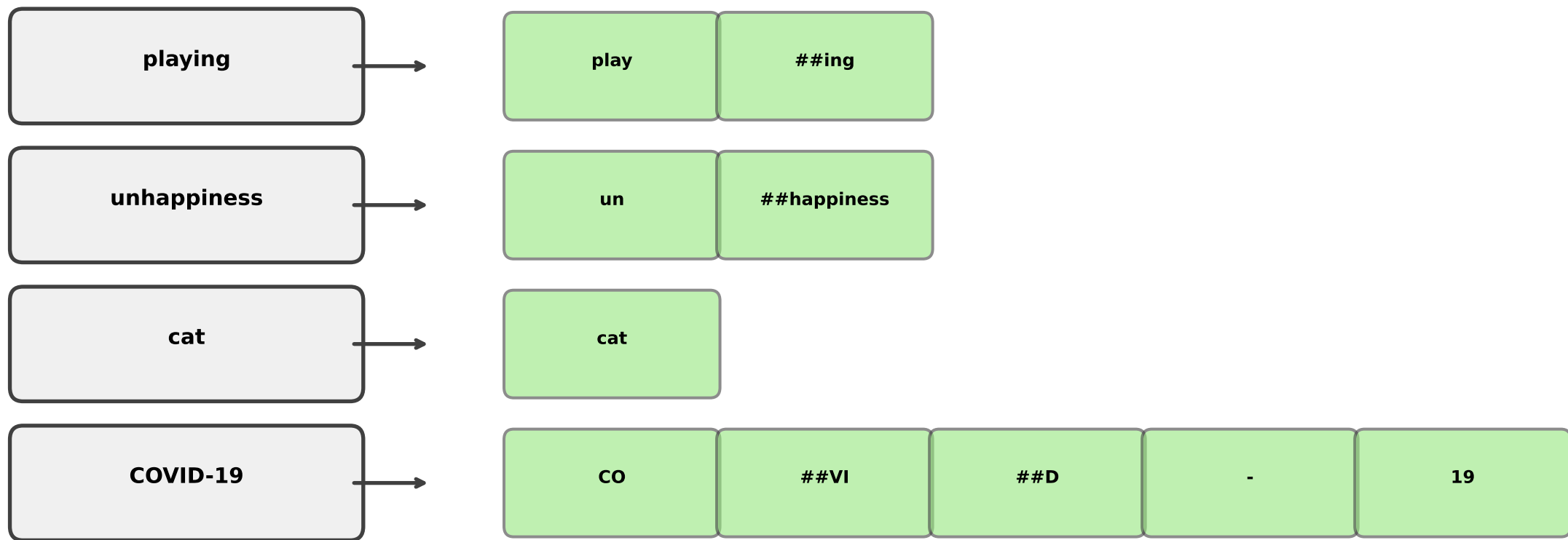


## WordPiece Tokenization: Handling Rare Words



*## prefix indicates continuation of previous subword*