State-of-the-art MT systems use so called log-linear model, which combines several components to predict the probability of the translation of a given sentence. Each component has its weight in the log-linear model. These weights are generally trained to optimize BLEU, but there are many alternative automatic metrics and some of them correlate better with human judgments than BLEU. We explore various metrics (PER, WER, CDER, TER, BLEU and SemPOS) in terms of correlation with human judgments. Metric SemPOS is examined in more detail and we propose some approximations and variants. We use the examined metrics to train Czech to English MT system using MERT method and explore how optimizing toward various automatic evaluation metrics affects the resulting model.