The Transformer model is a very recent, fast and powerful discovery in neural machine translation. We experiment with multi-task learning for enriching the source side of the Transformer with linguistic resources to provide it with additional information to learn linguistic and world knowledge better. We analyze two approaches: the basic shared model with multi-tasking through simple data manipulation, and multi-decoder models. We test joint models for machine translation (MT) and POS tagging, dependency parsing and named entity recognition as the secondary tasks. We evaluate them in comparison with the baseline and with dummy, linguistically unrelated tasks. We focus primarily on the standard-size data setting for German-to-Czech MT. Although our enriched models did not significantly outperform the baseline, we empirically document that (i) the MT models benefit from the secondary linguistic tasks; (ii) considering the amount of training data consumed, the multi-tasking models learn faster; (iii) in low-resource conditions, the multi-tasking significantly improves the model; (iv) the more fine-grained annotation of the source as the secondary task, the higher benefit to MT.