



## Introduction

### What is Semantic Prosody (SP)?

- A collocational meaning formed through the co-occurrence of a linguistic unit and a consistent series of collocates.
- It is not part of semantic meaning and not necessarily related to it.

被 BEI+verb { Chinese passive structure  
Marker BEI has no concrete meaning  
**NEGATIVE** semantic prosody

BE+past participle { English passive structure  
**NEUTRAL** semantic prosody

### What problem does SP cause in Machine Translation (MT)?

- Literal translation may convey correct meaning but wrong semantic prosody.
- Ignoring semantic prosody may produce unnatural translation.

BE passive — ANY context

BEI passive — BAD event

Trained Human translators can notice the difference of semantic prosody and translate correctly, but MT models are ignorant of this problem.



Teach MT models about SP through **finetuning**

## Datasets

### BEI dataset — English-Chinese sentence pairs

- **Positive evidence:** 476 sentence pairs in which English BE passives are translated to Chinese BEI passives and express negative content. They are selected to reinforce the relation between BEI passives and negativity.

(1) 家珍被拖出去时，双手紧紧捂着凸起的肚子，那里面有我的儿子呵。

As Jiazhen **was carried out**, her hands firmly clasped her protruding belly, which held my son.

- **Negative evidence:** 424 English BE passives translated to active sentences. This subset is intended to attenuate the degree of correspondence between the two passives, so that models may use more active voice when translating BE passives into Chinese.

(2) 你在我朋友家里是待你同儿子一样的。

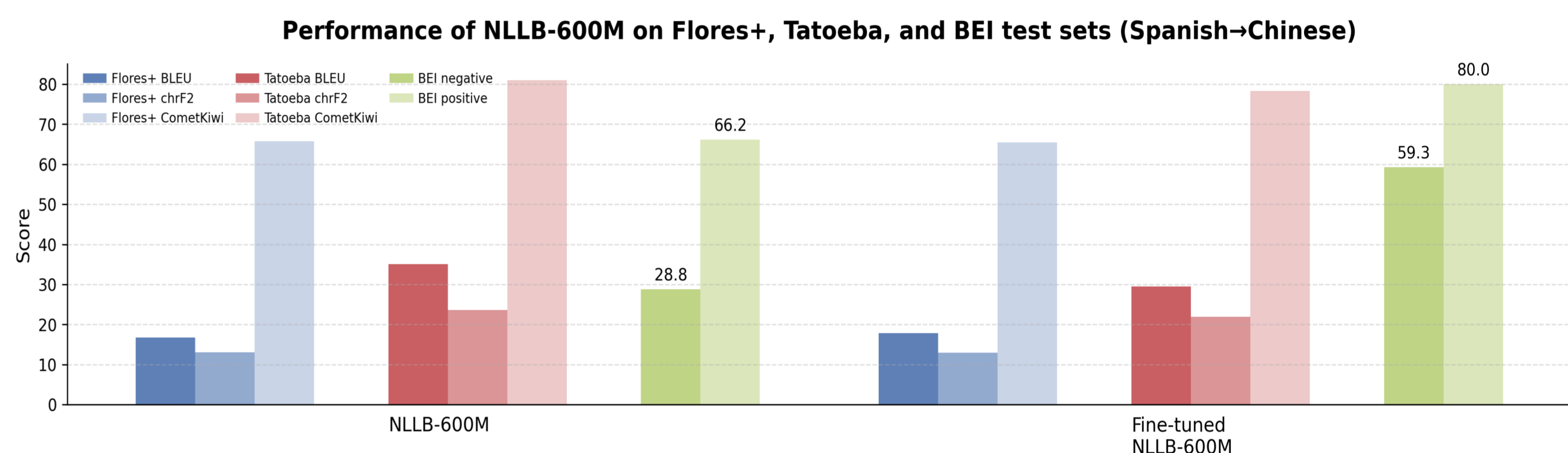
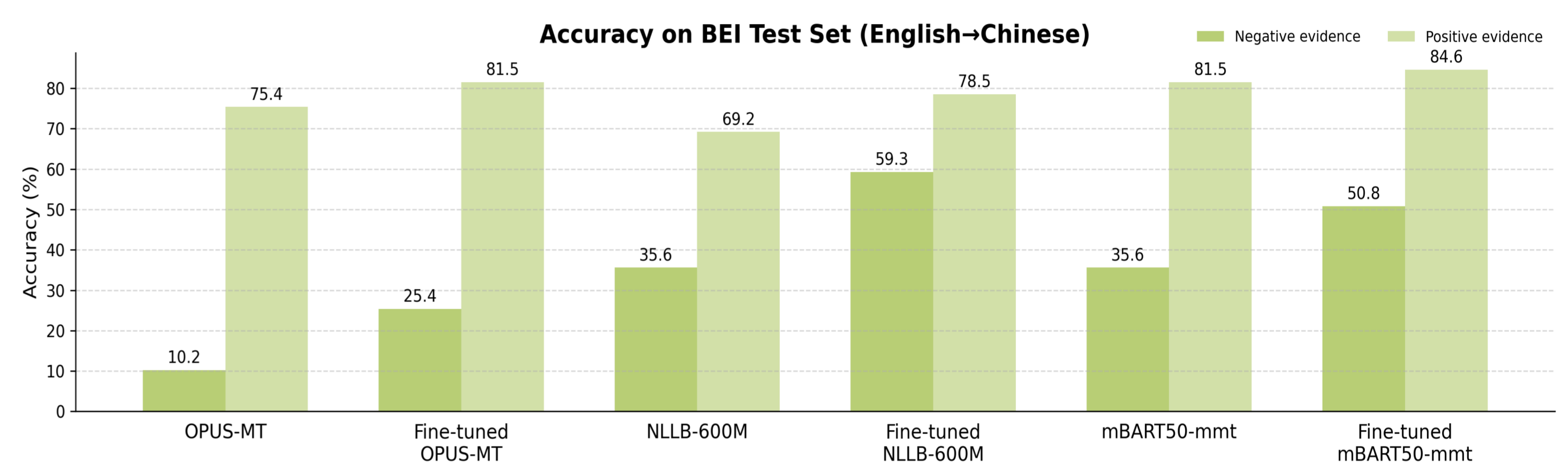
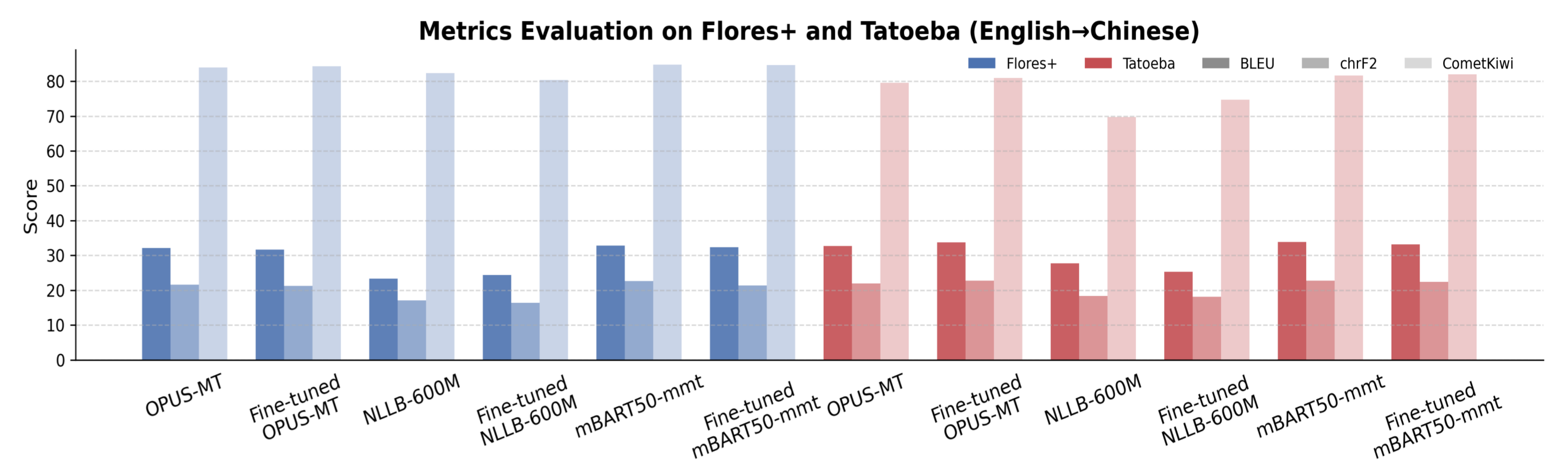
“In my friend’s house they **treated** you as a son.”

You **were treated** as a son in my friend’s house.

## Results

### Models finetuned:

- Helsinki-NLP/opus-mt-en-zh
- facebook/nllb-200-distilled-600M
- facebook/mbart-large-50-many-to-many-mmt



- ◆ After fine-tuning, the models **maintained their original accuracy translating general text** from English to Chinese.
- ◆ All three models achieved **higher accuracy in using BEI passives** when translating BE passives after fine-tuning.
- ◆ This knowledge of semantic prosody is **transferable** for multilingual models when translating passive sentences of other Languages (such as Spanish) into Chinese.

### One example of improvement

NLLB-600M: 在我朋友的家里，你被当作儿子。  
“In my friend’s house, you **were treated** as a son.”

NLLB-600M fine-tuned: 你在我朋友的家里就像一个儿子一样。  
“In my friend’s house you **were** like a son.”

## Future work

- Other ways of translating passives into Chinese (such as topic sentence).
- Teaching models about different semantic prosodies of multiple linguistic units simultaneously.
- Does models treat adjectival and verbal passives differently?
- Multilingual fine-tuning on multilingual models.