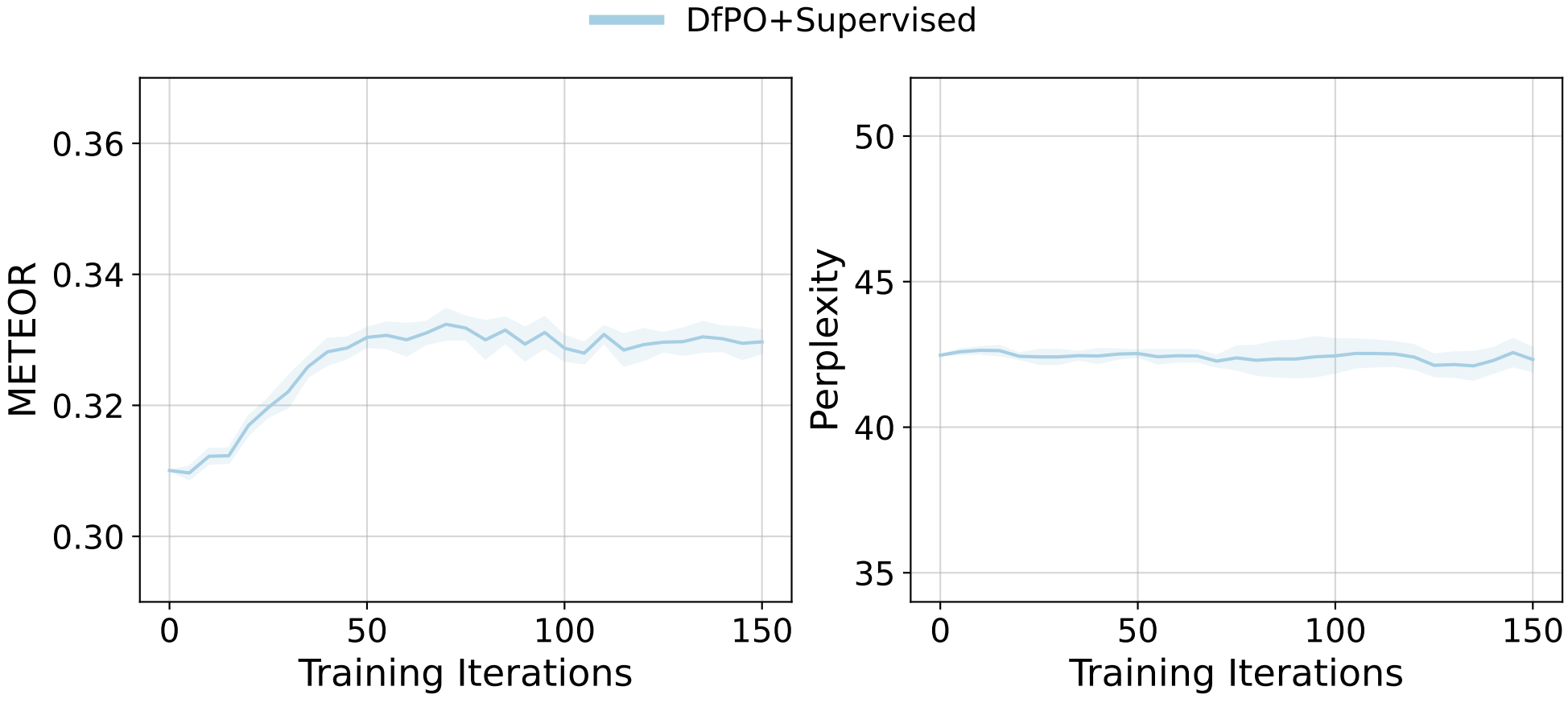

DfPO+Supervised
METEOR
0.36
0.34
0.32
0.30
0
50
100
150
Training Iterations
Perplexity
50
45
40
35
0
50
100
150
Training Iterations