gpt_4o_audio_preview: Comprehensive Configuration Comparison
Transcript Type vs. Configuration

Transcript Type (y-axis)
- Asr
- Groundtruth

Configuration (x-axis)
- Separate_Separate

Transcript Type	Separate_Separate
Asr	71.5%
Groundtruth	66.0%

Accuracy (colorbar): 0.6, 0.7, 0.8