Loss plot for word-based training