

# Multilingual Pre-training with Language and Task Adaptation for Multilingual Text Style Transfer

## Introduction

### Formality Transfer

Inf.: *that is just my gut feeling* ↔ For.: *That is my personal opinion.*

- **GYAFC**: English informal-formal pairs
- **XFORMAL**: Multilingual informal-formal pairs, for EVALUATION

**mBART-50** A pre-trained model helps preserve content

**Language** Language underrepresentation in pre-training

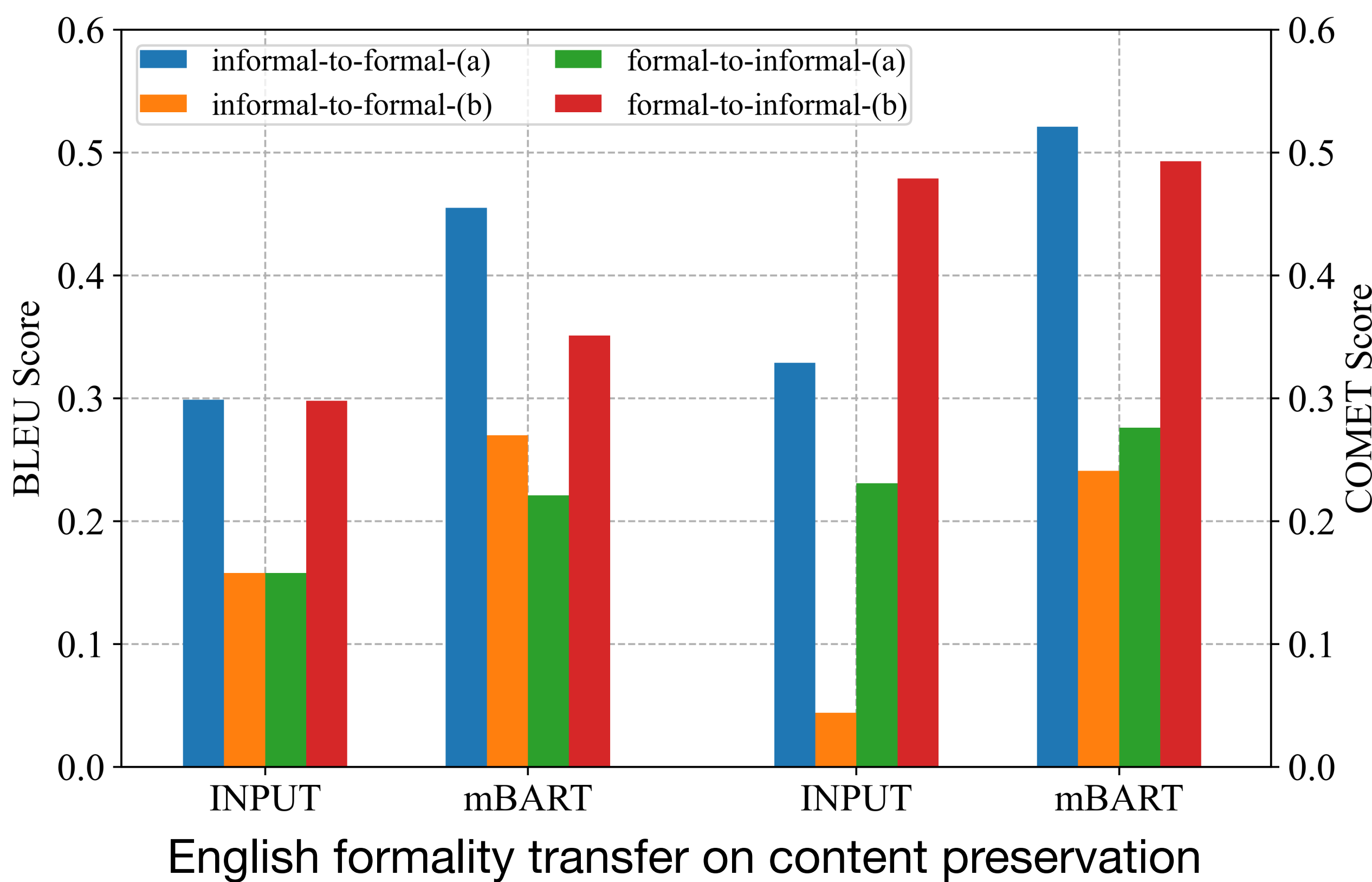
**Task Adaptation** Adapt model to task using English data

## Analysis

### Transfer Directions

MODEL	ITALIAN		FRENCH		PORTUGUESE	
	BLEU	COMET	BLEU	COMET	BLEU	COMET
Informal-to-Formal (setting (a))						
INPUT	0.176	0.078	0.198	-0.019	0.244	0.217
M1.1	0.179	0.170	0.234	0.133	0.269	0.282
Formal-to-Informal (setting (b))						
INPUT	0.174	0.364	0.196	0.277	0.243	0.463
M1.1	0.194	0.326	0.201	0.239	0.226	0.371

Setting (a) uses the original test, and (b) uses the test set of the opposite direction, swapping sources and references



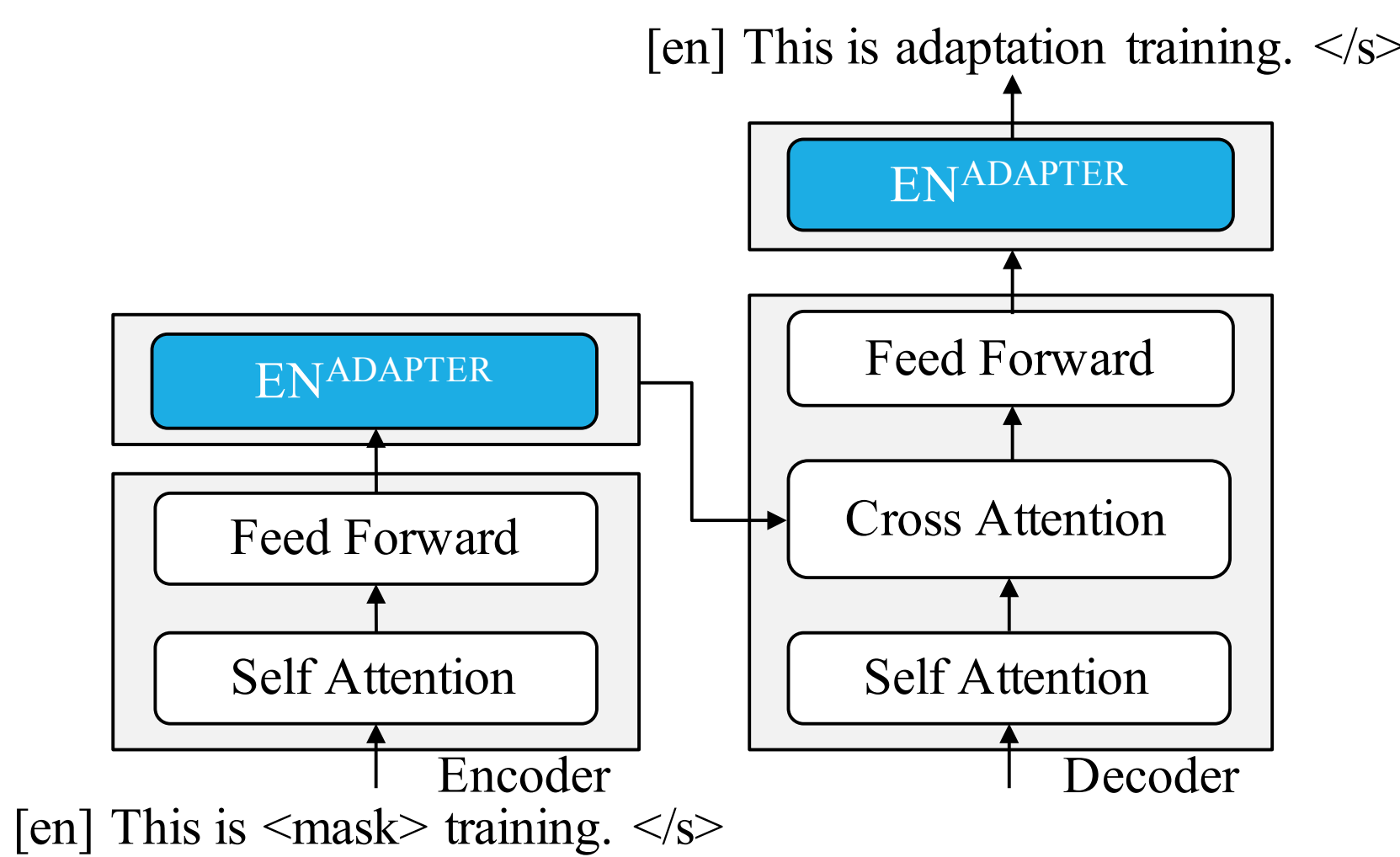
## Conclusion

- Machine Translated Data
  - Achieve SOTA results for informal-to-formal, but the results for the opposite direction are worse
- Language and Task Adaptation Strategies
  - Outperform classic IBT-based approaches
  - This can be applied to other tasks and other languages
- Formal-to-informal is harder than the opposite direction

## Methodology

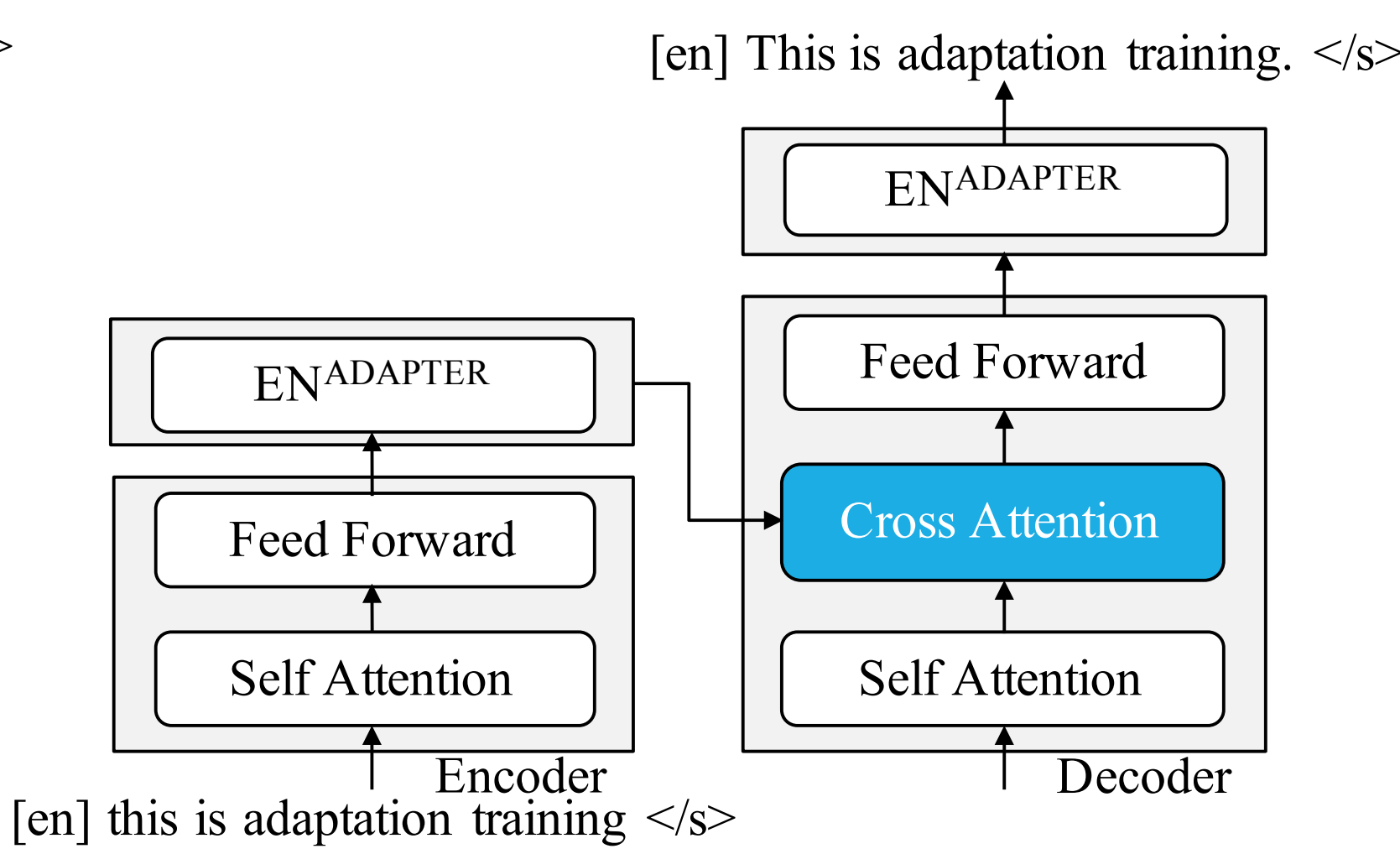
### Adaptation Training

#### Language Adaptation Module



Language adaptation training with monolingual data (news-crawl IT, FR, PT)

#### Task Adaptation Module



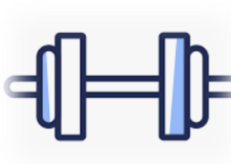
Task adaptation training with English parallel data

## Evaluation

### Automatic Metrics



Content Preservation: **BLEU**



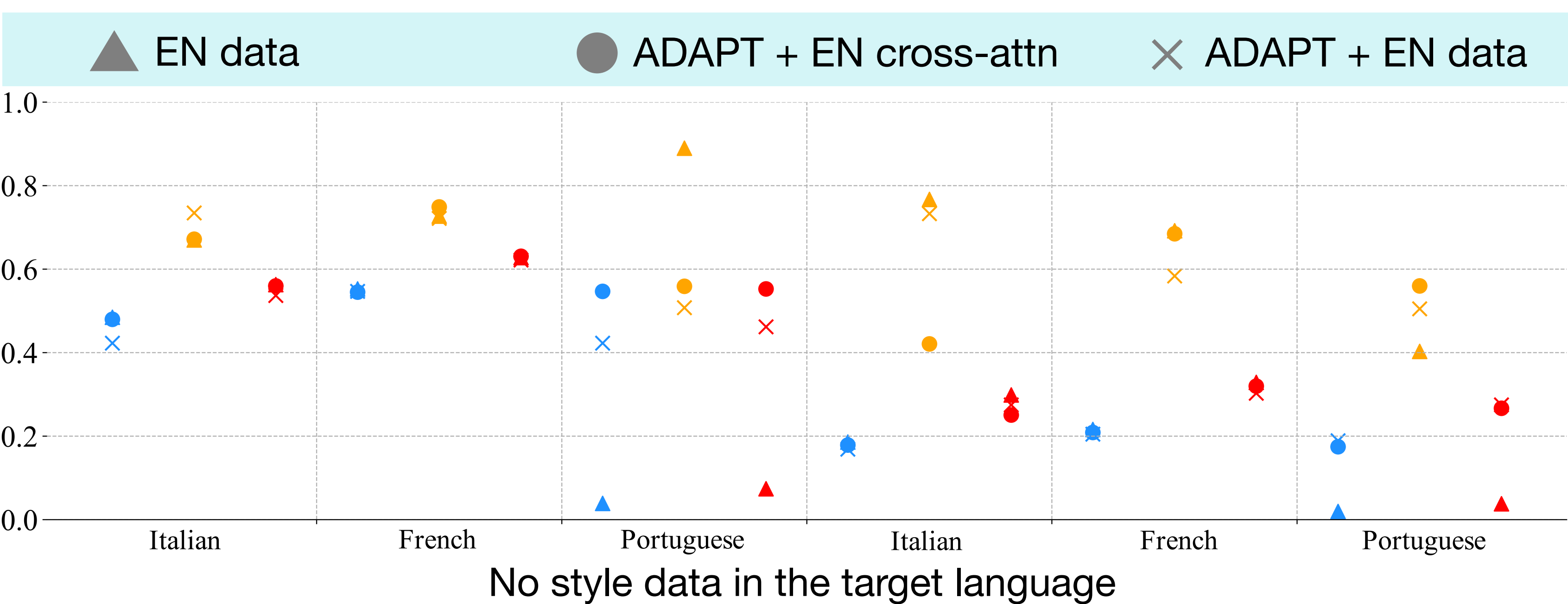
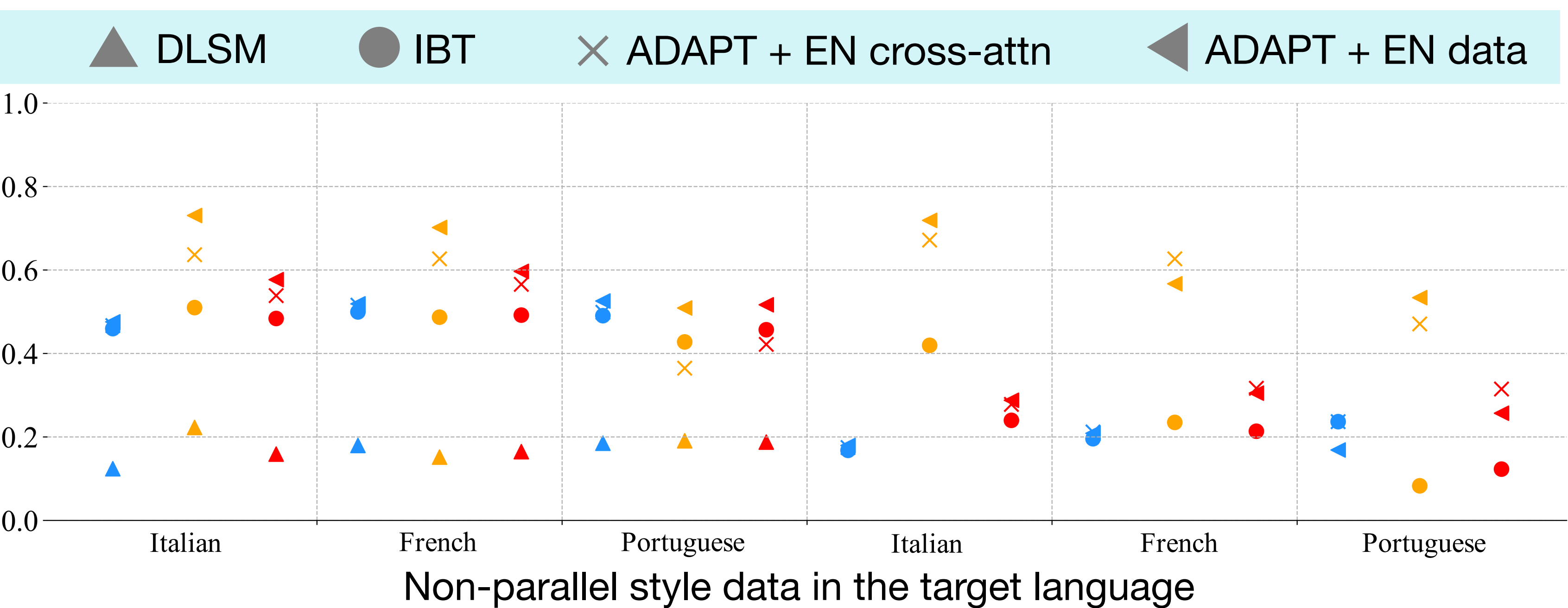
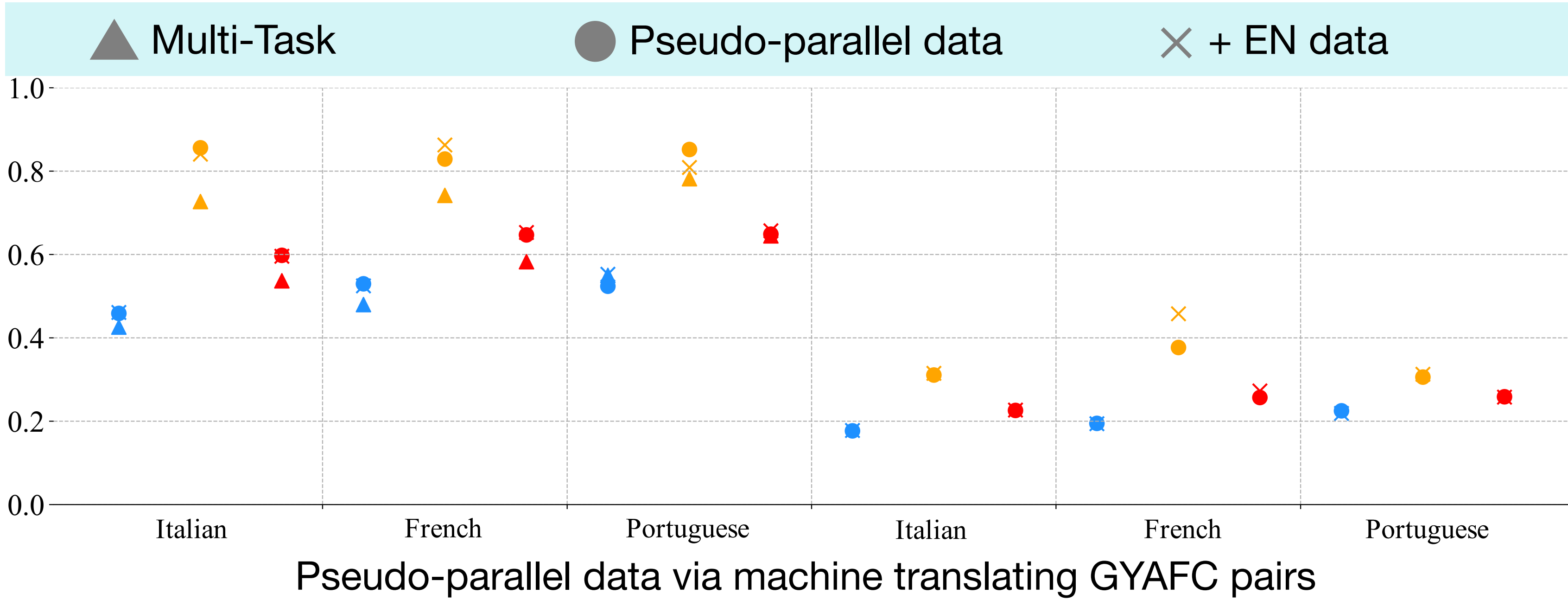
Style Strength: **ACC**  
(Style accuracy evaluated by Style Classifier)



Overall Score: **HM**  
(The harmonic mean of style accuracy and BLEU score)

#### Informal-to-Formal

#### Formal-to-Informal



Paper



Code



@HuiyuanLai @\_atoral @MalvinaNissim