Model performance in CER (%) against each dialect with pre-training