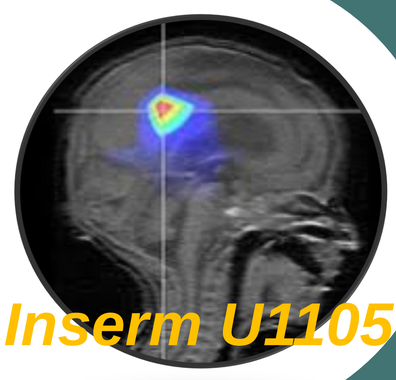


Investigating Infants' Responses to Infant-Directed Communication Using Eye-Tracking

Betania Y. Georlette[1], Mohammadreza Edalati[1], Nolane C. Richard[1], Rafael Román-Caballero[2], Laurel J. Trainor[2], Miriam D. Lense[3], Barbara Tillmann[4], Fabrice Wallois[1], Sahar Moghimi[1]

[1] Groupe de Recherches sur l'Analyse Multimodale de la Fonction Cérébrale (GRAMFC, Inserm UMR1105), Université de Picardie, 80054 Amiens, France
[2] Department of Psychology, Neuroscience, and Behaviour, and McMaster Institute for Music and the Mind, McMaster University, Hamilton, Ontario, Canada
[3] Department of Otolaryngology - Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, Tennessee 37203, United States of America
[4] Laboratory for Research on Learning and Development, CNRS–UMR 5022, Université Bourgogne Europe, 21000 Dijon, France

1 Introduction & research goals

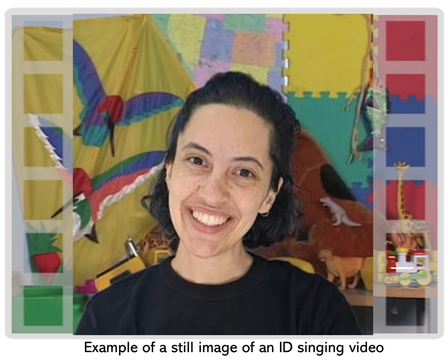
- **Infant-directed (ID) communication,** including singing and speaking, is a universal caregiver-infant interaction.
  It is a rich multimodal experience, auditory and visual, that not only captures infants' attention but also supports early auditory, social, and cognitive development[1]. To better understand the factors that guide infants' gaze in these interactions, this study examines specific visual and auditory cues using the following approaches:

- Exploring what drives infants' gaze during naturalistic ID singing beyond the beat – including visual and auditory features.

- Using Temporal Response Function (TRF) modelling to assess how different auditory and visual cues predict gaze behaviour.

2 Materials & Methods

We measure **visual behaviour** via **eye-tracking** in 50 term-born infants in the age range of 5 to 19 months while they were exposed to videos of ID communication (singing and speaking). Two TRF models were created, taking auditory and visual cues as inputs to predict gaze velocity, showing how these features influence infants' gaze behaviour.


Example of a still image of an ID singing video

Participants :

n = 50 full-term born

- From 5 mo to 19 mo

Stimuli :

n = 23 videos

- From 8 women
- 9 ID songs
- 9 ID speeches
- 5 cartoons
- Duration : 10 min 45 s

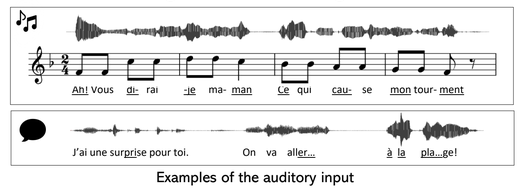
Outline of the protocol

Visual and auditory inputs:

Head movements velocity

Example of head movements during ID communication

Spectral flux



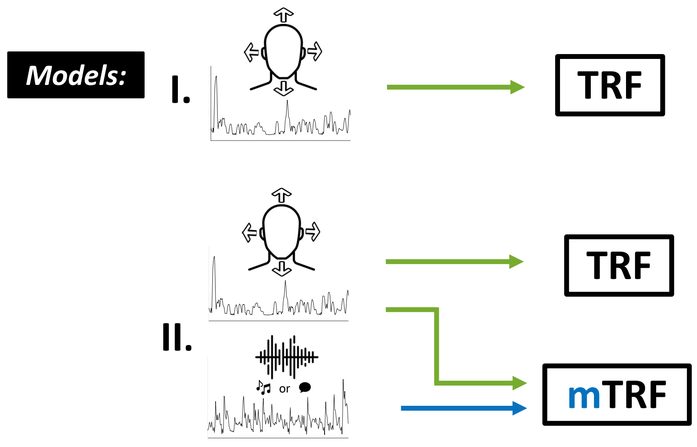
Examples of the auditory input

Output:

Gaze velocity


Examples of the gaze trajectory output in a singing video input

Models:



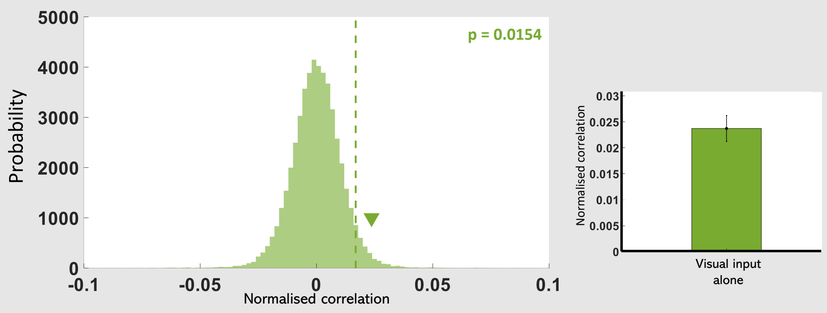
Overview of the TRF and mTRF models used by step to predict gaze velocity from head movement and spectral flux inputs. We first modelled gaze velocity using only visual inputs (TRF), and then included both visual and auditory inputs (mTRF) to calculate the correlation between predicted and observed gaze velocity.

3 Results

Singing

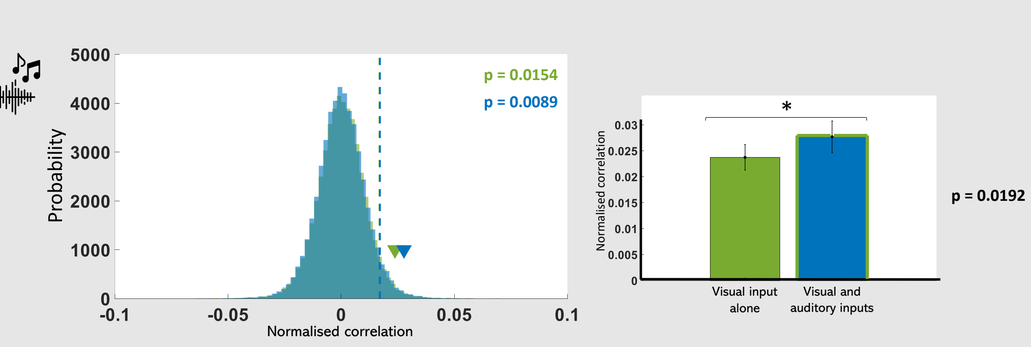
The curves represent the data permuted 1000 times which gives us a random distribution. The dashed line is the chance level with the 95% confidence interval.

I.



The green arrow, representing the normalised correlation between the recorded gaze velocity and the predicted gaze velocity using only visual input, is above chance level ▶ ***The visual inputs does predict the gaze velocity significantly.***

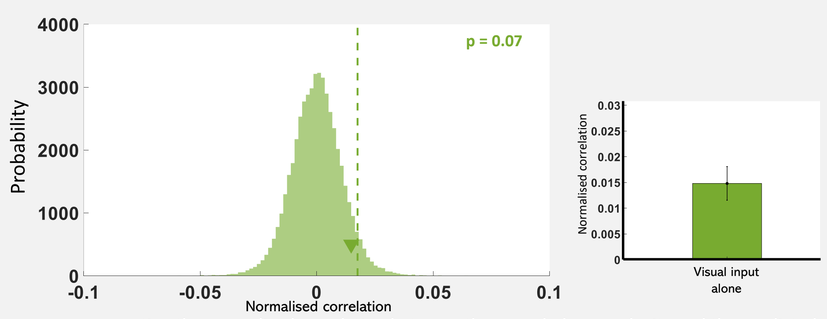
II.



The blue arrow, representing the normalised correlation between the recorded gaze velocity and the predicted gaze velocity using visual *and* auditory inputs, is further than the green one ▶ ***Adding the auditory input improves the gaze velocity prediction.***

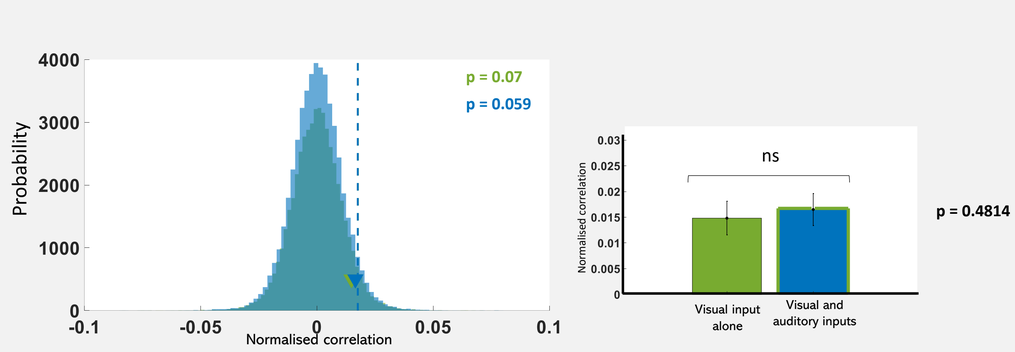
Speaking

I.



The green arrow, representing the normalised correlation between the recorded gaze velocity and the predicted gaze velocity using only visual input, is right below chance level ▶ ***The visual inputs predict the gaze velocity only marginally***

II.



The blue arrow, representing the normalised correlation between the recorded gaze velocity and the predicted gaze velocity using visual *and* auditory inputs, is still below chance level ▶ ***Adding the auditory input marginally improves the gaze velocity prediction.***

4 Conclusion

- Visual cues significantly modulates infant ocular movement. Combined with auditory rhythmicity, the effect is enhanced in the singing context.
- It is not the case in the speaking context. The less pronounced regularity in the speech signal, in comparison to the songs, seems to provide fewer temporal anchors for the infants in addition to the visual information.

Next steps :

- Age-related dynamics in gaze behaviour.
- Extend to preterm infants response by comparing them with full-term infants.

Reference :

[1]M.D. Lense, S. Shultz, C. Astésano, & W. Jones, Music of infant-directed singing entrains infants' social visual behavior, Proc. Natl. Acad. Sci. U.S.A. 119 (45) e2116967119, https://doi.org/10.1073/pnas.2116967119 (2022).

5 Contact information

Inserm UMR1105, Groupe de Recherches sur l'Analyse Multimodale de la Fonction Cérébrale, CURS, Avenue Laennec, 80036 Amiens Cedex, France.
Tel : +33 (0)3 22 66 78 50, web : https://gramfc.u-picardie.fr
Email : betania.georlette@etud.u-picardie.fr