gpt_4o_audio_preview: Pronunciation Comparison Accuracy
Asr - Aggregate Few-shot - Concatenated Test