



Image Model predictions:

bowl, broccoli, carrot, dining table

Language Model predictions for the last word:

fork, spoon, bowl

Generated caption: A plate of food with broccoli and a *fork*.