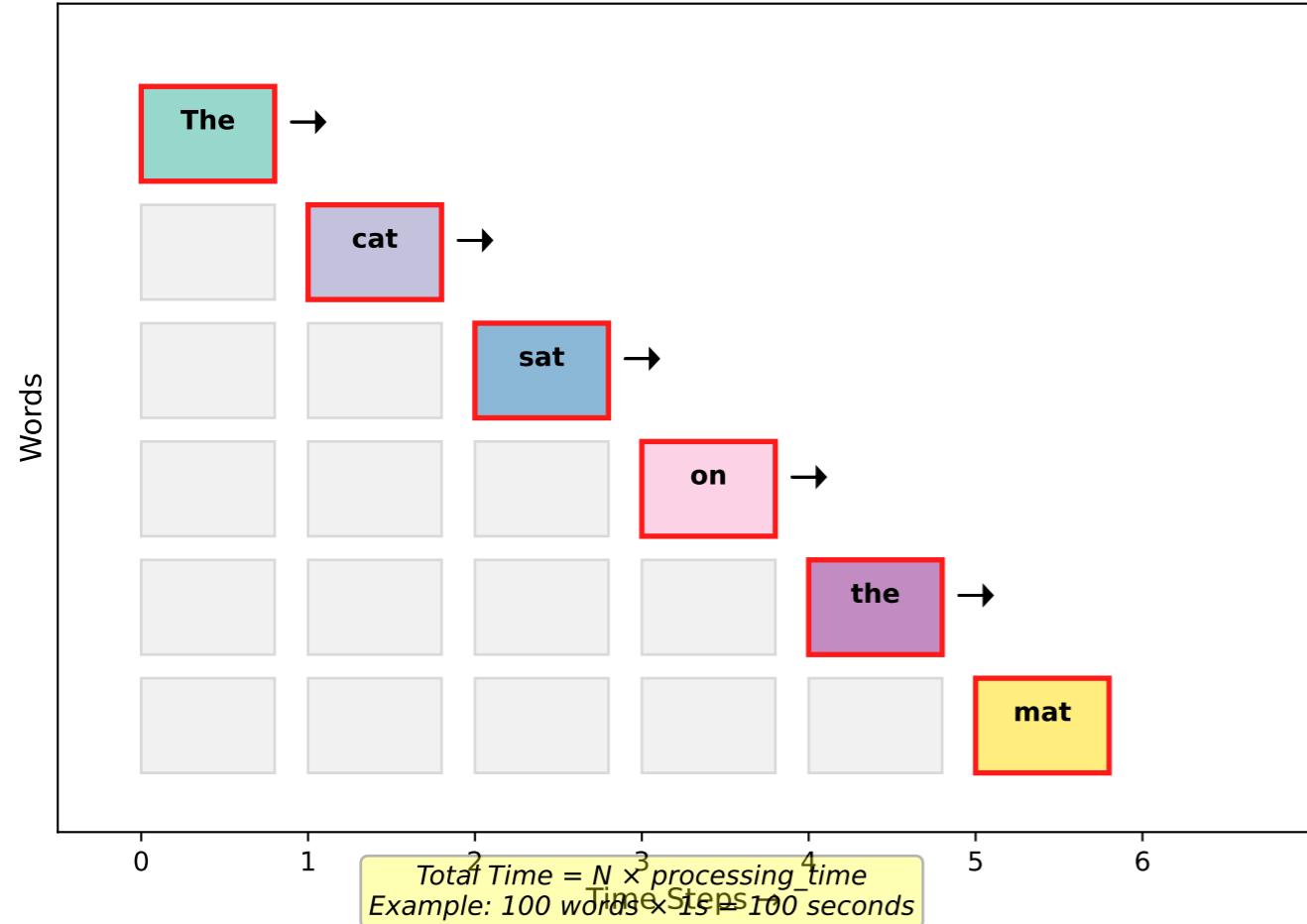
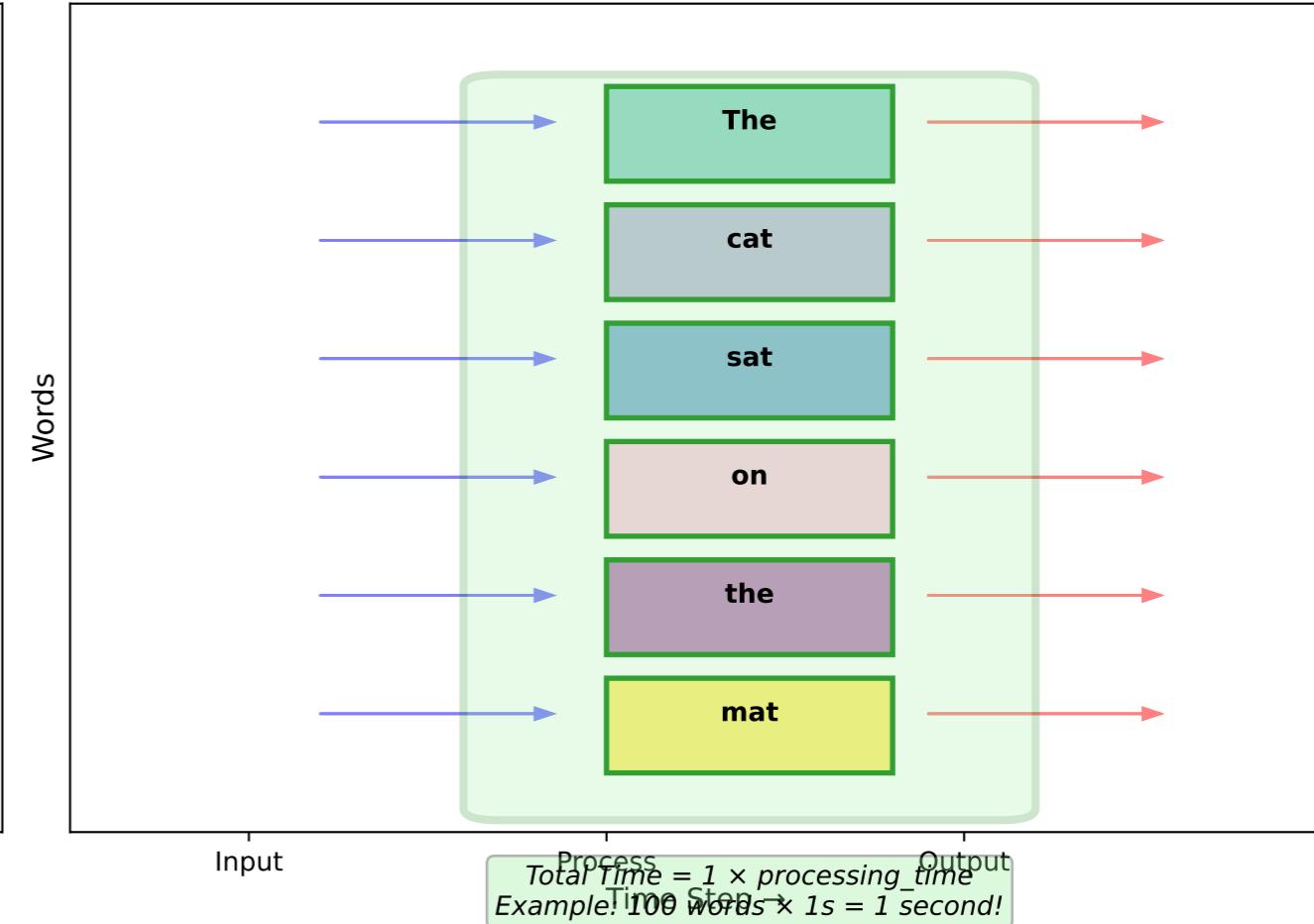


Sequential Processing (RNN) 100 words = 100 time steps



Parallel Processing (Transformer) 100 words = 1 time step!



90x Speedup through Parallelization!