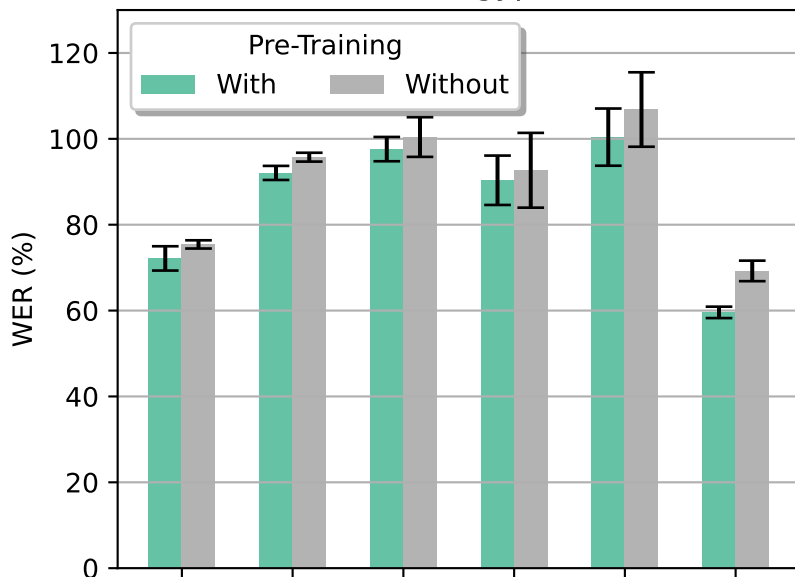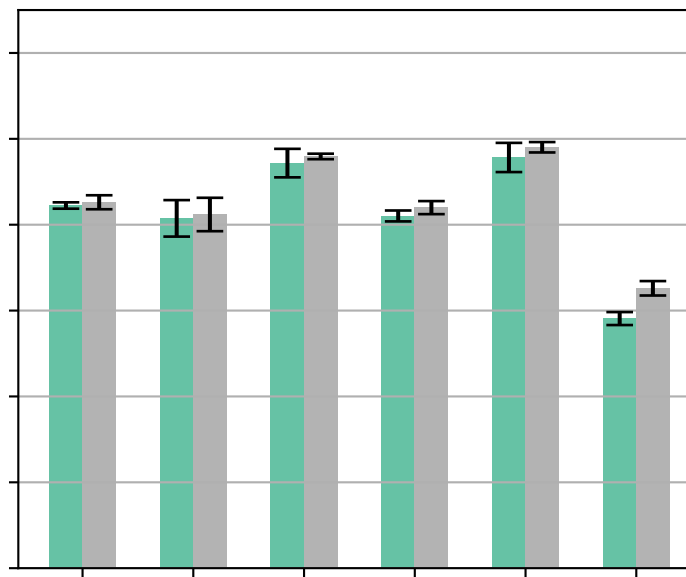Model performance on each test set with and without pre-traini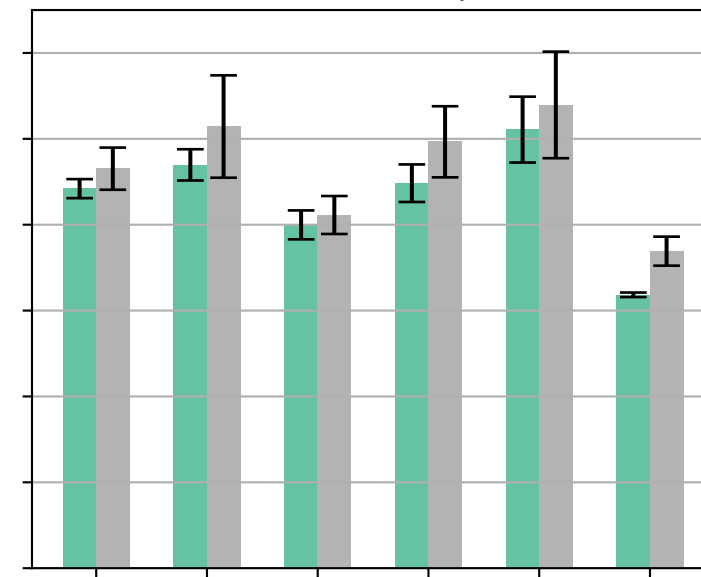ng