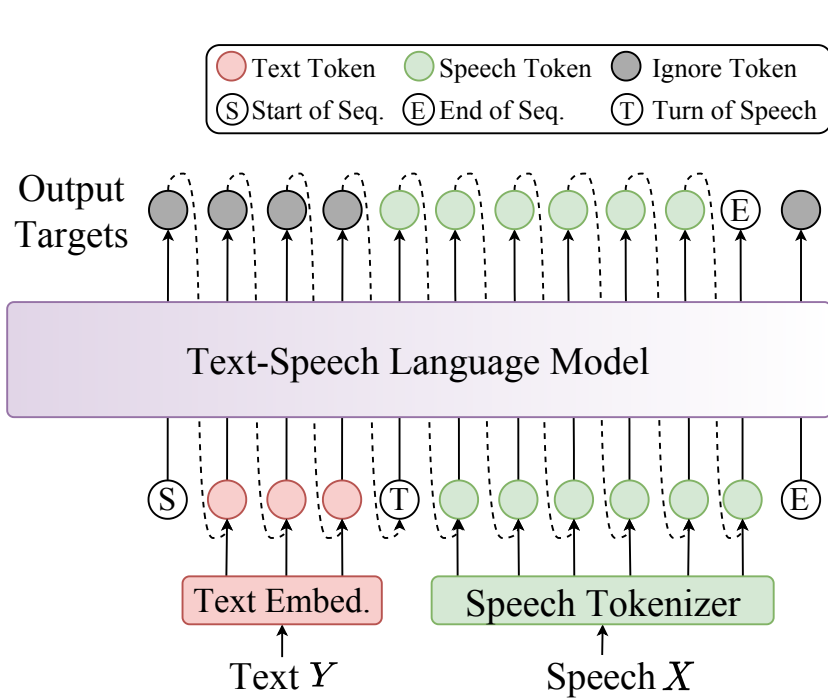
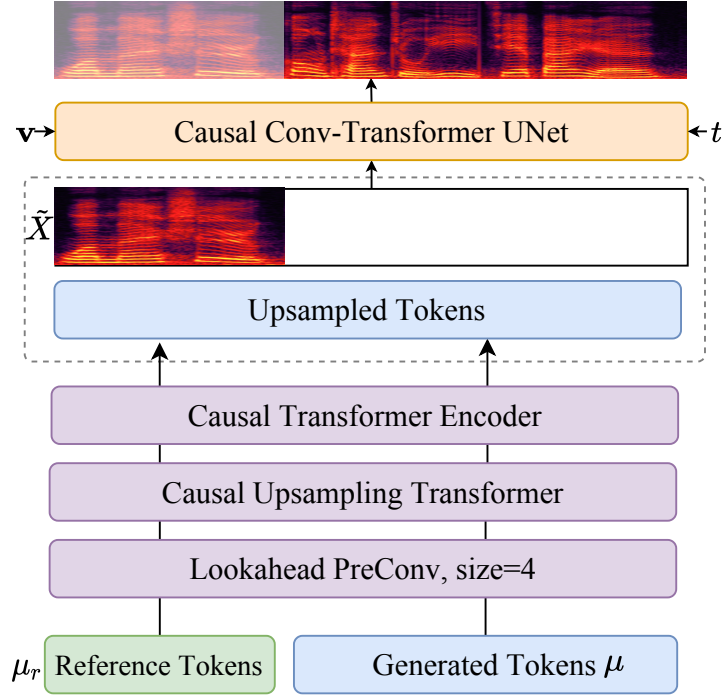


(a) Supervised semantic codec



(b) Text-Speech Language Model



(c) Causal Flow Matching