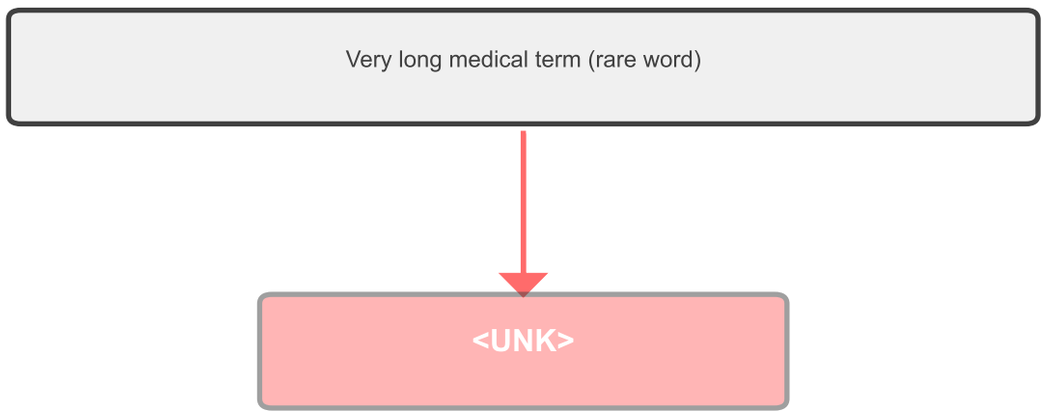
# Rare Word Handling: Word-Level vs Subword

## Word-Level: Information Loss

**Input: "pneumonoultramicroscopicsilicovolcanoconiosis"**

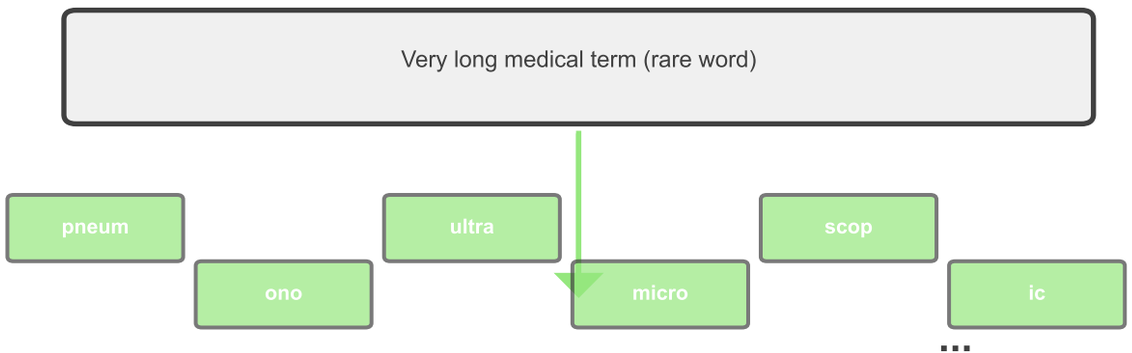



**PROBLEM: All information lost!**

Model has no idea what
the original word meant

## Subword (BPE): Meaning Preserved

**Input: "pneumonoultramicroscopicsilicovolcanoconiosis"**




**SUCCESS: Meaning preserved!**

Model can understand:
"pneum" (lung), "micro" (small),
"scop" (viewing), etc.