BPE Merge Operation: Step-by-Step Example

Step 1: Count all adjacent pairs

low

lower

newest

Step 2: Most frequent pair

'e s': 2

'l o': 2

'o w': 2
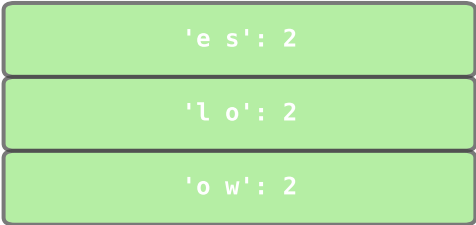
Step 3: Merge chosen pair (es)

e s	→	es

Step 4: Update corpus with new token

low

lower

new **es** t

Key Insight:

BPE learns subword units
from frequency patterns

No linguistic knowledge needed!