

Heatmap visualization of the model's output probabilities for the German sentence "Ich möchte Ihnen, Herrn Kommissar Patten, für diese klare Antwort danken." The x-axis represents the input tokens (including <START>, <END>, and a blank token) and the y-axis represents the output tokens. The color scale ranges from 0.0 (dark purple) to 1.0 (yellow). The highest probability (1.0) is for "patten" given "kommissar". Other high probabilities (around 0.6-0.8) are seen for "ich" given <START>, "möchte" given "ich", and "antwort" given "klare".

