

C	W	S	ar	he	es	it	et	fi	nl	no	hr	ru	average
MONO			76.3	80.2	83.7	83.3	70.4	70.8	77.3	80.8	76.8	82.3	78.2
✗	✓	ID	76.3	80.3	84.2	84.5	72.1	72.5	78.8	81.4	77.6	82.8	79.1
✗	✓	✓	76.4	80.4	84.1	84.4	71.9	72.0	78.7	81.5	78.0	82.8	79.0
✗	ID	ID	76.2	80.1	84.2	84.8	71.8	72.4	78.6	81.6	77.4	82.9	79.0
✗	ID	✓	76.4	80.4	84.2	84.3	72.2	72.3	78.3	81.6	77.4	82.8	79.0
✗	✗	ID	76.5	80.6	84.1	84.3	71.8	72.0	78.5	81.5	77.7	82.9	79.0
✓	ID	✓	76.2	79.8	84.4	84.7	72.4	71.5	79.2	81.6	76.9	82.8	78.9
ID	✗	✓	76.3	80.2	84.0	84.4	72.8	71.7	78.6	82.1	77.2	82.3	78.9
✗	✗	✓	76.6	80.3	84.0	83.7	71.5	72.9	78.3	81.5	77.4	82.8	78.9
ID	ID	ID	76.3	79.9	84.1	84.4	72.1	71.3	79.6	81.4	77.1	82.5	78.9
✓	ID	ID	76.2	80.1	84.3	84.5	72.5	71.8	78.7	81.1	76.7	82.6	78.8
ID	✗	✗	76.6	80.3	84.1	84.3	71.8	71.7	78.3	81.5	77.2	82.5	78.8
ID	✗	ID	76.2	79.9	84.1	84.3	71.9	72.0	78.5	81.7	77.2	82.3	78.8
✗	✗	✗	76.6	80.3	84.0	84.2	71.2	72.1	78.0	81.6	77.3	82.6	78.8
✓	✗	✓	76.1	80.1	84.0	84.1	72.1	72.0	78.2	81.8	76.8	82.6	78.8
✓	✗	ID	76.5	80.0	84.0	84.4	71.9	71.7	78.5	81.7	76.7	82.3	78.8
✓	✓	ID	76.6	80.1	84.0	84.6	72.1	71.0	78.3	81.0	76.9	82.8	78.7
ID	ID	✓	76.1	80.1	84.1	84.6	72.1	71.2	78.0	81.4	77.0	82.7	78.7
✓	✗	✗	76.4	80.3	84.3	84.0	72.3	71.0	78.3	81.3	77.0	82.3	78.7
✓	✓	✓	76.2	80.1	84.0	84.2	72.1	71.4	78.7	81.1	77.0	82.5	78.7
✗	ID	✗	76.6	80.1	84.1	84.3	71.7	71.6	77.6	81.0	77.0	82.6	78.7
ID	✓	✓	76.2	79.9	83.8	84.5	72.4	70.3	78.1	81.0	77.2	82.6	78.6
✗	✓	✗	76.3	79.9	83.9	84.4	72.4	71.3	77.4	80.7	76.9	82.5	78.6
ID	✓	ID	76.0	80.0	83.8	84.3	71.7	70.9	78.3	81.0	76.9	82.5	78.5
✓	✓	✗	76.1	79.7	83.8	84.5	71.9	70.4	77.8	81.1	76.5	82.3	78.4
✓	ID	✗	76.0	79.3	84.1	84.4	71.5	71.3	77.7	80.6	76.7	82.5	78.4
ID	✓	✗	76.1	79.9	83.9	84.4	71.1	70.6	77.8	80.9	77.0	82.0	78.4
ID	ID	✗	75.9	79.5	84.1	84.4	72.1	70.5	77.4	80.6	77.0	82.2	78.4

Table 5: Performance (LAS; in %) across select sharing strategies ranked by average performance. MONO is our single-task baseline; **W** refers to sharing the word parameters, **C** refers to sharing the character parameters, **S** refers to sharing the MLP parameters. ✓ refers to hard sharing, ID refers to soft sharing, using an embedding of the language ID and ✗ refers to not sharing.

C	W	S	he	no	fi	hr	ru	es	it	et	nl	ar	average
MONO			80.2	80.8	70.8	76.8	82.3	83.7	83.3	70.4	77.3	76.3	78.2
✗	✗	✓	80.3	81.5	71.9	77.6	82.7	84.0	83.8	72.5	78.7	76.4	78.9
✗	✗	✗	80.3	81.7	71.9	77.7	82.5	84.0	84.2	71.8	78.5	76.5	78.9
✗	ID	ID	80.3	81.1	72.1	77.7	82.7	84.2	84.2	72.4	77.9	76.0	78.9
✗	✗	ID	79.7	81.5	72.2	77.1	82.7	83.8	84.0	72.3	78.6	76.5	78.8
ID	✗	ID	80.3	81.6	71.8	77.5	82.6	84.1	83.8	71.8	78.3	76.2	78.8
ID	✗	✗	80.0	81.6	71.5	77.2	82.7	83.9	84.0	71.1	79.0	76.4	78.7
✓	✗	ID	80.3	81.5	71.5	77.5	82.7	83.7	83.9	71.6	77.9	76.7	78.7
ID	✗	✓	80.4	81.4	71.6	77.3	82.6	83.9	84.1	71.9	77.7	76.4	78.7
✗	✓	ID	80.5	81.2	72.2	77.0	82.5	84.0	83.8	71.5	78.3	76.1	78.7
✗	ID	✓	80.6	81.1	71.9	77.1	82.7	84.2	84.0	71.9	77.1	76.2	78.7
✓	✗	✓	80.3	81.1	71.3	77.0	82.6	84.0	84.2	71.8	77.8	76.3	78.6
✗	✓	✓	80.6	81.1	71.6	76.8	82.5	83.7	83.9	71.4	78.1	76.3	78.6
✓	✗	✗	80.1	80.9	71.4	76.8	82.9	83.9	84.3	70.9	78.0	76.5	78.6
✗	ID	✗	80.3	81.3	71.2	77.4	81.9	84.2	83.9	70.7	77.6	75.8	78.4
✗	✓	✗	79.6	80.9	71.9	76.9	82.2	83.7	83.8	70.9	77.0	76.4	78.3
ID	✓	ID	80.3	81.0	70.5	76.5	82.3	83.7	83.6	71.4	77.8	76.1	78.3
ID	✓	✓	80.1	80.8	70.4	77.0	82.2	83.8	83.8	71.0	77.6	76.2	78.3
✓	✓	✓	80.5	80.9	69.8	76.6	82.3	83.7	84.0	70.6	77.4	76.2	78.2
✓	ID	ID	80.3	80.7	70.2	76.1	82.1	83.8	83.8	70.8	77.6	76.2	78.2
✓	✓	ID	79.8	80.9	71.0	76.2	82.1	83.7	83.6	70.9	77.3	76.0	78.2
ID	ID	✓	80.0	80.8	69.8	76.2	82.2	83.8	84.3	70.7	77.2	76.2	78.1
ID	ID	ID	79.8	80.5	70.1	76.6	82.1	83.9	83.8	70.6	77.2	76.3	78.1
✓	ID	✓	80.3	81.1	70.2	76.1	82.2	84.1	83.7	70.3	76.9	76.0	78.1
✓	✓	✗	80.4	80.3	70.1	76.6	82.0	83.7	83.2	69.3	76.7	76.2	77.8
✓	ID	✗	79.6	80.6	69.4	76.7	81.7	83.8	83.4	69.2	77.6	76.2	77.8
ID	✓	✗	79.6	80.1	69.9	76.6	81.8	83.5	82.9	69.2	77.6	76.3	77.7
ID	ID	✗	79.8	80.6	69.2	76.7	81.4	83.8	83.2	69.4	76.6	76.0	77.7

Table 6: Unrelated language pairs. Performance (LAS; in %) across select sharing strategies ranked by average performance. MONO is our single-task baseline; **W** refers to sharing the word parameters, **C** refers to sharing the character parameters, **S** refers to sharing the MLP parameters. ✓ refers to hard sharing, ID refers to soft sharing, using an embedding of the language ID and ✗ refers to not sharing.