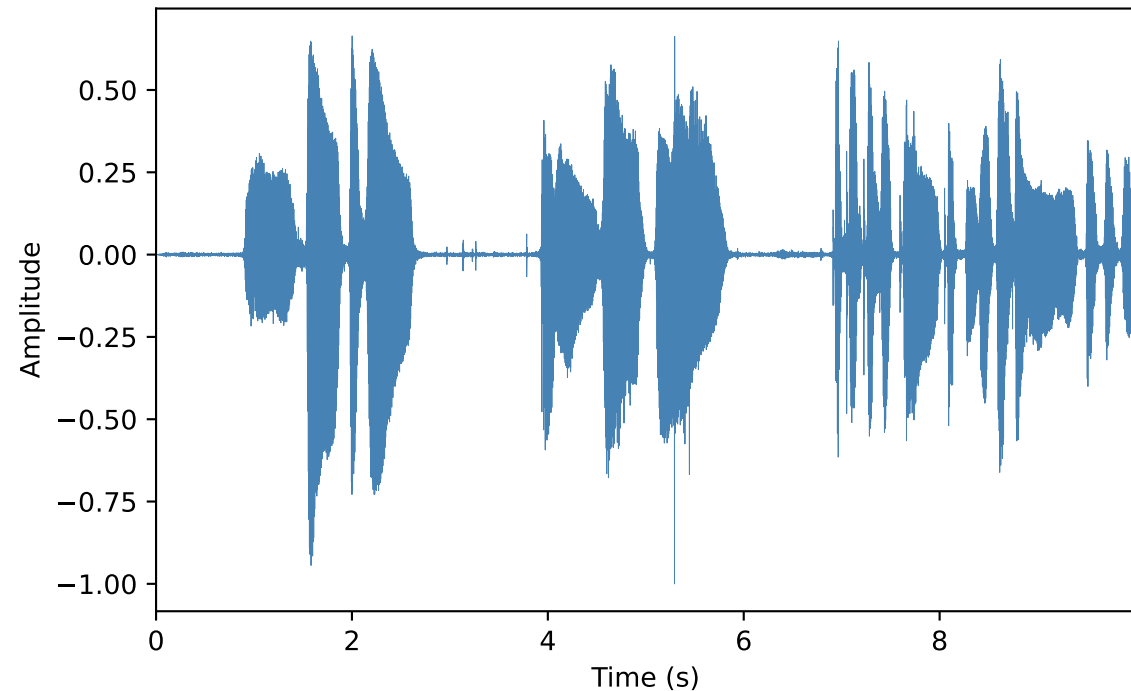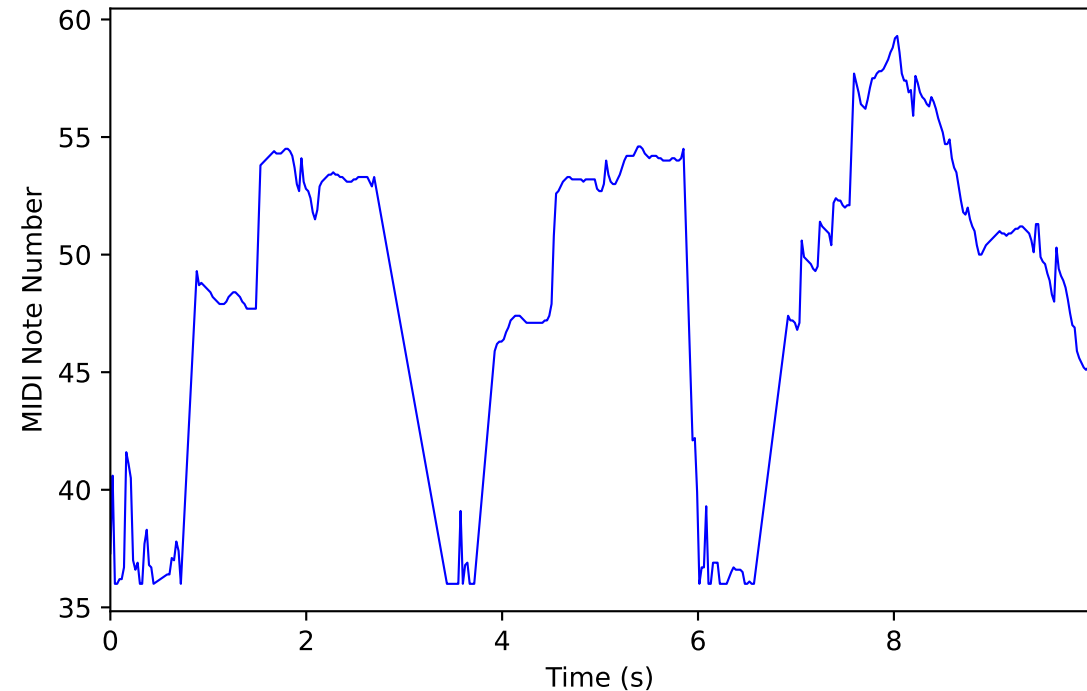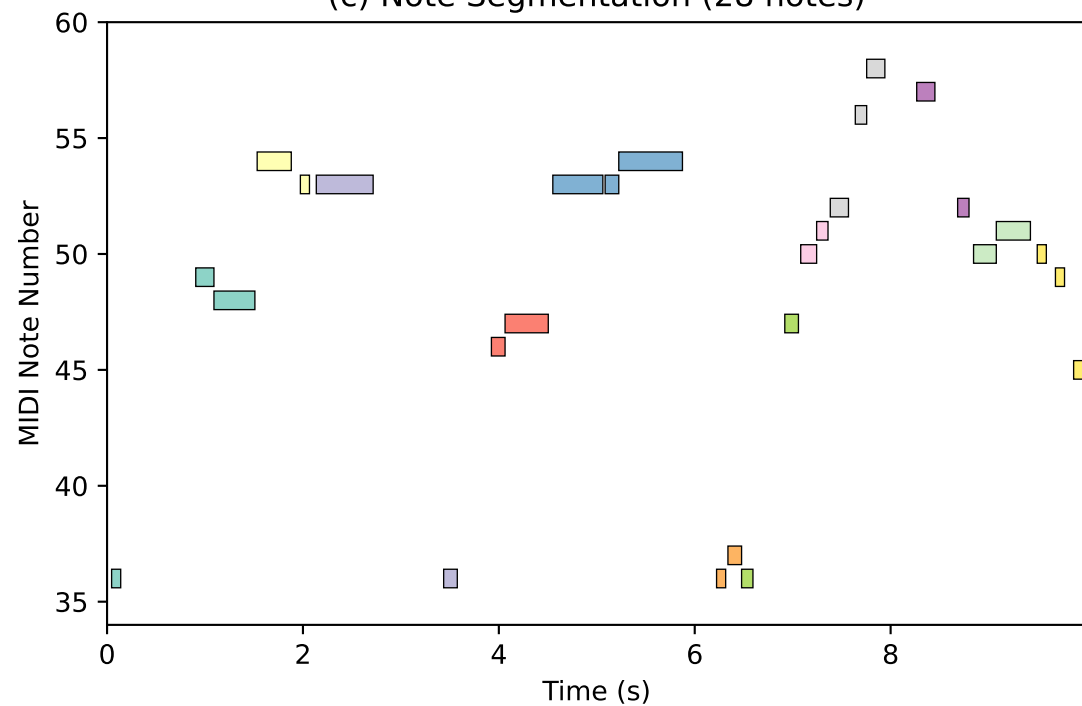(a) Preprocessed Input Waveform

(b) Extracted Pitch Contour

(c) Note Segmentation (28 notes)

(d) Generated Output Waveform (Style: lofi)