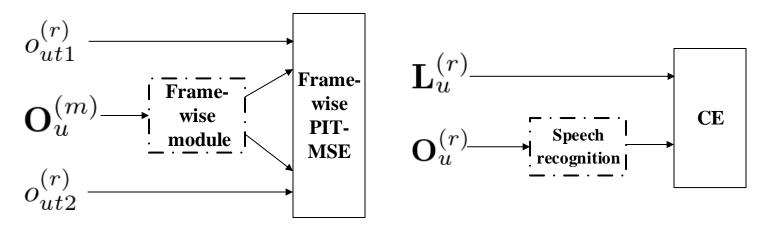
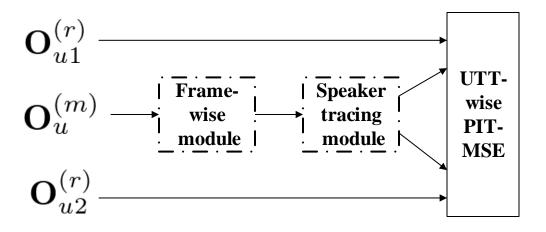


(a) PIT-ASR Joint Training

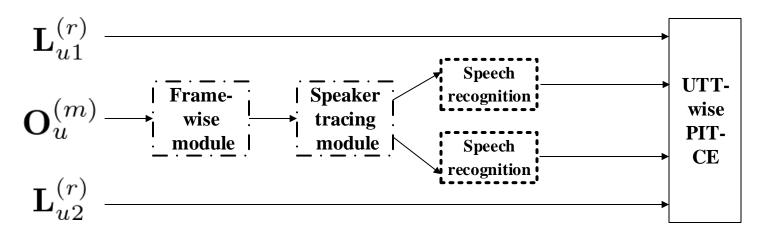


(b) Frame-wise Interpreting

(d) Speech Recognition



(c) Speaker Tracing



(e) Final Joint Training