

Transformer test loss with Gaussian init vs. mimetic init.

