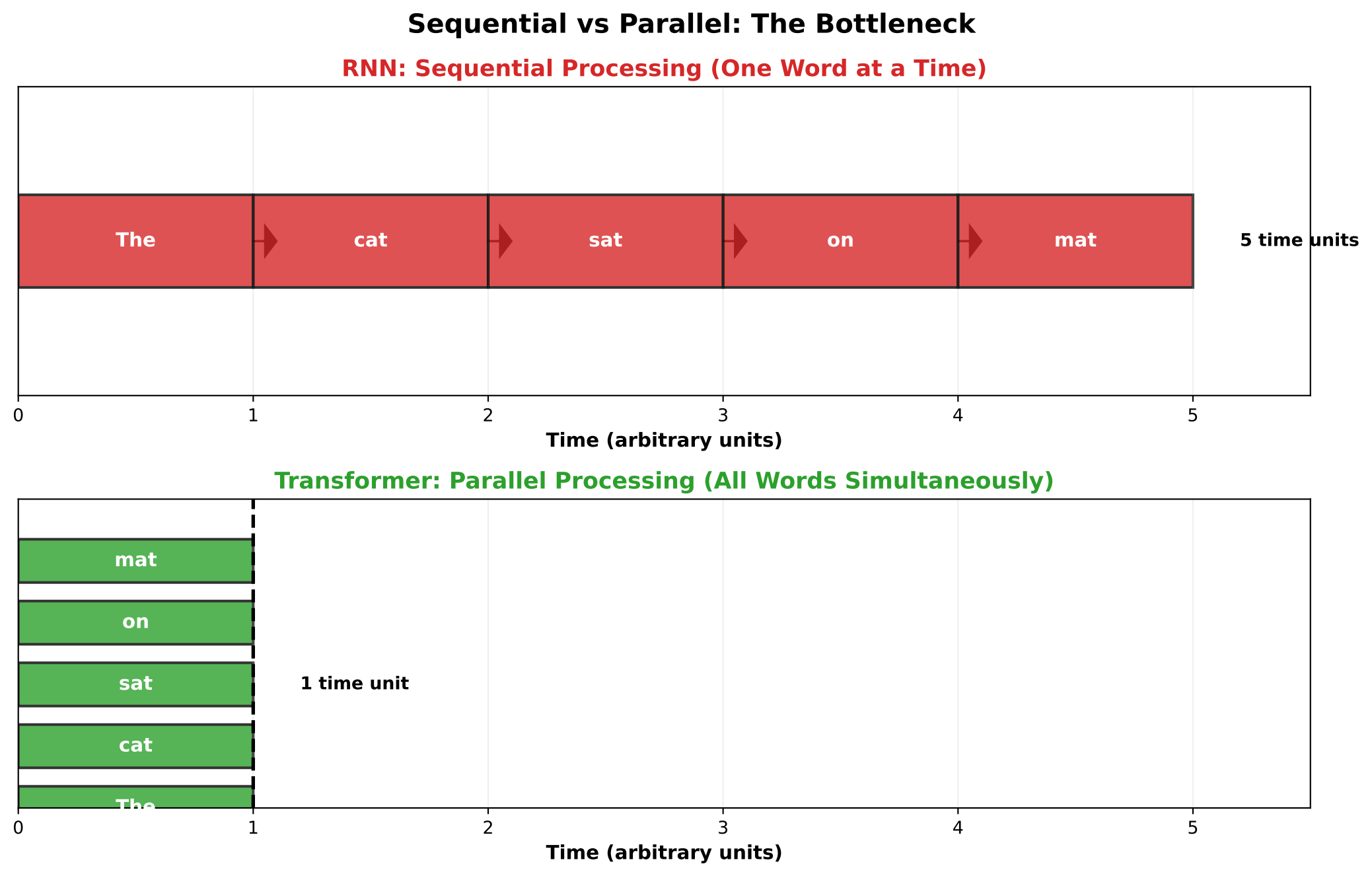
Sequential vs Parallel: The Bottleneck
RNN: Sequential Processing (One Word at a Time)
The
cat
sat
on
mat
5 time units
0
1
2
3
4
5
Time (arbitrary units)
Transformer: Parallel Processing (All Words Simultaneously)
mat
on
sat
cat
The
1 time unit
0
1
2
3
4
5
Time (arbitrary units)