







**Scene:** The scene takes place in a hospital, where Monica and Phoebe are visiting a guy who is in a coma. Monica is struggling to understand why she reacted so strongly to the guy, even though he's unconscious. Phoebe tries to lift the guy's spirits ... ..





Phoebe, what are you doing?





Maybe nobody's tried this





I wish we at least knew his name... Look at that face. I mean, even sleeping, he looks smart. I bet he's a lawyer.



Yeah, but did you see the dents in his knuckles? That means he's artistic.



Okay, he's a lawyer, who teaches sculpting on the side. And- he can dance!



Oh! And, he's the kinda guy who, when you're talking, he's listening, y'know, and not saying 'Yeah, I understand' but really wondering what you look like naked."



**Eye: Read role profiles & the dialogue, scene descriptions.**

**Phoebe:** Phoebe Buffay is the quirkiest and most unconventional member of the Friends group. Her personality is a unique blend of optimism, eccentricity, and kindness,... ..

**Scene:**.....  
**Dialogue:**  
Monica: ....  
Phoebe: ...  
Monica: ...  
Ross: ...



**Ear: Who said what, and how they said it**

**Monica:** Phoebe, what are you doing?  
**Emotion:** Concerned, slightly exasperated.....

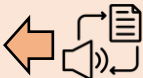
**Phoebe (Me):** Maybe nobody's tried this  
**Monica:** I wish we at least knew his name...  
**Emotion:** Concerned, slightly exasperated.....

**Phoebe (Me):** Yeah, but did you see the dents in his knuckles? That means he's artistic.  
**Monica:** Okay, he's a lawyer, who teaches sculpting on the side. And- he can dance!  
**Emotion:** Amused, optimistic, and slightly fanciful, with a hint of romanticism

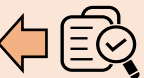


**Mouth: Speak the utterance with the intended emotion.**

Powered by retrieval argument generation and text-to-speech synthesis.



**Generation:**  
Prompting LLMs for speech synthesis via the retrieved audio sample



**Retrieve:**  
Based on the target emotion state generated by Brain, fetch a previous audio sample expressing a similar feeling.



**Brain: The delivery of my next utterance is shaped by my role, others' dialogue and emotions, and the scene context I observe.**

I am an actor, destined to portray {...}.  
I am performing in the scene {...}.  
The dialogue I have heard so far is: {Speaker: ..., Line: ..., Emotion: ...}  
My next utterance to deliver is: {...}.  
The emotion I should convey in delivering this utterance is: ?