

Attention to "penguins" before Divergence vs Non-Divergence Tokens
(N=100 samples)

