gpt_4o_audio_preview: Pronunciation Comparison Accuracy Groundtruth - Aggregate Few-shot - Concatenated Test Number of Shots (k)

