

A Survey on Zero Pronoun Translation

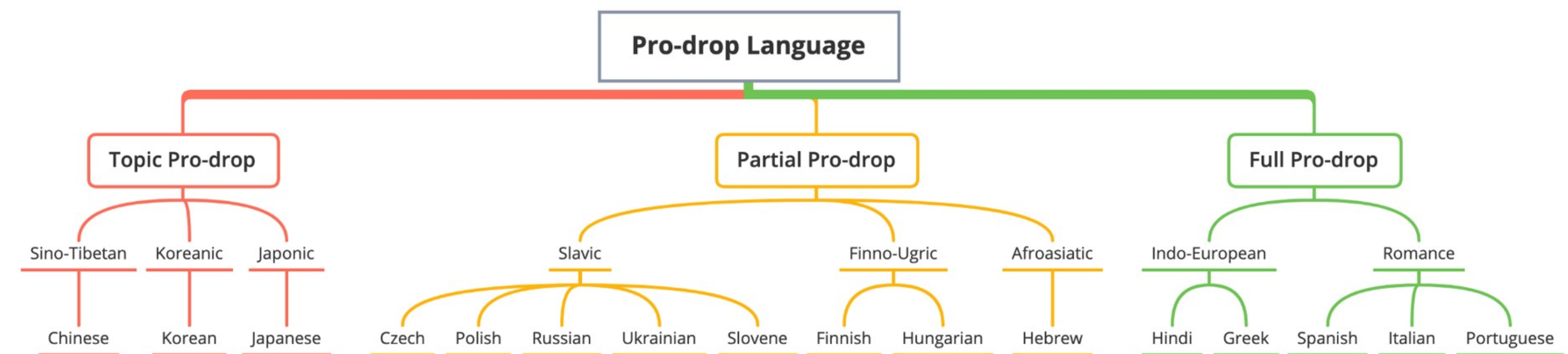
Longyue Wang Siyou Liu Mingzhou Xu Linfeng Song Shuming Shi Zhaopeng Tu

Tencent AI Lab

INTRODUCTION

- **Zero pronoun** (ZP) is a complex phenomenon that appears frequently in **pronoun-dropping languages** (pro-drop) such as Chinese, Hungarian
- Since recovery of such **ZPs** generally fails, this poses difficulties for NLP tasks, especially for **machine translation**
- This leads to severe problems: **incompleteness** & **incorrectness**
- This **survey paper** highlights the major works that have been undertaken in **zero pronoun translation** (ZPT) from 4 perspectives

- **Evolution**
- **Dataset**
- **Approach**
- **Evaluation**



KO	A: 너 <u>이것</u> 필요하니? B: 필요해.	HU	A: látjátok a <u>macskát</u> ? B: látjuk.	HI	A: तुमने <u>नाद्या को</u> खाना ? B: हाँ दे दिया। .
EN	A: Do you need <u>this</u> ? B: (I) need (it).	EN	A: Do (you) see the <u>cat</u> ? B: (We) see (it).	EN	A: Did you give the food to <u>Nadya</u> ? B: Yes, (I) gave (her) (food).
OT	A: Do you need this? B: I <u>need</u> .	OT	A: Do you see the cat? B: We <u>see</u> .	OT	A: Did you <u>eat</u> Nadya? B: Yes <u>given</u> .

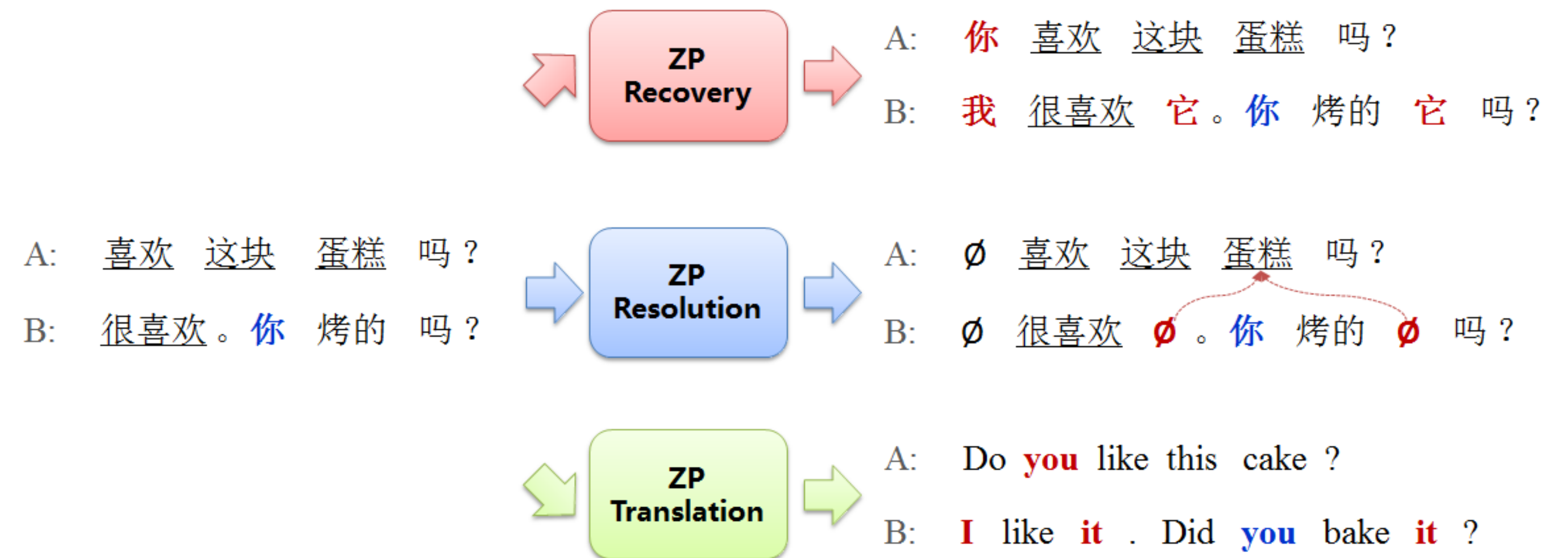
EVOLUTION

Overview

- **Zero Pronoun Resolution.** The task contains **three steps**: ZP detection, anaphoricity determination and reference linking.
- **Zero Pronoun Recovery.** Given a source sentence, this aims to insert omitted pronouns in **proper positions** without changing the original meaning.
- **Zero Pronoun Translation.** When pronouns are omitted in a source sentence, ZPT aims to generate ZPs in **target translation**.

Discussions and Findings

- **From Intermediate to End.**
- **From Separate To Unified.**



DATASETS

Overview

- **Modeling ZPs** has so far not been extensively explored in prior research, largely due to the **lack of publicly** available data sets.
- **Existing works** mostly focused on **human-annotated, small-scale and single-domain** corpora such as OntoNotes.

Discussions and Findings

- **Language Bias.**
- **Domain Bias.**
- **Become An Independent Research Problem.**
- **Coping with Data Scarcity.**

Dataset	Lang.	Anno.	Domain	Size	Task		
					Reso.	Reco.	Trans.
OntoNotes (Pradhan et al., 2012)	ZH	Human	Mixed Sources	42.6K	✓	✗	✗
OntoNotes (Aloraini and Poesio, 2020)	AR	Human	News	9.4K	✓	✗	✗
CTB (Yang and Xue, 2010)	ZH	Human	News	10.6K	✗	✓	✗
KTB (Chung and Gildea, 2010)	KO	Human	News	5.0K	✗	✓	✗
BaiduKnows (Zhang et al., 2019)	ZH	Human	Baidu Knows	5.0K	✗	✓	✗
TVsub (Wang et al., 2018a)	ZH, EN	Auto	Movie Subtitles	2.2M	✗	✗	✓
ZAC (Pereira, 2009)	PT	Human	Mixed Sources	0.6K	✓	✗	✗
Nagoya (Zhan and Nakaiwa, 2015)	JA	Auto	Scientific Paper	1.2K	✓	✗	✗
SKKU (Park et al., 2015)	KO	Human	Dialogue	1.1K	✓	✗	✗
UPENN (Prasad, 2000)	HI	Human	News	2.2K	✓	✗	✗
LATL (Russo et al., 2012)	IT, ES	Human	Europarl	2.0K	✓	✗	✓
UCFV (Bacolini, 2017)	HE	Human	Dialogue	0.1K	✓	✗	✗

APPROACHES

Overview

- **Pipeline**, where input sentences are labeled with ZPs using an **external ZP recovery system** and then **fed into** a standard MT model
- **Implicit**, where ZP phenomenon is **implicitly resolved** by modelling **document-level contexts**
- **End- to-End**, where ZP prediction and translation are **jointly learned** in an **end-to-end manner**

Discussions and Findings

- **Existing Methods Can Help ZPT But Not Enough.**
- **Pipeline Methods Are Easier to Integrate with NMT.**
- **Data-Level Methods Do Not Change Model Architecture.**
- **Multitask and Multi-Lingual Learning.**

Model	TVsub		BaiduKnows		Webnovel	
	BLEU	APT	BLEU	APT	BLEU	APT
Baseline (Vaswani et al., 2017)	29.4	47.4	12.7	25.4	11.7	30.9
Pipeline (Song et al., 2020)	29.8	49.5	13.2	56.4	11.6	32.0
Implicit (Ma et al., 2020)	29.8	53.5	13.9	26.3	12.2	35.3
End-to-End (Wang et al., 2018a)	30.0	52.3	12.3	30.4	12.0	33.4
ORACLE	32.8	86.9	14.7	88.8	12.8	85.1

EVALUATION METHODS

Overview

- Accuracy of ZP Recovery
- General Translation Quality
- Pronoun-Aware Translation Quality

Metric	T.S.	B.K.	I.H.	Ave.
BLEU	0.09	0.76	0.57	0.47
TER	0.41	0.01	0.26	0.23
METEOR	0.23	0.74	0.28	0.42
COMET	0.59	0.15	0.37	0.37
APT	0.68	0.76	0.58	0.67

Discussions and Findings

- General-Purpose Evaluation Are Not Applicable to ZPT.
- Human Evaluations Are Required as A Com- plement.
- The Risk of Gender Bias.

1. Out-of-Domain	INP.	[他的] _p 主要 研究 领域 为 ...
	NMT	The main research areas are ...
	ZPR	我 主要 研究 领域 为 ...
	ZPR+	My main research areas are ...
2. Error Propagation	INP.	如果 [你们] _s 见到 她 ...
	NMT	If you see her ...
	ZPR	如果 我 见到 她 ...
	ZPR+	If I see her ...
3. Multiple ZPs	INP.	[他] _s 好久没 ... [他] _s 怪 想念 的。
	NMT	for a long time did not ... strange miss.
	ZPR	我 好久没 ... 我 怪 想念 的。
	ZPR+	I haven't ... for a long time, I miss.