Pre-training Configuration	CER Top 1	Accuracy Top 1	Top 5
BORT-PR	0.083	0.869	0.931
BORT-SP	0.106	0.843	0.906
BORT-PR-SP	0.057	**0.901**	0.947
BORT-PR-NOISY	0.089	0.863	0.925
BORT-SP-NOISY	0.096	0.848	0.911
BORT-PR-SP-NOISY	0.060	0.895	0.947
BART-BASE	0.228	0.725	0.820

(a) "Easy" Task

Pre-training Configuration	CER Top 1	Accuracy Top 1	Top 5
BORT-PR	0.462	0.451	0.634
BORT-SP	0.526	0.401	0.579
BORT-PR-SP	0.447	*0.456*	0.641
BORT-PR-NOISY	0.452	*0.458*	0.640
BORT-SP-NOISY	0.469	0.446	0.625
BORT-PR-SP-NOISY	0.420	**0.467**	0.656
BART-BASE	0.606	0.363	0.533

(b) "Hard" Task

Table 3: Accuracies for each pre-trained model after fine-tuning to our two tasks. Bold font indicates accuracy was significantly different from all other models, with the exception of those italicized, according to McNemar's test.