Look it up: Bilingual and Monolingual Dictionaries Improve Neural Machine Translation

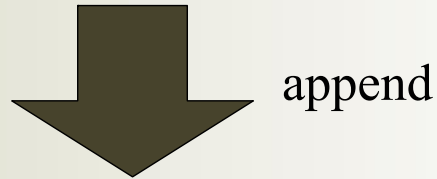## Xing Jie Zhong

Rare Word Problem
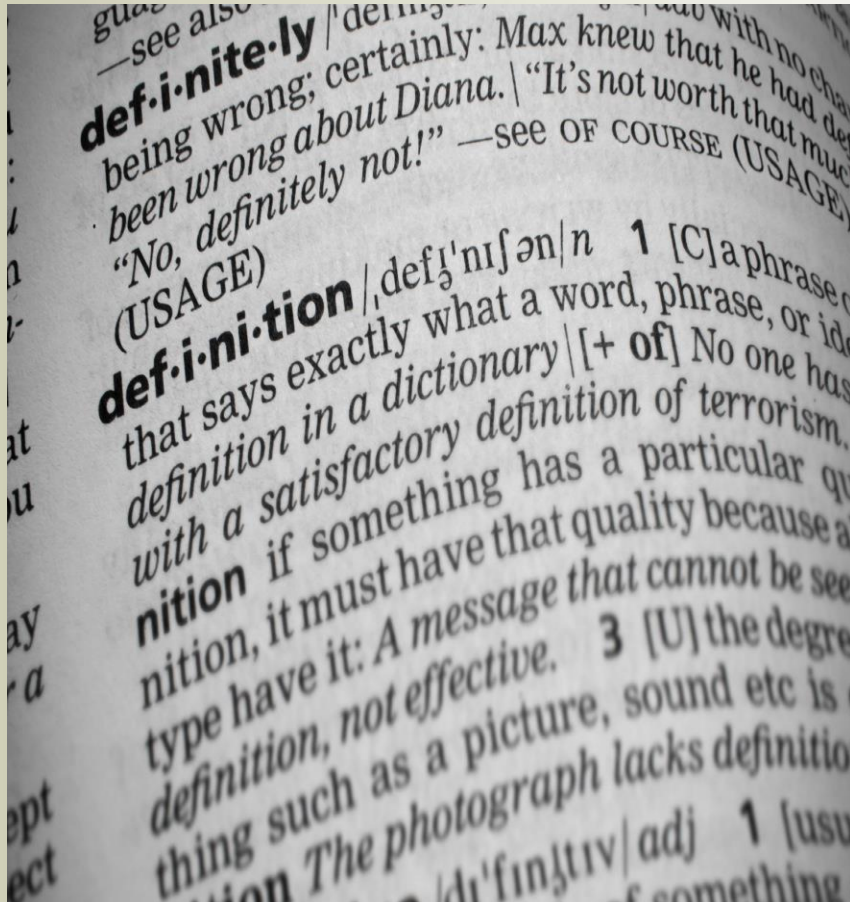
- Still a challenge in NMT

- Dictionaries:
  - Useful in rule-based MT and Statistical MT
  - Not commonly used in NMT

How Dictionaries Are Used

大家 都 知道 死海正在 死亡 Everyone knows the Dead Sea is dying

我哥哥听说我们做了火药 Well, my brother heard that we had made gunpowder

死海 Dead Sea

火药 gunpowder

append

大家 都 知道 死海正在 死亡 Everyone knows the Dead Sea is dying

我哥哥听说我们做了火药 Well, my brother heard that we had made gunpowder

死海 Dead Sea

火药 gunpowder

Our Approach
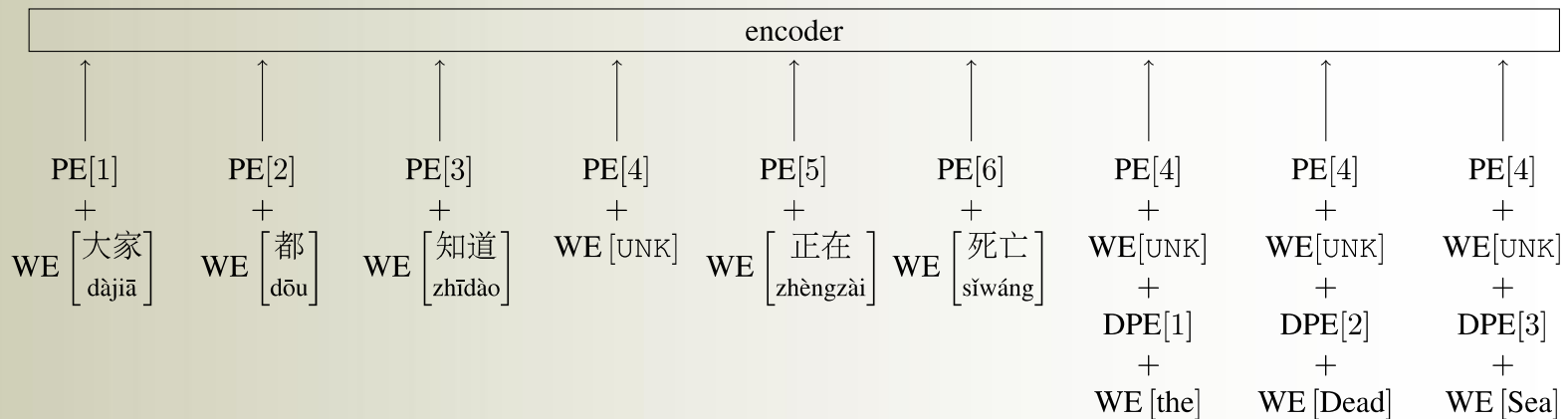
Original sentence: 大家 都 知道 死海(UNK) 正在 死亡

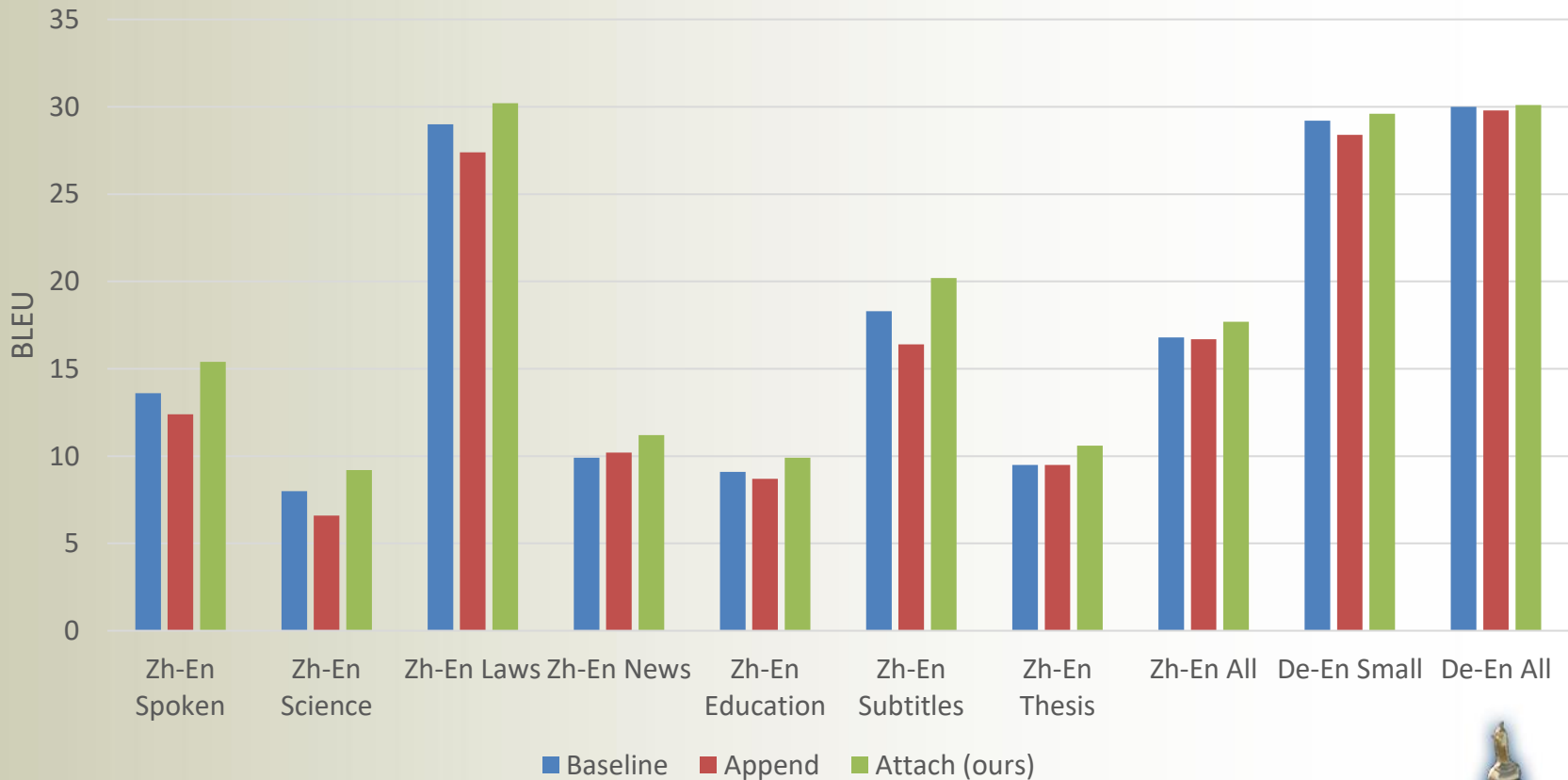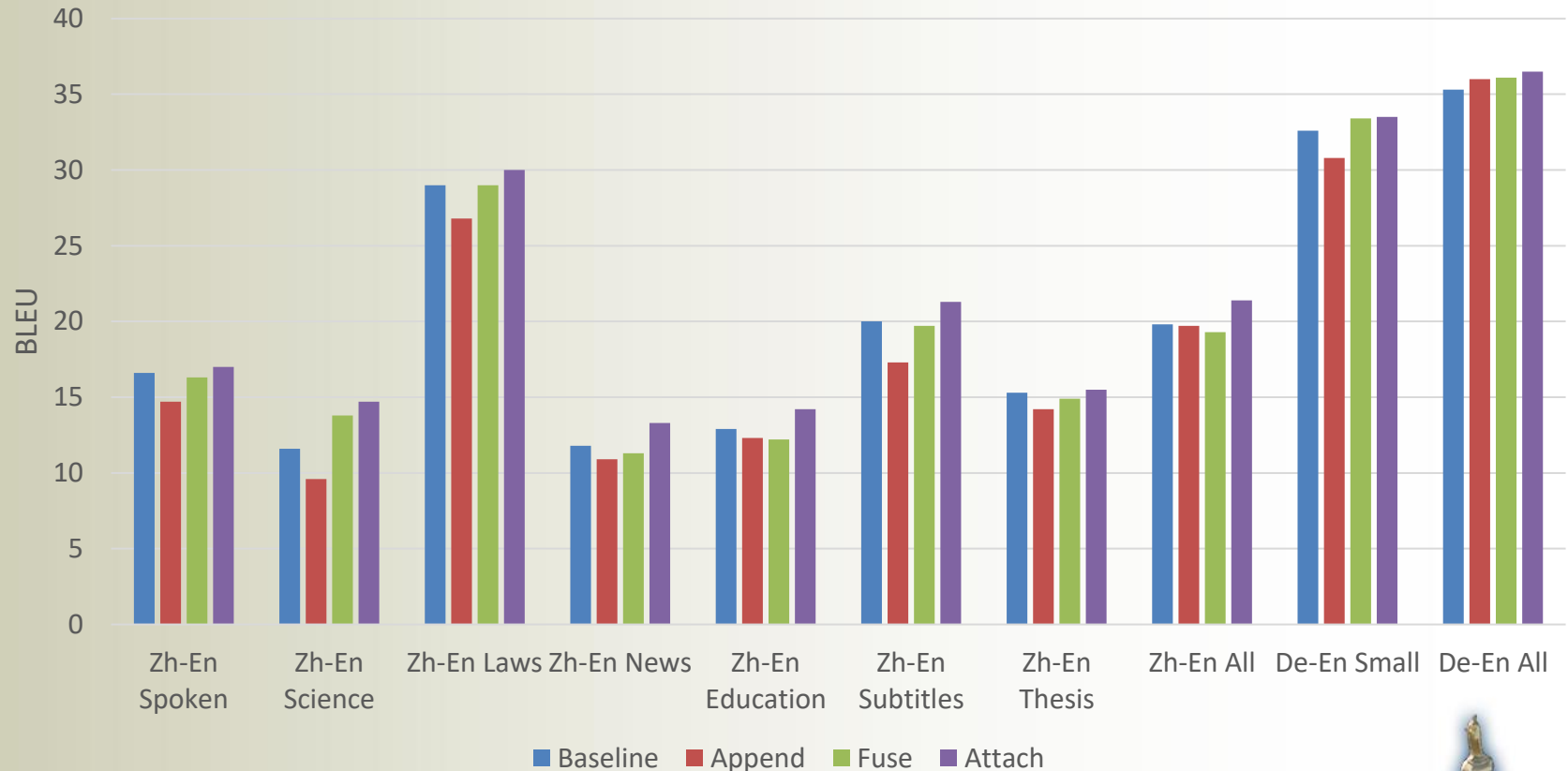Unknown word: 死海

Dictionary Entry: 死海: the Dead Sea

The College of Engineering at the University of Notre Dame

Results on Word Based Models

Word-based results in BLEU

Results on BPE Based Models

BPE-based results in BLEU

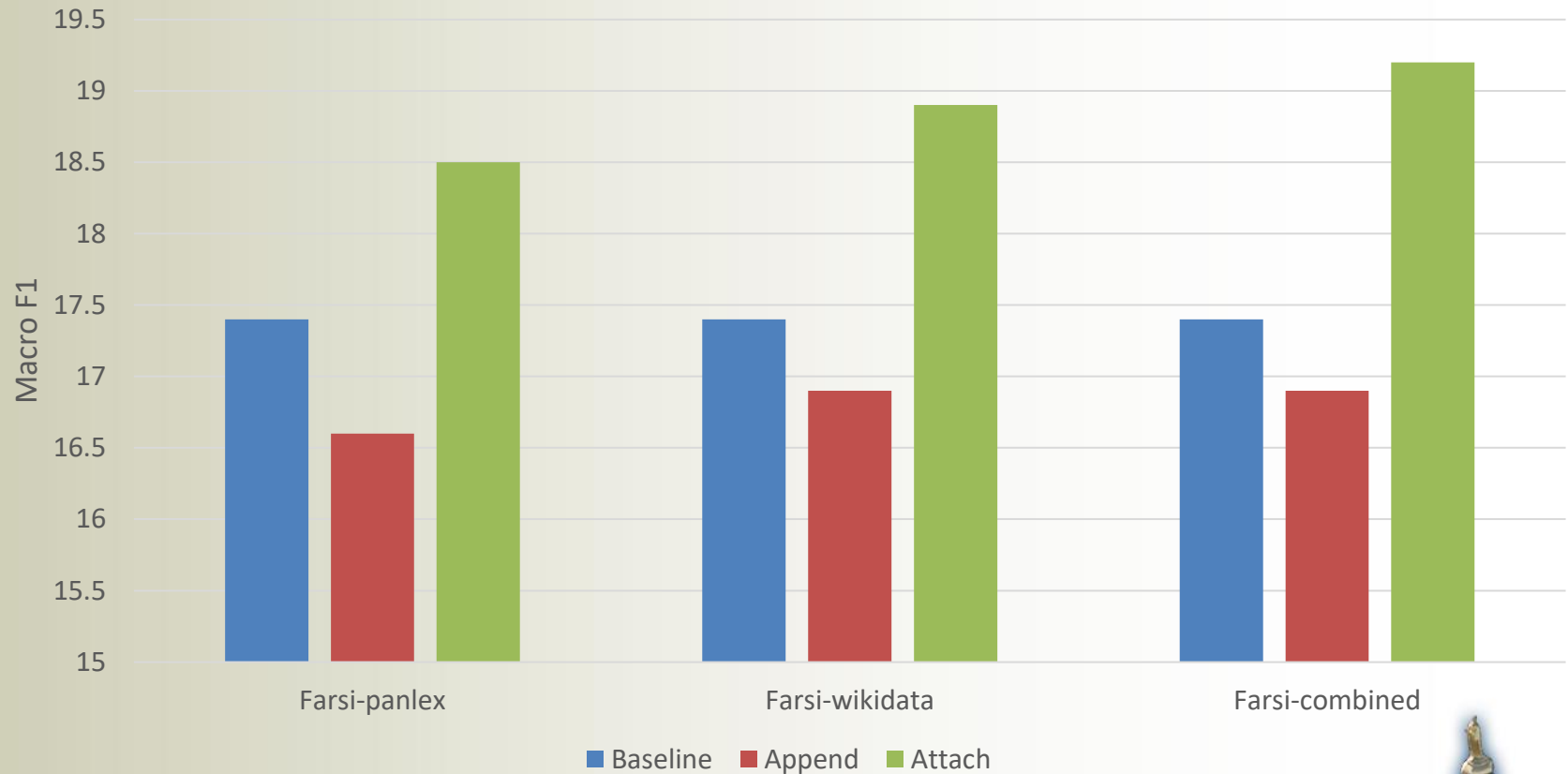The College of Engineering
at the University of Notre Dame

Macro F1 on BPE Based Models

BPE-based results in Macro F1

The College of Engineering at the University of Notre Dame
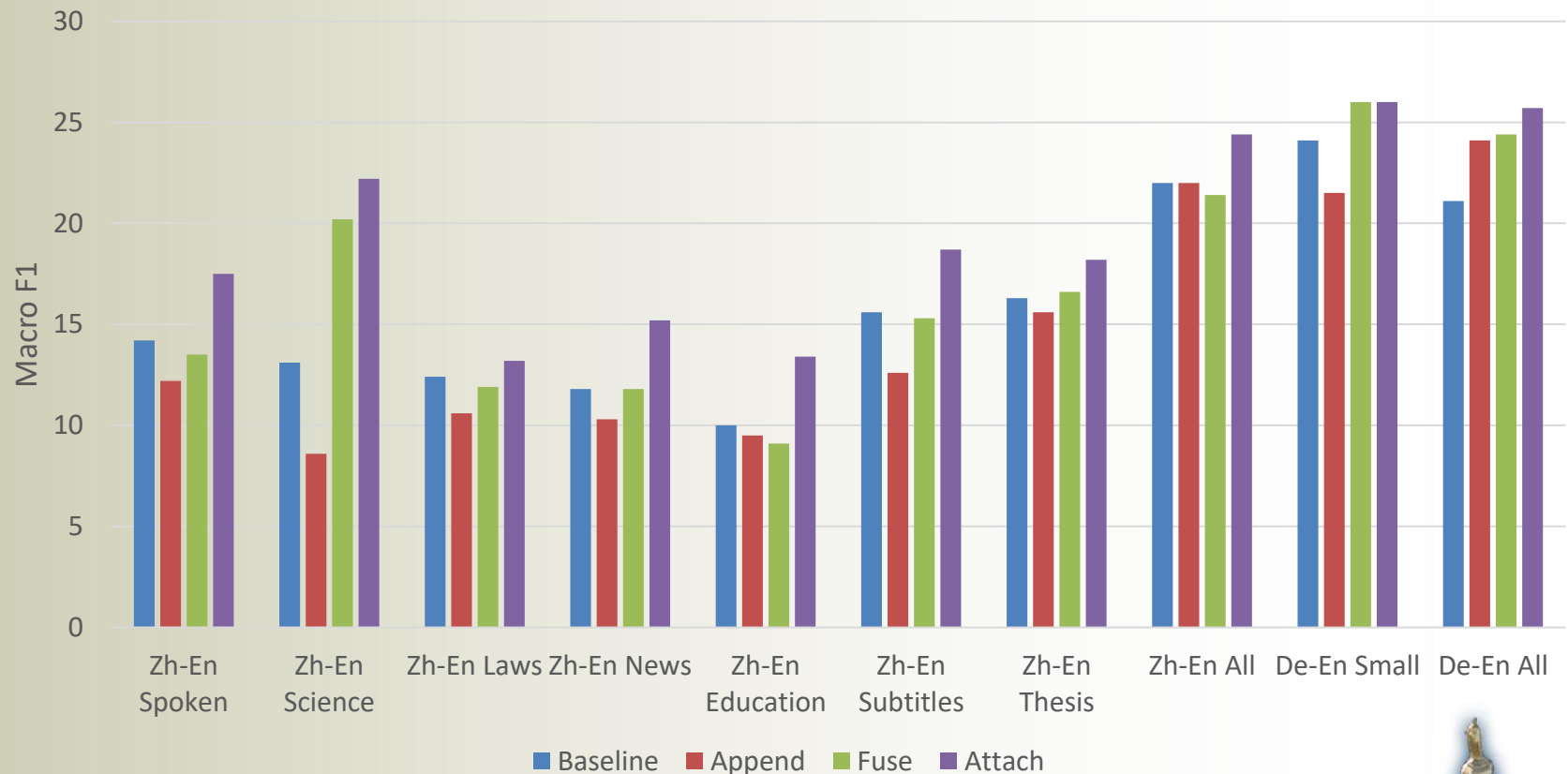
Macro F1 on BPE Based Models

BPE-based results in Macro F1

Example

Source sentence BPE: 我 哥哥 听说 我们 做 了 <mark>火@@ 药</mark> 。

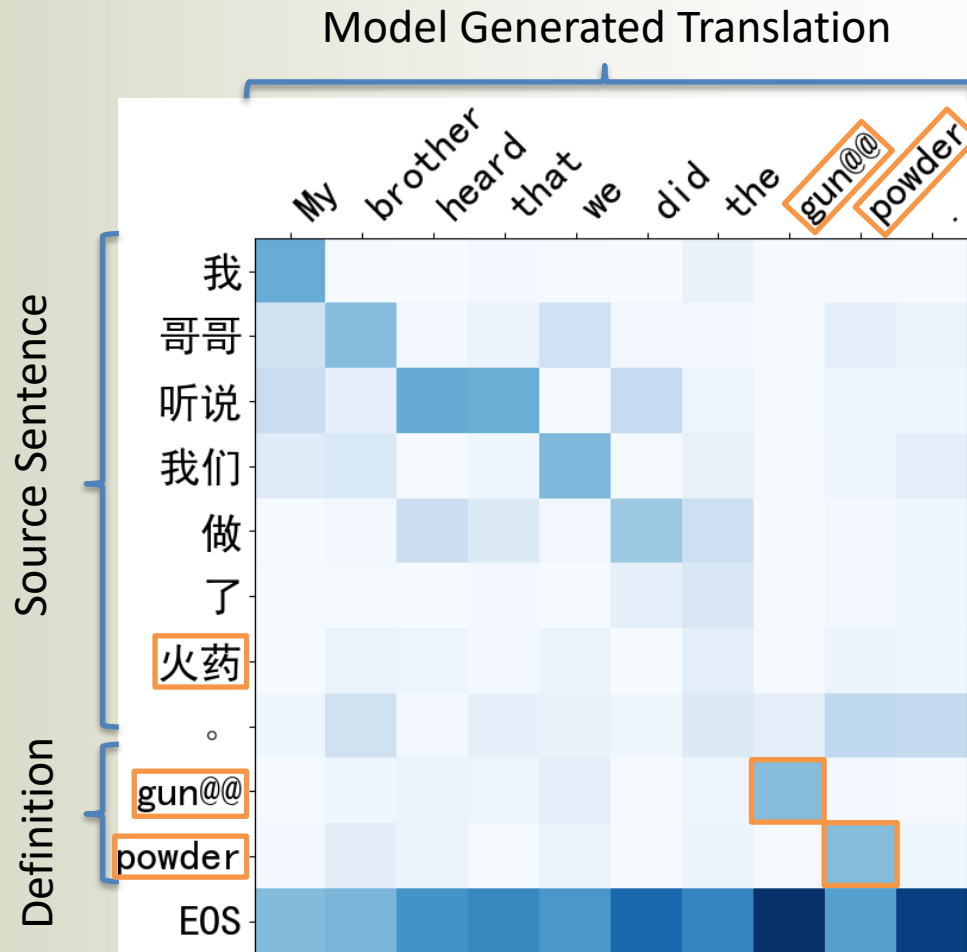Dictionary Entry: 火药: gunpowder

Reference: Well, my brother heard that we had made **gunpowder**.

BPE Baseline: My brother had heard that we did a **fire pills**.

Attach: My brother heard that we did the **gunpowder**.

Attention Visualization

The College *of* Engineering
at the University *of* Notre Dame

Conclusion

- Definitions not equal to translations.

- Don't force the NMT to generate definitions.

- Appending dictionaries may hurt the system.

- Attaching dictionaries improves translation quality and can work on monolingual dictionaries.

The College *of* Engineering
at the University *of* Notre Dame