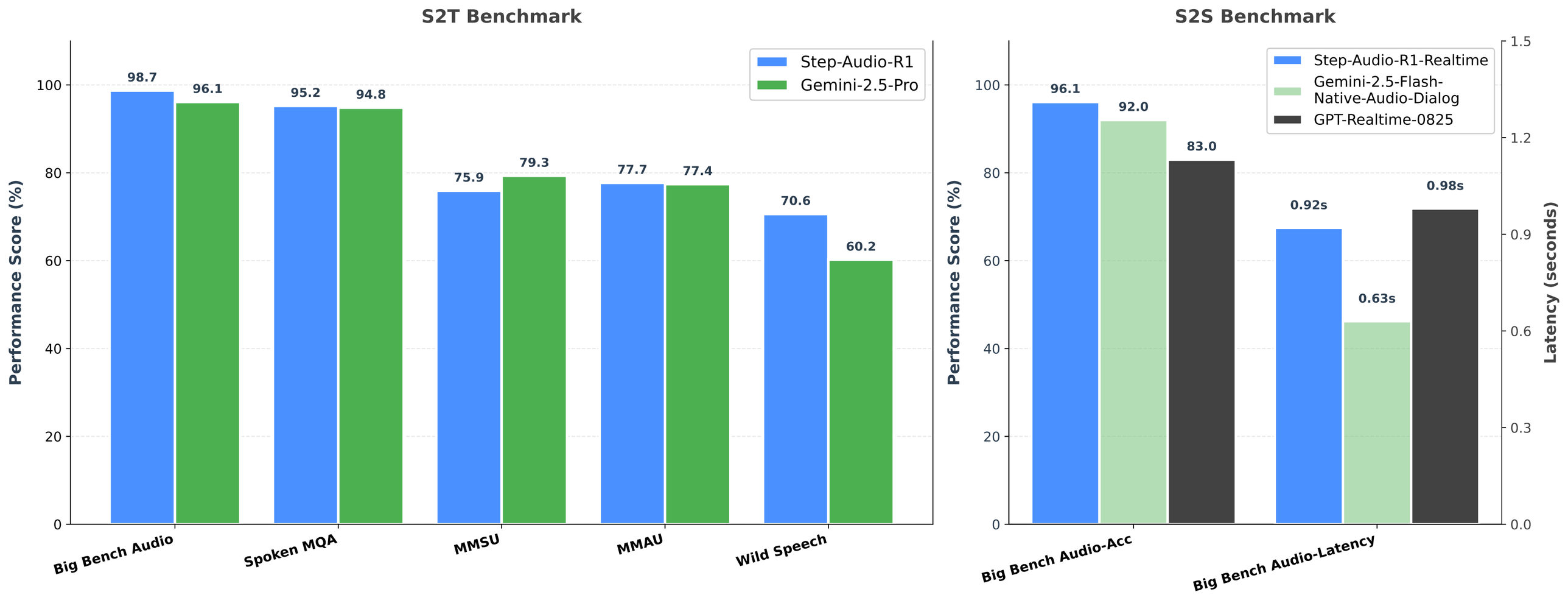
S2T Benchmark
Performance Score (%)
Step-Audio-R1
Gemini-2.5-Pro
98.7
96.1
95.2
94.8
75.9
79.3
77.7
77.4
70.6
60.2
Big Bench Audio
Spoken MQA
MMSU
MMAU
Wild Speech
S2S Benchmark
Performance Score (%)
Latency (seconds)
Step-Audio-R1-Realtime
Gemini-2.5-Flash-Native-Audio-Dialog
GPT-Realtime-0825
96.1
92.0
83.0
0.92s
0.63s
0.98s
Big Bench Audio-Acc
Big Bench Audio-Latency