

Question

What is the emotion of the speaker in the audio? After thinking and reasoning, choose the final answer from the following options: ['neutral', 'happy', 'sad', 'angry'].

Question prompt

+



Input audio

Groundtruth: happy

Response

Implicit Reasoning (IR)

✗ <answer> neutral </answer>

Explicit Unstructured Reasoning (EUR)

<think>

🤔 Based on the tone and rhythm of the voice, one can sense a certain impatience in the speaker's emotion. This emotion manifests in the speech as slightly *faster pace*, somewhat *sharp intonation*, and *certain words being pronounced less clearly*. From the sounds they emit, it can be inferred that they might be experiencing some inner struggle and emotional conflict. Therefore, I would describe this scenario as "angry."

</think>

✗ <answer> angry </answer>

Explicit Structured Reasoning (ESR)

<think>

<content>

🤔 The content of this speech is "*let's enjoy that party*" and "*don't forget to blow trumpets when the cake comes.*" From the content, it's evident that the speaker is *attending a party* and looking forward to cake and trumpets.

</content>

<acoustics>

🤔 Based on the voice's tone and rhythm, one can perceive that the speaker has *a higher pitch, faster pace, and louder volume*, which are vocal characteristics typical of excitement or happiness.

</acoustics>

<summary>

💡 *These all indicate that the speaker is enjoying the party and feeling happy.* This positive and confident emotion is likely an expression of happiness. Therefore, the corresponding emotion is likely "happy".

</summary>

</think>

✓ <answer> happy </answer>