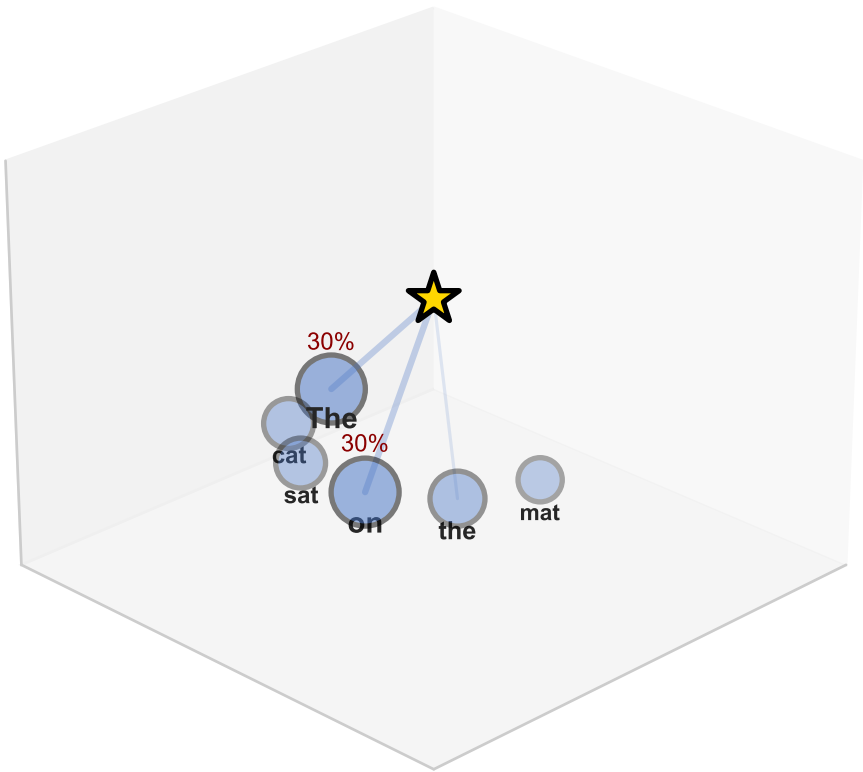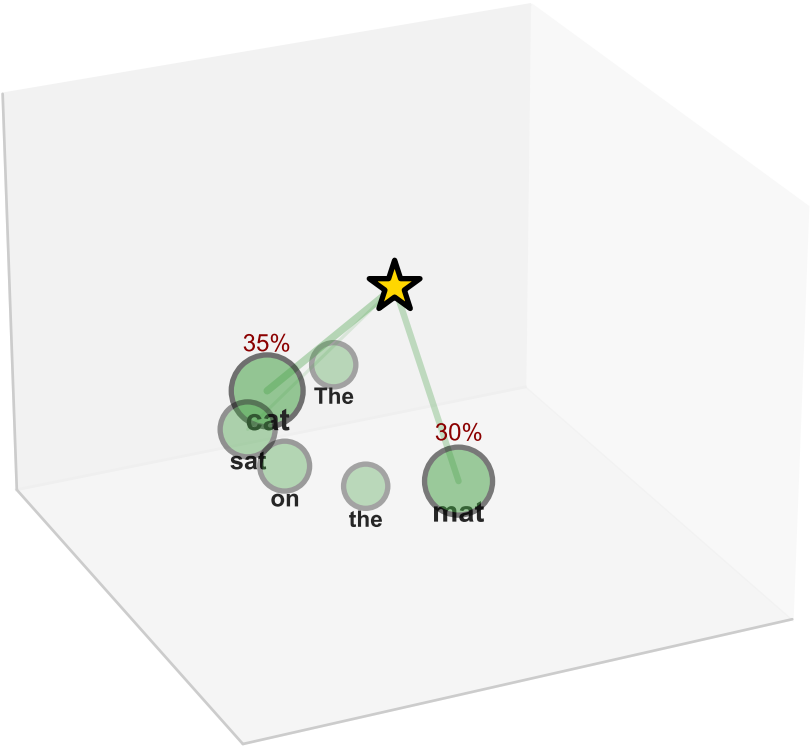# Multi-Head Attention: Four Different Perspectives on Same Sentence
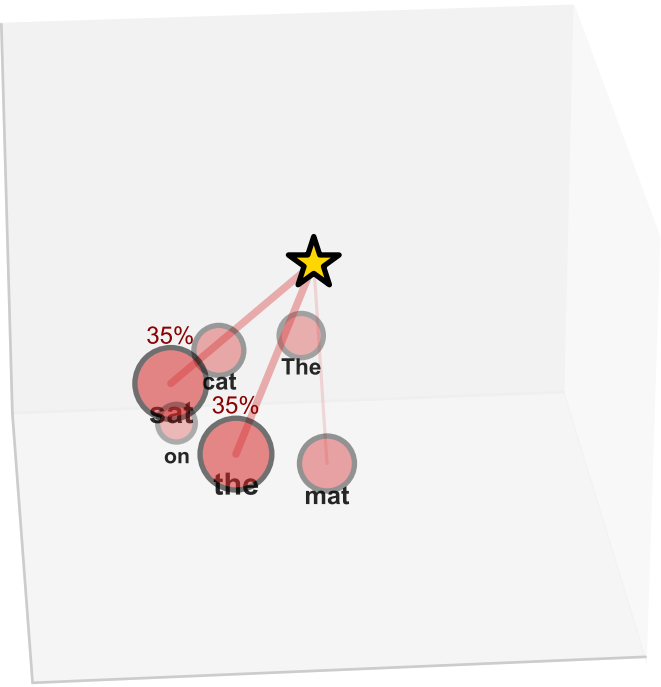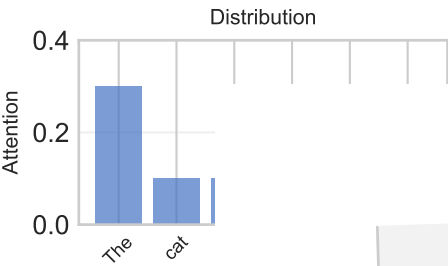
Grammar Head
Focuses on articles and prepositions

30%
30%
The
cat
sat
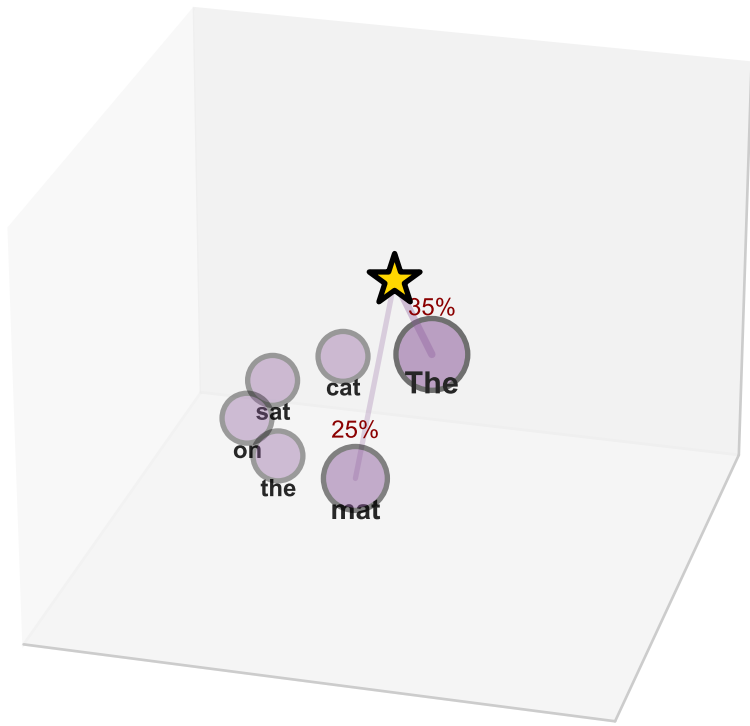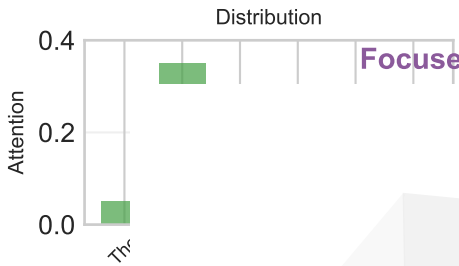on
the
mat

Semantic Head
Focuses on meaning relationships

35%
The
cat
sat
on
the
30%
mat

Distribution

Attention
0.4
0.2
0.0
The   cat

Position Head
Focuses on nearby words

35%
cat
The
sat
35%
on
the
mat

Distribution

Attention
0.4
0.2
0.0
Th

Global Head
Focuses on sentence boundaries

35%
cat
The
sat
on
the
25%
mat

Distribution

Attention
0.4
0.2
0.0
The   cat   sat   on   the   mat

Each head learns to focus on different linguistic patterns - combined they capture all aspects!

Distribution

Attention
0.4
0.2
0.0
The   cat   sat   on   the   mat