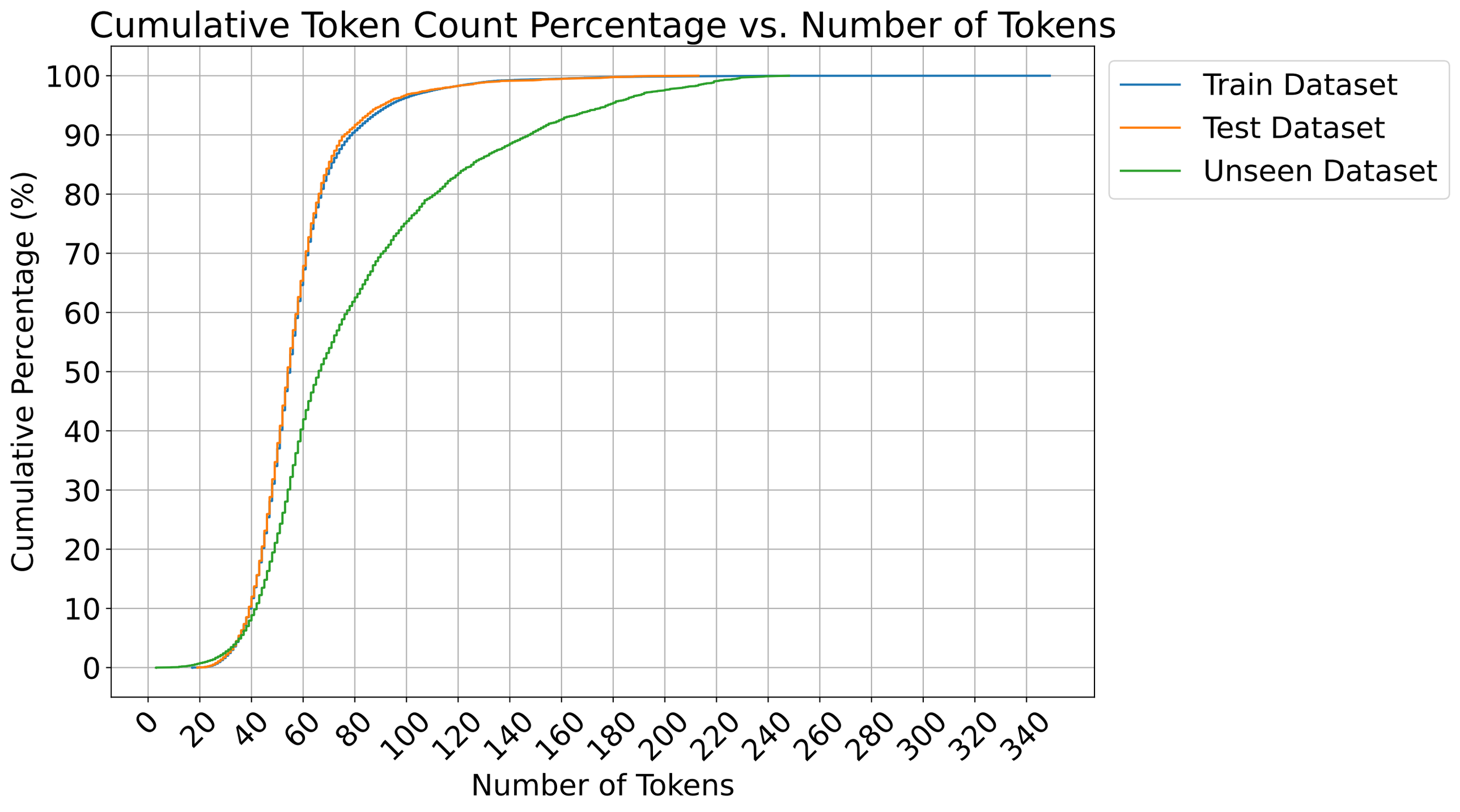
Cumulative Token Count Percentage vs. Number of Tokens
Cumulative Percentage (%)
Number of Tokens
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
Train Dataset
Test Dataset
Unseen Dataset