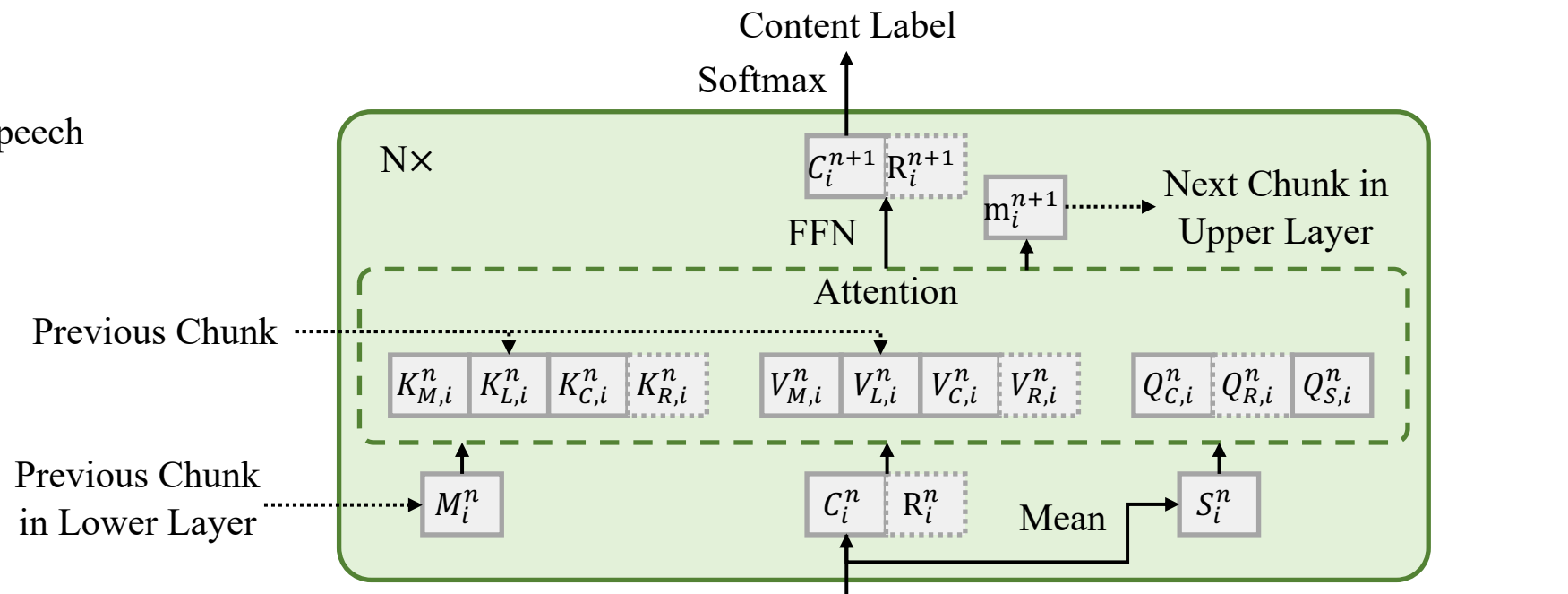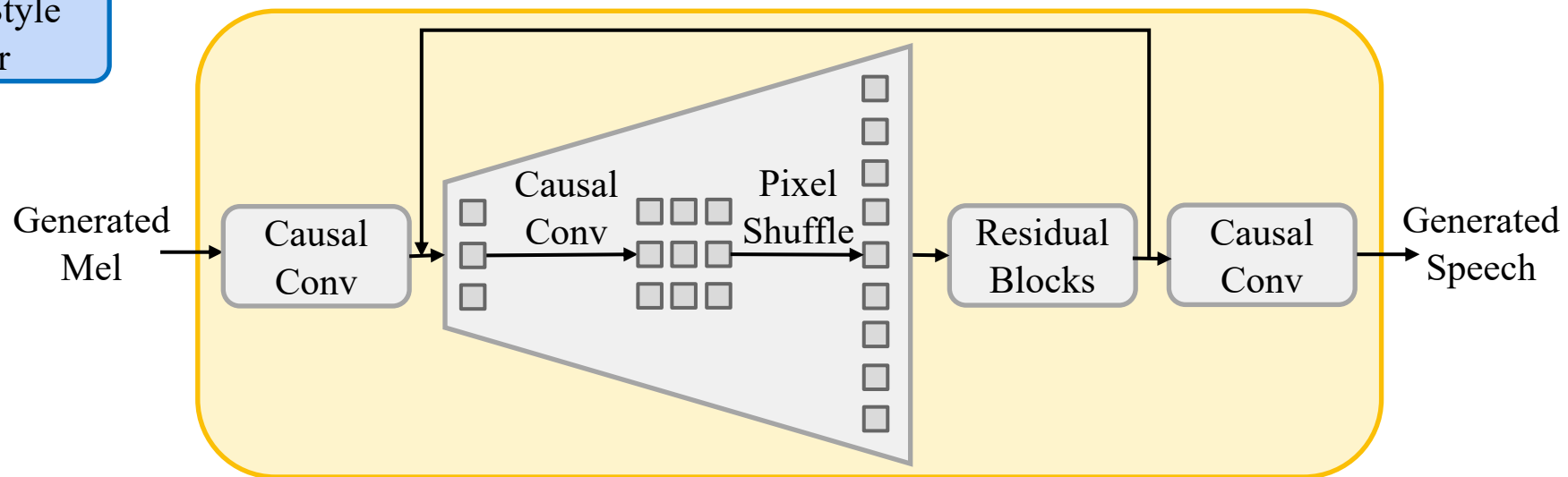(a) Overall Architecture

(b) Stream Content Extractor

(c) Causal Shuffle Vocoder