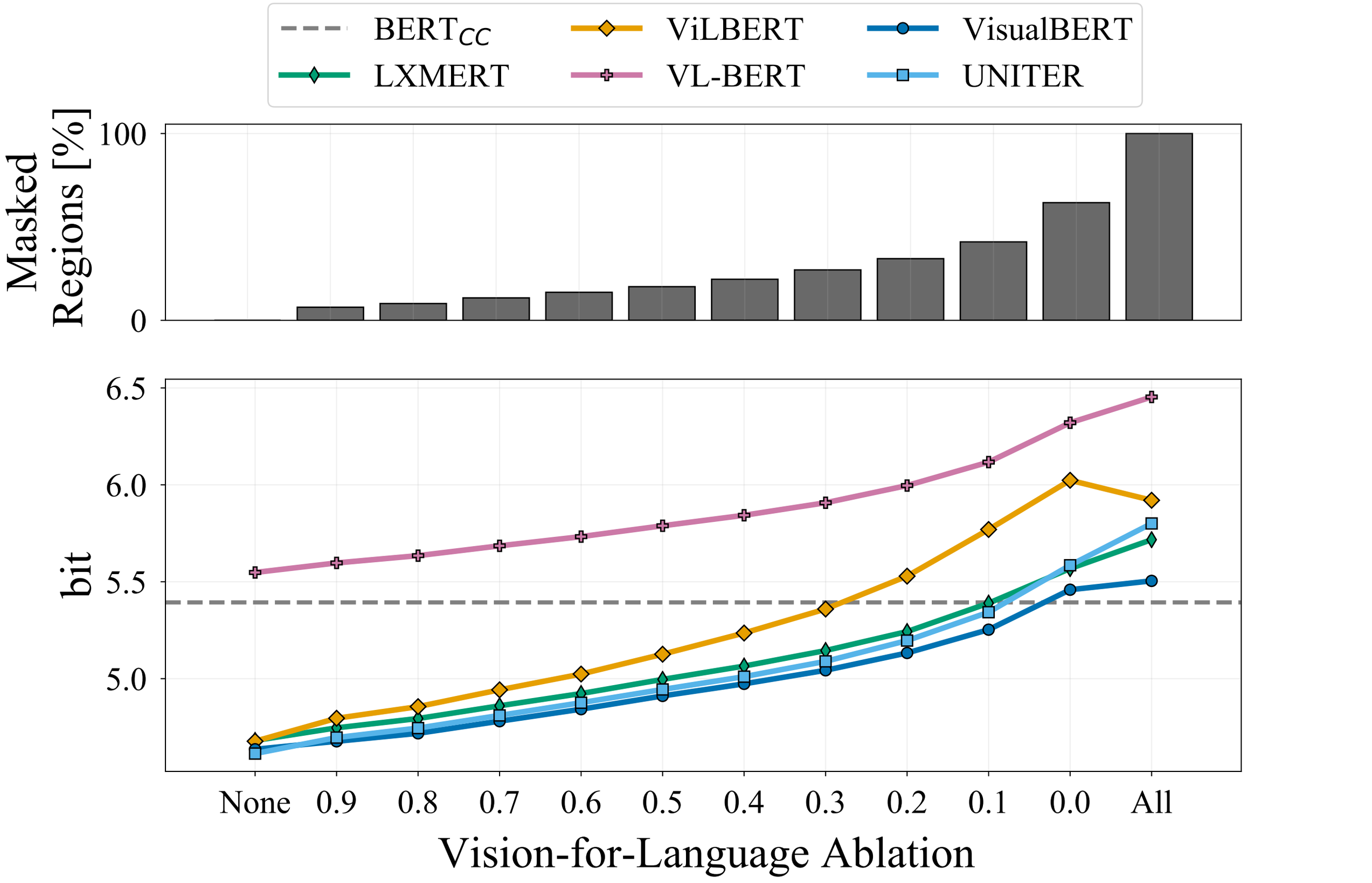
BERT$_{CC}$
ViLBERT
VisualBERT
LXMERT
VL-BERT
UNITER
Masked Regions [%]
100
0
6.5
6.0
5.5
5.0
bit
None
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
All
Vision-for-Language Ablation