Time-resolved discrimination of audio-visual emotion expressions

Falagiarda F.[2], Collignon O.[1,2]

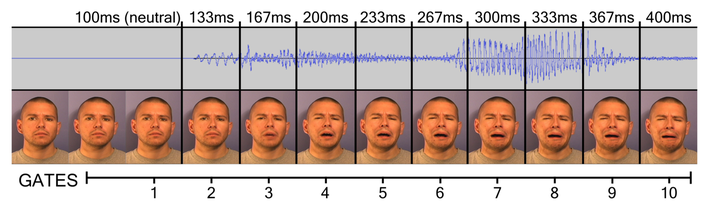
(1) Center for Mind/ Brain Sciences (CIMeC), University of Trento, Italy (2) Institute of Psychology (IPSY) and Institute of Neuroscience (IONS), University of Louvain

Introduction

- Facial and vocal expressions are the most important social stimuli in human interactions. They convey a plethora of information about our interlocutors like age, gender, identity, speech and emotional state
- The intrinsic dynamicity of facial expressions has been heavily overlooked in face perception research. Only a minority of studies has employed dynamic as well as multimodal stimulation[1]
- The current study employs a psychophysical constant stimuli gating[2] paradigm. The objective of the study is measuring whether discriminatory decisions are reached at different time points during the unfolding of these visual, auditory and audio-visual expressions.

Methods

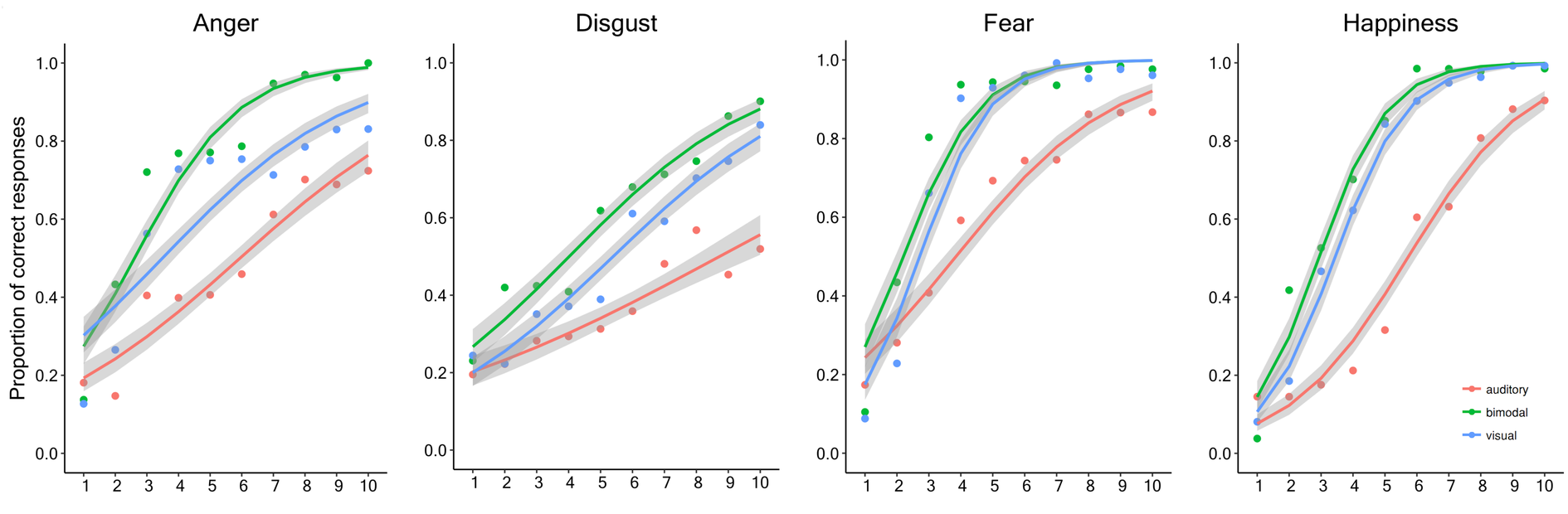
- Three modalities of dynamic stimulation
  - visual (facial expression)
  - auditory (vocal expression)
  - bimodal
- Five basic emotions
  - anger, disgust, fear, happiness and sadness
- Ten gates[2]
  - shortest: 100ms
  - longest: 400ms
  - gate size: 33.37ms



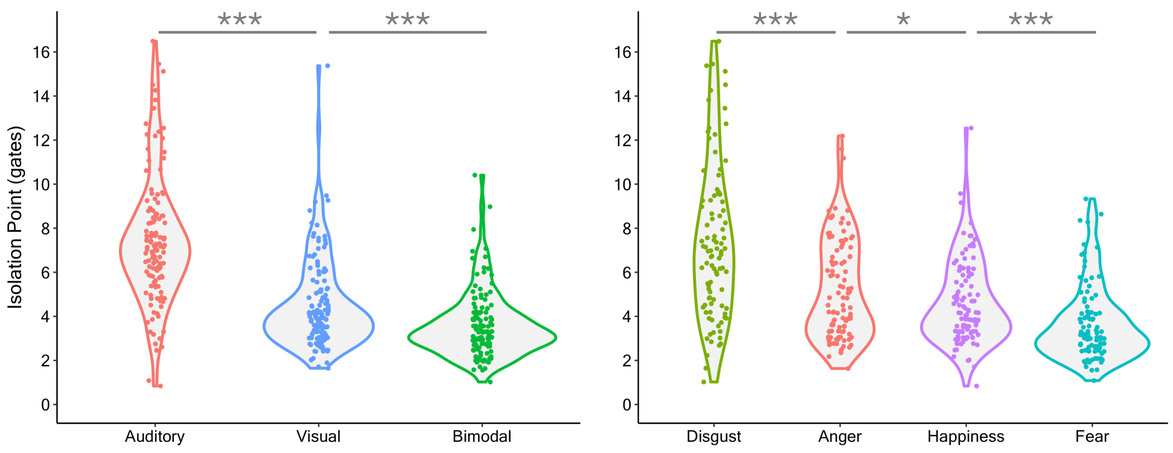
- Task: report the emotion expressed in the stimulus
- 36 healthy volunteers

Results

Psychometric curves:
maximum likelihood logistic curves fitted to the proportion of correct responses as a function of the stimulus duration (gate). Thresholds (Isolation Points) are extracted from the curves at 0.6 proportion of correct responses.



Violin plots:
Isolation points values extracted from the fitted functions.



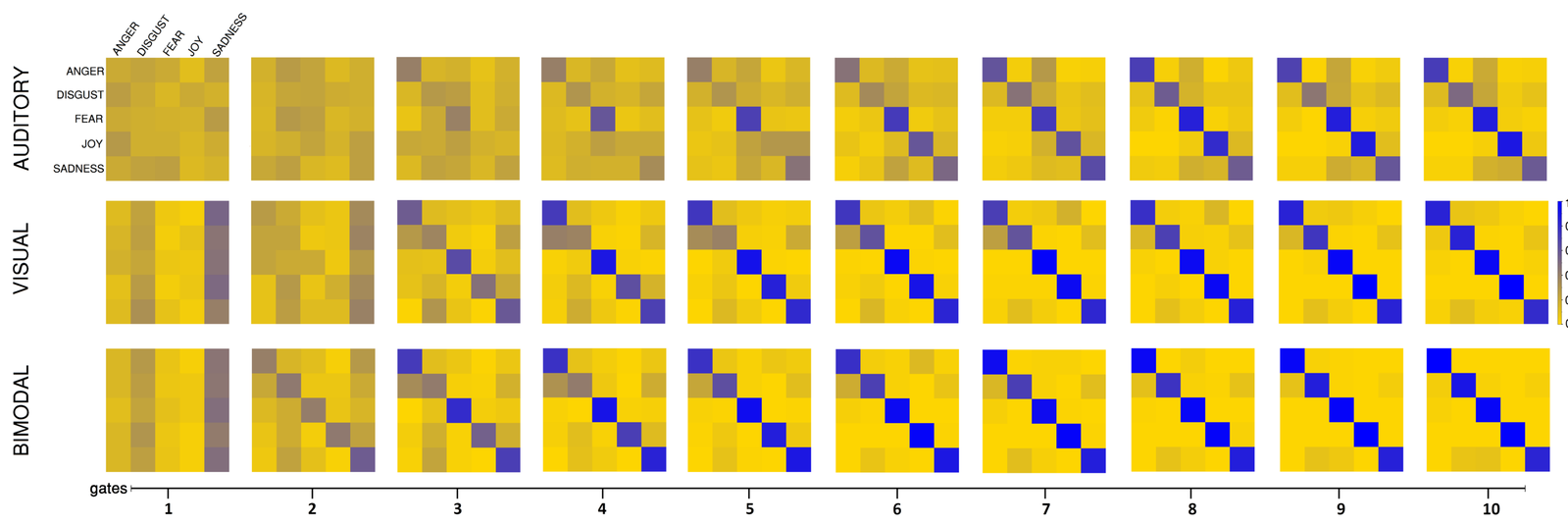
We observed:

- The Isolation Point (IP) is reached earliest in the bimodal condition, then in vision and finally in audition: IP (bimodal) < IP (visual) < IP (auditory)
- The IP is also reached earliest for fear, followed by happiness, anger, and finally disgust: IP (fear) < IP (happiness) < IP (anger) < IP (disgust)

*** p < .001; ** p < .01; * p < .05

Generalized Linear Mixed Model (marginal R^2=.62; conditional R^2=.67): Isolation Point ~ modality * emotion + (1|subj)

Confusion matrices:
emotions labeled on the rows are the content of the stimuli, emotions on the columns are the subjects' responses. The diagonals therefore represent correct responses, while the off-diagonal data constitutes the "confusion patterns".



We observed:

- The confusion patterns in the bimodal condition can be explained both by the visual confusions and, to a lesser extent, the auditory confusions.

Linear regression: bimodal ~ visual + auditory, R^2 = 0.23; visual estimate = 0.477; auditory estimate = 0.55; all ps < .0001

A strong bias to report "*sadness*" for neutral facial stimuli (see matrices of gate 1 in the visual and bimodal conditions) did not allow curve fitting for *sadness* data

Conclusions

- In a unimodal context, emotional information accumulates faster (lower IP) in vision than in audition. Additionally, bimodal confusion patterns can be predicted by visual confusions to a wider extent than auditory confusion patterns. Both these results suggest that facial expressions constitute a more reliable source when discriminating emotions than vocal expressions.
- In multimodal perception, information accumulates faster and allows to reach an earlier decision compared to either unimodal conditions for all investigated emotions
- The different emotion expressions investigated need different amounts of exposure time in order to be efficiently discriminated. *Fear*, often suggested to be "special" among the basic emotions[3,4], is in the present study the expression that leads to the earliest discriminatory decision.

References

[1] Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. Brain research, 1242, 126–135

[2] Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. Perception & Psychophysics, 28, 267–283

[3] Milders, M., Sahraie, A., Logan, S., & Donnellon, N. (2006). Awareness of faces is modulated by their emotional meaning. Emotion, 6(1), 10–17

[4] Yang, E., Zald, D., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. Emotion, 7(4), 882–886

Acknowledgments

This study was supported by the European Research Council starting grant (MADVIS grant #337573), awarded to Olivier Collignon.

Corresponding author: federica.falagiarda@uclouvain.be