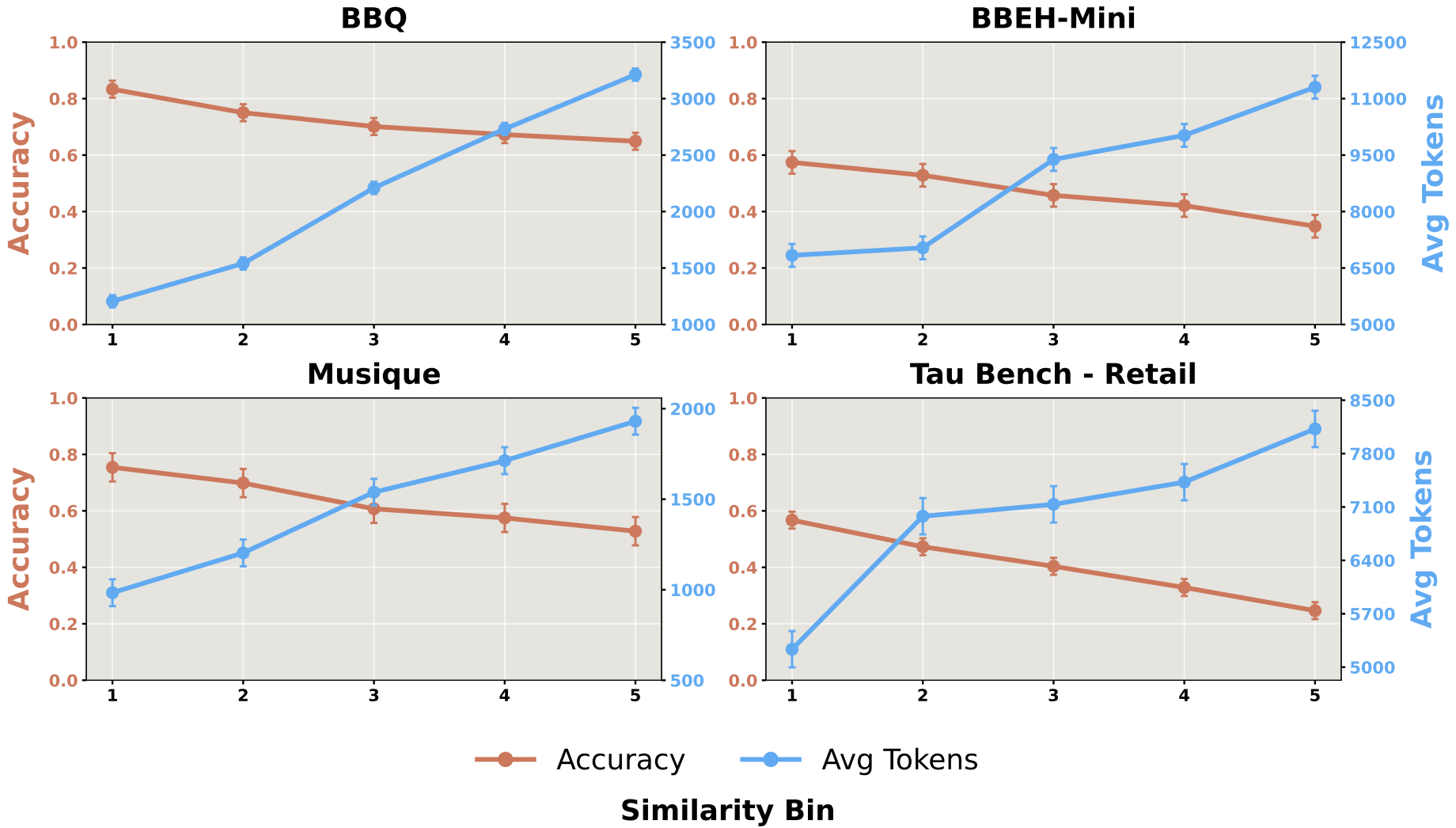

BBQ
BBEH-Mini
Musique
Tau Bench - Retail
Accuracy
Avg Tokens
Accuracy
Avg Tokens
Similarity Bin