

Supplemental Material

In this appendix we cover: a) hyper-parameters, b) error frequencies c) precision and recall results for the three main adaptation experiments.

A Hyper-parameters

For pre-training the general purpose model:

```
-epochs 15
-word_vec_size 500
-encoder_type brnn
-decoder_type rnn
-enc_layers 3
-dec_layers 3
-batch_size 296
-rnn_size 500
-dropout 0.1
-learning_rate 0.001
-optim adam
-start_decay_at 6
-report_every 200
-max_grad_norm 1.0
```

For fine-tuning the models on “in-domain” data:

```
-epochs 25
-word_vec_size 500
-encoder_type brnn
-decoder_type rnn
-feat_merge concat
-enc_layers 3
-dec_layers 3
-batch_size 128
-rnn_size 500
-dropout 0.1
-learning_rate 0.00025
-optim adam
-start_decay_at 16
-report_every 50
-max_grad_norm 1.0
```

B Error Frequencies

L1-Level	# errors
CN-C1	12.13
FR-B1	13.17
DE-B1	12.48
IT-B1	12.13
PT-B1	13.14
ES-A2	17.33
ES-B1	13.28
ES-B2	12.53

Table 1: The numbers of errors per 100 words for the L1-Level testsets.

C Precision and Recall Results

	A2			B1			B2		
	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	34.9	20.1	30.4	39.0	24.5	34.9	38.1	21.6	33.1
Random	57.3	29.9	48.4	57.5	28.8	47.9	51.4	25.1	42.5
Adapted	60.4	37.0	53.6	58.1	29.4	48.6	51.9	24.9	42.7

	C1			C2		
	P	R	F0.5	P	R	F0.5
None	37.8	20.9	32.5	39.0	20.5	33.0
Random	51.7	23.0	41.4	50.8	20.5	39.2
Adapted	54.3	23.9	43.3	54.5	20.8	41.1

Table 2: Adaptation by proficiency level.

	Chinese			Italian			Portuguese			Spanish		
	P	R	F0.5	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	42.6	22.7	36.2	33.4	19.6	29.3	36.9	20.2	31.7	36.5	21.6	32.1
Random	54.3	26.8	45.0	55.2	25.7	44.9	55.3	26.2	45.2	57.5	28.2	47.6
Adapted	54.4	28.9	46.2	57.5	26.8	46.8	56.9	28.8	47.6	59.5	30.1	49.8

	French			German			Russian			SwissGerman		
	P	R	F0.5	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	37.8	21.2	32.7	35.1	22.2	31.4	41.1	23.8	35.8	35.4	20.9	31.1
Random	54.9	25.9	44.9	55.0	25.5	44.7	54.0	27.6	45.3	54.4	25.2	44.2
Adapted	56.8	26.3	46.1	58.3	26.7	47.1	56.7	29.5	47.8	57.8	27.2	47.1

	Arabic			Greek			Polish			Turkish		
	P	R	F0.5	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	43.0	24.9	37.5	36.4	23.1	32.7	40.3	25.2	36.0	40.0	24.4	35.4
Random	54.9	28.5	46.3	56.3	27.2	46.4	54.5	28.7	46.2	55.7	29.0	47.0
Adapted	55.8	31.3	48.3	58.4	29.8	49.0	57.0	30.2	48.4	58.7	32.7	50.6

Table 3: Adaptation by L1

	Chinese-B2			Chinese-C1			French-B1		
	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	41.4	23.9	36.1	39.9	18.6	32.5	36.4	21.0	31.8
Random	51.2	25.6	42.7	49.9	20.9	39.1	54.8	26.7	45.3
Adapted Level	51.9	26.1	43.4	52.2	22.0	41.0	55.7	27.9	46.5
Adapted L1	52.1	27.4	44.1	51.3	22.6	40.9	56.4	27.2	46.5
Adapted L1 & Level	53.5	28.4	45.5	52.9	24.8	43.1	57.6	29.0	48.1

	German-B1			Italian-B1			Portuguese-B1		
	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	35.3	21.2	31.2	32.1	18.8	28.1	36.2	20.6	31.4
Random	56.5	26.5	46.1	54.7	24.0	43.5	55.1	26.2	45.2
Adapted Level	57.0	27.4	46.9	56.4	25.3	45.3	56.0	27.0	46.1
Adapted L1	59.2	27.5	48.1	58.6	25.5	46.5	55.2	28.0	46.2
Adapted L1 & Level	60.9	29.5	50.2	58.6	26.6	47.3	57.5	28.7	47.9

	Spanish-A2			Spanish-B1			Spanish-B2		
	P	R	F0.5	P	R	F0.5	P	R	F0.5
None	32.8	19.7	28.9	35.8	22.1	31.9	38.9	22.1	33.7
Random	58.7	31.8	50.2	55.6	27.9	46.4	54.4	25.1	44.1
Adapted Level	62.7	40.8	56.6	56.8	28.8	47.5	54.0	24.8	43.7
Adapted L1	61.3	36.1	53.8	56.4	29.2	47.6	54.4	25.6	44.4
Adapted L1 & level	63.7	43.2	58.2	57.5	30.3	48.8	56.0	26.1	45.6

Table 4: Adaptation by L1 and proficiency level. Comparison with adaptation by Level only or by L1 only.