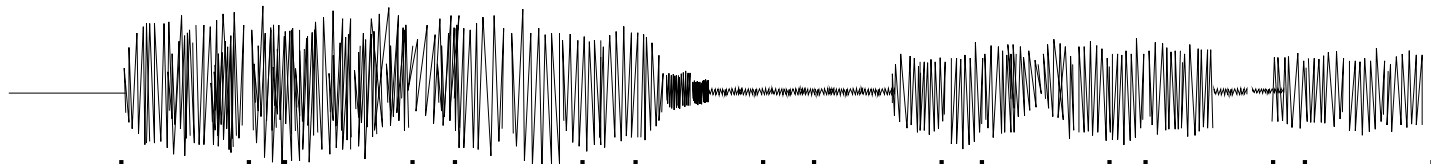


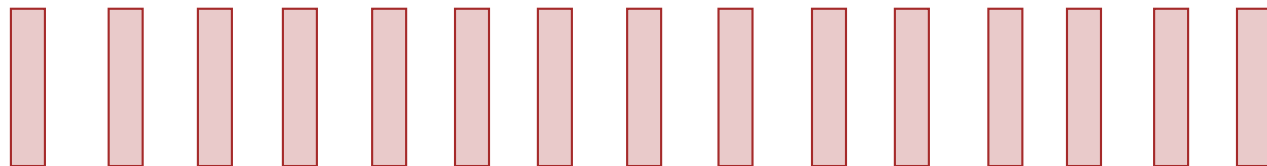
Speech Waveform



Feature Extraction
(Signal Processing)

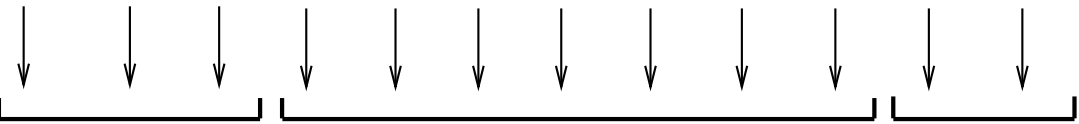


Spectral Feature Vectors



<i>ay</i>	0.70	<i>ay</i>	0.80	<i>ay</i>	0.80	<i>n</i>	0.50	
<i>aa</i>	0.22	<i>aa</i>	0.12	<i>aa</i>	0.12	<i>en</i>	0.20	■ ■ ■
<i>ax</i>	0.04	<i>ax</i>	0.04	<i>ax</i>	0.04	<i>m</i>	0.12	
<i>eh</i>	0.03	<i>eh</i>	0.03	<i>eh</i>	0.03	<i>em</i>	0.11	
...	

Phone Likelihoods P(olq)



i

need

a

...

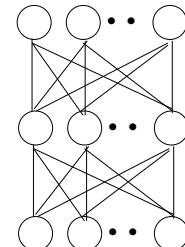
Phone Likelihood Estimation (Gaussians or Neural Networks)

Phone Likelihoods P(olq)

Decoding (Viterbi or Stack Decoder)

Words

Neural Net



N-gram Grammar

	<i>dog</i>	<i>cat</i>
<i>the</i>	0.1	0.2
<i>a</i>	0.3	0.1

HMM Lexicon

