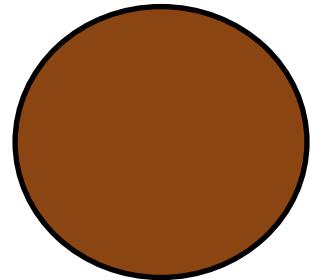


Figure 10: Visual Grounding Examples
(28,489 words grounded, 9.8% coverage)

"dog"

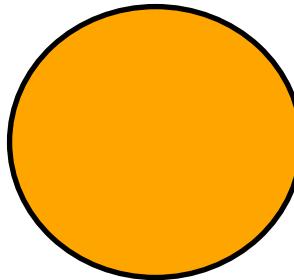
✓



*Direct grounding
from COCO*

"cat"

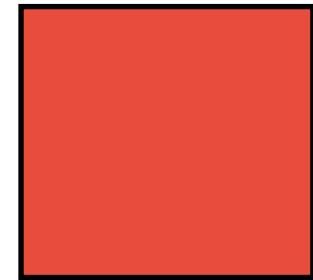
✓



*Direct grounding
from COCO*

"red"

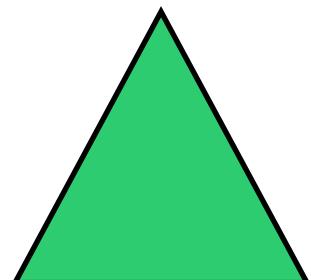
✓



*Color attribute
grounded*

"running"

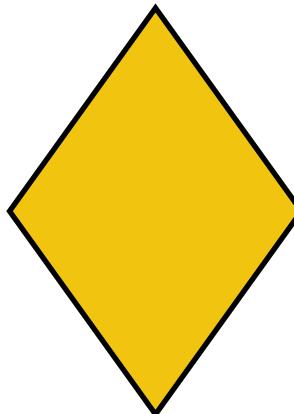
✓



*Action propagated
from motion*

"happy"

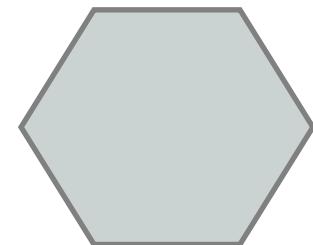
✓



*Abstract: 3-hop
from "smile"*

"justice"

✗



*Ungrounded
(no visual ref)*