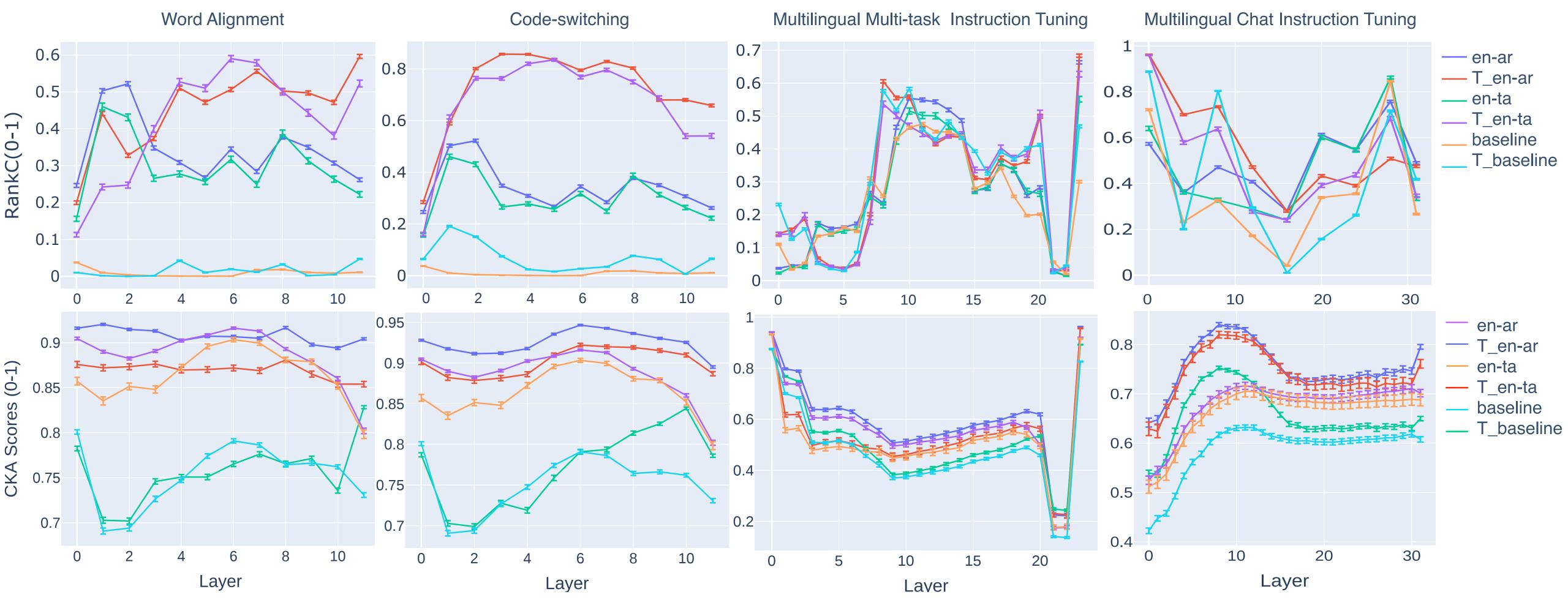
Findings

Cross-lingual supervision can alleviate the consistency bottleneck to enhance alignments between coreferential entities, which can be achieved by training with an explicit alignment objective or a code-switching objective. On the other hand, parallel samples providing cross-lingual generalization supervision offer limited gains to consistency.



Method: We examine tuned models that show promising performance in multilinguality.