FONDAZIONE BRUNO KESSLER

eCREAM enabling Clinical Research in Emergency and Acute care Medicine

UNIVERSITÀ DI PADOVA

RESEARCH AND ED PRACTICE **GET TOGETHER**

Pedro Campana
psouzacampana@fbk.eu
15 October 2026



Multimodal Sentiment Analysis: What Can We Learn About the Doctor–Patient Relationship from a Recorded Consultation?

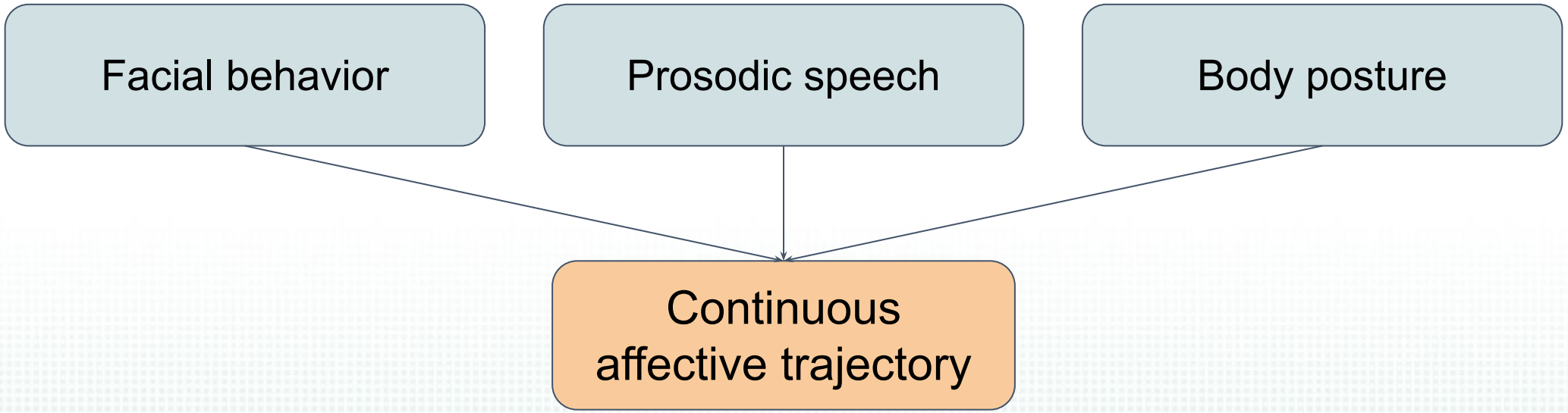
INTRODUCTION

Motivation

- Emotions play an important role in clinical consultations:
  - Help the doctor identify moments of confusion or distress that may not be verbally expressed
  - Support more adaptive communication strategies by revealing how patients react to specific explanations or decisions
  - Ensure emotional experience is recognized, even when patients struggle to articulate it
- Current research mostly reduces patient affect to discrete labels that miss how emotions evolve during the encounter

Goal

- Use **non-verbal modalities** to build continuous affective trajectories that describe how emotional intensity changes during a clinical consultation
- Measure changes in emotional intensity over time and examine how they relate to clinically important events



Research Challenges

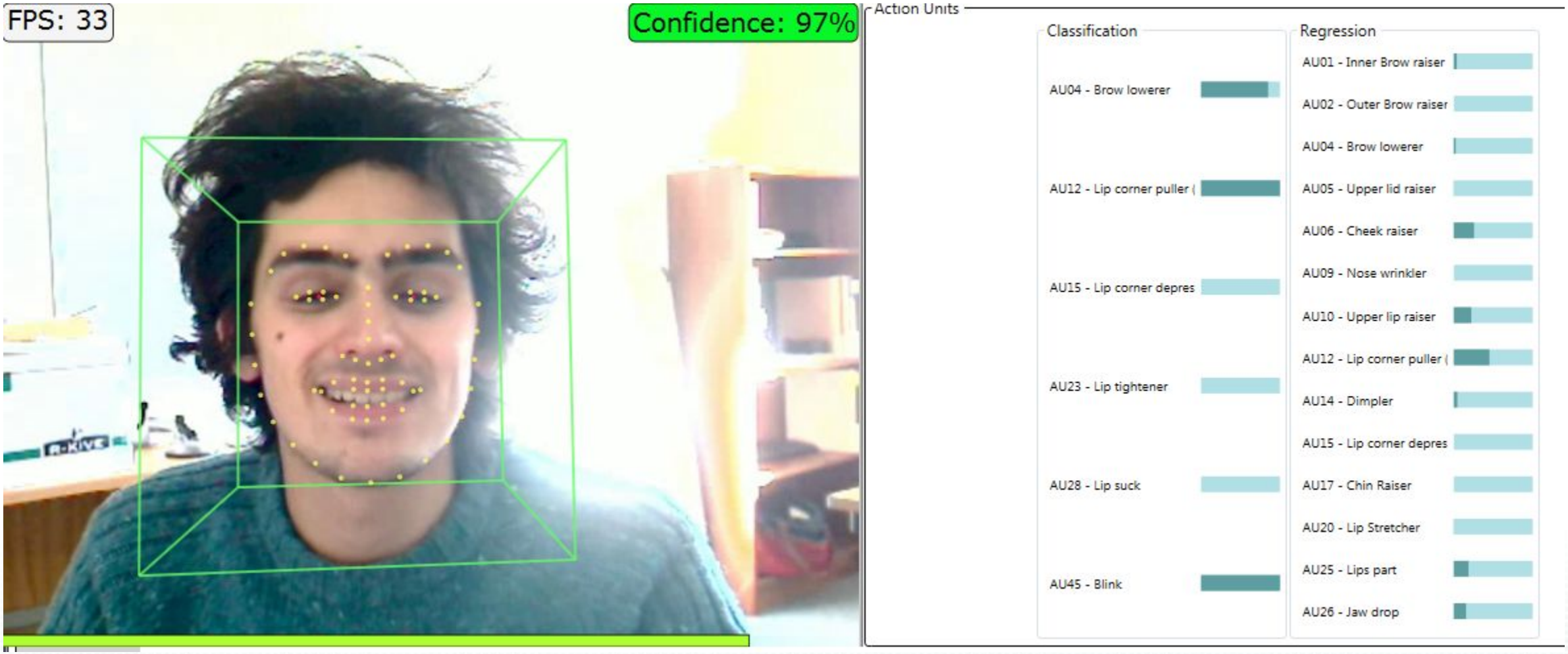
- Patients can react differently to the same clinical event
- Recurrent consultations reduce emotional impact
- Modalities have different limitations and can reveal peaks at different times
- Consultations can vary in duration and clarity

Videos

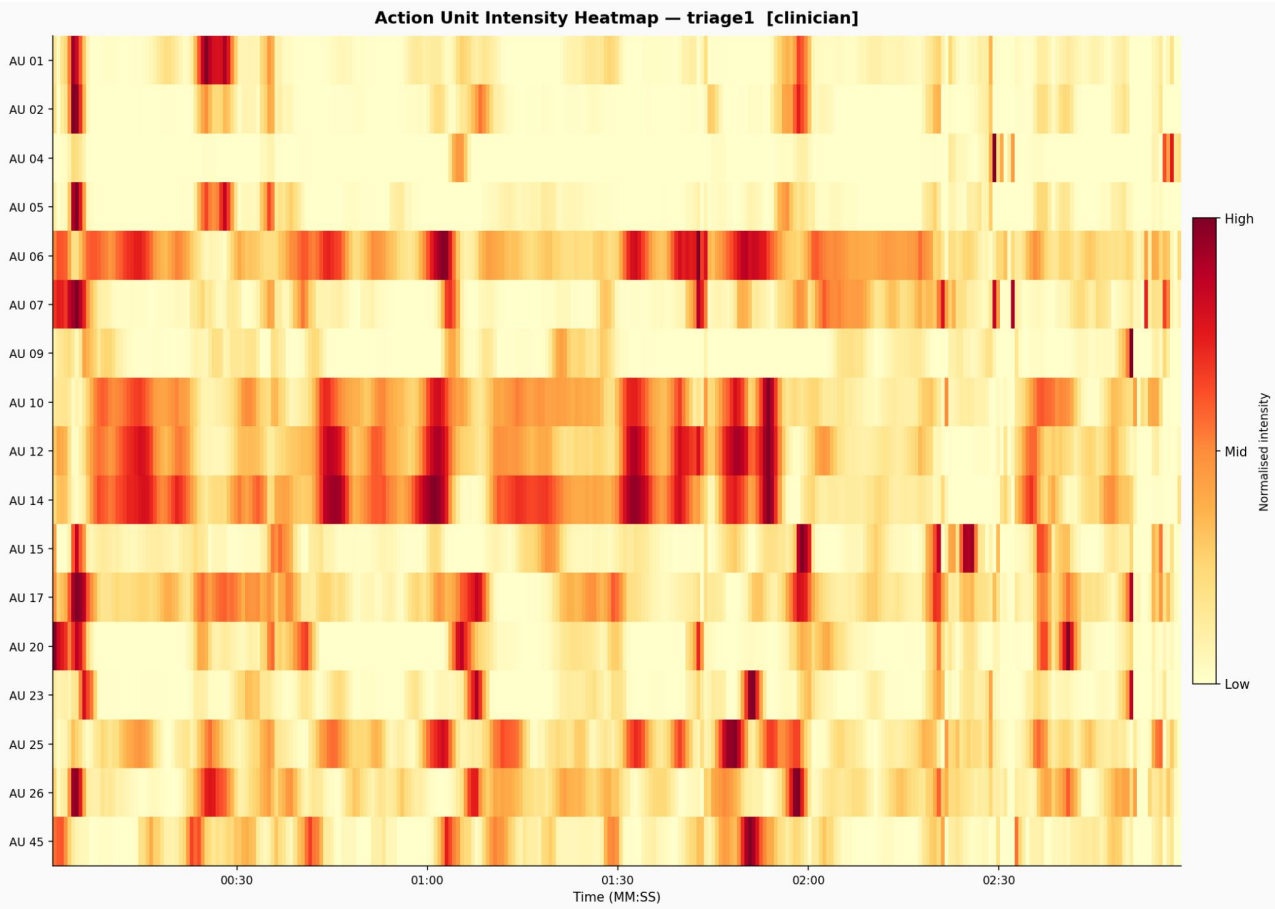


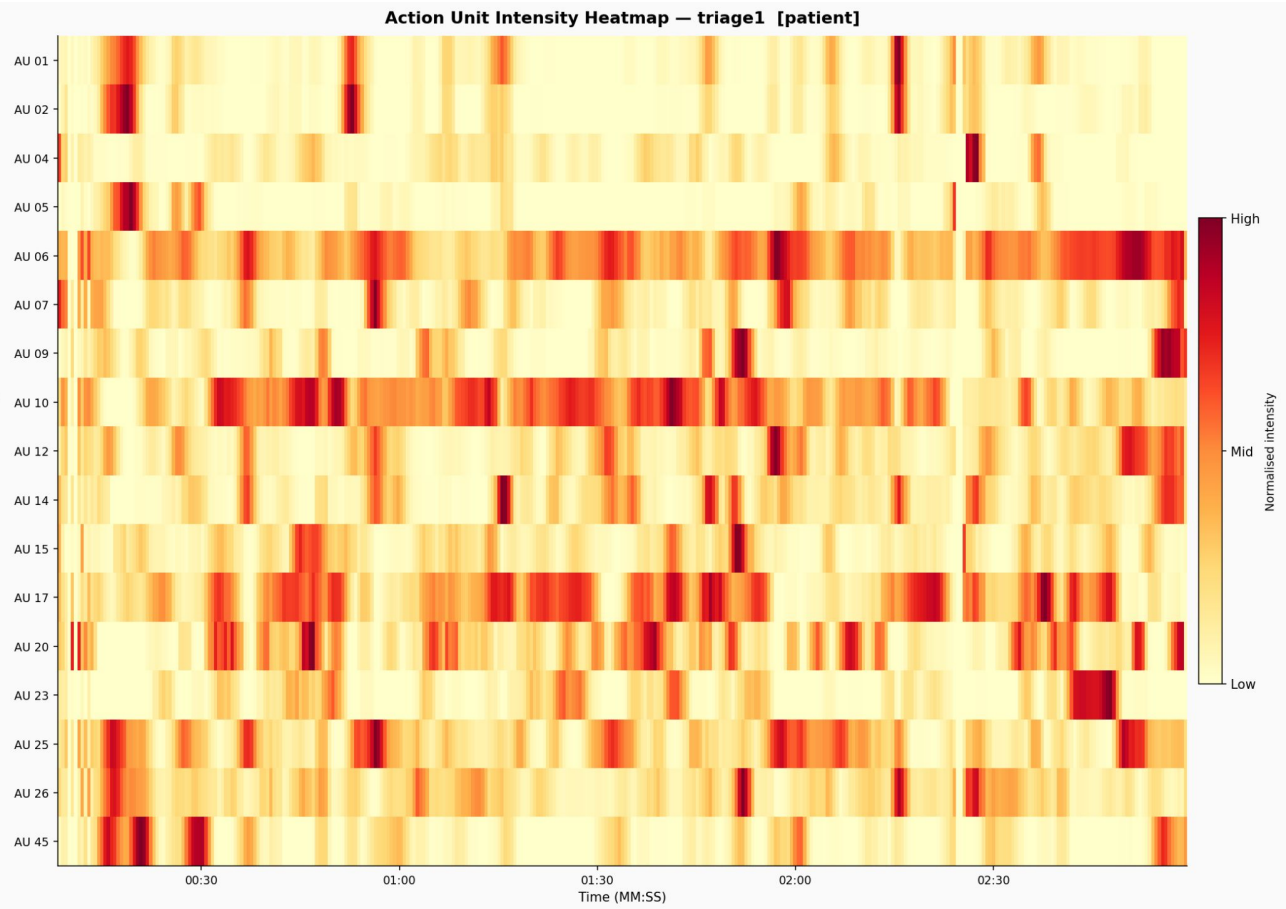
FACIAL BEHAVIOR

OpenFace 2.2.0



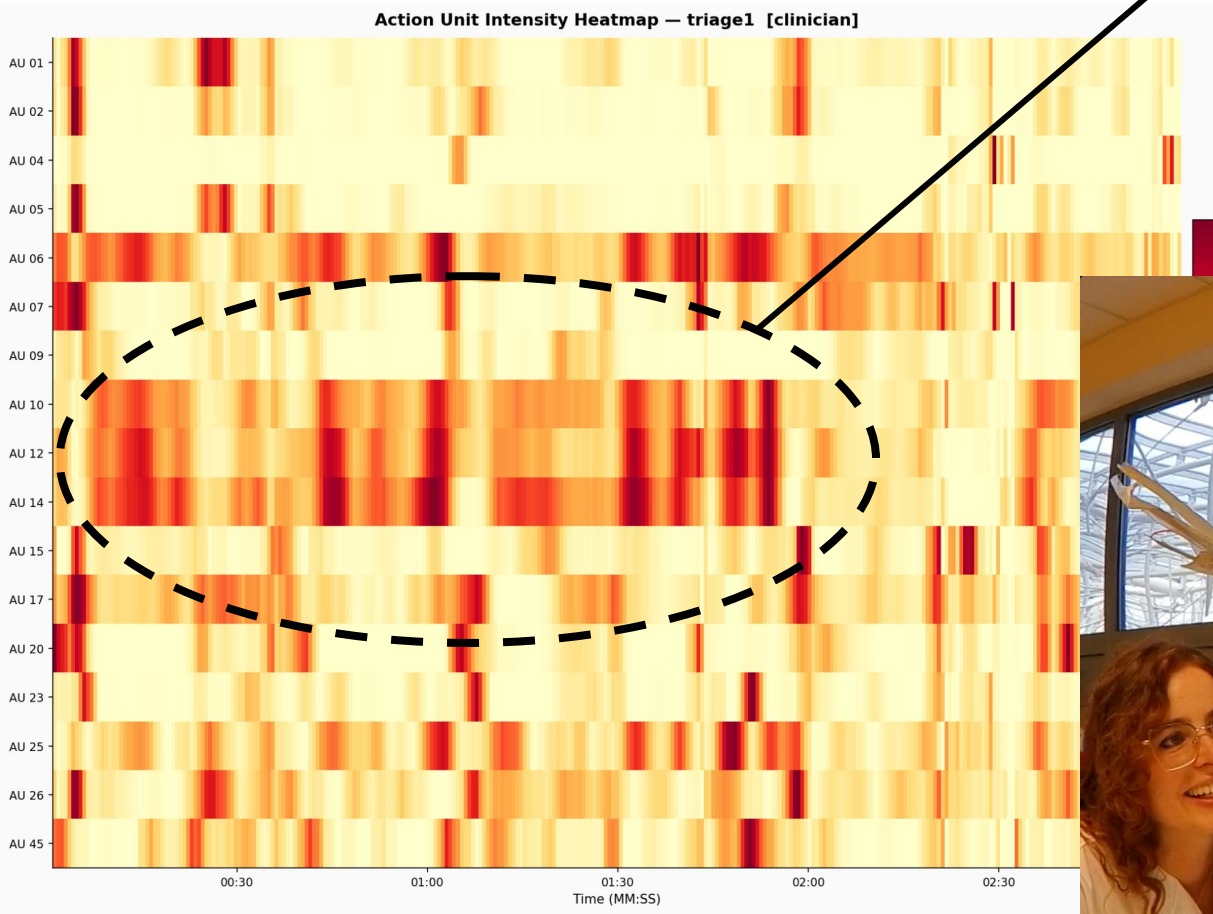
Facial Action Units

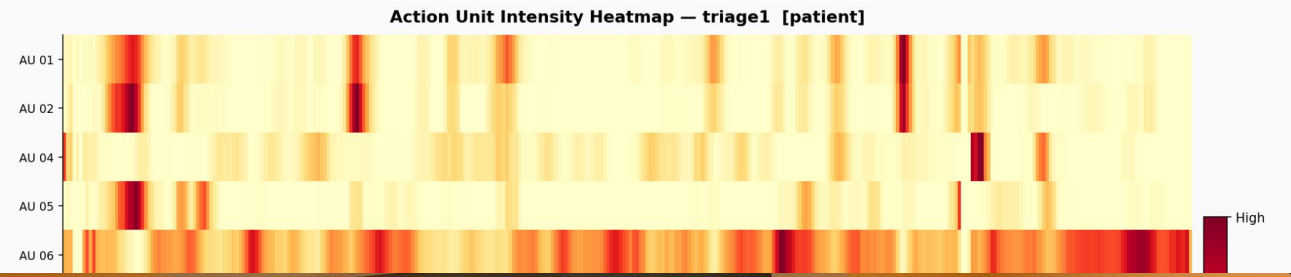




Facial Action Units

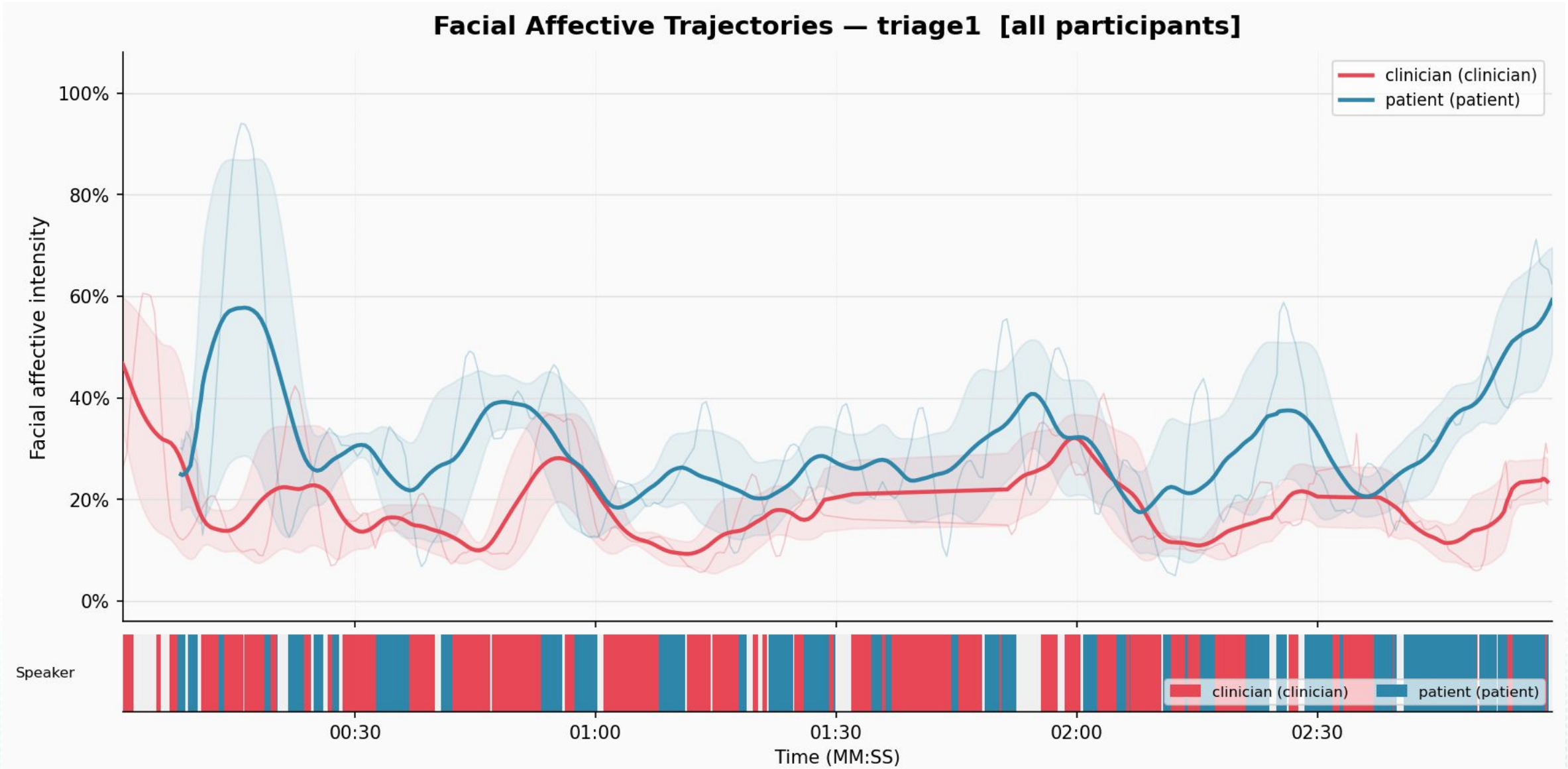
Clinician smiles more
(polite and professional composure)







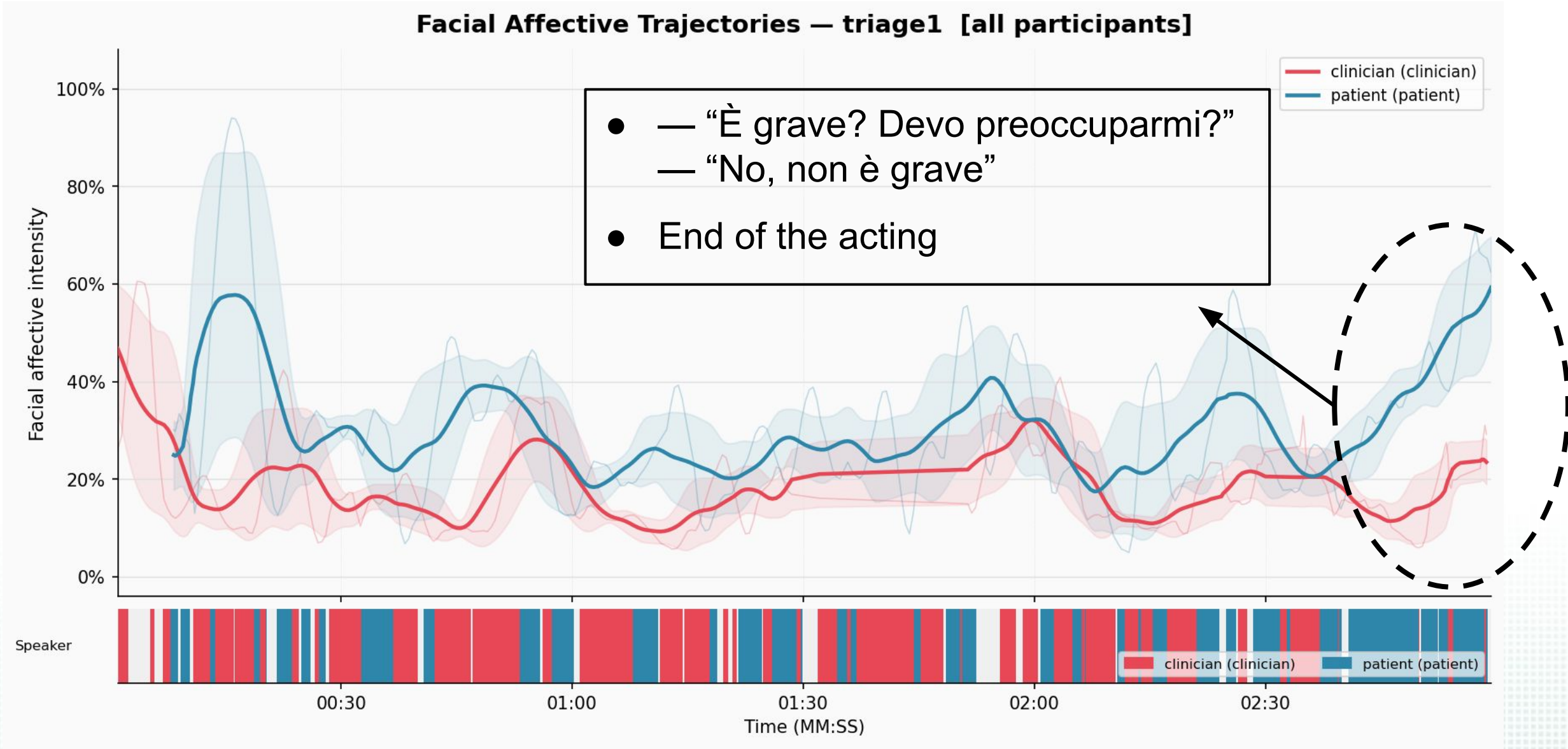
Facial Trajectory



Facial Trajectory



Facial Trajectory



PROSODIC SPEECH

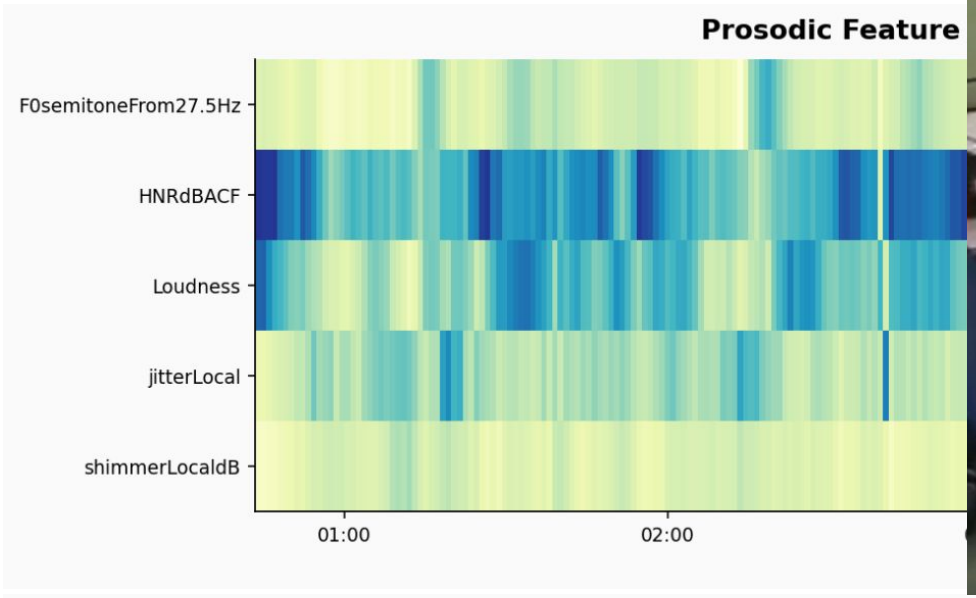
openSMILE



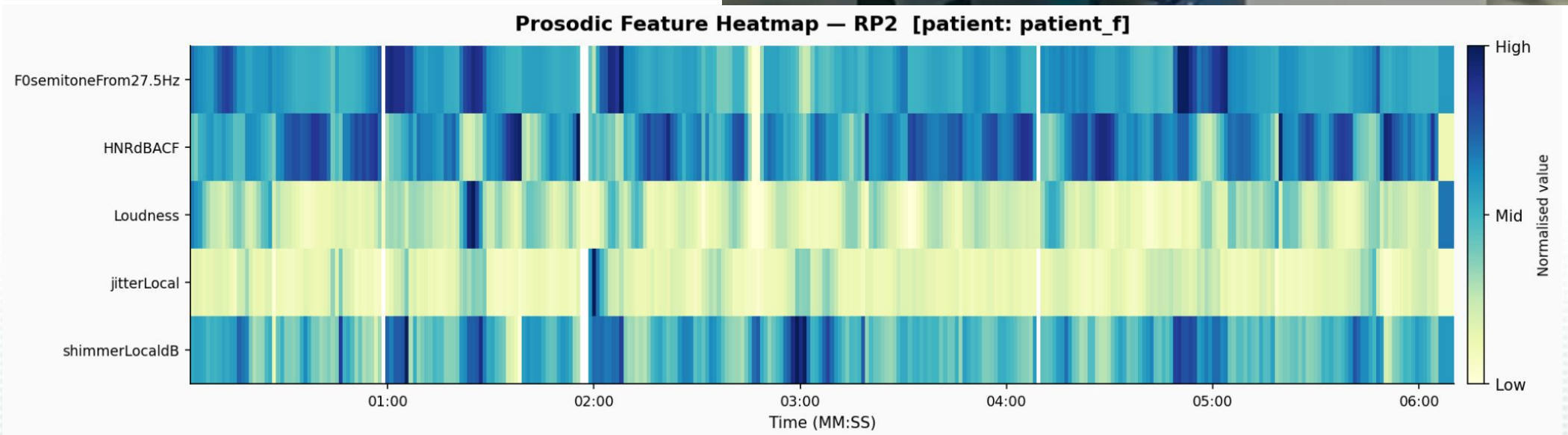
Features used:

- Pitch
- Harmonics-to-noise ratio: noise, breathiness, or roughness?
- Loudness: perceived intensity or energy of the speech signal
- Jitter: variations in pitch (higher jitter → less stable or rough voice)
- Shimmer: variations in amplitude (higher shimmer → rough voice)

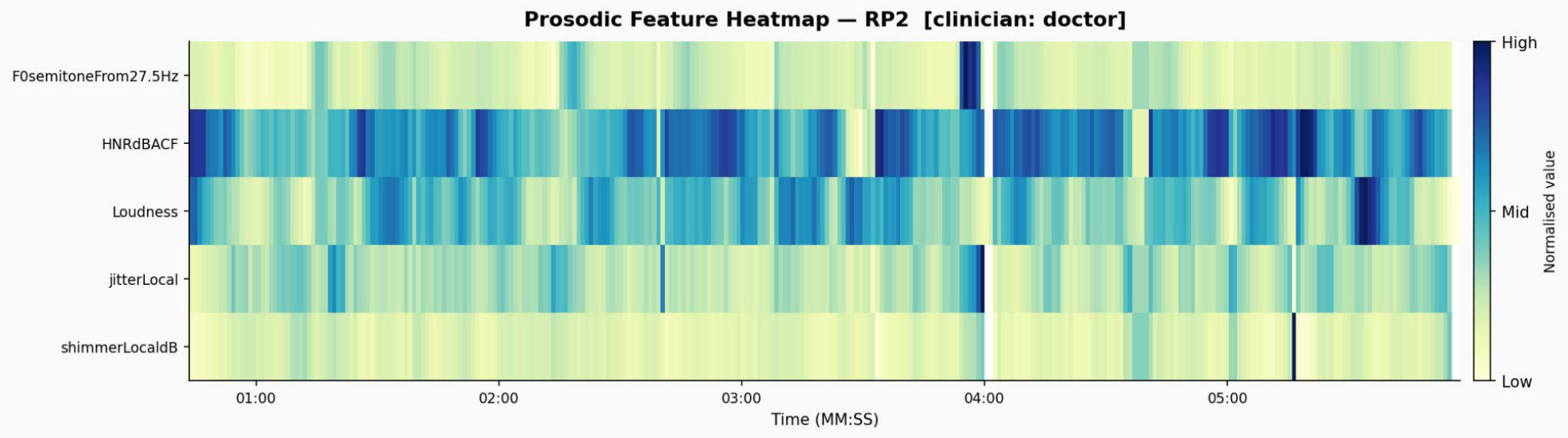
Feature Heatmap

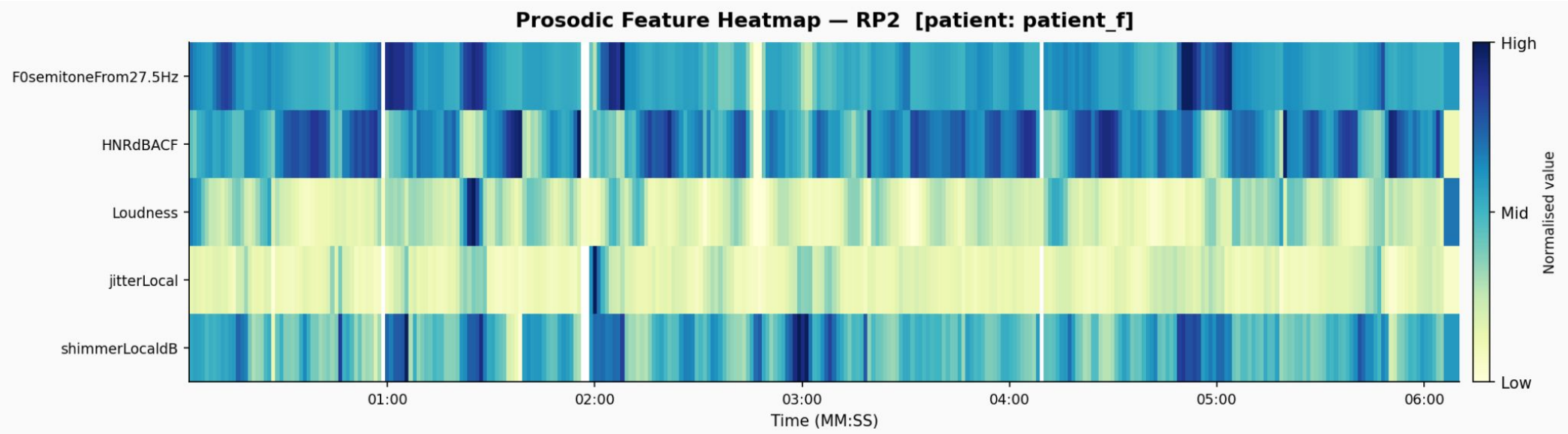




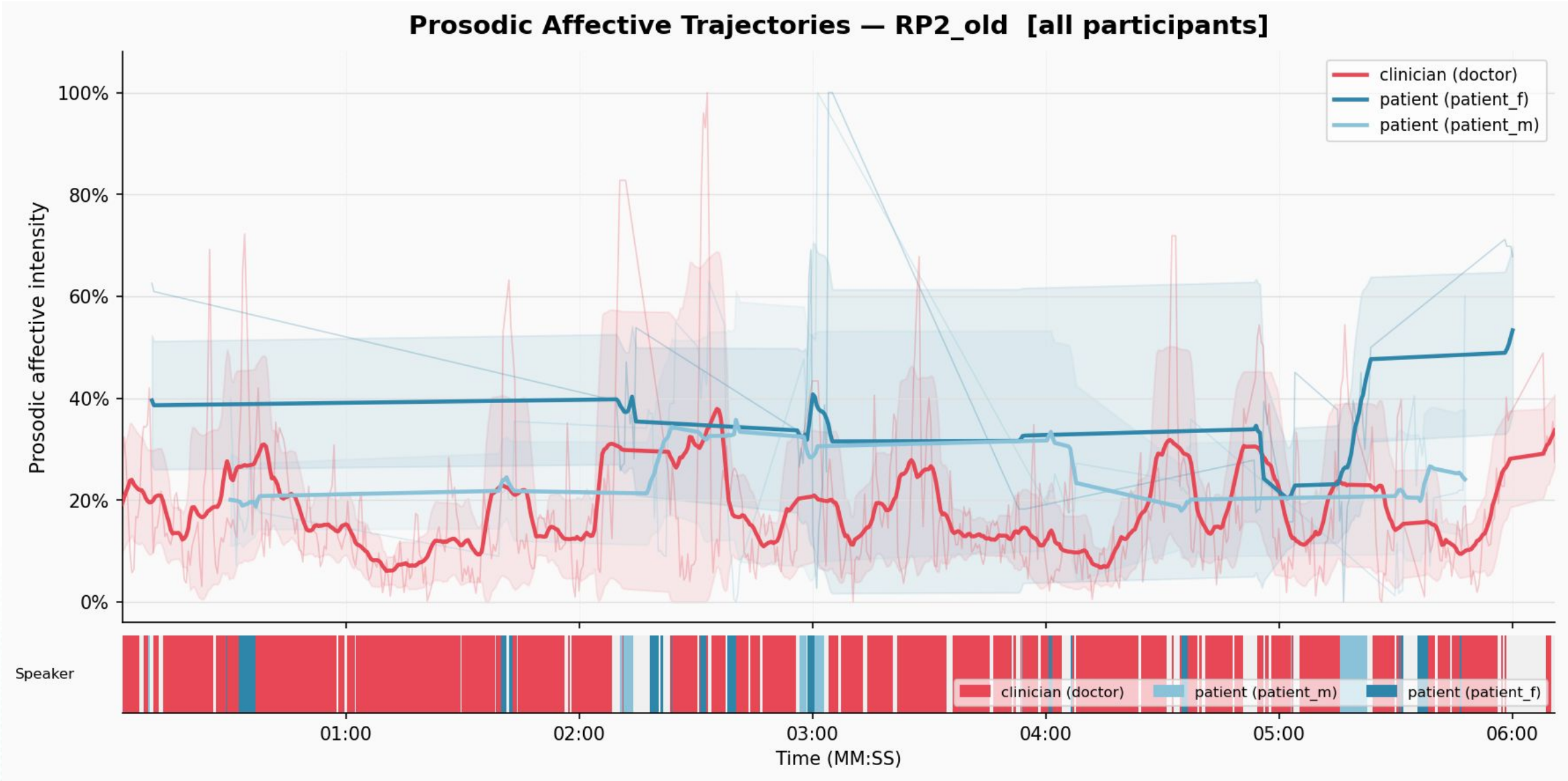


Feature Heatmap





Speech Trajectory

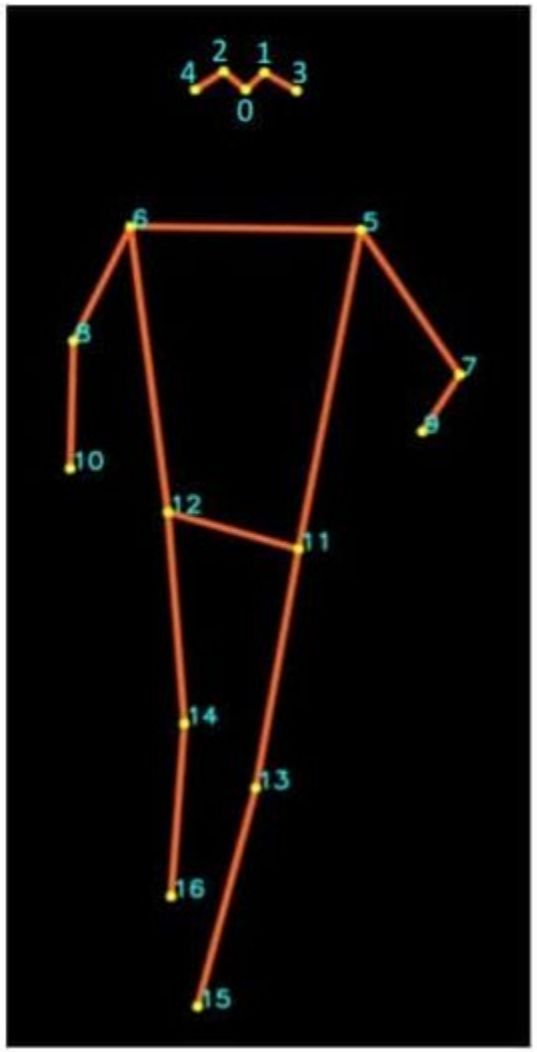


BODY POSTURE

YOLO v8 (Keypoint Extraction)

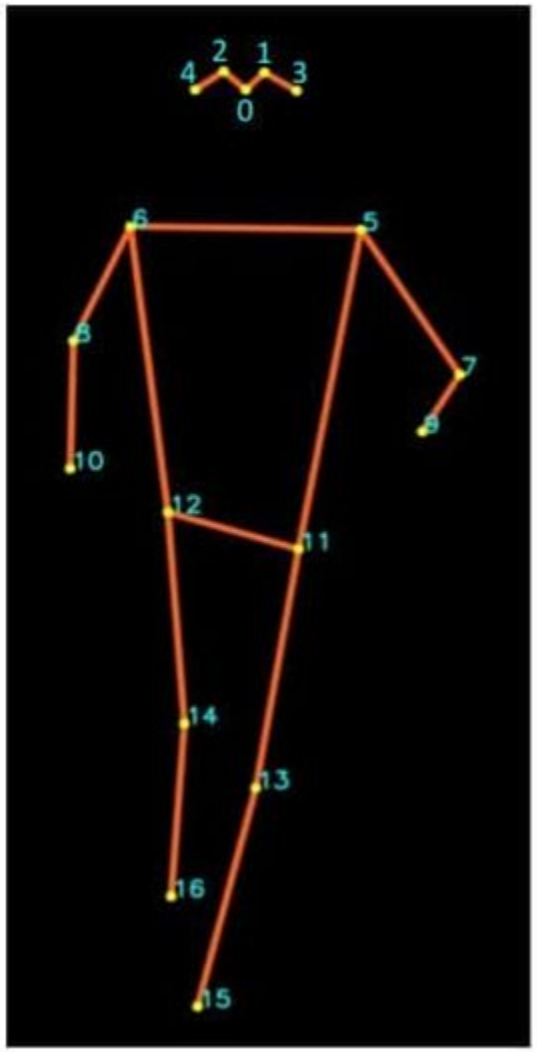


Index	Key point
0	Nose
1	Left-eye
2	Right-eye
3	Left-ear
4	Right-ear
5	Left-shoulder
6	Right-shoulder
7	Left-elbow
8	Right-elbow
9	Left-wrist
10	Right-wrist
11	Left-hip
12	Right-hip
13	Left-knee
14	Right-knee
15	Left-ankle
16	Right-ankle



Posture Features

- Head inclination degree
- Shoulder asymmetry
- Postural shift velocity
- Wrist-to-nose distance
- Wrist velocity

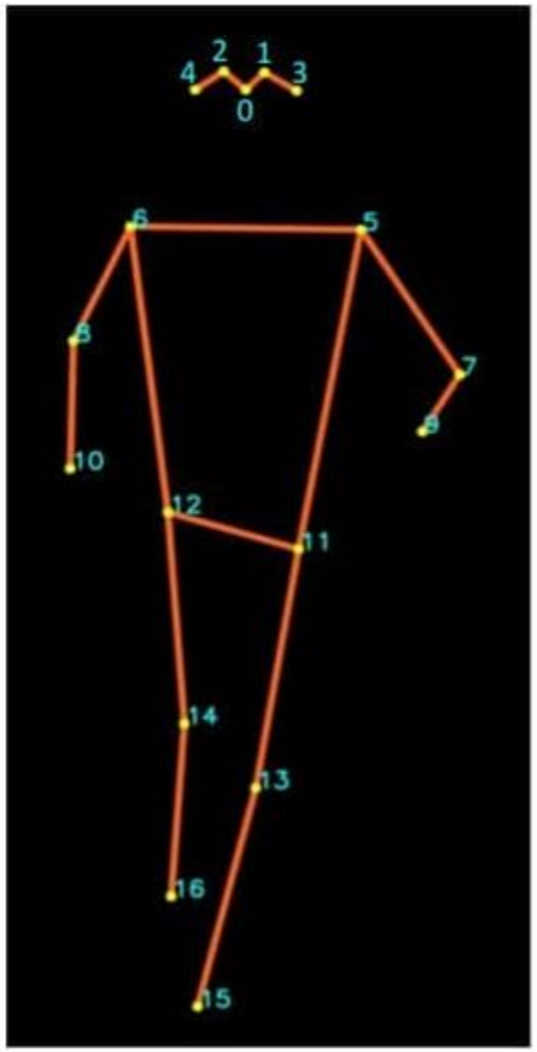


Posture Features

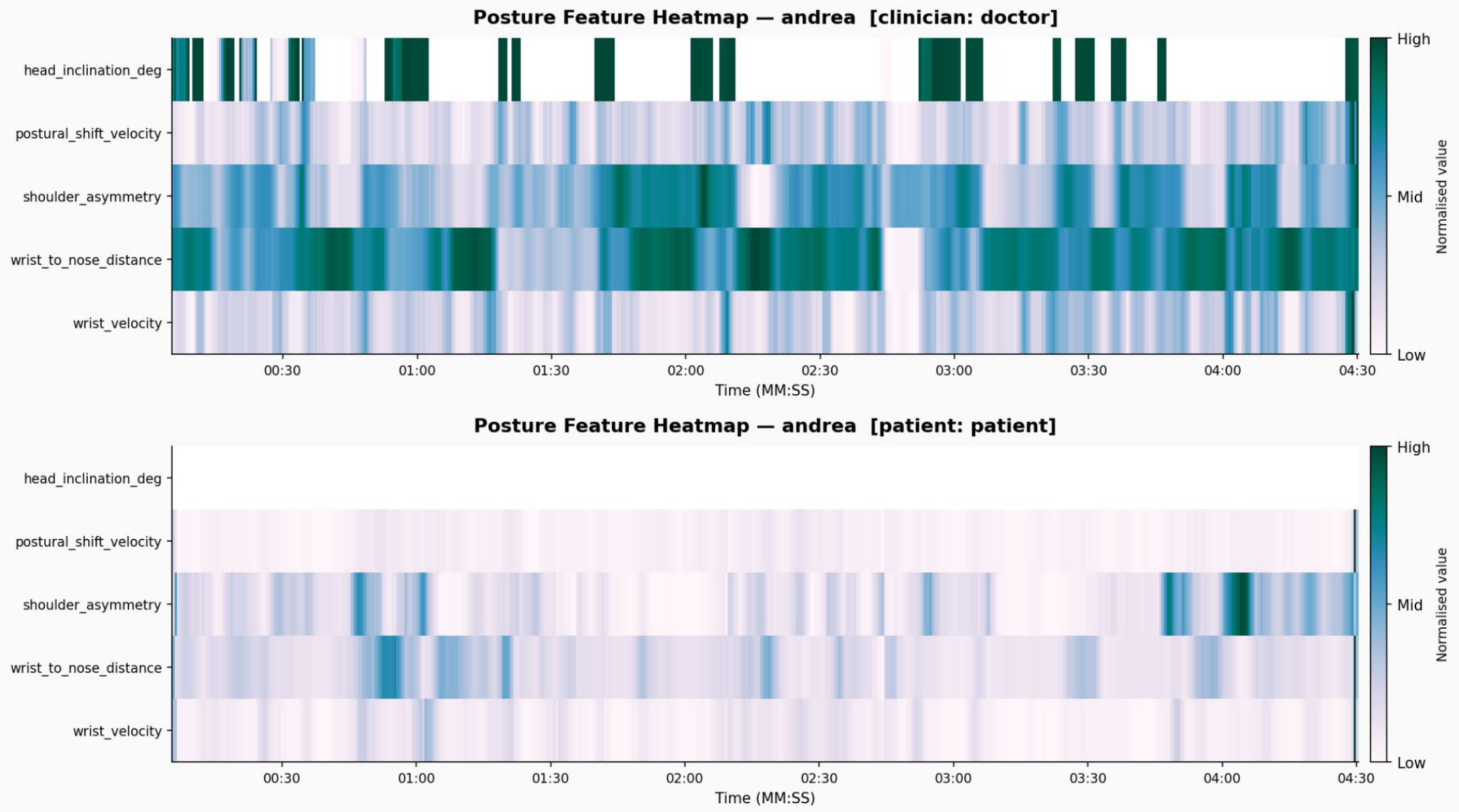
- Head inclination degree
- Shoulder asymmetry
- Postural shift velocity
- Wrist-to-nose distance
- Wrist velocity

$$v_{\text{posture},t} = \frac{\|\mathbf{h}_t - \mathbf{h}_{t-1}\|}{\Delta t \; SW_t}$$

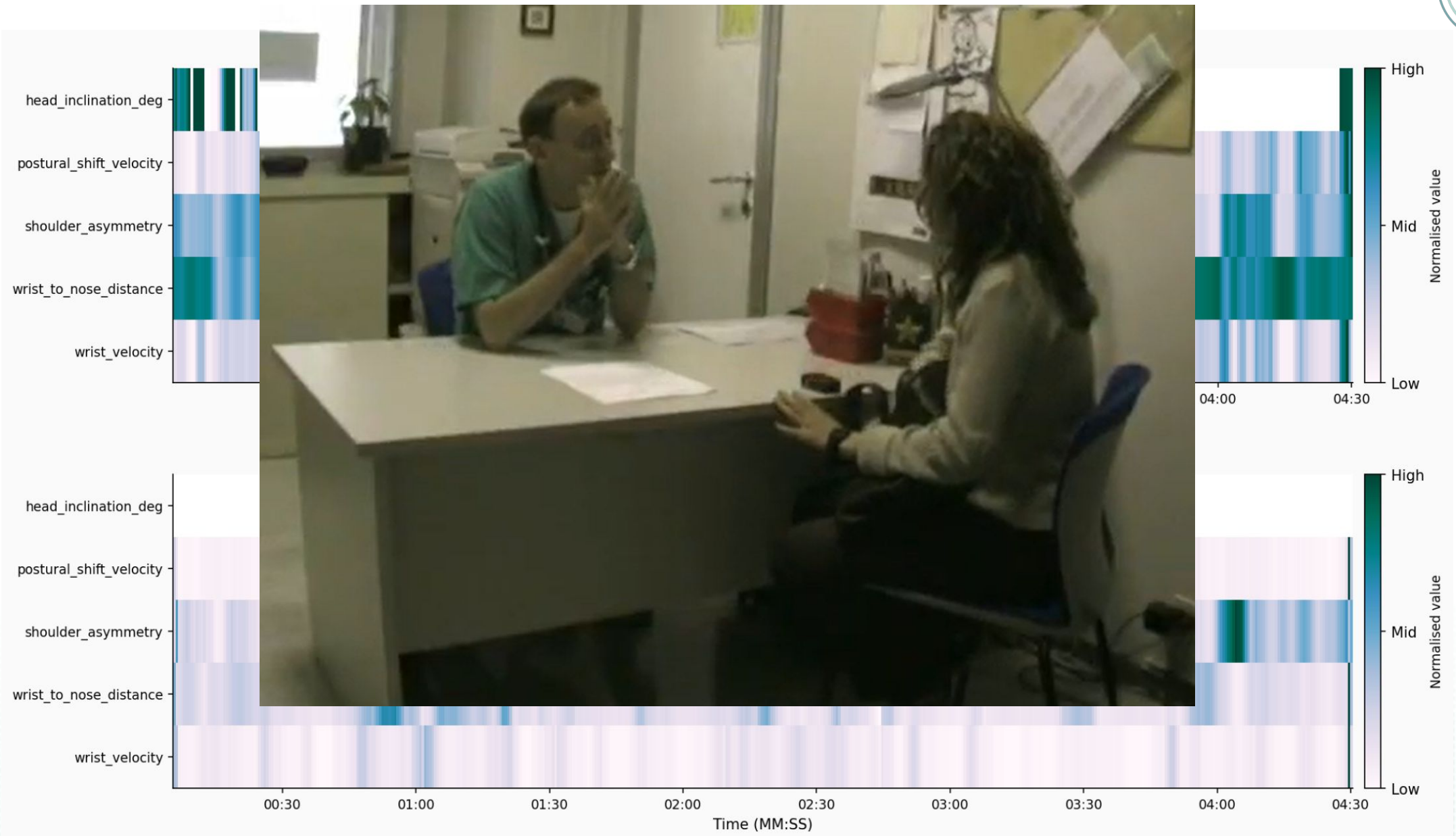
$$v_{\text{wrist},t} = \frac{1}{N} \sum_{i \in \{\text{left,right}\}} \frac{\|\mathbf{W}_{i,t} - \mathbf{W}_{i,t-1}\|}{\Delta t \; SW_t}$$



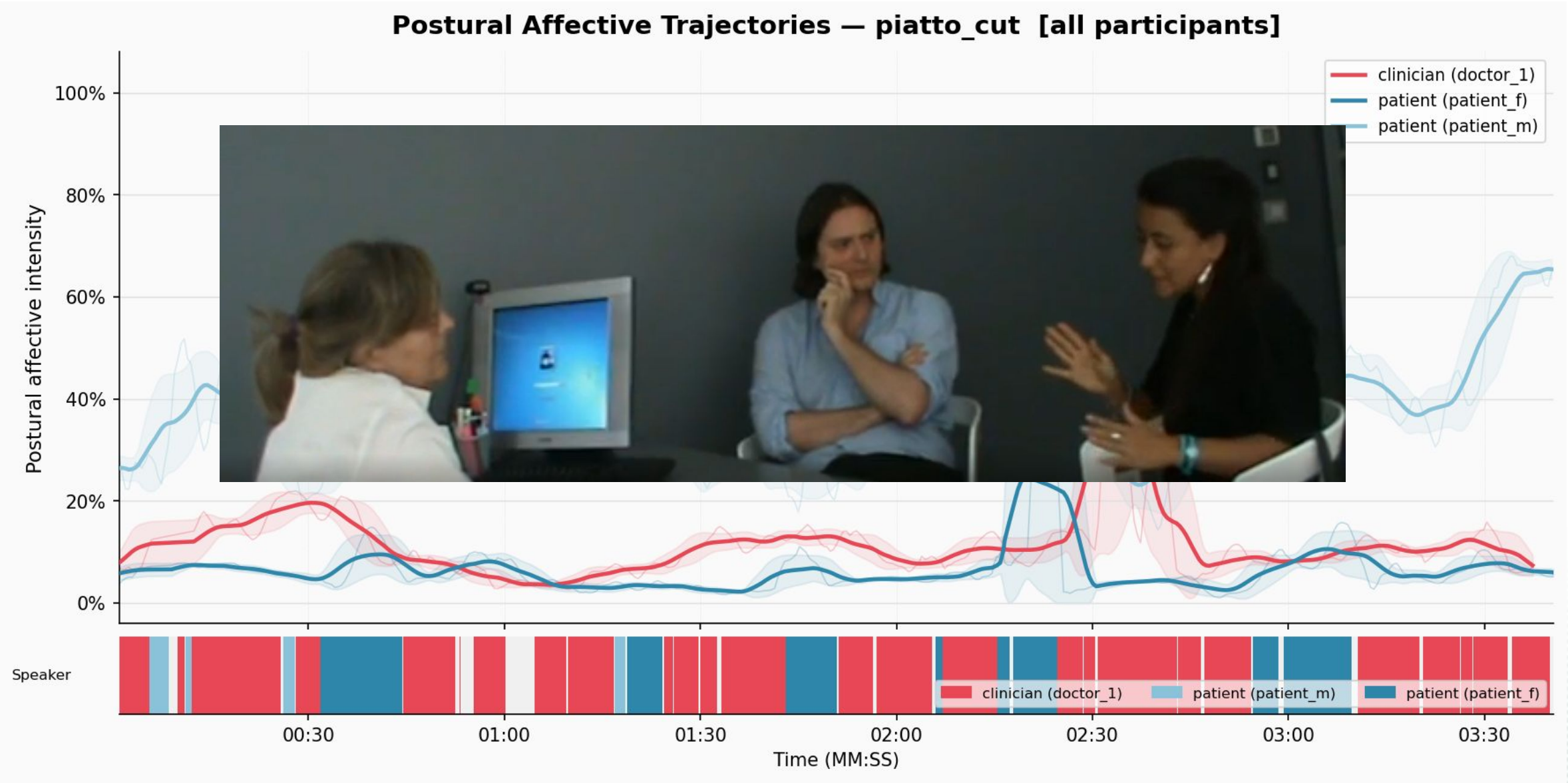
Feature Heatmap



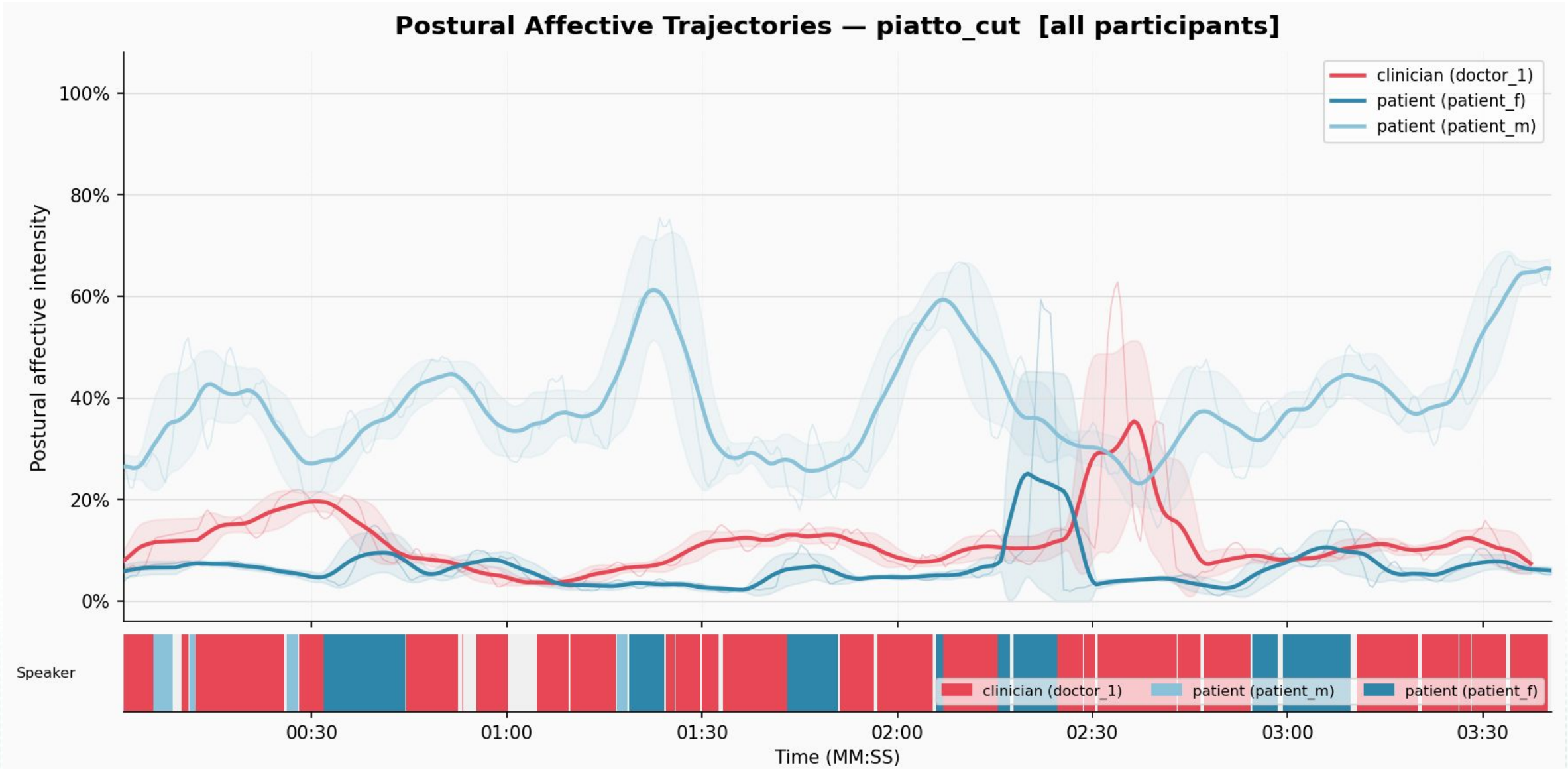
Feature Heatmap



Postural Trajectory



Postural Trajectory



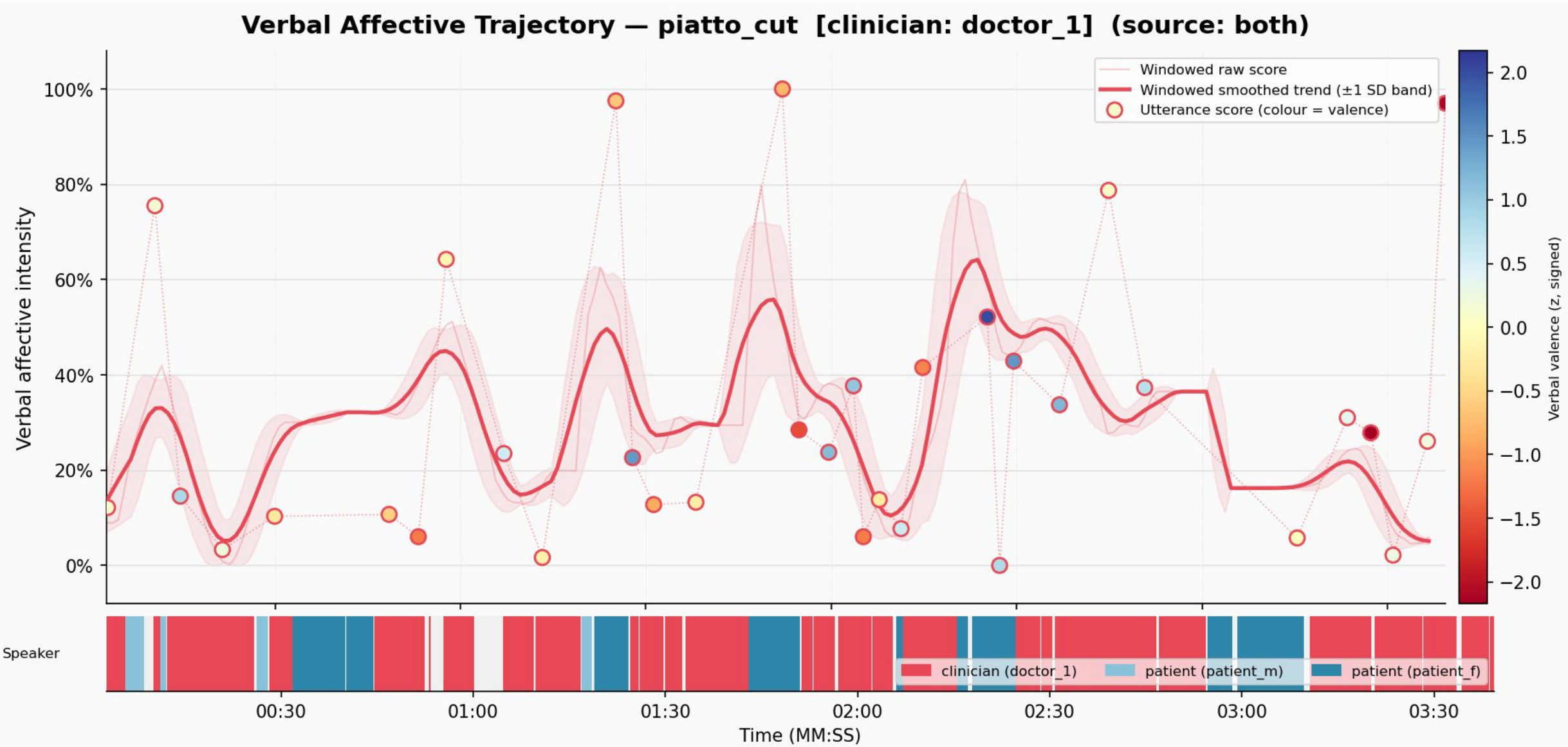
VERBAL TRAJECTORY

Valence and arousal

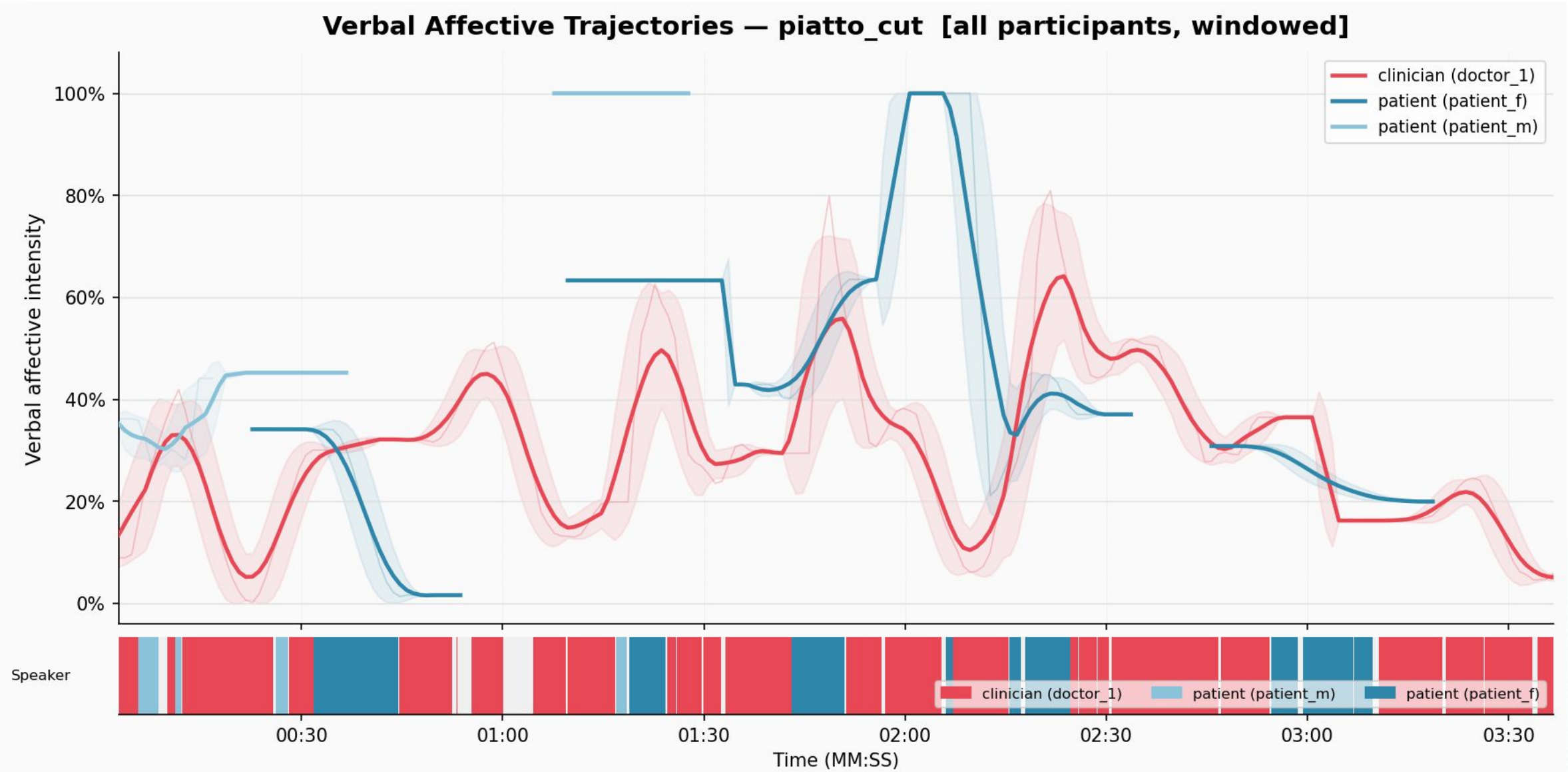
- Match each **utterance**'s words against a **lexicon** to extract word-level affective features like **valence** (positive—negative) and **arousal** (intensity)


```
{
  "start": 137.868,
  "end": 140.17,
  "text": "Sono arrabbiata con il mondo, con il destino.",
  "score": 0.571551,
  "valence": -0.589965
},
```

Verbal Trajectory



Verbal Trajectory

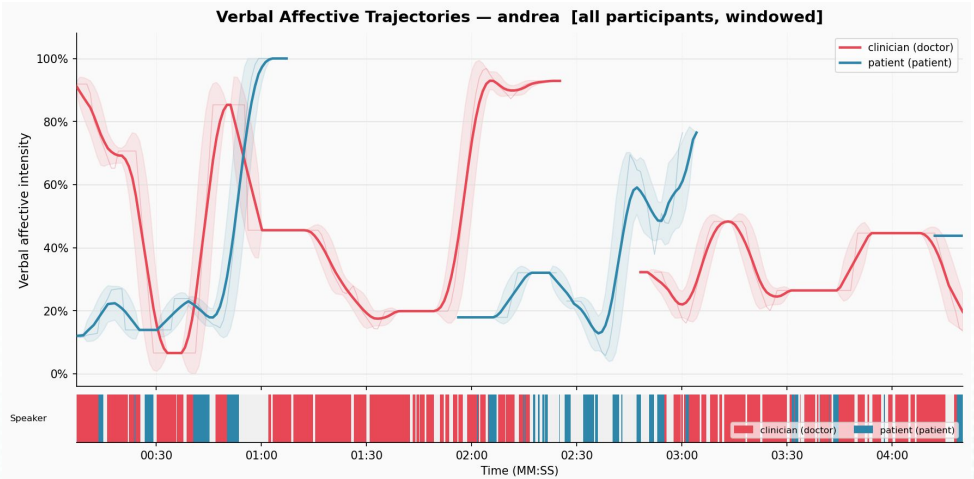


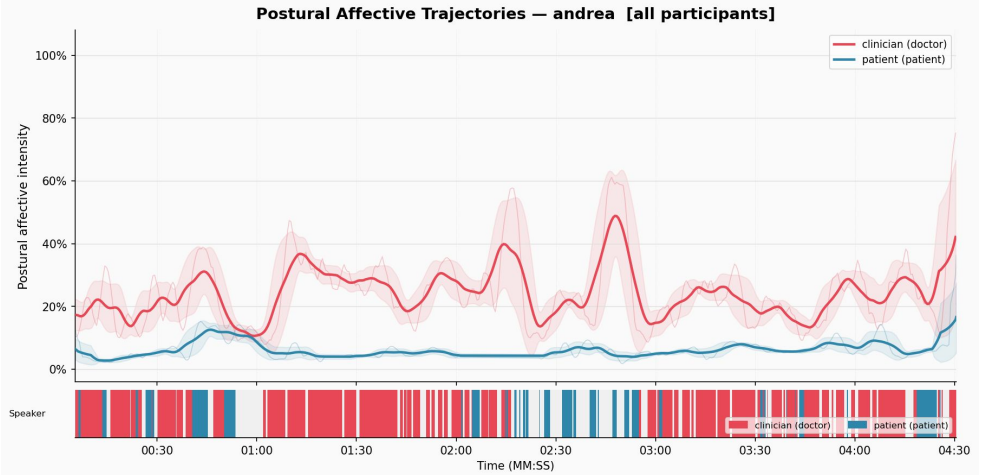


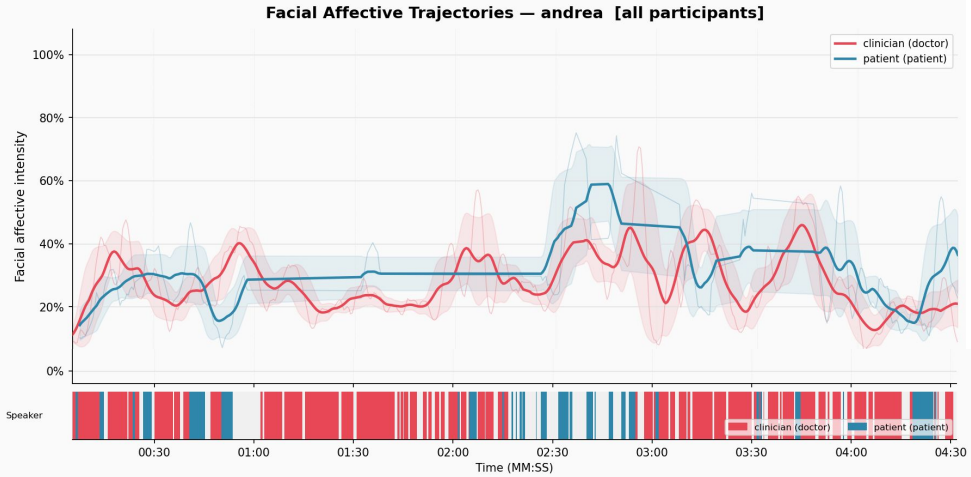
FULL EXAMPLE AND CONCLUSION

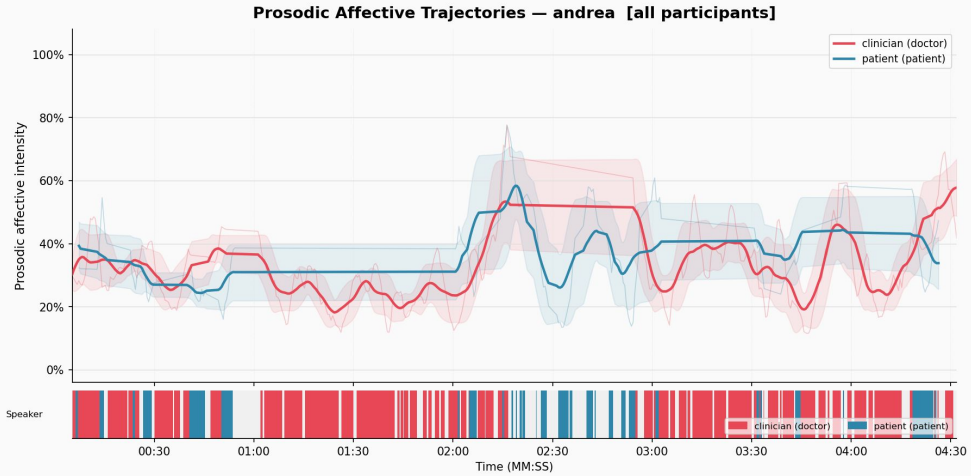
Full Example







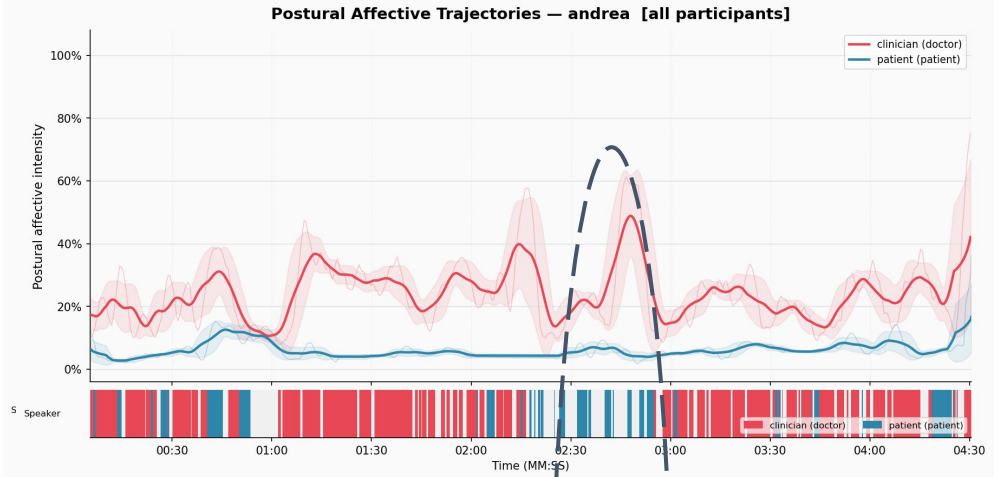


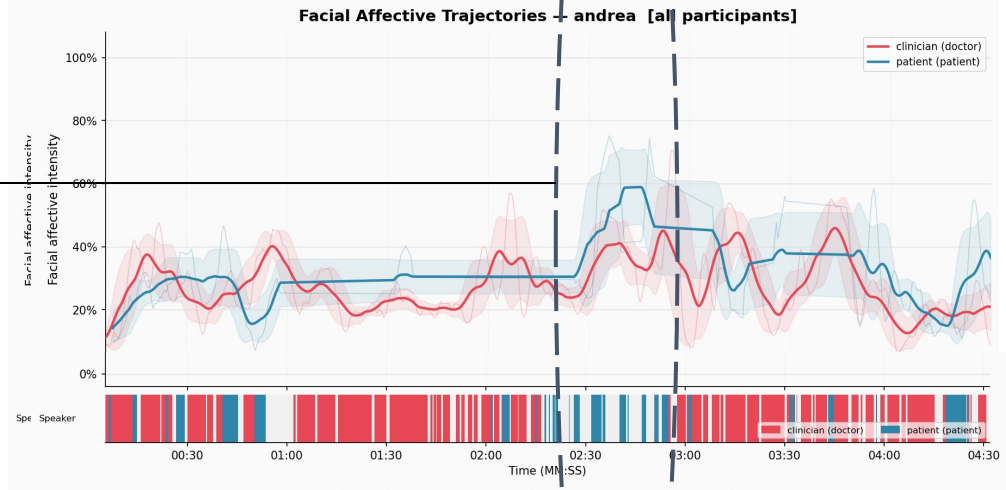


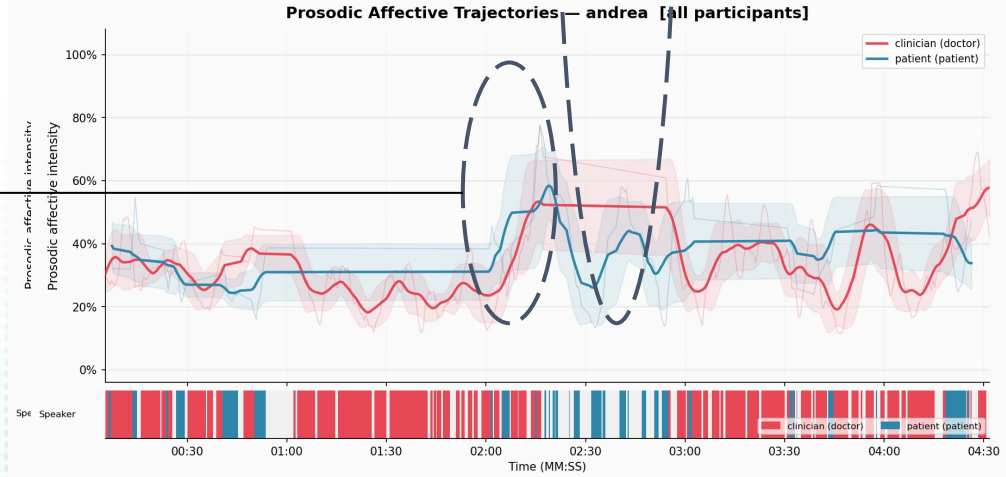
Action and reaction

- Patient speaking
- Highest peak in doctor's postural trajectory
- Highest peak in the patient's facial trajectory
- Peak in the patient's prosodic trajectory

- Doctor speaking
- Highest peak in the patient's prosodic trajectory







Conclusion

- All modalities are informative and relate to each other
- Continuous multimodal trajectories as a representation for studying affective dynamics in healthcare interactions
- Potential for improving doctor-patient relations and help medical students



eCREAM
enabling Clinical Research
in Emergency and Acute care Medicine



THANK YOU FOR YOUR ATTENTION

psouzacampana@fbk.eu

Looking forward