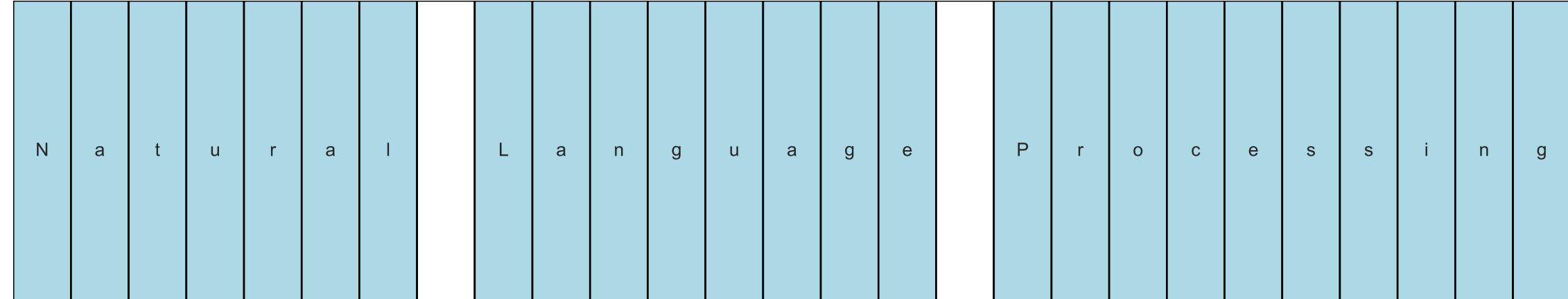
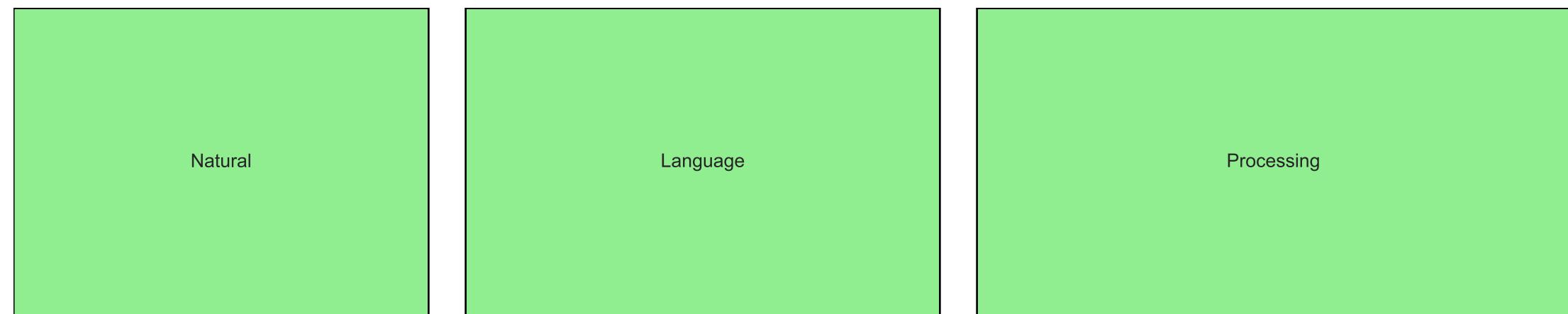


Character-level Tokenization



Word-level Tokenization



Subword (BPE) Tokenization

