Token Similarity vs. Prediction Probability