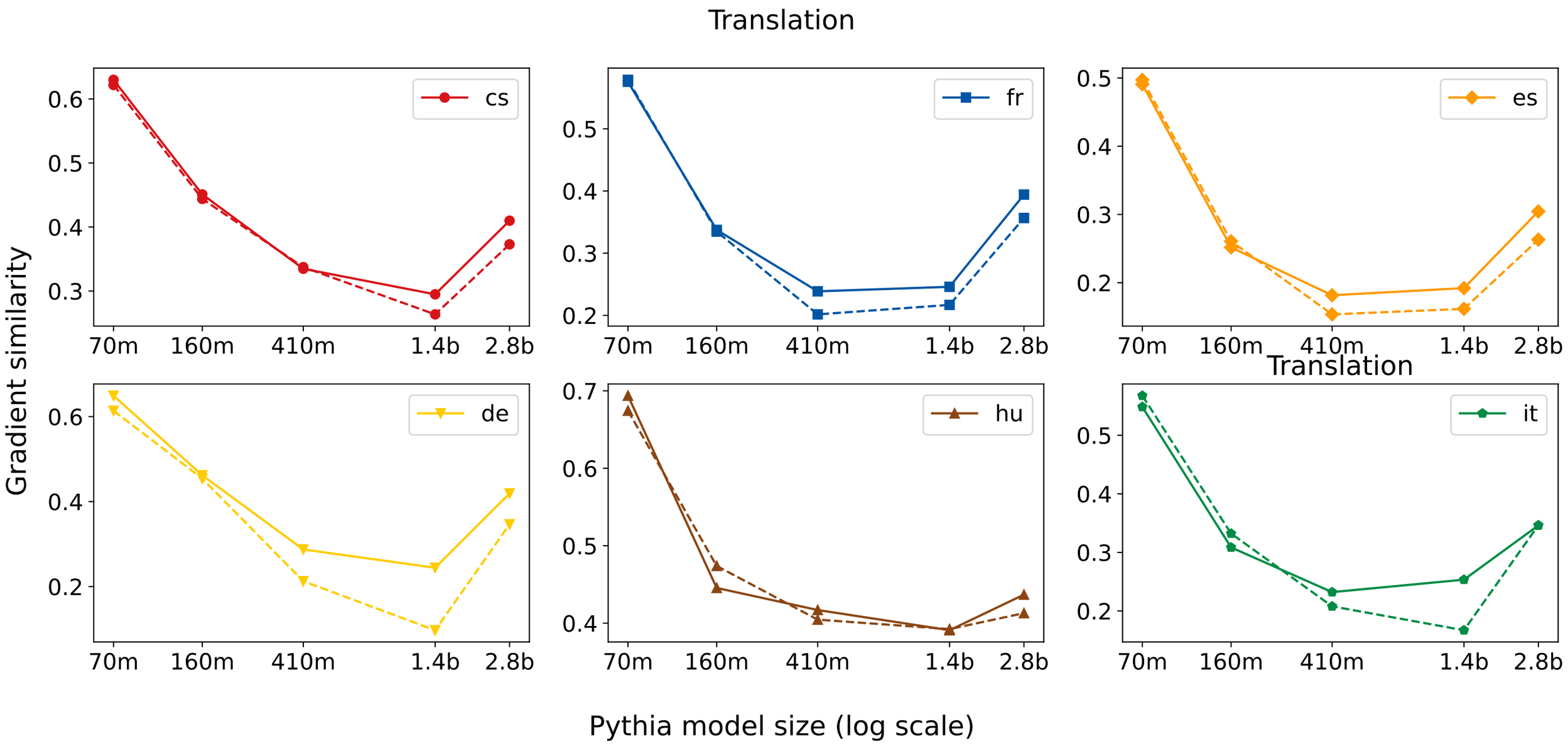
Translation
cs
fr
es
de
hu
it
Translation
Gradient similarity
Pythia model size (log scale)
70m
160m
410m
1.4b
2.8b