gpt_4o_audio_preview: Pronunciation Comparison Accuracy
No Transcript - Aggregate Few-shot - Separate Test