

# LLM's Assessment for actual conversation content

---

Currently no LLM has the functionality to do this task by itself.

Here is the output that we want to get to:

- 1. **\*\*Tone and Vocal Quality (SER)**
  - Tech Recruiter: Warmth: 9, Confidence: 8, Sincerity: 9
  - CEO: Warmth: 7, Confidence: 9, Sincerity: 8
  - Overall Rating: 8.5
- 2. **Pacing (Praat)**
  - Tech Recruiter: Pacing Score: 8
  - CEO: Pacing Score: 9
  - Overall Rating: 8.5
- 3. **Use of Silence (VAD)**
  - Tech Recruiter: Effective Use of Silence: 7
  - CEO: Effective Use of Silence: 8
  - Overall Rating: 7.5
- 4. **Audibility and Clarity (PESQ)**
  - Tech Recruiter: Audibility: 9, Clarity: 9
  - CEO: Audibility: 8, Clarity: 9
  - Overall Rating: 8.8
- 5. **Interruption Handling (Rasa)**
  - Tech Recruiter: Interruption Handling: 9
  - CEO: Interruption Handling: 8
  - Overall Rating: 8.5
- 6. **Verbal Affirmations (Wit.ai)**
  - Tech Recruiter: Verbal Affirmations Count: 4, Appropriateness: 9
  - CEO: Verbal Affirmations Count: 3, Appropriateness: 8
  - Overall Rating: 8.5
- 7. **Non-Verbal Sounds (AudioSet)**
  - Tech Recruiter: Impact of Non-Verbal Sounds: 7
  - CEO: Impact of Non-Verbal Sounds: 8
  - Overall Rating: 7.5
- 8. **Call Technical Quality (WebRTC)**
  - Tech Recruiter: Technical Quality: 9
  - CEO: Technical Quality: 9
  - Overall Rating: 9

Certainly, Dev Master CP. Below are the markdown tables for each of the criteria you've specified.

Here are some models that's been trained for each of the criteria:

- 1. Tone and Vocal Quality (Why: Emotional Impact)

Model	Description	Confidence
-------	-------------	------------

Model	Description	Confidence
DeepMoji	Uses deep learning to understand emotional content, could be adapted for voice.	7/10
OpenSMILE	Feature extraction for speech analysis, including emotional tone.	8/10
SER (Speech Emotion Recognition)	General approach for recognizing emotions in speech.	9/10

## 2. Pacing (Why: Comprehension and Comfort)

Model	Description	Confidence
Praat	Analyzes speech rate and other prosodic features.	8/10
HTK (Hidden Markov Toolkit)	Can be used for speech rate analysis.	7/10
Tempo-CNN	A CNN-based model for tempo estimation.	6/10

## 3. Use of Silence (Why: Thoughtfulness and Respect)

Model	Description	Confidence
VAD (Voice Activity Detection)	Detects voice and silence segments.	9/10
pyAudioAnalysis	Includes silence removal as a feature.	8/10
WebRTC VAD	A VAD implemented in WebRTC project.	7/10

## 4. Audibility and Clarity (Why: Basic Communication)

Model	Description	Confidence
PESQ (Perceptual Evaluation of Speech Quality)	Measures speech quality.	9/10
POLQA	Another standard for speech quality assessment.	8/10
MOSNet	Deep learning-based model for speech quality assessment.	7/10

## 5. Interruption Handling (Why: Respect and Flow)

Model	Description	Confidence
Dialogflow	Can be customized to handle interruptions.	7/10
Rasa	Open-source conversational AI, can be customized.	8/10

Model	Description	Confidence
Microsoft Bot Framework	Can be adapted for interruption handling.	6/10

## 6. Verbal Affirmations (Why: Engagement)

Model	Description	Confidence
Wit.ai	Natural Language Processing for recognizing affirmations.	8/10
LUIS (Language Understanding Intelligent Service)	Can be trained for this.	7/10
Snips NLU	On-device NLU engine.	6/10

## 7. Non-Verbal Sounds (Why: Subtext)

Model	Description	Confidence
AudioSet	A dataset labeled for sound events, can be used to train models.	8/10
ESC-50	Dataset for environmental sound classification.	7/10
SoundNet	Trained on a large amount of unlabeled video data.	6/10

## 8. Call Technical Quality (Why: Seamless Interaction)

Model	Description	Confidence
Jitter Buffer Algorithms	For handling packet loss and jitter.	8/10
WebRTC	Built-in features for call quality.	9/10
Asterisk	Open-source telephony engine, can be customized.	7/10