Accuracy Analysis:
We compared the Whisper transcription to the human-verified ground truth and calculated Word Error Rate (WER). For our audio, WER was **1.68%**, with **2 substitutions**, **0 insertions**, **0 deletions**, and **117 correct words** out of **119 reference words**. Overall accuracy was **98.32%**.

Cost Analysis:

We calculated the estimated cost of transcribing our audio file using the Whisper API. Our audio was **50.58 seconds** long (about **0.843 minutes**), so at the current pricing of **$0.006 per minute**, the transcription cost is roughly **$0.00506 USD**.

We also created a Python script that can automatically calculate the cost for any audio file based on its exact duration in seconds. Using this, we estimated costs for various scenarios:

- 30 sec → $0.003
- 60 sec → $0.006
- 600 sec (10 min) → $0.060
- 3600 sec (1 hour) → $0.360
- 7200 sec (2 hours) → $0.720

This gives a clear picture of how costs scale with audio length and allows for easy budgeting when working with longer recordings.

Performance Insights:
Whisper performed very well, even though some parts of the audio were slightly unclear. Accuracy remained high, showing that Whisper can handle short to medium recordings effectively. The speed of transcription creation is also high.

Recommendations:
Whisper is accurate and cost-effective for short audio files. For longer recordings, the cost increases linearly, so it's important to use the cost calculation script to estimate expenses. We also recommend testing prompts if the audio is tricky or noisy to maximize transcription quality.