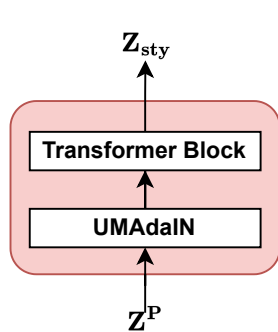
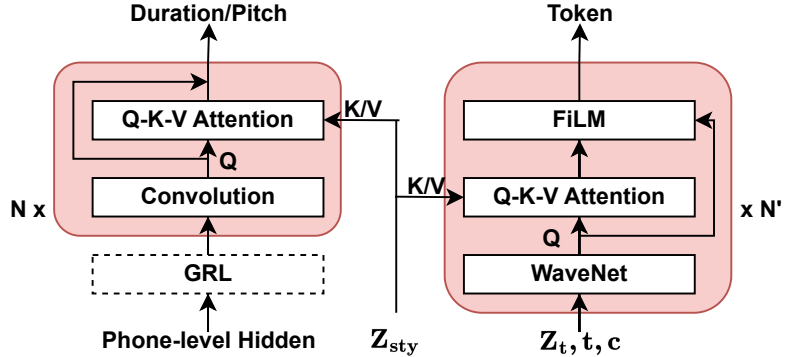


(a) Latent Encoder



(b) Prompt Encoder



(c) Variance Adaptor and SoundStorm Decoder