LRE faithfulness in GPT2-xl occupation genderadjective comparative verb past tense adjective superlative country largest city name birthplace country language country capital city substance phase of matter object superclass name religion name gender-fruit outside colorcharacteristic genderplays pro sportadjective antonymoccupation age person native language fruit inside color person university person lead singer of band word sentiment person plays instrument-word first lettercountry currency word last letter landmark on continent work location president birth year president election year food from country person occupation task person typecity in country superhero personperson sport positions star constellation name landmark in countryproduct by companysuperhero archnemesis task done by toolcompany hq-pokemon evolution person mother person father company CEO 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 0.0 Faithfulness