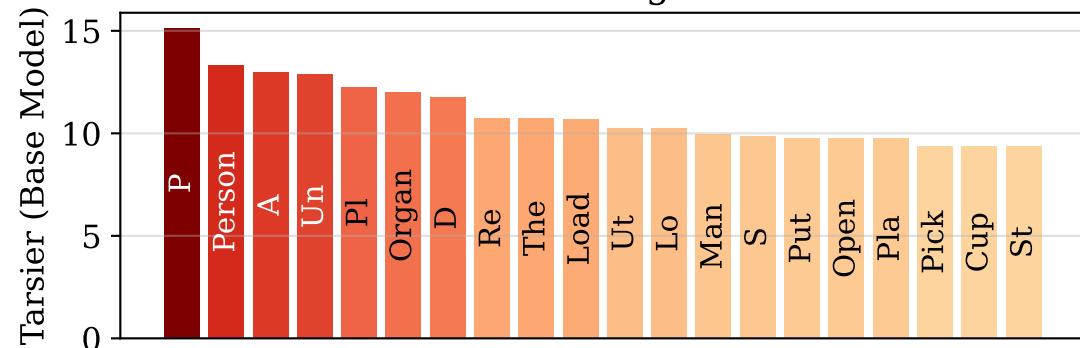


Input video with caption: put-down plate



Video logits



Text logits

