Caption input:
A *long vase with* **red roses** *on a nightstand*

Equally condition on caption and masked RGB