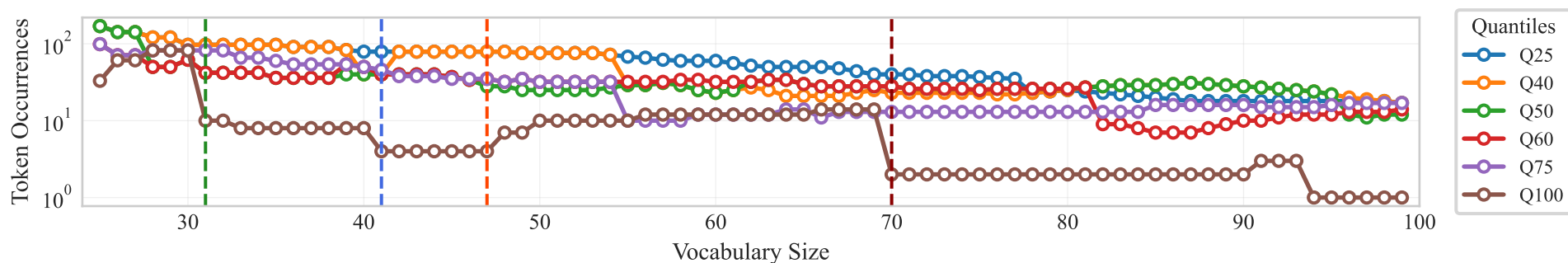
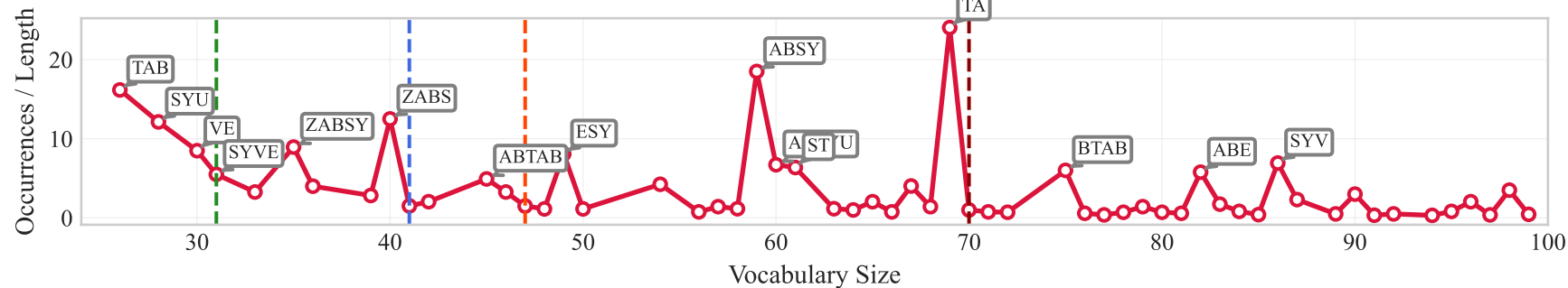


Tokenization Analysis: BPE

Token Occurrences by Length Quantile



Token Frequency Normalized by Length



Shannon Entropy of Token Distribution

