多言語対話における応答先と応答内容の選択

佐藤 元紀[1], 大内 啓樹[1,2], 坪井 祐太[3]

1 : 奈良先端科学技術大学院大学 (NAIST) 松本研究室

2 : 理化学研究所 革新知能統合研究センターAIP

3 : 株式会社Preferred Networks

2018/3/15 NLP2018

発表の概要

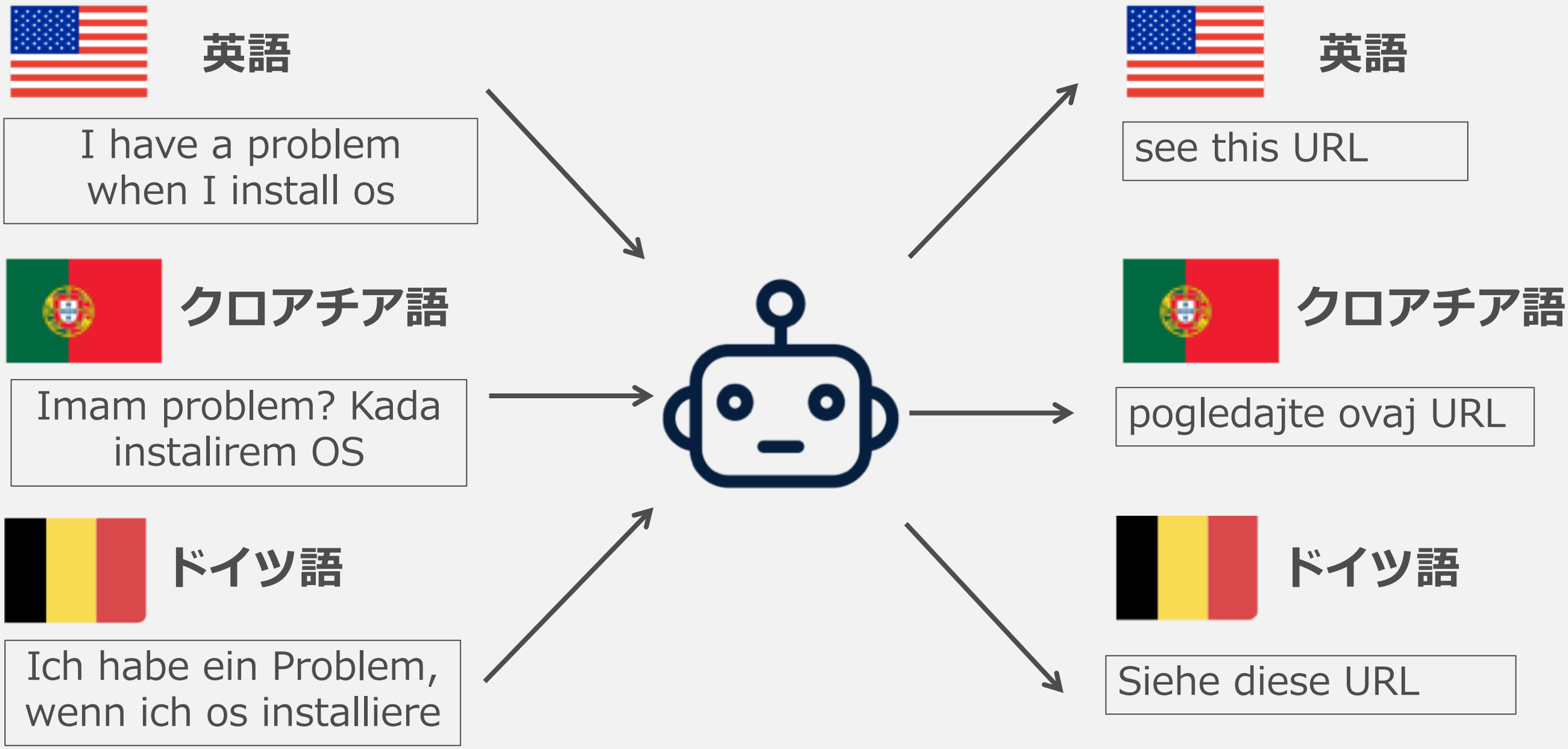
1. 本研究が扱う問題
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. タスク定式化
3. 提案手法
4. 実験
5. まとめと今後の課題

発表の概要

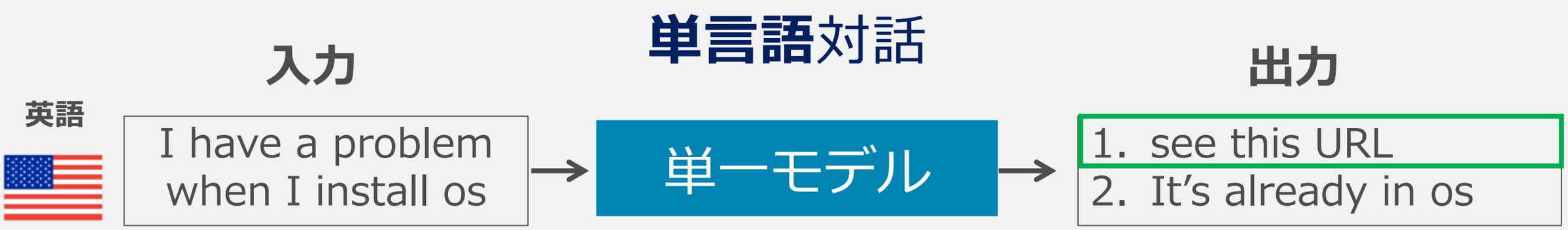
1. **本研究が扱う問題**
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. タスク定式化
3. 提案手法
4. 実験
5. まとめと今後の課題

本研究が扱う問題

- **複数の言語**を扱う対話　：　多言語対話タスクを提案



単言語対話



- 各言語について, **それぞれモデルを学習する**
- **大量の学習データ**を用意できる言語(**英語, 中国語**)が主に取り扱われる

問題点

- そもそも**マイナー言語** (少量の学習データ)は扱われない
- **1つの言語のみ**しか扱えない

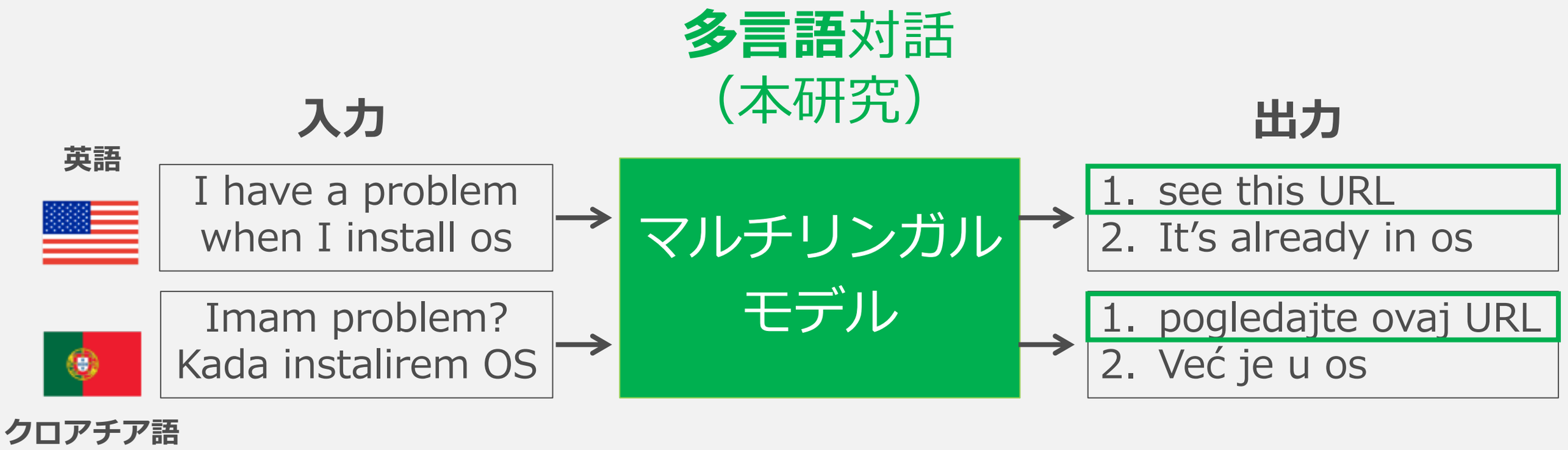
多言語対話

- **複数の言語**を扱う対話
- **少量の学習データ**の言語（クロアチア語 etc）も扱う

問題点

- **High-resource (高資源)から Low-resource(低資源)の知識の転移をどのように行うか**

単言語対話と多言語対話



対話生成のアプローチ

- **生成型** (Generation-based)
  - 適切な応答文を**生成**する (例: RNN)
  - **人手での評価**はコストが高い

- **検索型** (Retrieval-based)
  - リポジトリの**応答候補**から適切な**応答を選択**
  - **Accuracy**で評価するため**簡単に評価できる**

本研究の立ち位置

	単言語	多言語
生成型アプローチ	[Shang et al.,2015] [Vinyals and Le,2015]	-
検索型アプローチ	[Ji et al.,2014] [Lowe et al., 2015] [Ouchi and Tsuboi, 2016]	[本研究]

- **[Shang et al.,2015]** "Neural responding machine for short-text conversation", IJCNLP 2015.
- **[Vinyals and Le, 2015]** "A neural conver-sational model", arxiv 2015.
- **[Ji et al.,2014]** "An information retrieval approach to short text conversation", arxiv 2014.
- **[Lowe et al., 2015]** "A large dataset for research in unstructured multi-turn dialogue systems.", *SIGDIAL.*
- **[Ouchi and Tsuboi, 2016]** Addressee and Response Selection for Multi-Party Conversation, EMNLP 2016

本研究の貢献

1. 多言語応答選択**タスクの定式化**

2. 多言語に対応可能な**モデルの提案**

3. 多言語対話**データセットの作成・公開**
   - Ubuntu Dialog Corpus
   - 5言語（英語,イタリア語,クロアチア語, ポルトガル語, ドイツ語）

発表の概要

1. 本研究が扱う問題
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. **タスク定式化**
3. 提案手法
4. 実験
5. まとめと今後の課題

タスク定義

- 検索型アプローチの中でも特に「**ARS**」と呼ばれるタスクを多言語に拡張する

- **[Ouchi and Tsuboi, EMNLP 2016]**で提案されたタスク
  - **応答先・応答内容選択** (Addressee and Response Selection; **ARS**)

- ARSを多言語化した**Multilingual-ARS**を提案する

タスク定式化：ARS

- **[Ouchi and Tsuboi, EMNLP2016]**で提案されたタスク
- 応答先・応答内容選択(Addressee and Response Selection; **ARS**)

発話者　　応答先　　応答文

User	Addressee	Utterance
User 1	-	I have a problem when I install …
SYSTEM	-	did you set initial params ?
User 2	-	Show the error message, and …
User 1	*SYSTEM*	how ?
User 1	User 2	ok just a moment !
SYSTEM	[To Whom?]	[What?]
	1. User 1	1. see this URL : http://xxxx
	2. User 2	2. It ’s already in os

- 対話のコンテキストから,
  - 誰に対する応答か (**応答先**; Addressee)
  - どんな応答文か（**応答内容**; Response）　　を当てるタスク

タスク定式化：ARS

入力

- 発話者 a_{res}
- 発話履歴内の
  ユーザー集合A(C)
- 応答内容候補集合

  $R=\{r_1,\cdots,r_{|R|}\}$

出力

- 応答先a
- 応答内容r

$$\text{入力}：\boldsymbol{x}=(a_{\text{res}},\ \mathcal{C},\ \mathcal{R})$$

$$\text{出力}：\boldsymbol{y}=(a,\boldsymbol{r})$$

発話者	応答先	応答文
User	Addressee	Utterance
User 1	-	I have a problem when I install …
SYSTEM	-	did you set initial params ?
User 2	-	Show the error message, and …
User 1	*SYSTEM*	how ?
User 1	User 2	ok just a moment !
SYSTEM	[To Whom?]	[What?]

1. User 1 — 1. see this URL : http://xxxx
2. User 2 — 2. It ’s already in os

評価

- 3つの**正解率** (**ADR-RES**, **ADR**, **RES**)

1. 応答先と応答内容の両方(**ADR-RE**)
2. 応答先のみ(**ADR**)
3. 応答内容のみ(**RES**)

タスク定式化: Multilingual-ARS

- 対象とする**言語集合K**の各言語kに対して
  学習データが与えられているとする:

学習

$$\mathcal{D}_{\text{train}}^{(k)} = \{ \boldsymbol{x}_i^{(k)}, \boldsymbol{y}_i^{(k)} \}_1^{N^{(k)}}, \ k \in \mathcal{K}$$
$$\mathcal{D}_{\text{train}} = \bigcup_k \mathcal{D}_{\text{train}}^{(k)}$$

評価

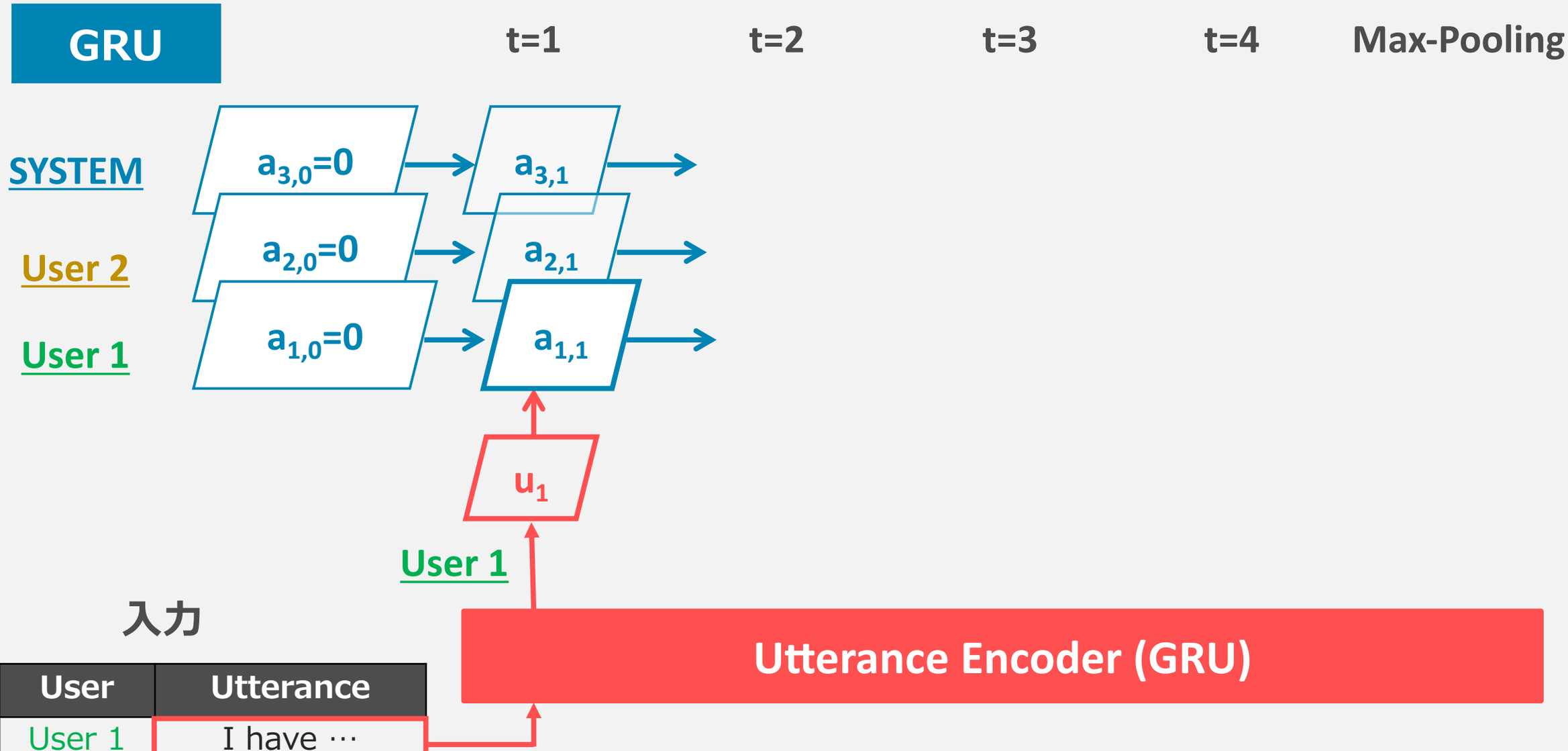
$$\text{ADR-RES} = \frac{\sum_k \text{ADR-RES}^{(k)}}{|\mathcal{K}|}$$

- 対象とする全言語におけるマクロ平均を用いる

発表の概要

1. 本研究が扱う問題
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. タスク定式化
3. **提案手法**
4. 実験
5. まとめと今後の課題

Dynamic Model [Ouchi and Tsuboi]



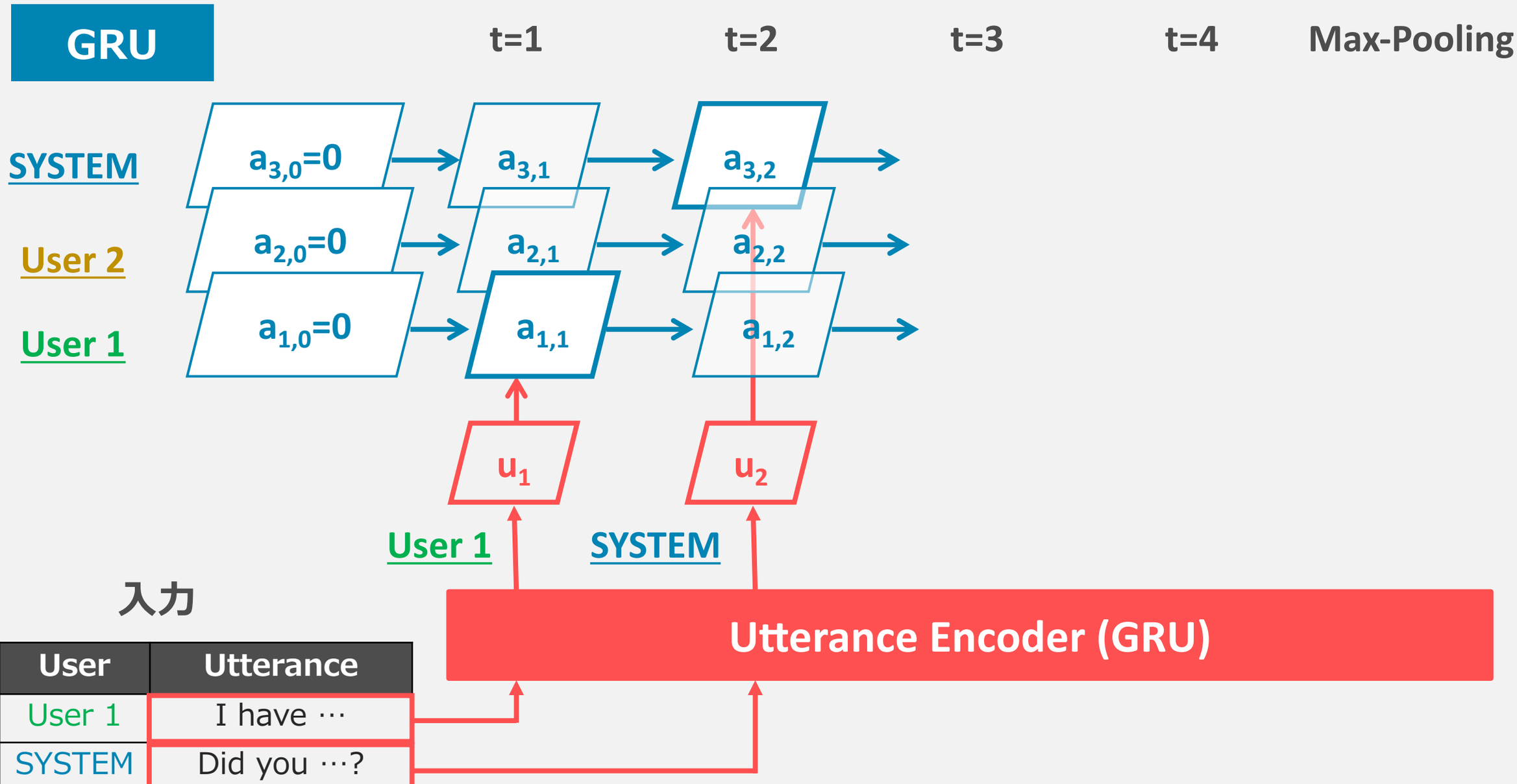
入力

t	User	Utterance
1	User 1	I have …
2	SYSTEM	Did you …?
3	User 2	Please show …
4	User 1	OK. …
5	SYSTEM	[ ]

Response Candidate
1. It seems that…
2. How about… ?

応答文候補

Dynamic Model [Ouchi and Tsuboi]



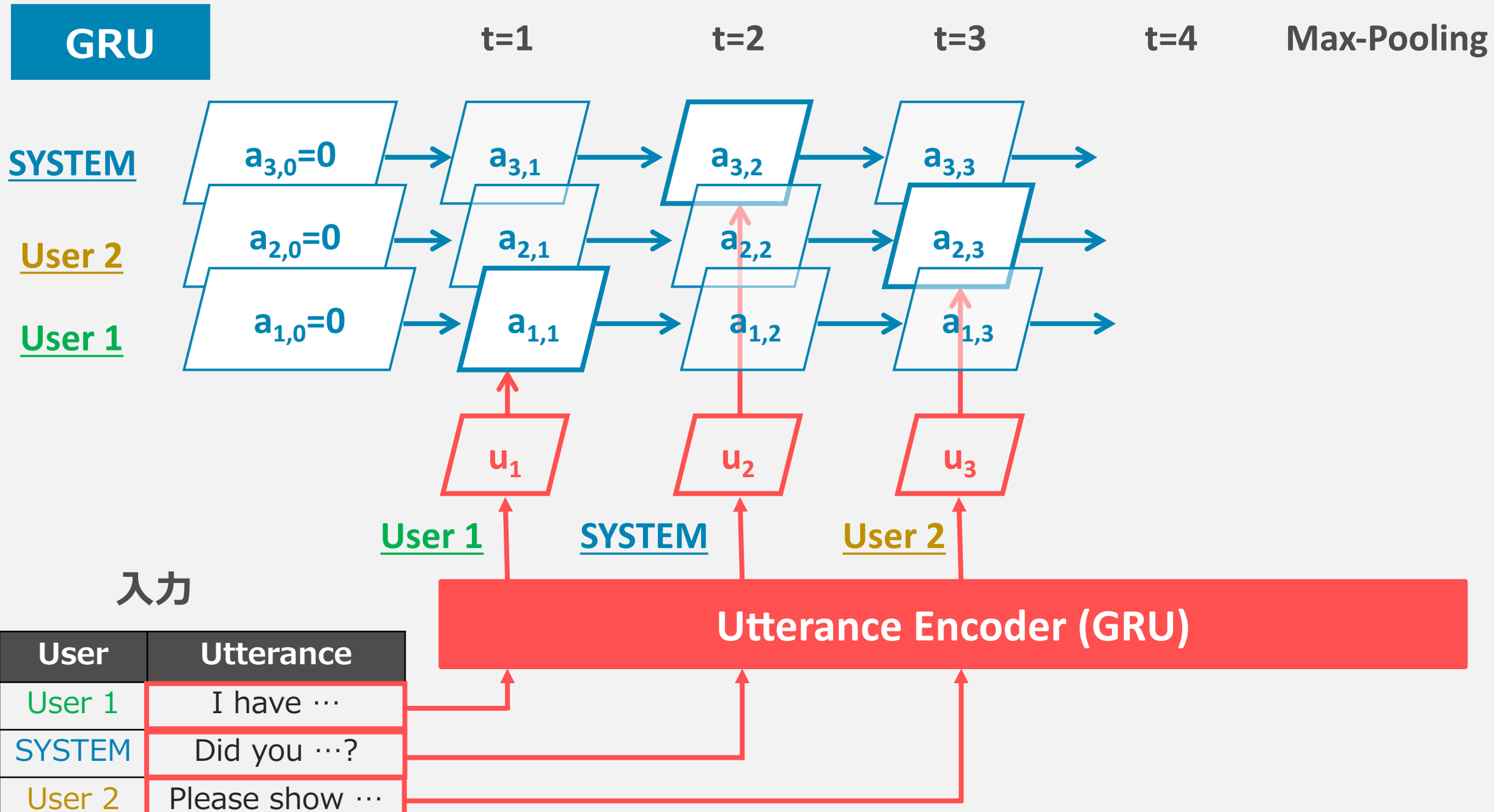
入力

t	User	Utterance
1	User 1	I have …
2	SYSTEM	Did you …?
3	User 2	Please show …
4	User 1	OK. …
5	SYSTEM	[ ]

Response Candidate
1. It seems that…
2. How about… ?

応答文候補

Dynamic Model [Ouchi and Tsuboi]



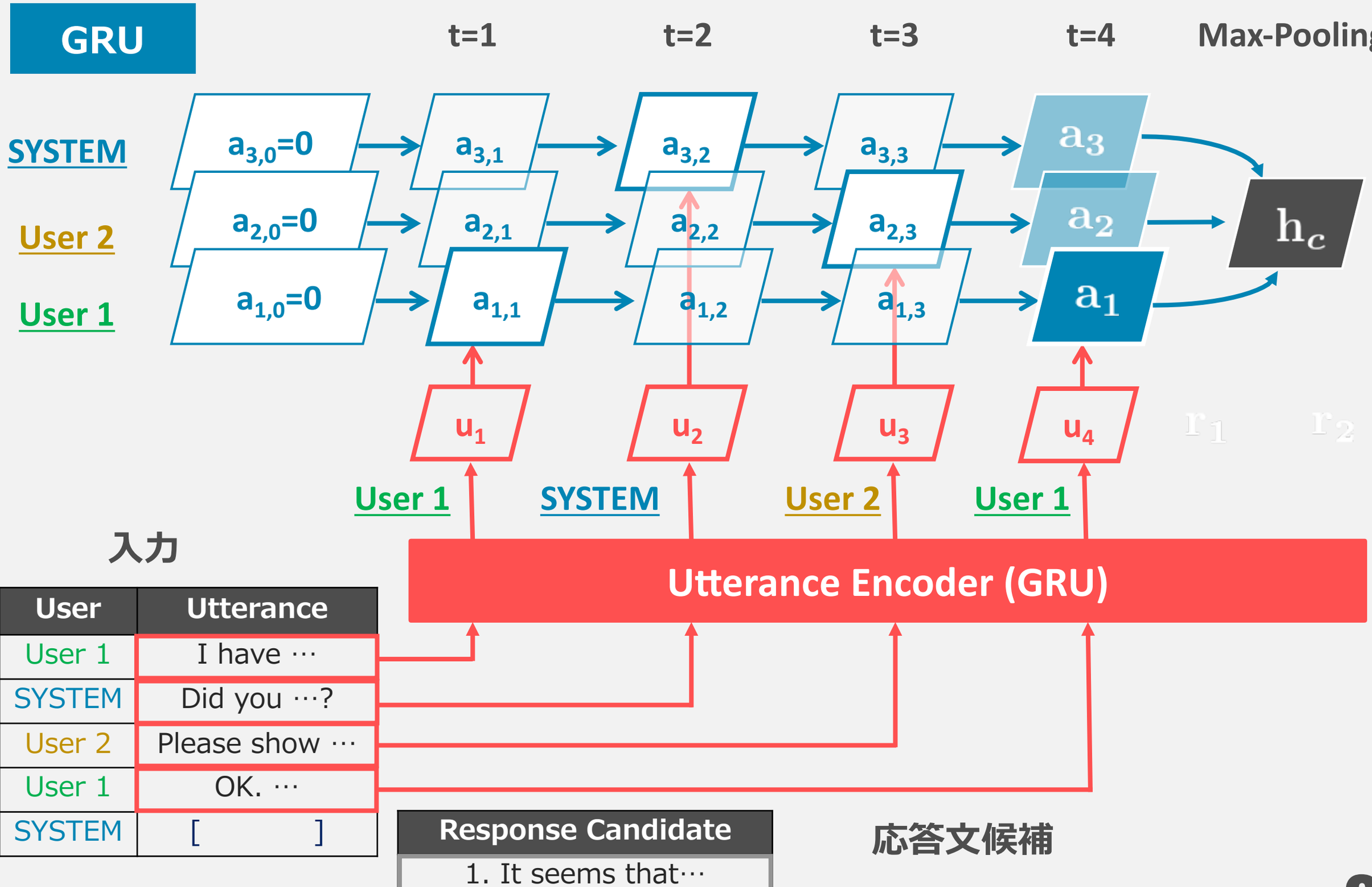
入力

t	User	Utterance
1	User 1	I have …
2	SYSTEM	Did you …?
3	User 2	Please show …
4	User 1	OK. …
5	SYSTEM	[ ]

Response Candidate
1. It seems that…
2. How about… ?

応答文候補

Dynamic Model [Ouchi and Tsuboi]



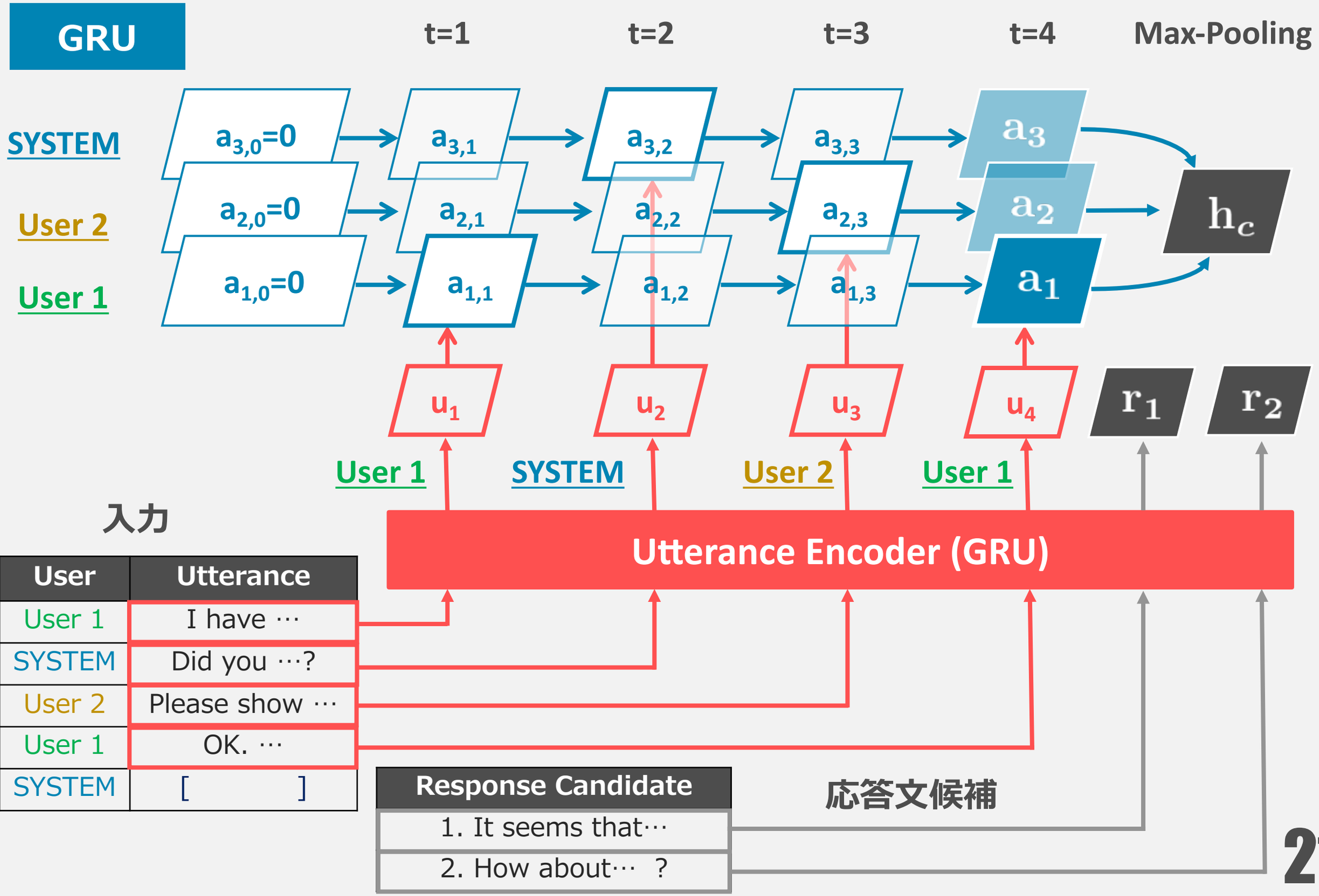
入力

t	User	Utterance
1	User 1	I have …
2	SYSTEM	Did you …?
3	User 2	Please show …
4	User 1	OK. …
5	SYSTEM	[ ]

Response Candidate
1. It seems that…
2. How about… ?

応答文候補

Dynamic Model [Ouchi and Tsuboi]



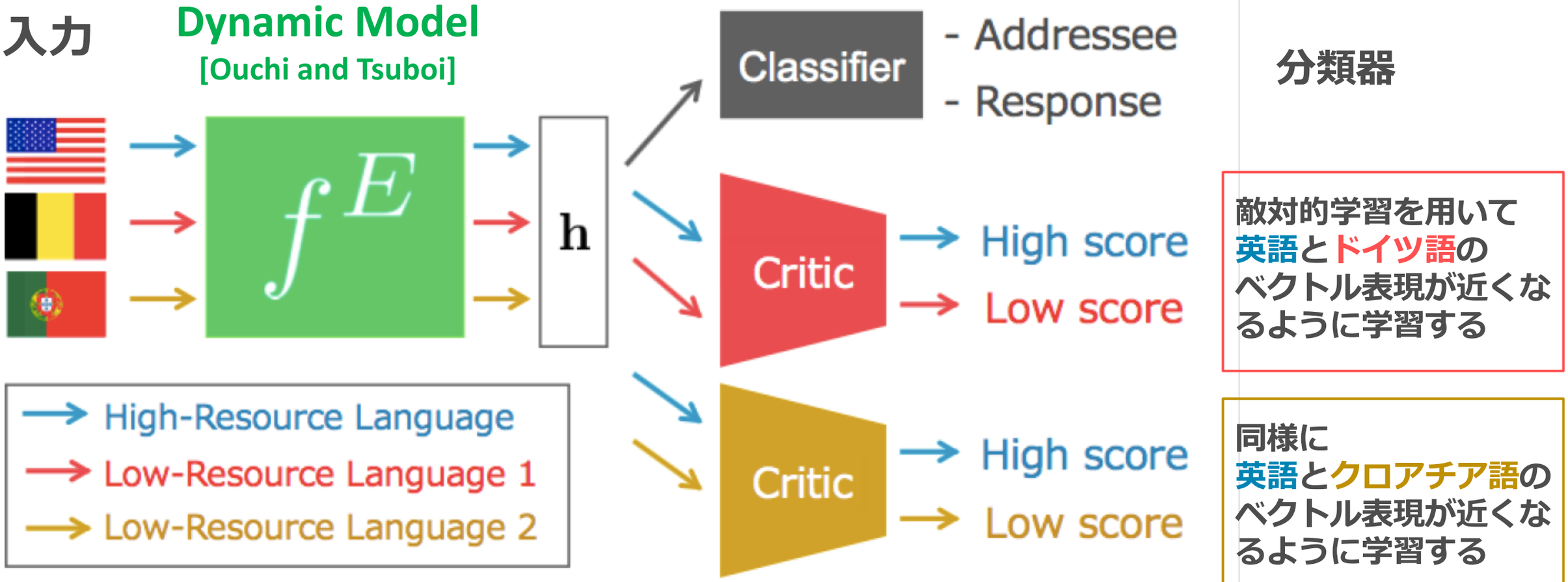
入力

t	User	Utterance
1	User 1	I have …
2	SYSTEM	Did you …?
3	User 2	Please show …
4	User 1	OK. …
5	SYSTEM	[ ]

Response Candidate
1. It seems that…
2. How about… ?

応答文候補

提案手法: WGANを多言語に拡張



高資源(High-Resource)言語：英語
低資源(Low-Resource) 言語：ドイツ語, クロアチア語

Wasserstein-GAN を用いて, **言語間**の特徴量分布を**近付けさせる.**

多言語に対応するために**複数のCritic**に拡張した.

発表の概要

1. 本研究が扱う問題
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. タスク定式化
3. 提案手法
4. **実験**
5. まとめと今後の課題

データセット・実験設定

データセット：

Language	Dataset Train	Dev	Test
English (en)	665.6 k	45.1 k	51.9 k
Italian (it)	38,511	2,561	3,873
Croatian (hr)	11,387	512	1,145
German (de)	5,500	354	569
Portuguese (pt)	5,951	285	975

- **5つの言語**についてデータセットを作成.
- **Ubuntu Dialog Corpus**からコーパスを収集した.
- **英語**のデータは**大量**に手に入るが,
  その他の言語は**少量の学習データ**のみ.

実験設定：

Two-Language Adaptation

２つの言語ペアでモデルを学習し,
マクロ平均で評価する

- （英語, イタリア語）
- （英語, クロアチア語）
- （英語, ドイツ語）
- （英語, ポルトガル語）

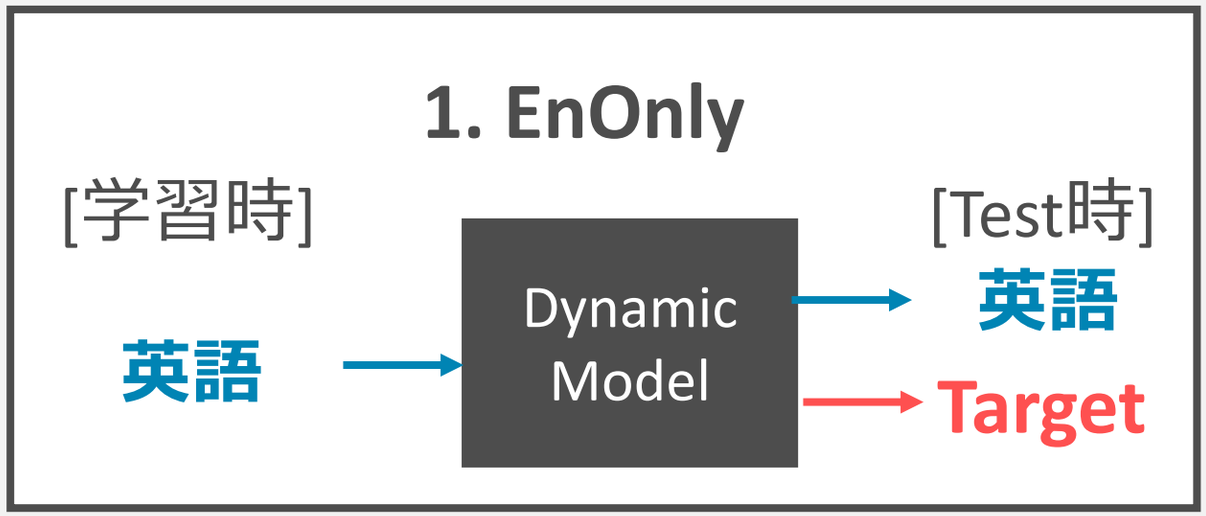
Five-Language Adaptation

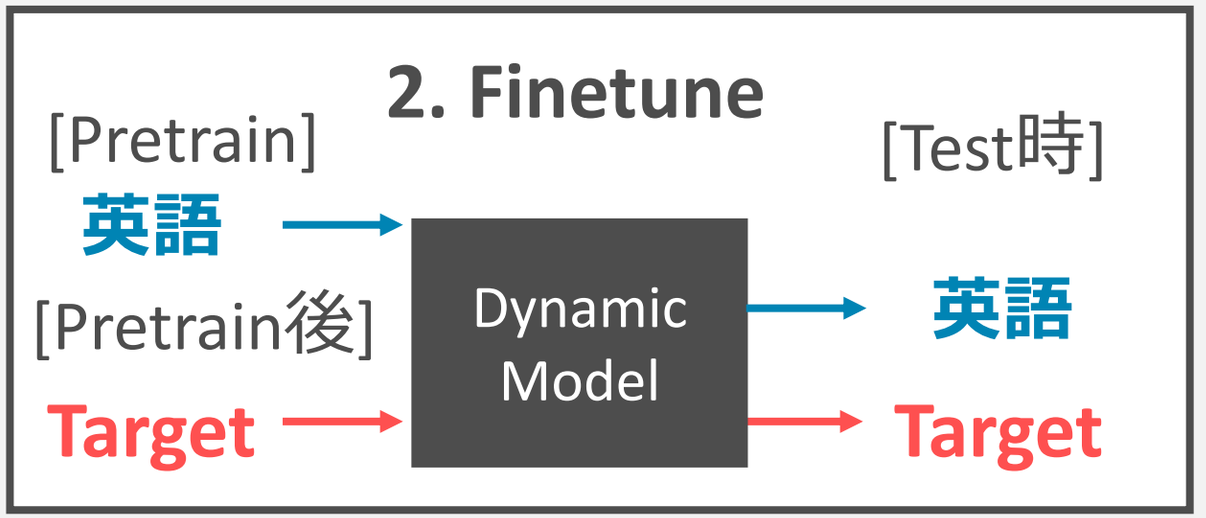
５つの言語ペアでモデルを学習し,
マクロ平均で評価する

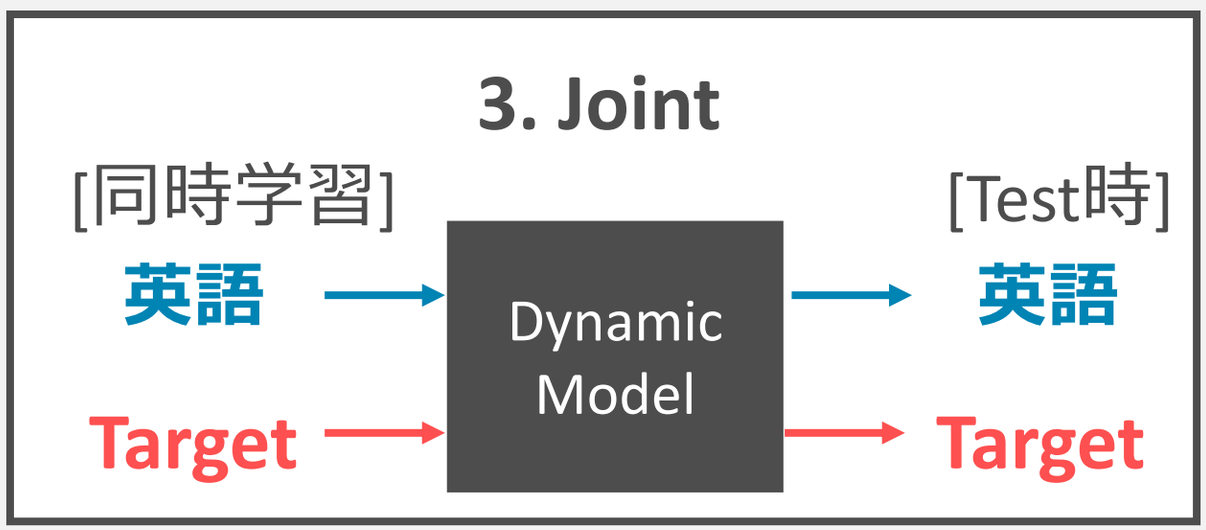
- （英語, イタリア語, クロアチア語,
  ドイツ語, ポルトガル語）

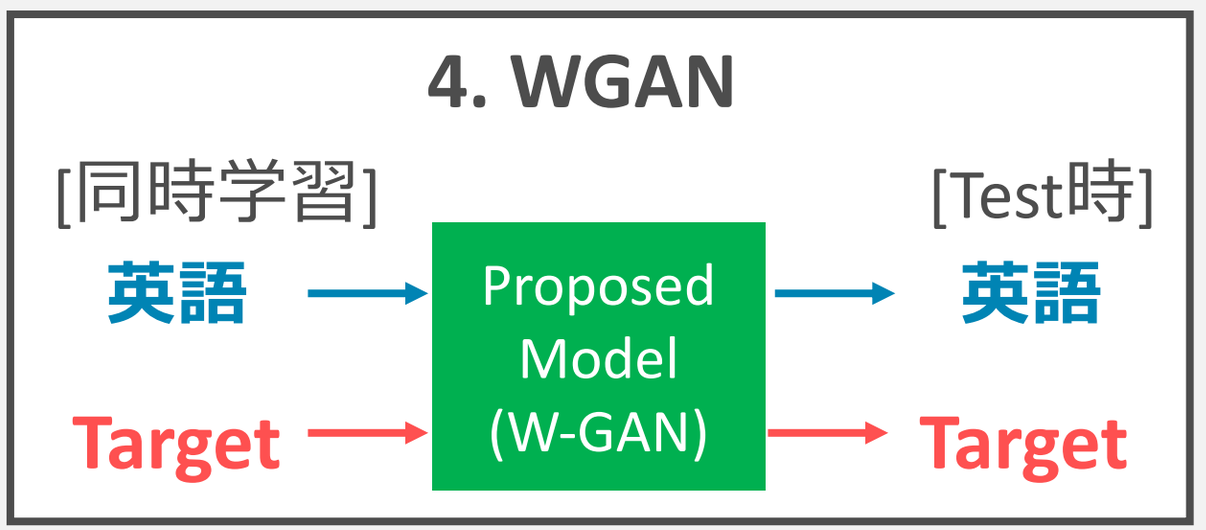
比較手法

1. **EnOnly** ：**Source(高資源言語)**で訓練したモデル（再学習なし）
2. **Finetune**：**Target(低資源言語)**データのみでFinetune(再学習あり)
3. **All** ：**Source** +**Target**で同時学習
4. **WGAN** ：**Source** +**Target**で敵対学習（提案手法）









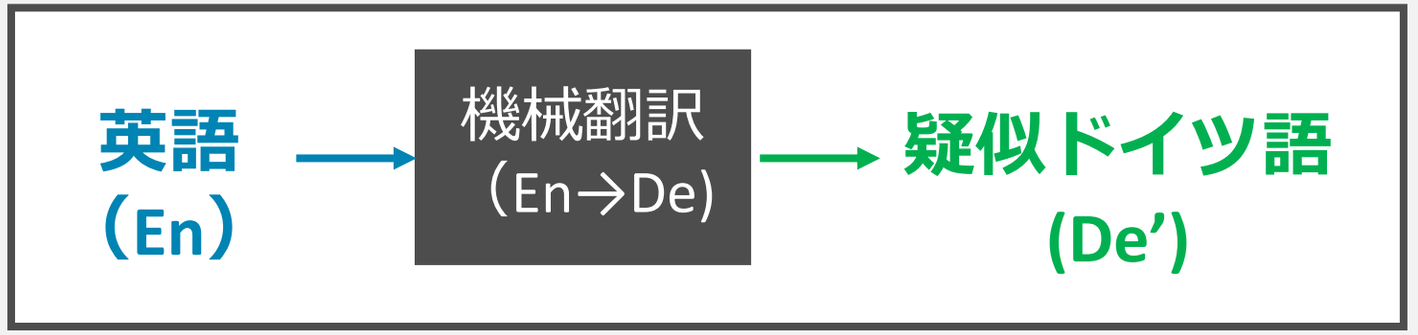
※全言語を**共通ベクトル空間**に埋め込んだ**訓練済みMultiCCA**を
入力ベクトルとして用いた（**学習時から固定**）

実験結果

Setting	Method	$\|\mathcal{R}\| = 2$ ADR–RES	ADR	RES	$\|\mathcal{R}\| = 10$ ADR–RES	ADR	RES
Two Language Adaptation	CHANCE	2.30	4.59	50.00	0.46	4.59	10.00
	TF-IDF	38.32	60.29	64.25	13.99	60.29	23.84
	ENONLY	46.77	67.62	67.90	19.88	65.86	27.83
	FINETUNE	50.98	68.79	72.60	24.30	68.89	32.96
	JOINT	53.37	69.75	74.94	26.60	69.75	35.59
	WGAN	**54.14**	**70.07**	**75.63**	**27.23**	**69.76**	**36.11**
Five Language Adaptation	TF-IDF	39.04	63.10	62.06	13.12	63.10	20.64
	ENONLY	41.55	66.48	61.75	13.05	63.77	19.38
	JOINT	50.69	69.00	72.18	22.80	69.18	31.11
	WGAN	**52.11**	**69.74**	**73.34**	**23.39**	**69.35**	**31.88**

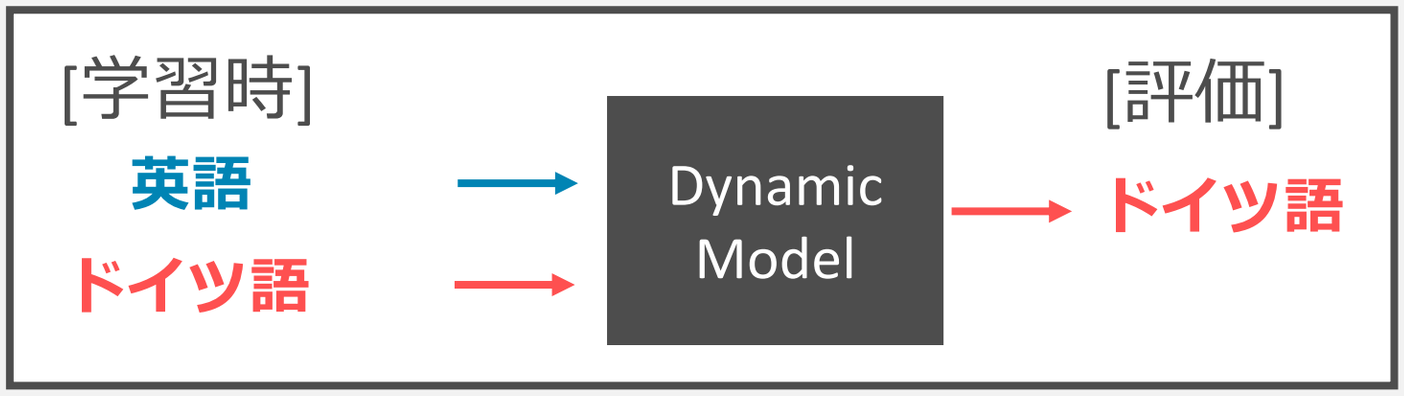
- ２言語/５言語において提案手法が最も高いスコアを得ることができた
- W-GANを用いることで、言語間で共通のベクトルが学習できた

機械翻訳によるData Augmentation

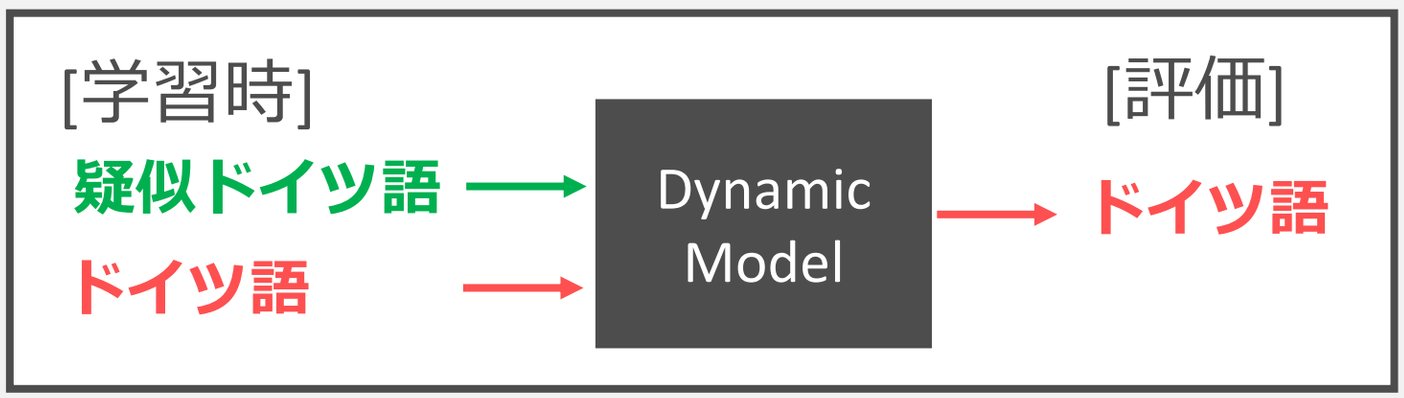


- OpenNMTの訓練済みニューラル機械翻訳のモデルを用いた
- 「英語→ドイツ語」で擬似ドイツ語データを作る.

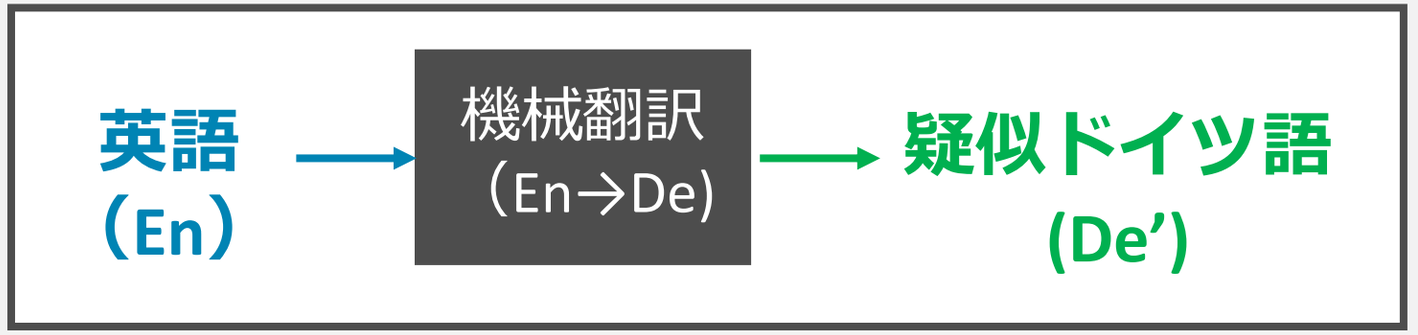
Data Augmentationなし：

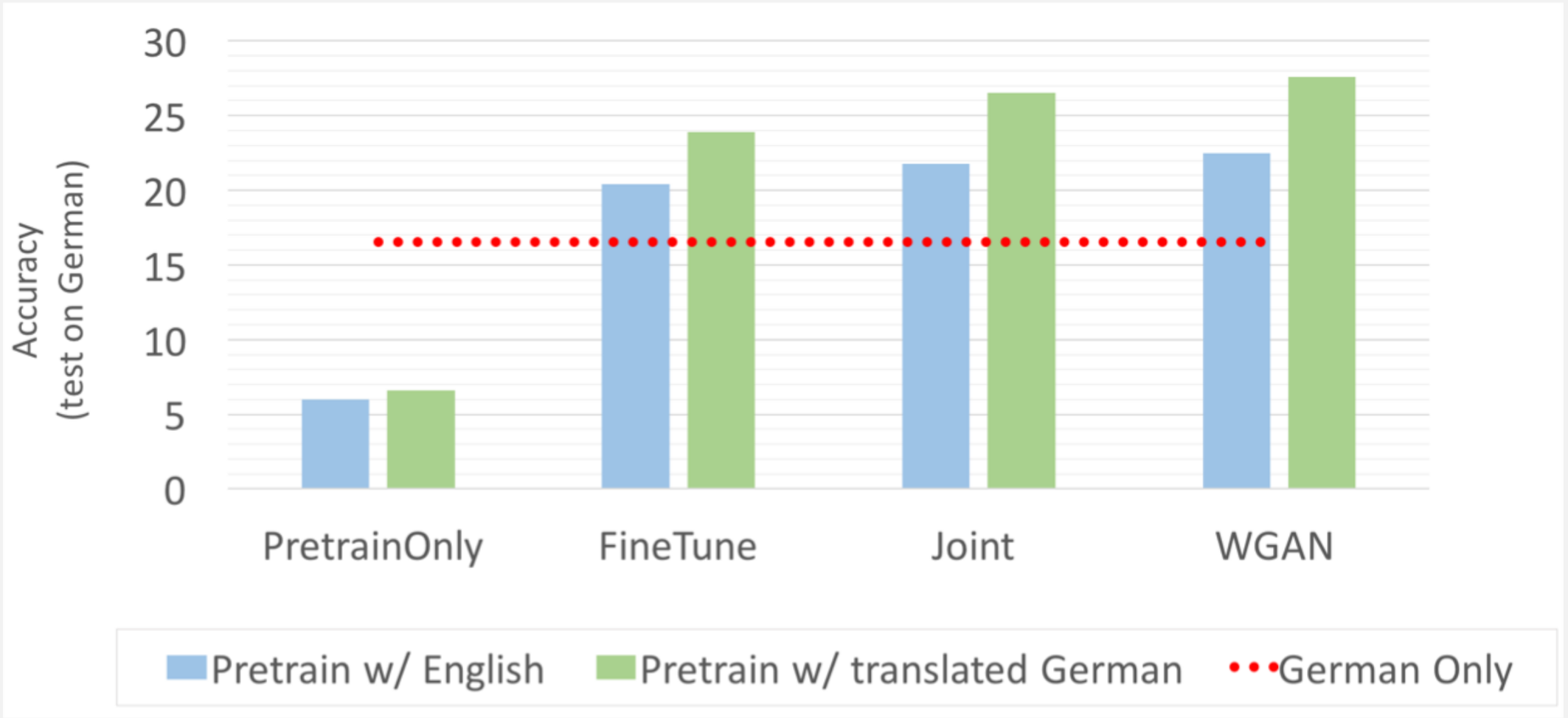


Data Augmentationあり：

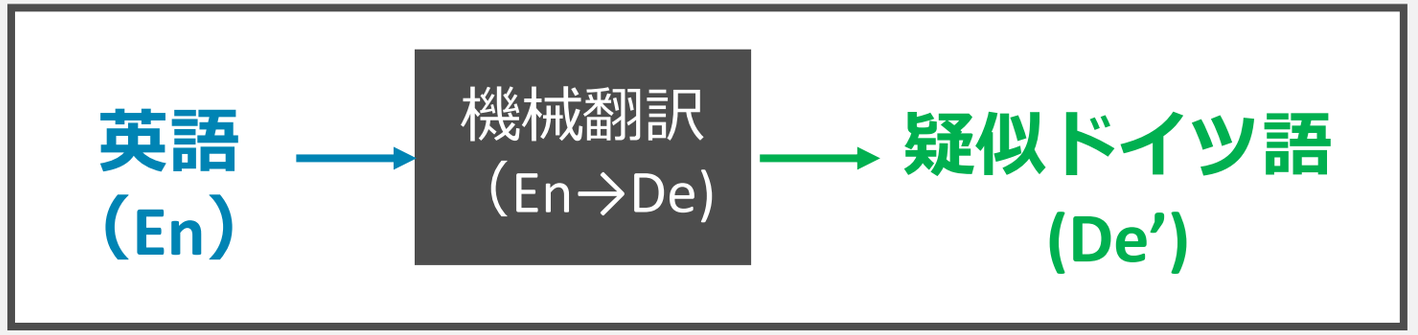


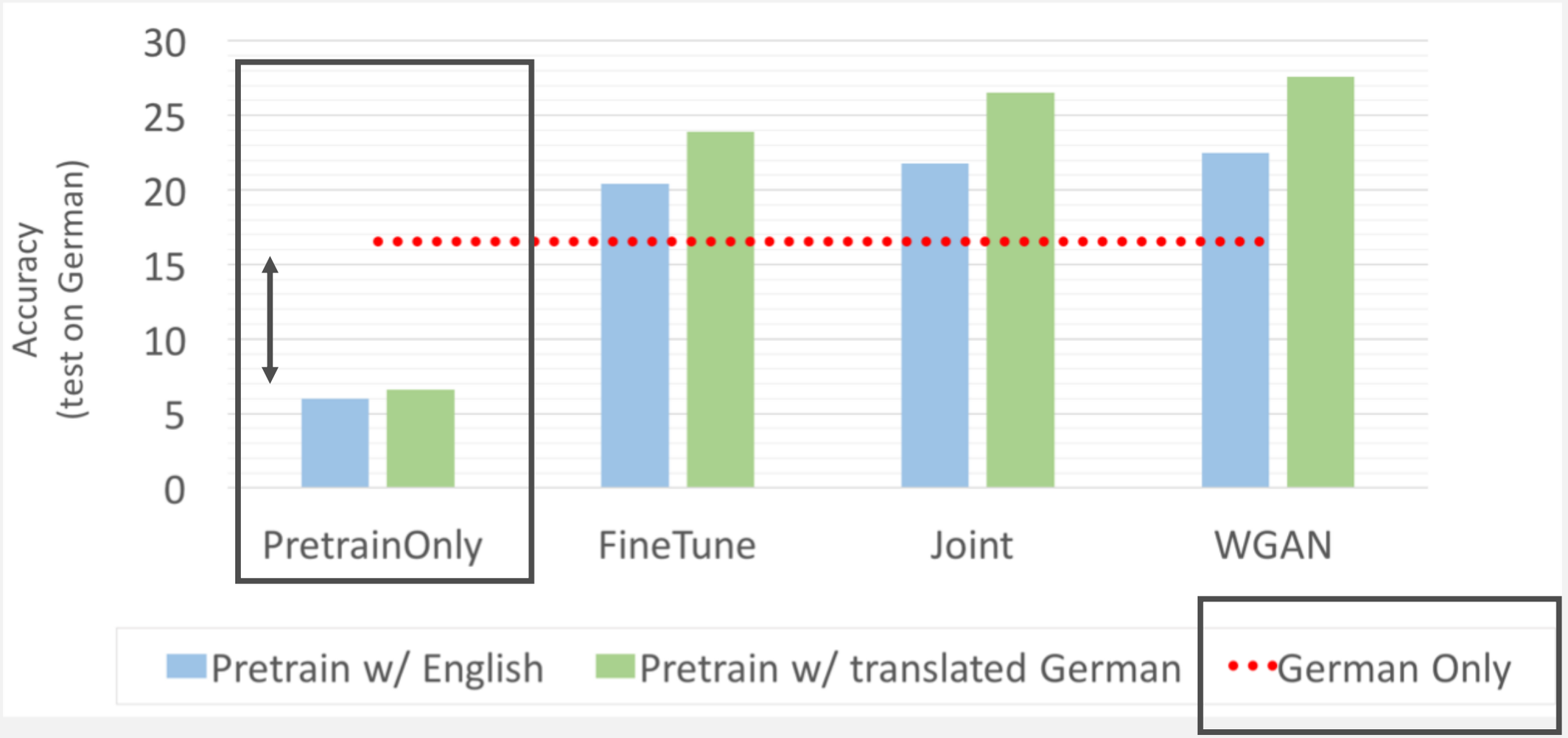
機械翻訳によるData Augmentation





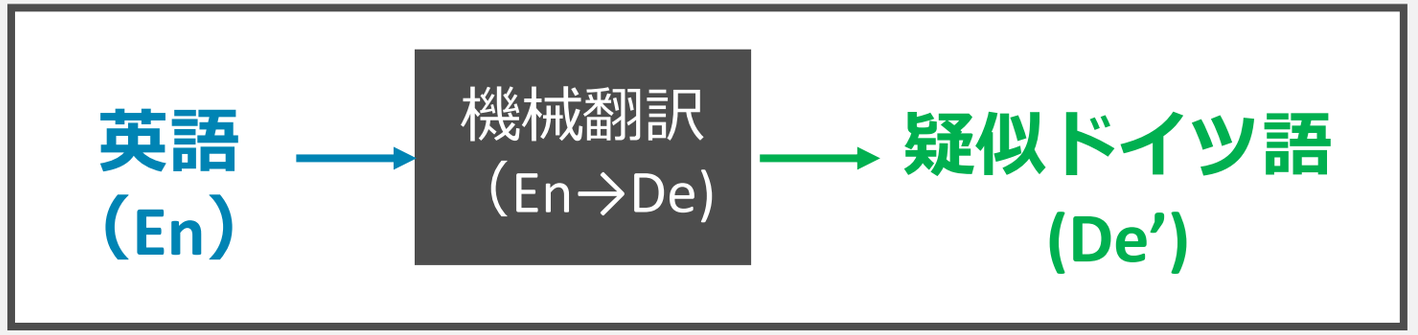
機械翻訳によるData Augmentation

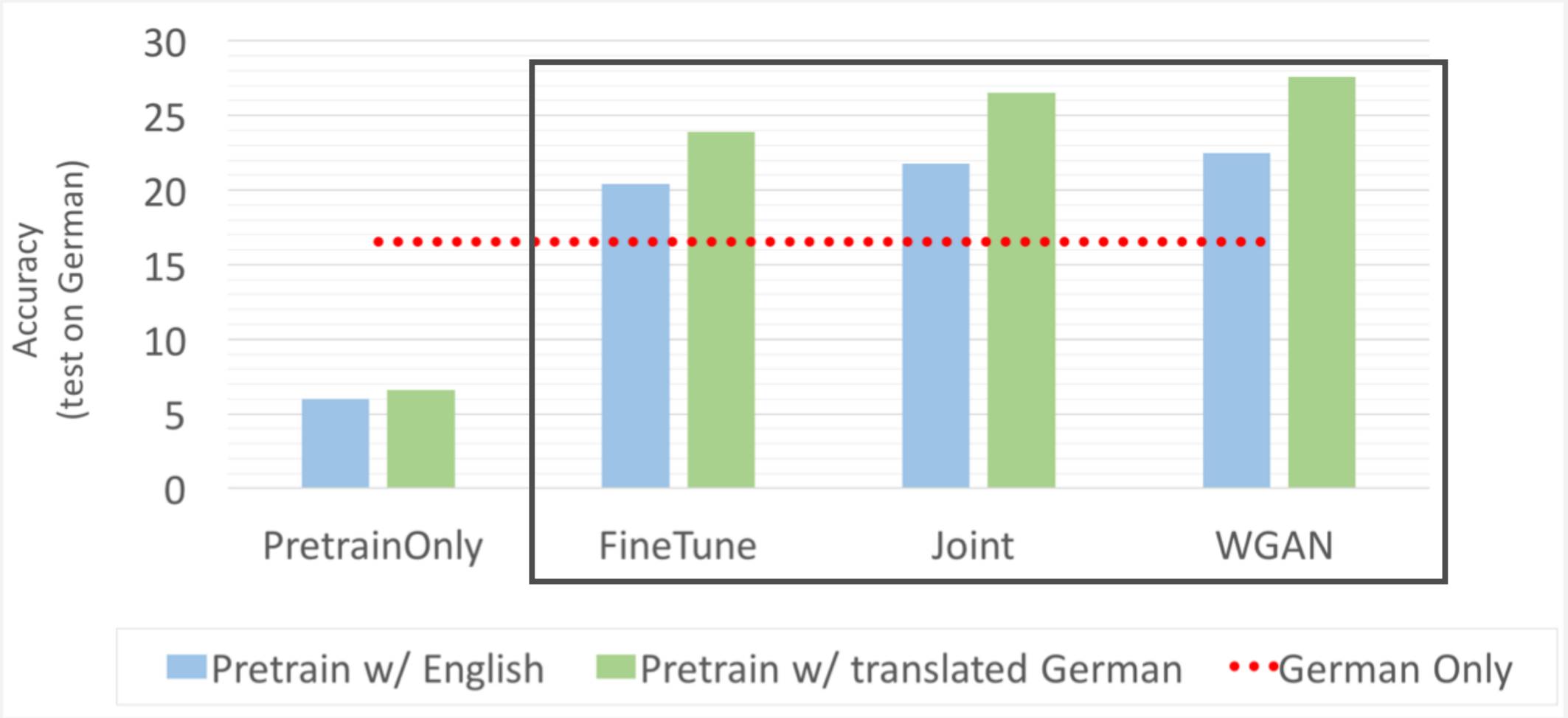




- ドイツ語を学習に使わなかった場合，ドイツ語のみを学習に使った方が性能が高い

機械翻訳によるData Augmentation





- 学習に英語ではなく疑似ドイツ語を用いた方が性能が上がる
  （NMTのモデルがある場合）

発表の概要

1. 本研究が扱う問題
   - 単言語/多言語
   - 応答生成 (Generation-/ Retrieval- Based)
2. タスク定式化
3. 提案手法
4. 実験
5. **まとめと今後の課題**

まとめ・今後の課題

まとめ：

- 多言語応答選択**タスクの定式化**
- 多言語に対応可能な**モデルの提案**
- 多言語対話**データセットの作成・公開**
  - http://sato-motoki.com/projects/
- W-GANを用いた提案手法の有効性を示した.
- 機械翻訳によるData Augmentationの効果を示した.

今後の課題：

- ドイツ語以外でのData Augmentationの効果
- より学習データが少ない言語での実験

タスク定式化：ARS

- 応答先・応答内容選択(Addressee and Response Selection; ARS)

$$\text{入力}：\boldsymbol{x} = (a_{\text{res}},\ \mathcal{C},\ \mathcal{R})$$

$$\text{出力}：\boldsymbol{y} = (a, \boldsymbol{r})$$

- **入力**:
  - 応答したユーザーa_{res}
  - 発話履歴に含まれるユーザー集合A(C)
  - 応答内容候補集合$R=\{r_1,\cdots,r_{|R|}\}$

- **出力**
  - 応答先a
  - 応答内容r

- **評価**
  - 3つの正解率(ADR-RES, ADR, RES)
  1. 応答先と応答内容の両方(ADR-RE)
  2. 応答先のみ(ADR)
  3. 応答内容のみ(RES)　　の正解率(Accuracy)で評価.