Visualization of deep neural network models with Transformer architecture is generally a very demanding task which is usually solved by visualizing attention blocks and monitoring which words these block focus on. However, Transformer models have many layers and there are multiple attention heads on each layer. Therefore, each head may attend to different linguistic features. In this work, we focus on developing an application that is designed to visualize the behaviour of GPT-2 language models more clearly. We propose four visualization methods that examine the dependencies of generated words on previous words in the text. We monitor these dependencies by removing one of the words in the previously generated text or replacing it with a similar word and then observing changes of the probability of the generated word. We show the results of our methods produced on the GPT-2 Medium model and formulate hypotheses with the aim to explain them.