

Why it Helps? Not Really Vocabulary (1/2)

| | Length | BLEU Components | BP |
|-----------|--------|--------------------|-------|
| Base ENET | 35326 | 48.1/21.3/11.3/6.4 | 0.979 |
| ENRU+ENET | 35979 | 51.0/24.2/13.5/8.0 | 0.998 |
| ENCS+ENET | 35921 | 51.7/24.6/13.7/8.1 | 0.996 |

(The reference length in the matching tokenization was 36062.)

- Child models produce longer outputs \Rightarrow lower brevity penalty.
- But n -gram precisions also better.

| 1-gram present in | ENRU+ENET | ENCS+ENET |
|-------------------|---------------------|---------------------|
| Child, Base, Ref | 15902 (44.2 %) | 15924 (44.3 %) |
| Child only | 9635 (26.8 %) | 9485 (26.4 %) |
| Child, Base | 7209 (20.0 %) | 7034 (19.6 %) |
| Child, Ref | 3233 (9.0 %) | 3478 (9.7 %) |
| Total | 35979 (100.0 %) | 35921 (100.0 %) |

- The 3k better toks are regular ET words, not NEs or numbers.