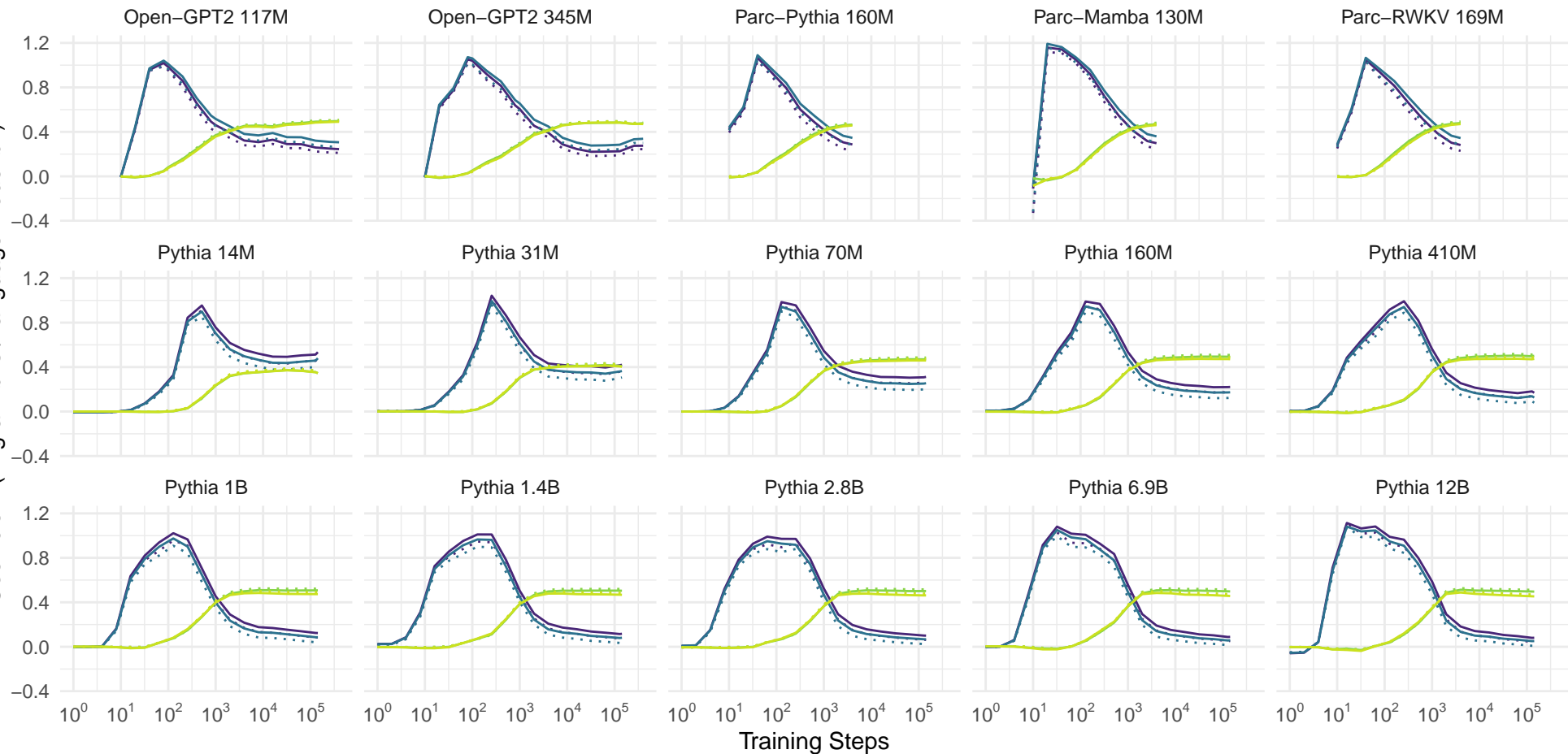


Coefficient (n-gram bits / language model bit)



Similarity Metric

- Semantic Similarity (Wiki Unweighted)
- Semantic Similarity (CC Unweighted)

Heuristic

- 1-gram Logprobs (Matched)
- 1-gram Logprobs (Unmatched)
- 5-gram Logprobs (Matched)
- 5-gram Logprobs (Unmatched)