XNLI on Gemini with Word Substitution Attack - Accuracy per Language 1.0 8.0 Accuracy 0.47 ± 0.08 0.46 ± 0.09 0.46 ± 0.08 0.46 ± 0.08 0.44 ± 0.08 0.43 ± 0.08 0.4 0.2

