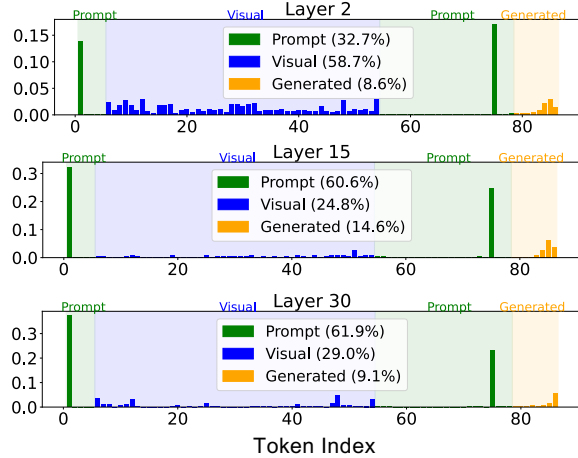
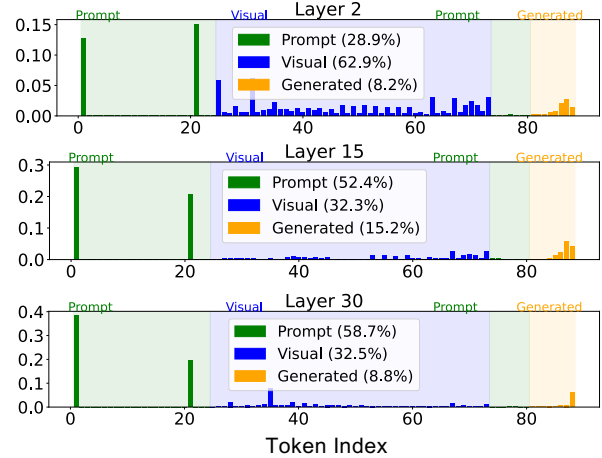


(a) Baseline (Frozen LLM)



(b) Remove tokens of "BOS" and "."



(c) Reposition visual tokens