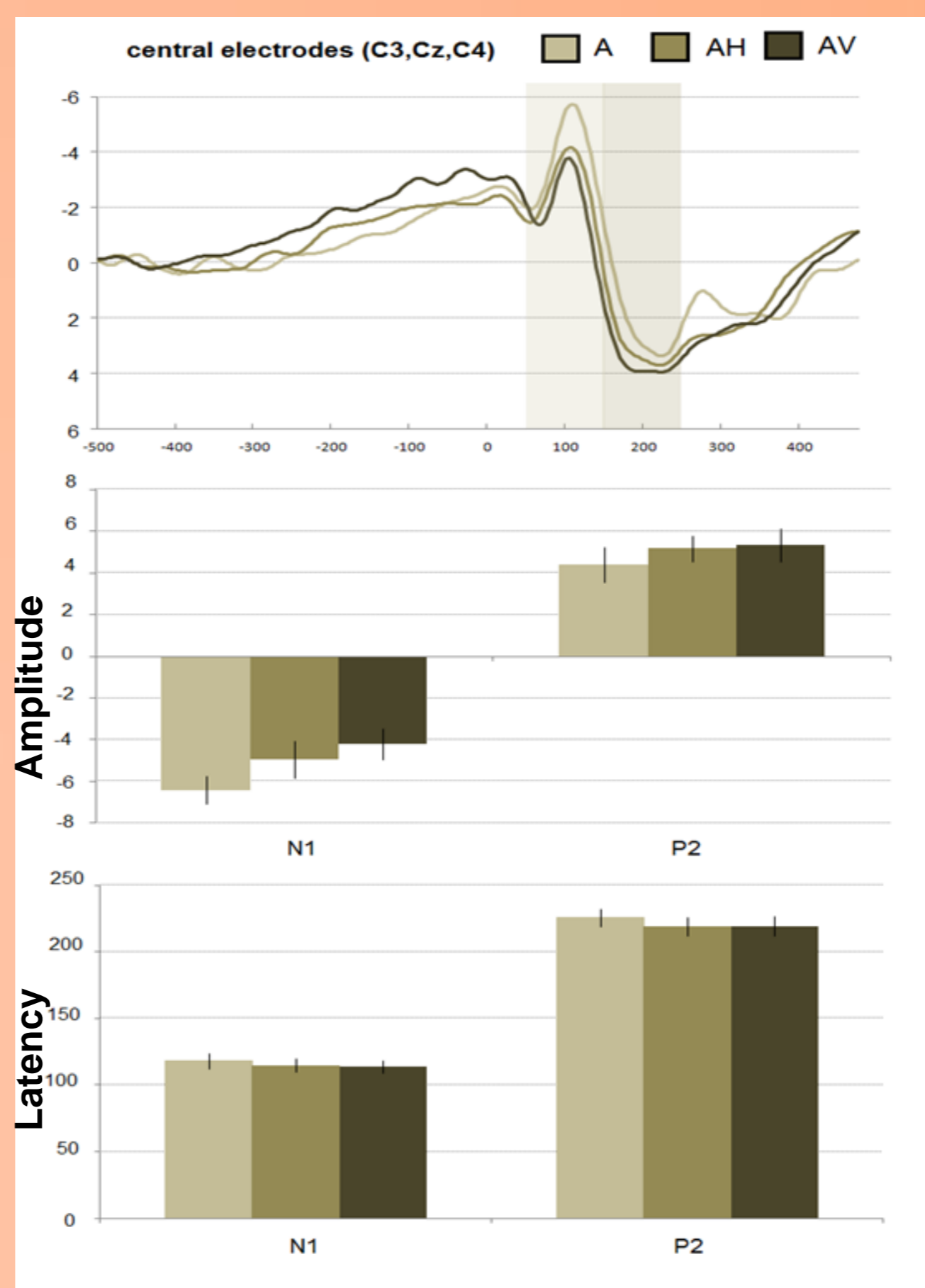
Audio-Haptic : Study 1

EEG : early integration and predictability of speech stimuli during audio-visual (AV; familiar modality) and audio-haptic (AH; unfamiliar modality) perception

Hyp : N1/P2 amplitude reduction and latency facilitation for both AV and AH modalities compared to auditory-only speech



Methods : 16 participants
80 trials per modality and per syllable (/pa/, /ta/, /ka/)
3 conditions : Audio (A), Audio-Visual (AV) & Audio-Haptic (AH)
live dyadic interactions



N1 Amplitude:

A>AV=AH

N1 Latency :

A>AV=AH

P2 Amplitude :

A=AV=AH

P2 Latency :

A=AV=AH

→ Shortened latencies and reduced amplitude of early AEP observed during both audio-visual and audio-haptic speech perception compared to auditory speech perception = **evidence for early multisensory integrative mechanisms, including haptic modality.**

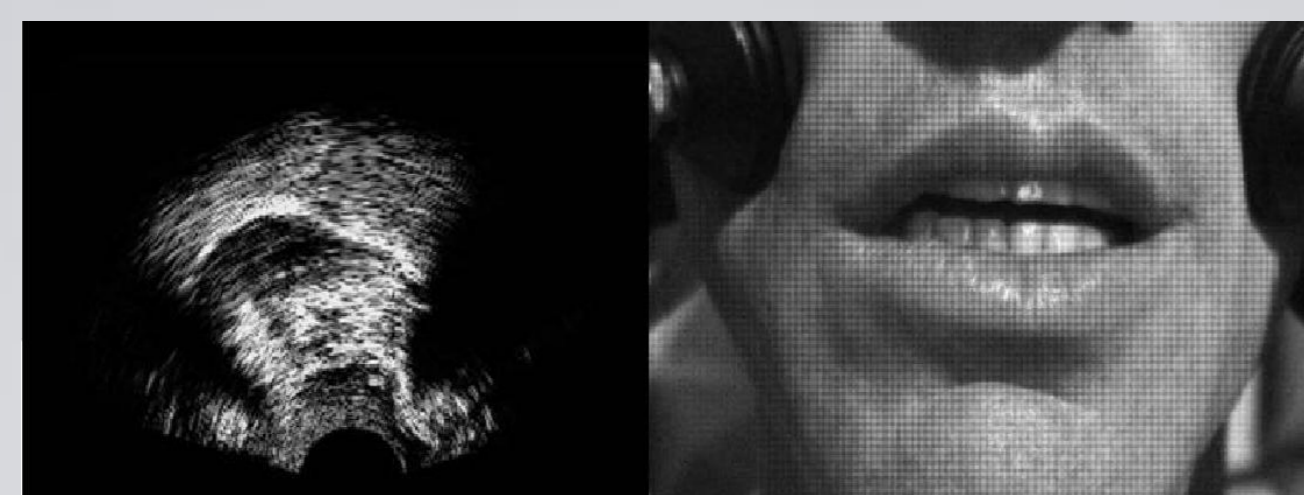
→ But no difference between /pa/, /ta/ and /ka/ syllables = Amplitude and latency reduction independent on the degree of visual and haptic recognition of the speech targets?

Ref.: Treille et al. (2014a). Haptic and visual information speed up the neural processing of auditory speech in live dyadic interactions. *Neuropsychologia*, 57, 71-77; Treille et al. (2014b). The sound of your lips: electrophysiological cross-modal interactions during hand-to-face and face-to-face speech perception. *Frontiers in psychology*, 5, Art-420..

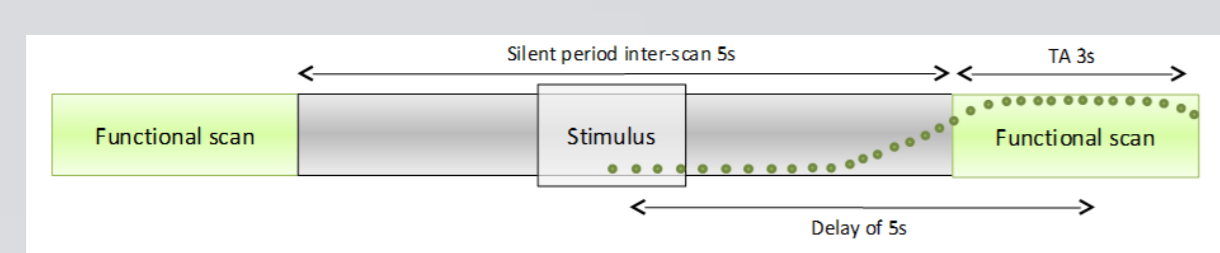
Audio-Visuo-Lingual : Study 2

fMRI : neural correlates of audio-visual-lingual (unfamiliar modality) and audio-visuo-facial (familiar modality) speech perception

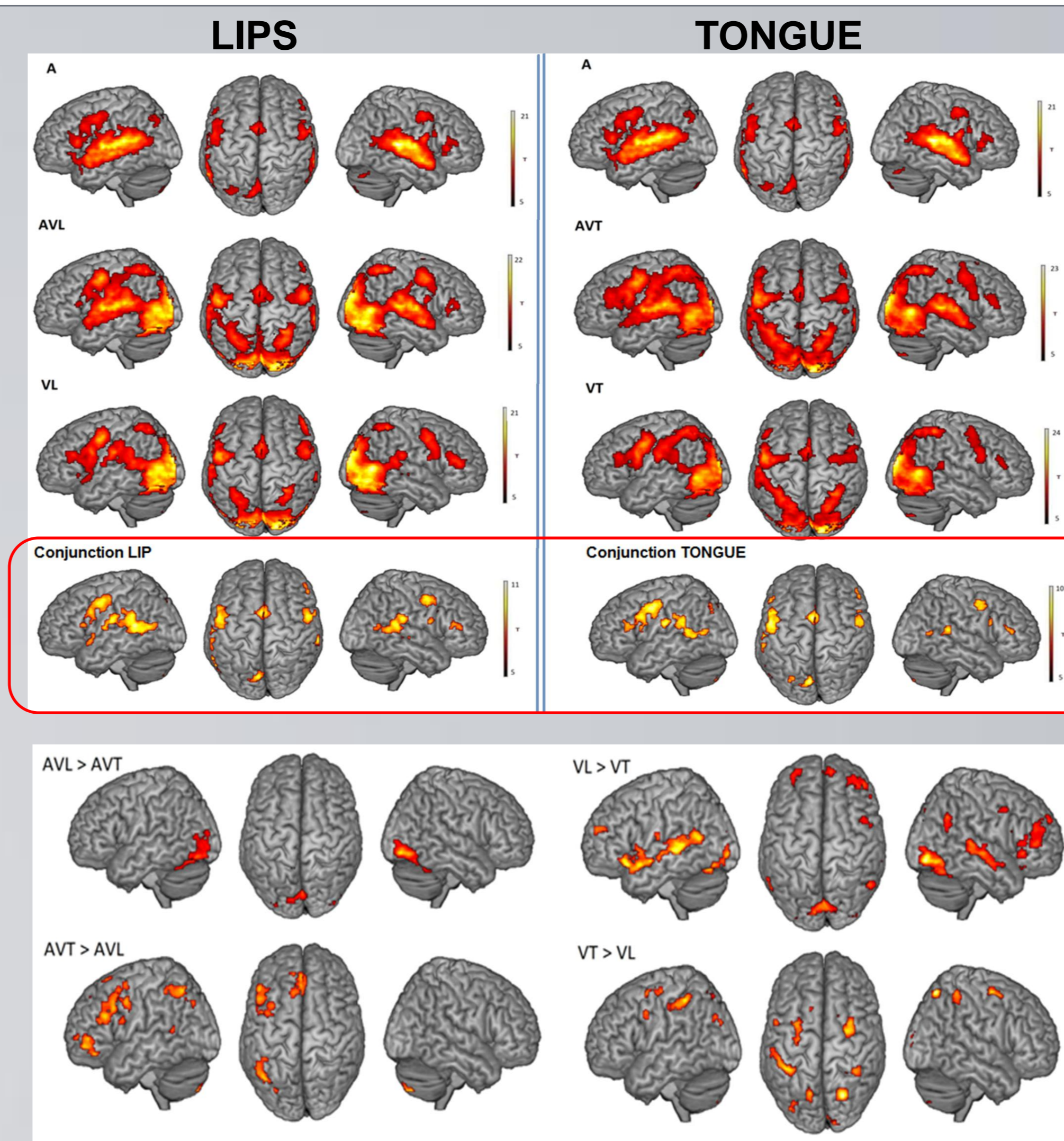
Hyp : natural temporal relationships of visuo-labial and visuo-lingual signals lead to the integration of auditory and visual speech signals



Method : 12 participants
24 trials per modality (/pa/, /ta/, /ka/ syllables)
5 conditions: Audio (A), Visual-lips (VL), Visual-tongue (VT), Audio-visual-lips (AVL) & Audio-visual-tongue (AVT)
passive fMRI session using sparse sampling.



Sparse sampling (1 trial)



Conjunction :

Motor + pSTS/STG are activated for all conditions and both lips and tongue movements.

Analysis by modality :

VL ≠ VT: -VL : auditory and visual cortices.

-VT : motor, premotor cortices and parts of the sensorimotor cortex.

→ In line with a sensorimotor nature of speech perception, common overlapping activity was observed for both facial and tongue-related speech stimuli in auditory and premotor cortices.

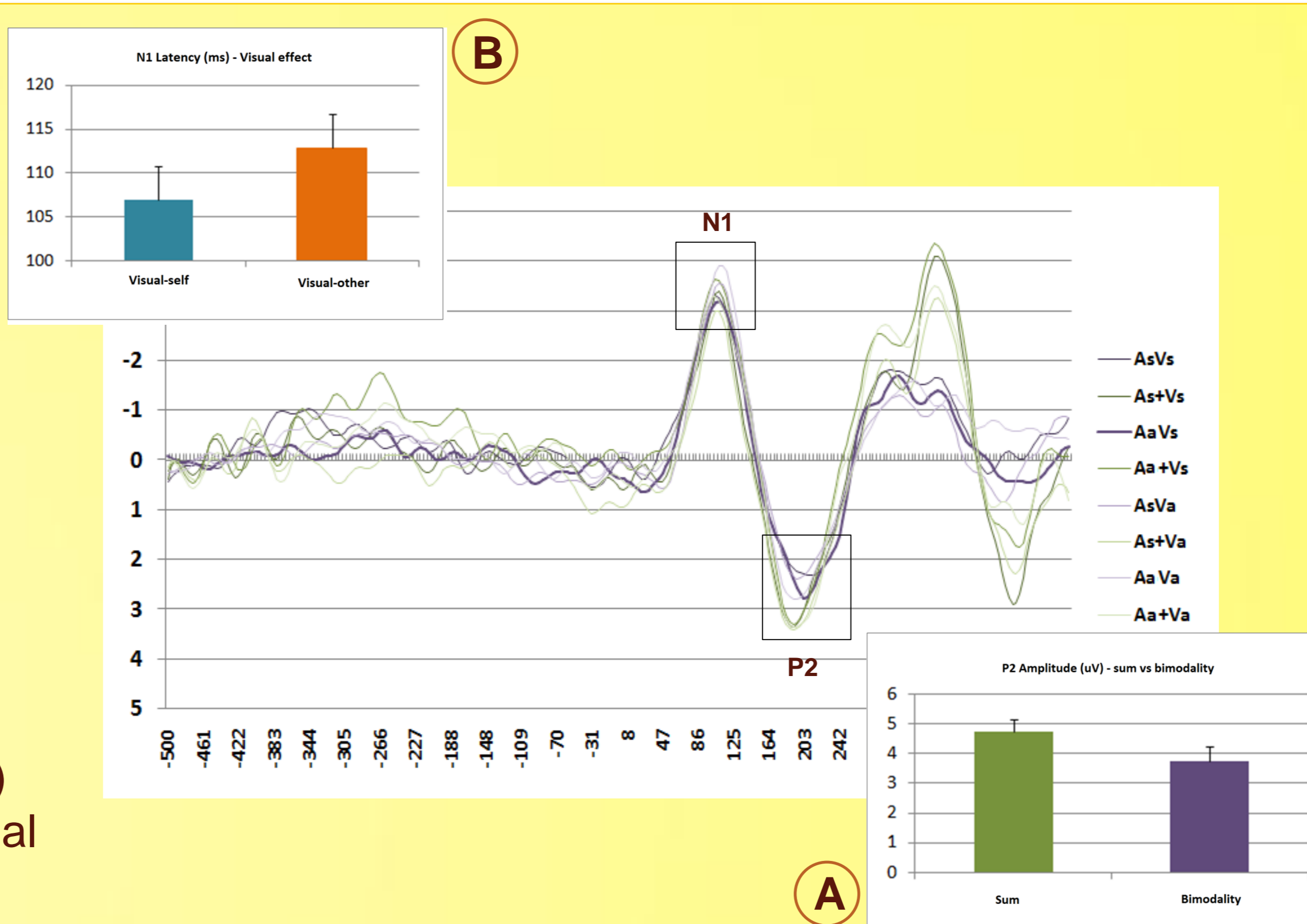
Self-Other : Study 3

EEG : mechanisms of early integration during the sight and listening of our own productions or those of somebody else

Hyp : modulation of the amplitude and/or latency of N1/P2 related to a possible self-influence during AV integration process



Method : 17 participants
72 trials per modality and speaker (self, other)
4 conditions (Audio (A), Visual (V), Audio-visual (AV), AV incongruent (AVi)) related to our own speech gestures or those of somebody else.



A) P2 Amplitude

AV < A + V

B) N1 latency

visual-self < visual-other

→ A self-advantage was observed with shortened latencies of early AEPs for self-related audio-visual speech stimuli.

Conclusion

Altogether our results provide evidence for multisensory interactions between auditory speech signals and their haptic, visuo-facial or visuo-lingual speech correlates. They further emphasize the multimodal nature of speech perception and demonstrate that multisensory speech perception is partly driven by sensory predictability and by the listener's knowledge of speech production.