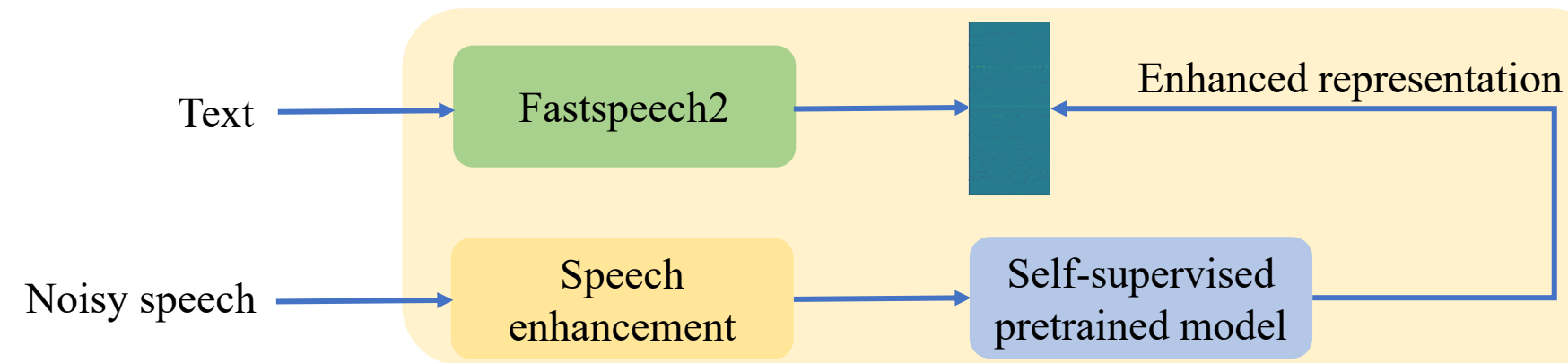
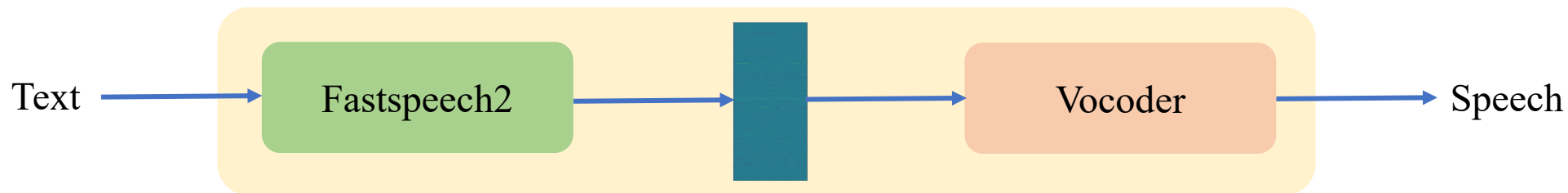


(a) Training Vocoder



(b) Training FastSpeech2



(c) Inference