

# Sequence to Sequence Learning with Neural Networks

Ilya Sutskever, et Oriol Vinyals, et Quoc V. Le/NIPS/2014

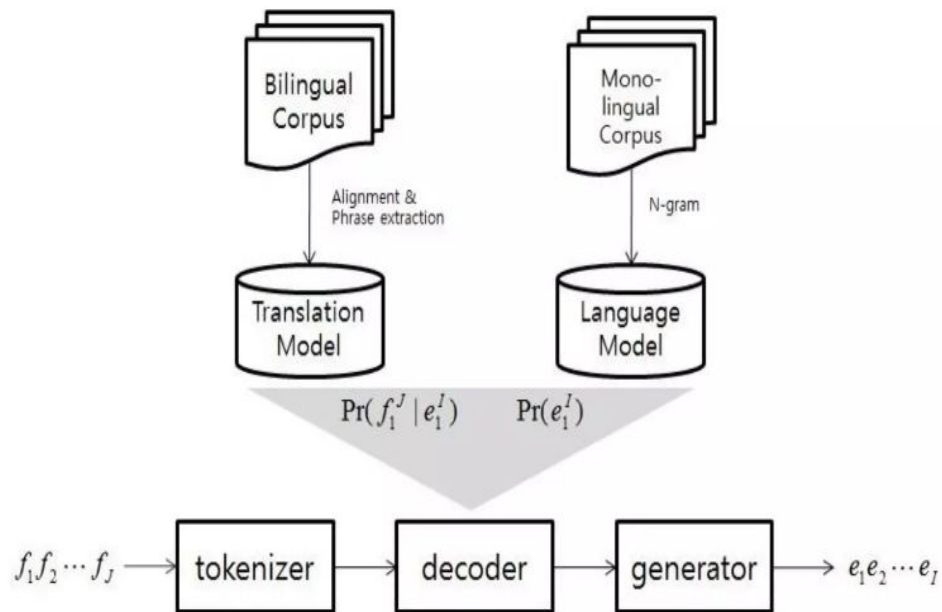
발표자: 나 리 나

# Language Model

- 단어 시퀀스에 확률을 할당(assign) 하는 모델
- 기계 번역

$P(\text{나는 사과를 좋아한다} | \text{I like apple}) > P(\text{나는 키위를 좋아한다} | \text{I like apple})$

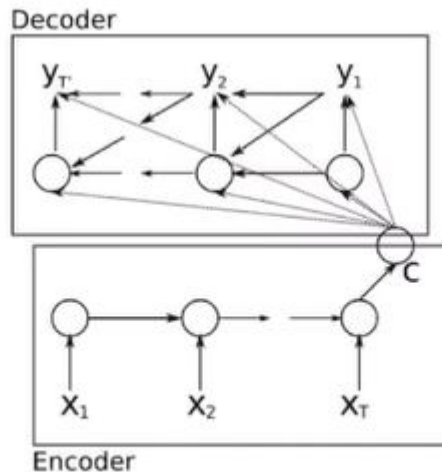
# Statistical Language Model (SMT)



- 카운트에 기반해 확률을 계산하는 모델
- 번역모델과 언어모델로 나누어 번역 수행

$$\operatorname{argmax}_y P(x|y)P(y)$$

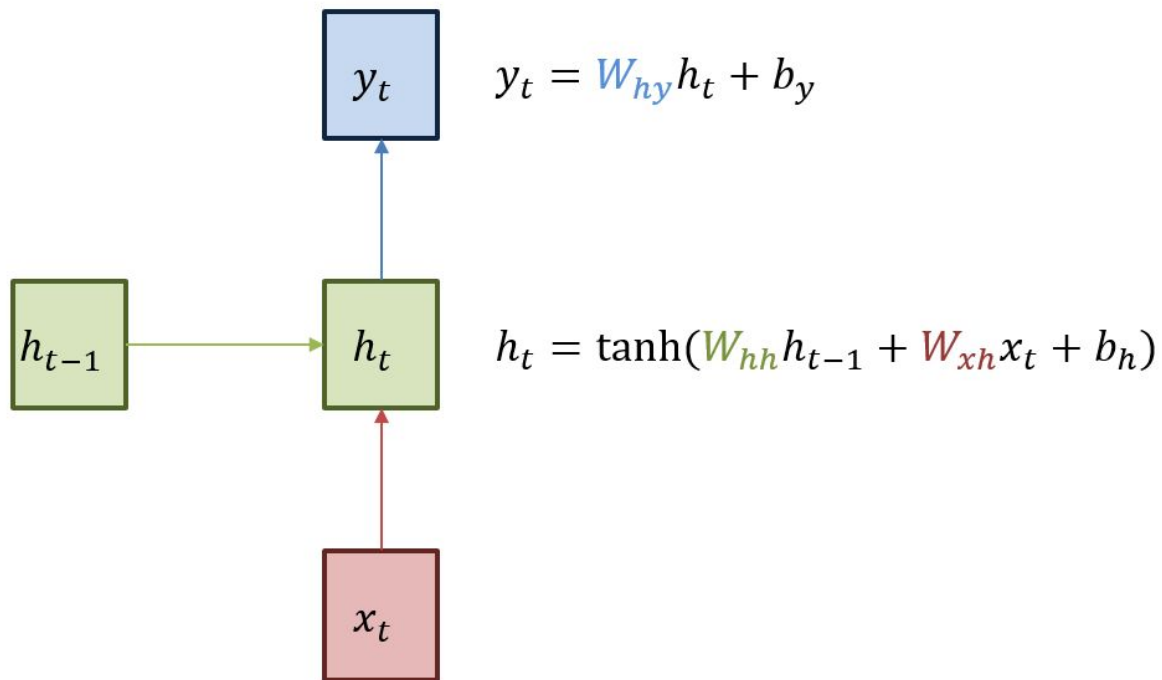
# Neural Machine Translation (NMT)



$$P(y|x) = P(y_1|x) P(y_2|y_1, x) P(y_3|y_1, y_2, x) \dots P(y_T|y_1, \dots, y_{T-1}, x)$$

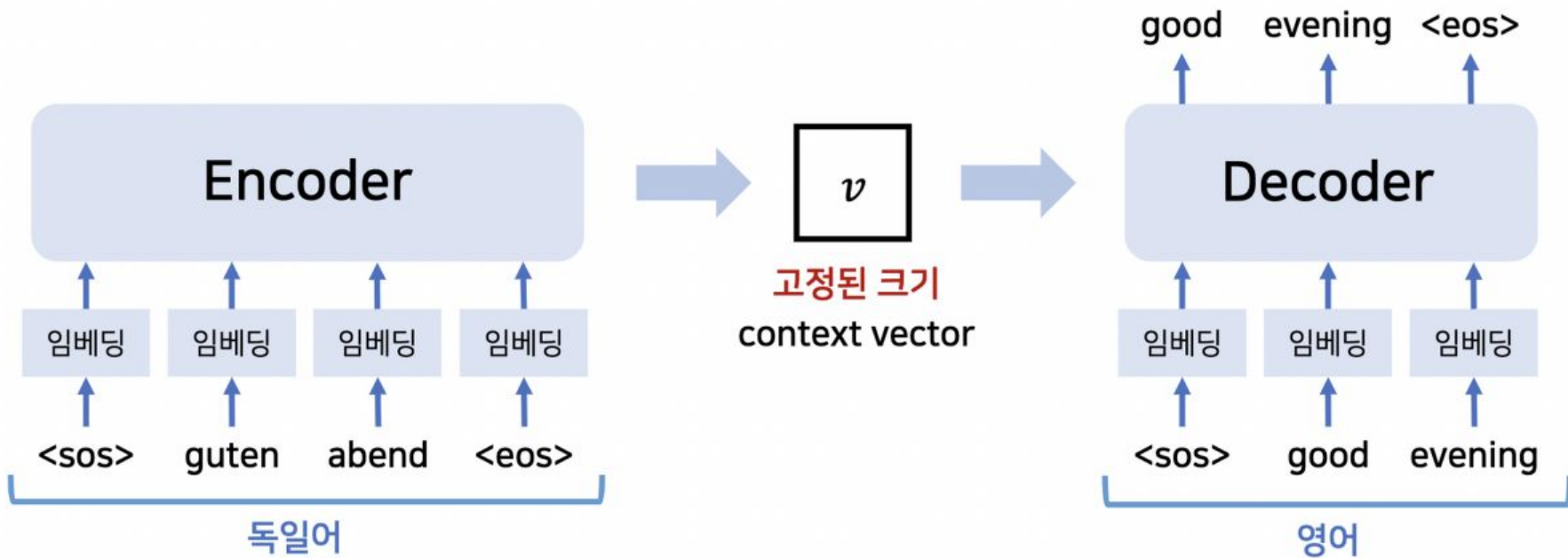
- **Encoder - Decoder** 구조

## RNN 기반 번역



- 입력과 출력의 크기가 같다고 가정

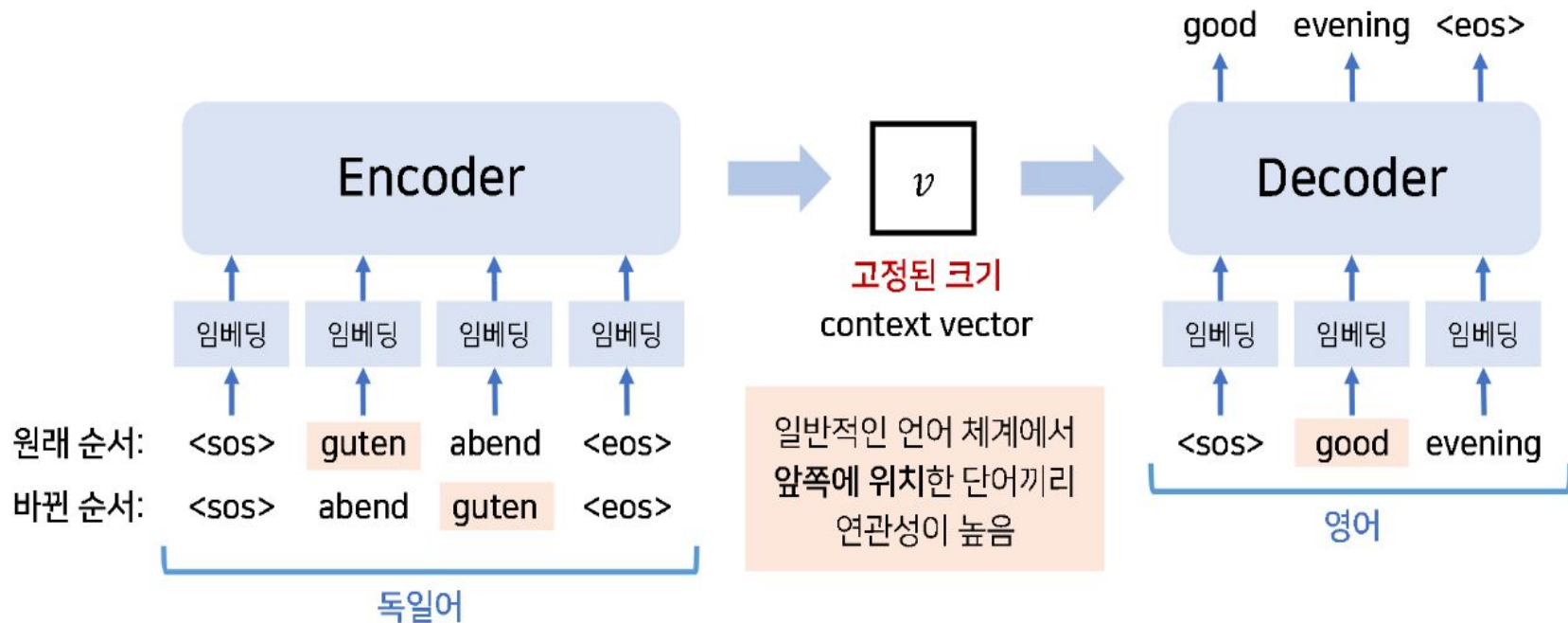
## RNN 기반의 Sequence to Sequence



- **Context Vector** = Encoder에 등장한 모든 단어의 정보 포함

Decoder RNN 셀의 첫번째 Hidden State에 사용

## Sequence to Sequence의 LSTM & 문장 순서 뒤집기



- 출력의 문장 순서는 바꾸지 않음