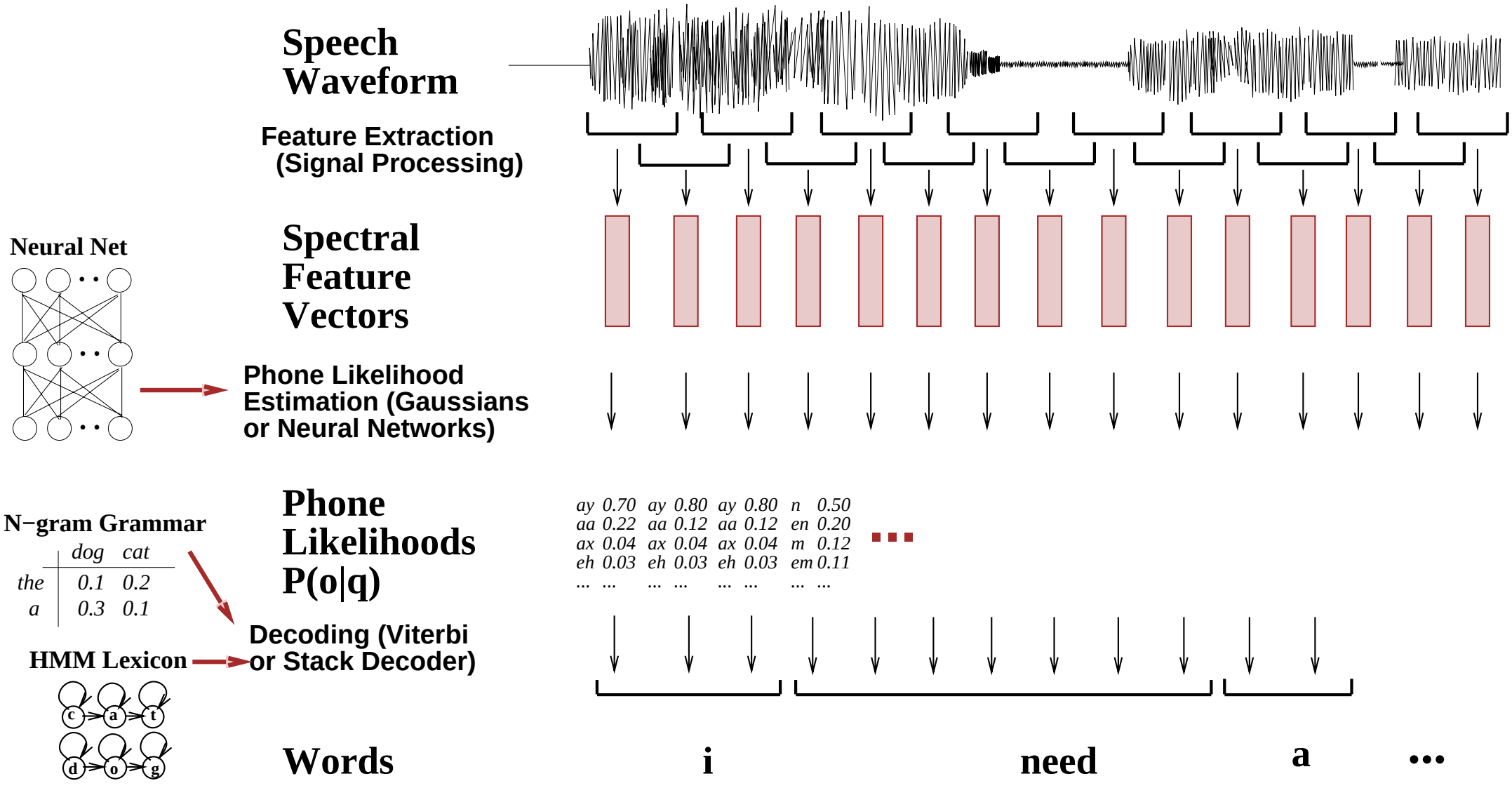
Speech Waveform
Feature Extraction (Signal Processing)
Spectral Feature Vectors
Neural Net
Phone Likelihood Estimation (Gaussians or Neural Networks)
Phone Likelihoods P(o|q)
N−gram Grammar
dog cat
the 0.1 0.2
a 0.3 0.1
HMM Lexicon
Decoding (Viterbi or Stack Decoder)
ay 0.70 ay 0.80 ay 0.80 n 0.50
aa 0.22 aa 0.12 aa 0.12 en 0.20
ax 0.04 ax 0.04 ax 0.04 m 0.12
eh 0.03 eh 0.03 eh 0.03 em 0.11
Words
i
need
a
•••