

# Pseudo Zero Pronoun Resolution Improves Zero Anaphora Resolution

Ryuto Konno<sup>1</sup>, Shun Kiyono<sup>2,3</sup>, Yuichiroh Matsubayashi<sup>3,2</sup>, Hiroki Ouchi<sup>4,2</sup>, Kentaro Inui<sup>3,2</sup> <sup>1</sup>Recruit Co., Ltd. <sup>2</sup>RIKEN <sup>3</sup>Tohoku University <sup>4</sup>Nara Institute of Science and Technology

#### Summary

- We tackled zero anaphora resolution (ZAR) in Japanese
- We proposed a new pretraining task and a new finetuning method
- The two proposals boost the SoTA performance of ZAR
- Our analysis provides new insights on the remaining challenges

#### Pretraining

Previous: Cloze Task

university [MASK]

No supervision on anaphoric relations

teachers [MASK]

The university has surveyed teachers' means ... most teachers use ...

Proposed: Pseudo Zero Pronoun Resolution (PZERO)

**Explicit supervision** on anaphoric relations

[MASK]

The university has surveyed **teachers'** means ... most **teachers** use ...

## Task: Zero Anaphora Resolution (ZAR)

In Japanese, Chinese, Korean, Italian, Spanish, ...

The criminal's weapon was found in the victim's room.

Recognizing the antecedents of zero pronoun

It seems that  $\phi$  used a hammer.

**Pronouns are omitted** → zero pronoun

Previous: Argument Selection with Label Probability (AS)

NOM The criminal's weapon ... that  $\phi$  used a hammer.

 ★ NOM = nominative ACC = accusative

- Pretrain-finetune discrepancy [Yang+'19]
  - [MASK] is not used
  - Last Layer is replaced/added
- Proposed: Argument Selection as PZERO (AS-PZERO)

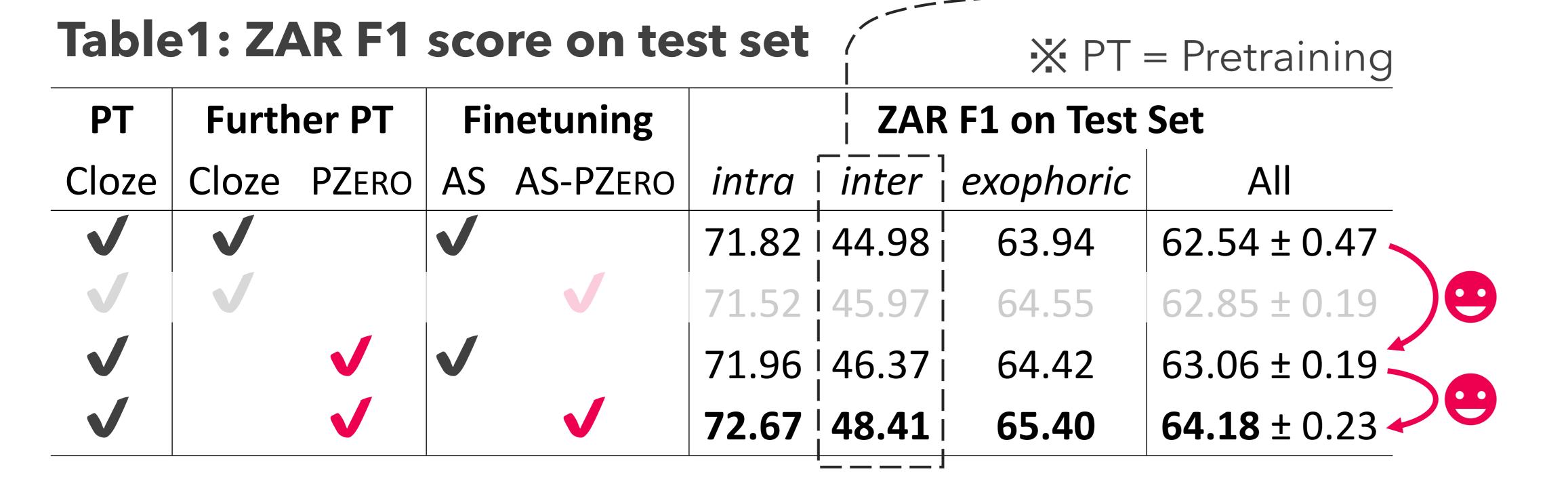
The **criminal**'s weapon ... that  $\phi$  used a hammer. [MASK] NOM used

**Smoother adaptation** 

Active / Passive / Causative

- Both training predict antecedents
- We can use the same network structure

## Results and analysis



#### **→** Table2: *inter* recall on dev set

\* Remaining challenge

Type of instances

Cloze + AS

PZERO + AS-PZERO

51.96 / 42.6 / 40.3

(I) Number of gold antecedents in input

35.96 / 53.1 Only one / More than one

39.57 / 54.1

(II) Distance of sentences between argument and the target predicate

48.86 / 37.7 / 40.5 One / Two / More than two

(III) Voice of the target predicate

45.02 / 25.2 / 60 **47.42 / 29.1 /** 60