Bucketing

Bucketing solves the problem of high variance of the length of sentences by putting them into buckets of different sizes.

Buckets for the length intervals (0, 5], [5, 10], (10, 20], …

Q: How are you?

A: I am fine

↓

Q: **[** PAD, "?", "you", "are", "How" **]**

A: **[** GO, "I", "am", "fine", ".", EOS, PAD, PAD, PAD, PAD **]**