Toshiba MT System Description for the WAT2015 Workshop

Satoshi SONOH
Satoshi KINOSHITA

Knowledge Media Laboratory,
Corporate Research & Development Center,
Toshiba Corporation.
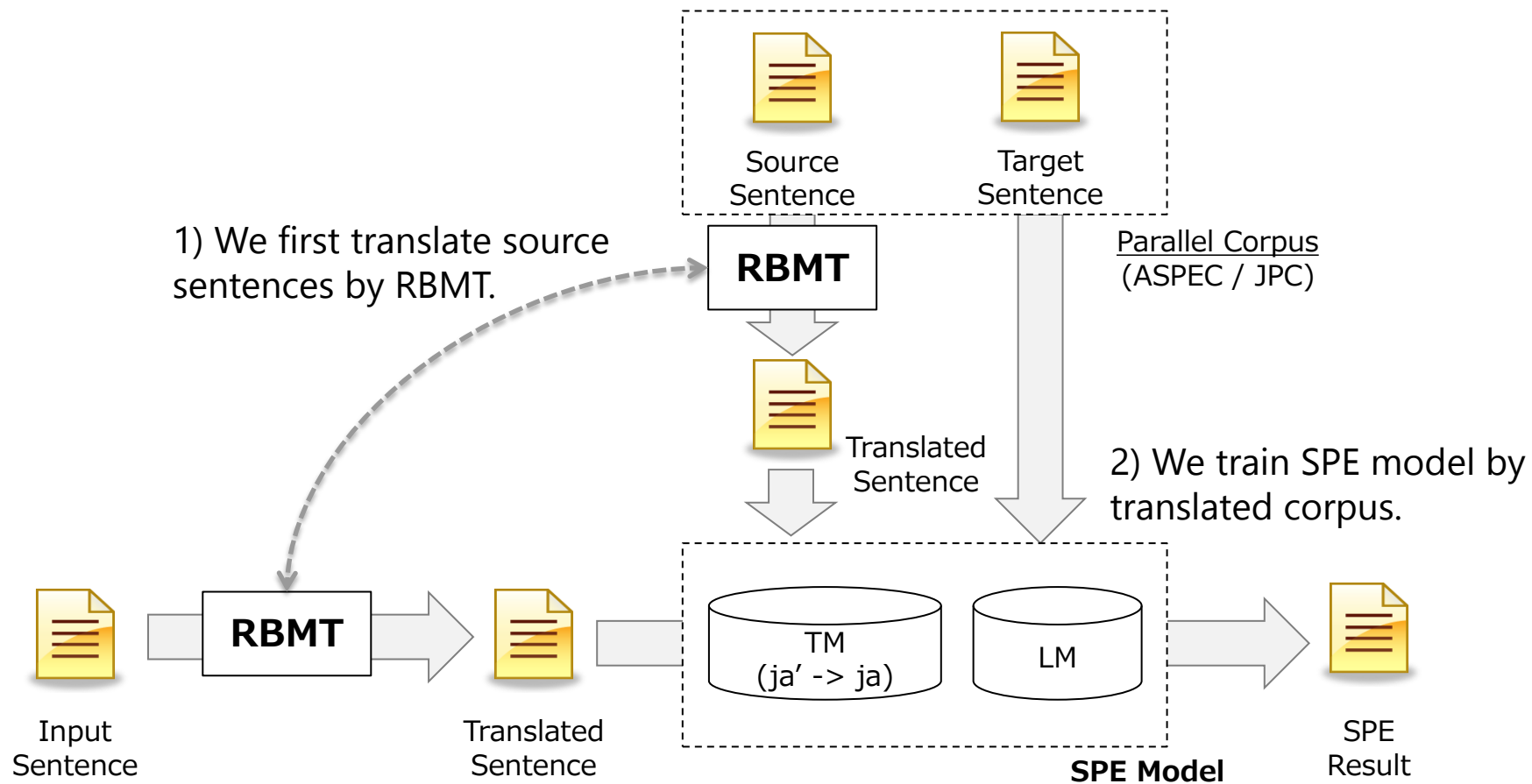
WAT 2015, Oct. 16, 2015 @ Kyoto

Motivations

- **Rule-Based Machine Translation (RBMT)**
  - We have been developed RBMT for more than 30 years.
  - Japanese⇔English, Japanese⇔Chinese, Japanese⇔Korean
  - Large technical dictionaries and translation rules

- **Pre-ordering SMT and Tree/Forest to String**
  - Effective solutions for Asian language translation (WAT2014)
  - But, **pre-ordering rules and parsers are needed**.

- **Our approach:**
  - Statistical Post Editing (SPE) (same as WAT2014)
    - Verify effectiveness in all tasks
  - System combination between SPE and SMT (new in WAT2015)

Statistical Post Editing (SPE)

Translating RBMT results to post-edited results.

1) We first translate source sentences by RBMT.

Source Sentence

Target Sentence

Parallel Corpus
(ASPEC / JPC)

RBMT

Translated Sentence

2) We train SPE model by translated corpus.

Input Sentence

RBMT

Translated Sentence

TM
(ja' -> ja)

LM

SPE Model

SPE Result

本发明具有以下效果。　　　本発明は以下効果を持っている。　　　本発明は以下の効果を有する。

Features of SPE

- **From RBMT's standpoint**
  - **Correct mistranslations / Translate unknown words**
    - Phrase-level correction (domain adaptation)
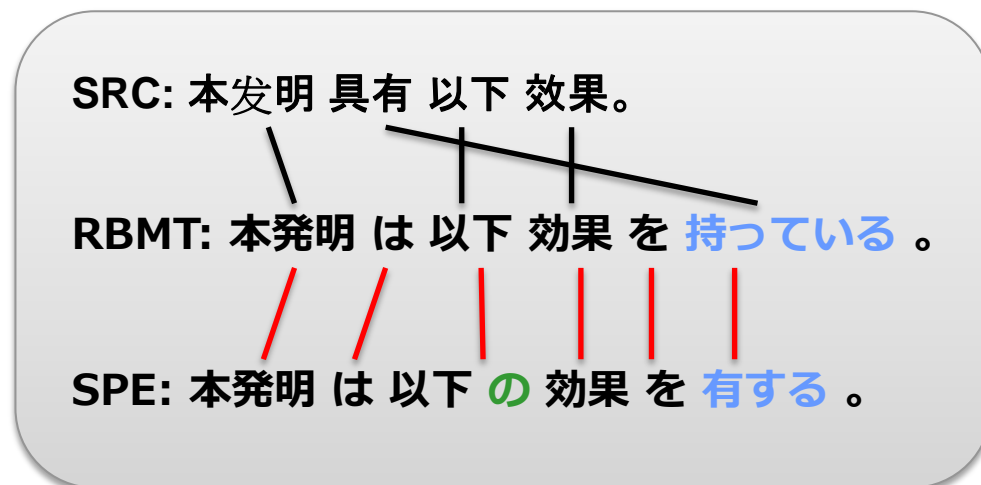 - **Improve fluency**
    - Use of more fluent expressions
    - Insertion of particles
  - Recover translation failure

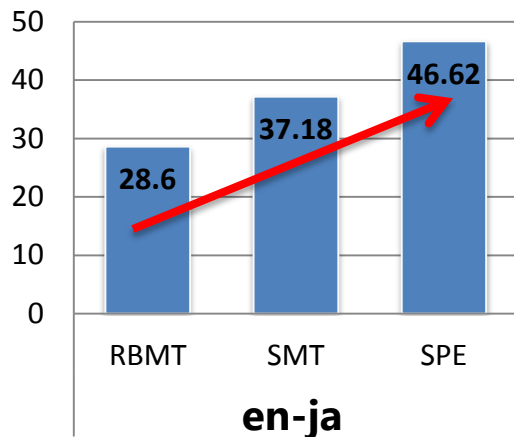- **From SMT's standpoint**
  - **Pre-ordering by RBMT**
  - Reduction of NULL alignment (subject/particle)
  - Use of syntax information (polarity/aspect)
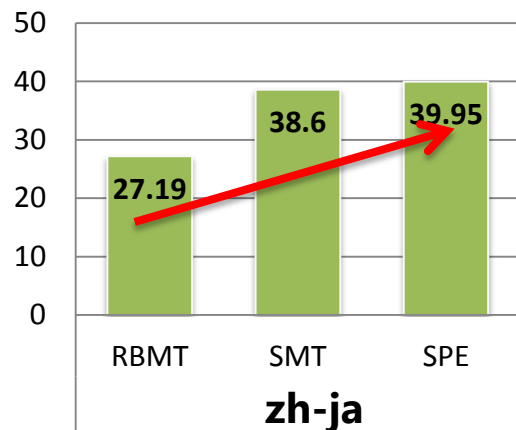 - Enhancement of lexicon

SRC: 本发明 具有 以下 效果。

RBMT: 本発明 は 以下 効果 を 持っている 。

SPE: 本発明 は 以下 の 効果 を 有する 。

SPE for Patent Translation

BLEU (en-ja)

	RBMT	SMT	SPE
	28.6	37.18	46.62

BLEU (zh-ja)

	RBMT	SMT	SPE
	27.19	38.6	39.95

BLEU (ko-ja)

	RBMT	SMT	SPE
	51.4	70.57	68.71

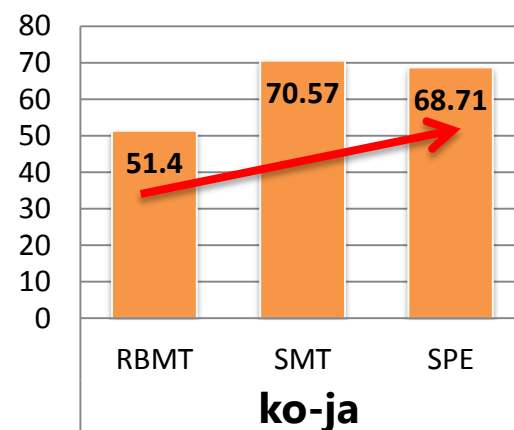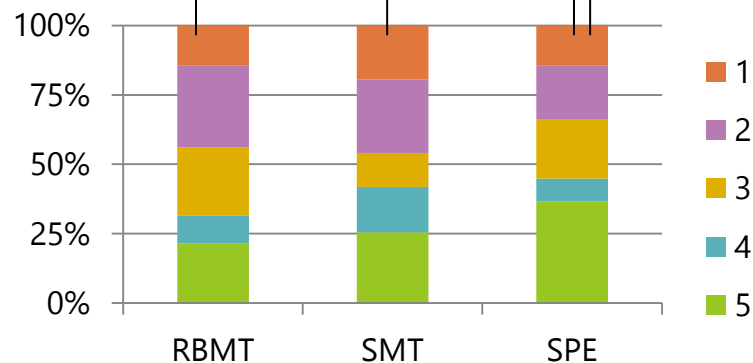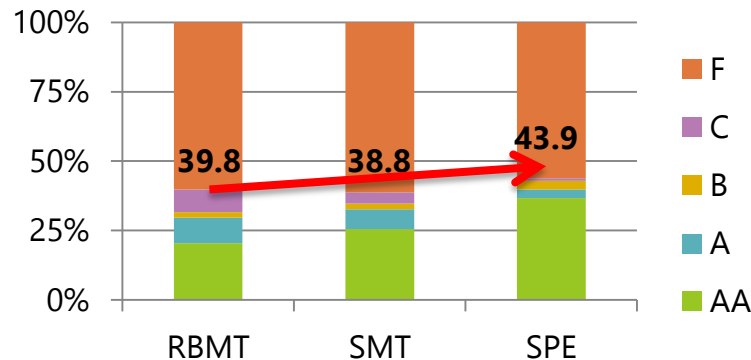**SPE shows:**
- **Better scores than PB-SMT in automatic evaluation**
- **Improvements of understandable level (>=C in acceptability)**

Adequacy — legend: 1, 2, 3, 4, 5

Acceptability — legend: F, C, B, A, AA; values: RBMT 39.8, SMT 38.8, SPE 43.9

Human evaluation for zh-ja

TOSHIBA
Leading Innovation >>>

System Combination

- **How combine systems?**
  - **Selection based on SMT scores and/or other features**.
  - Selection based on estimated score (Adequacy? Fluency? …)
    - Need data to learn the relationship…

- **Our approach in WAT2015:**
  - **Merge n-best candidates and rescore them.**
  - We used **RNNLM** for reranking.
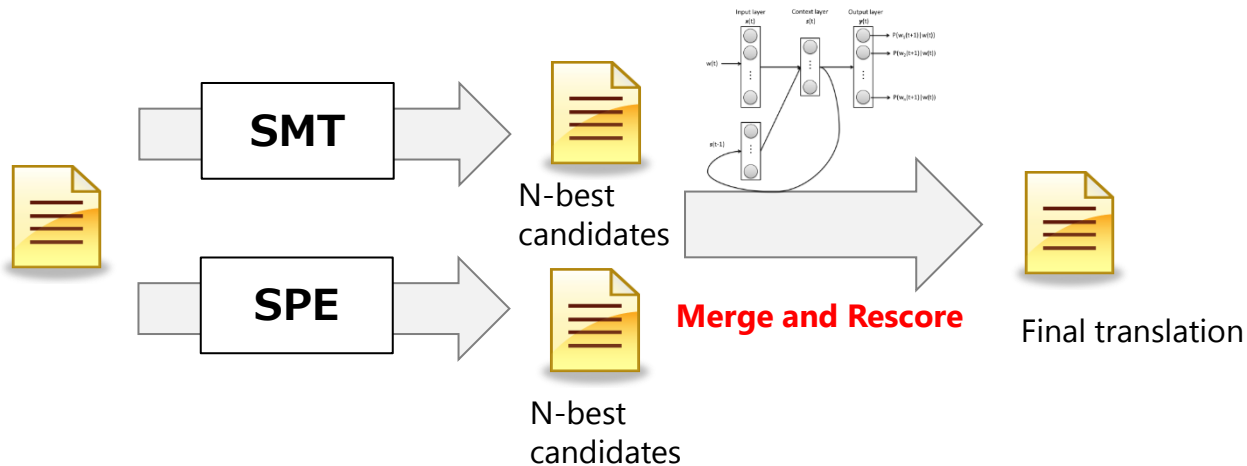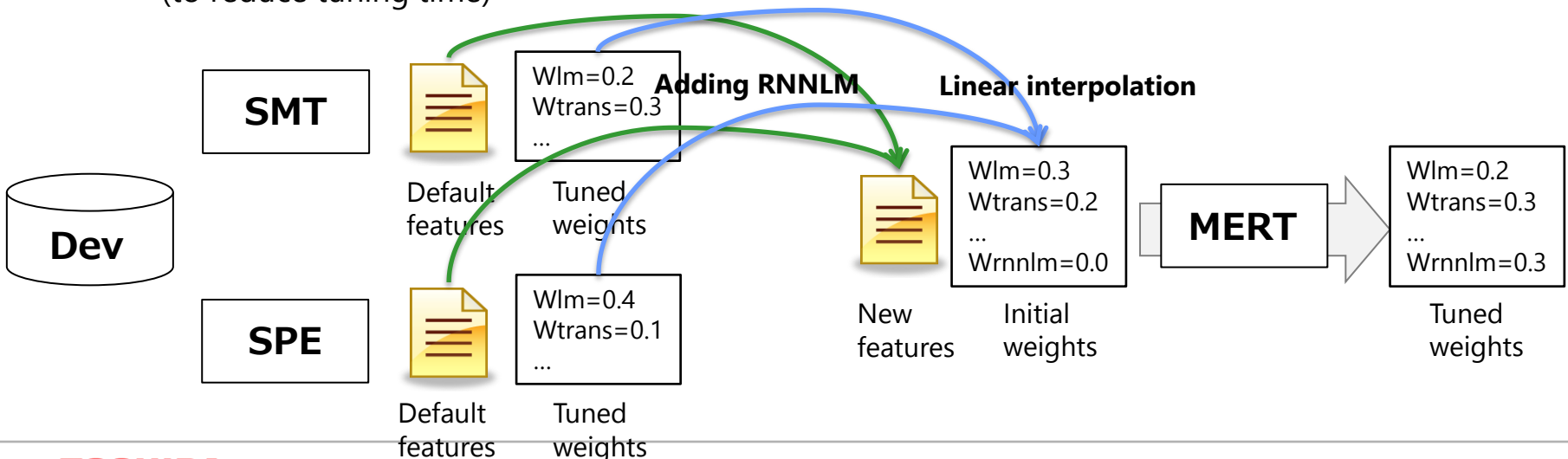
RNNLM reranking and Tuning

- **Reranking on the log-linear model**
  - Adding RNNLM score to default features of Moses.
  - RNNLM trained by rnnlm toolkit (Mikolov '12).
    - 500,000 sentences for each language
    - # of hidden layer=500, # of class=50
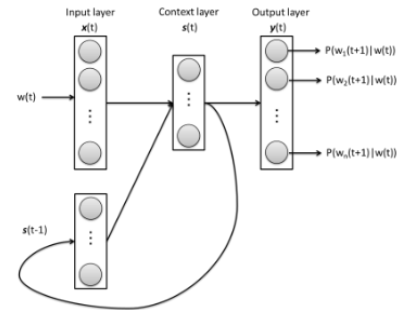
- **Tuning**
  - Using tuned weights without RNNLM, we ran only 1 iteration.
  (to reduce tuning time)

Experimental Results

BLEU for ASPEC

	SMT	SPE	COMB
ja-en	17.41	22.65	23.00 (+0.35)
en-ja	25.17	31.10	31.82 (+0.72)
ja-zh	28.20	29.48	29.60 (+0.12)
zh-ja	36.34	35.76	37.47 (+1.71)

BLEU for Patent

	SMT	SPE	COMB
JPCzh-ja	38.77	39.01	40.23 (+1.22)
JPCko-ja	70.17	68.47	70.4 (+1.93)

TOSHIBA
Leading Innovation >>>
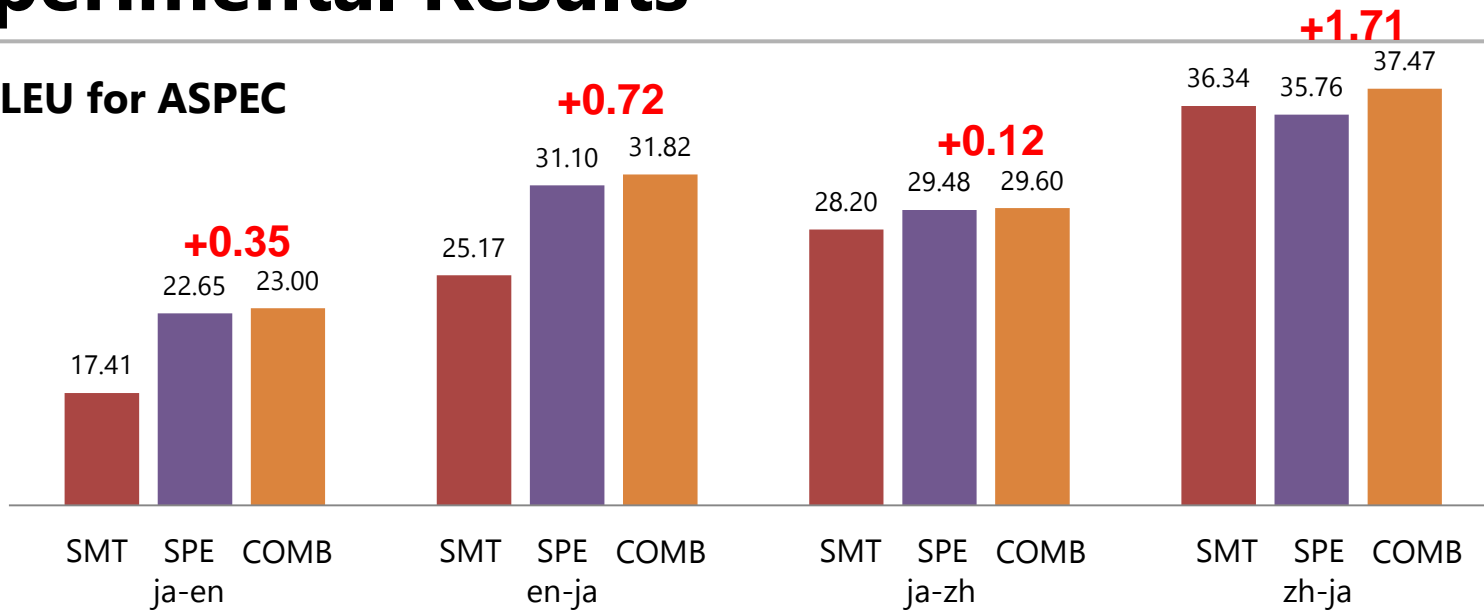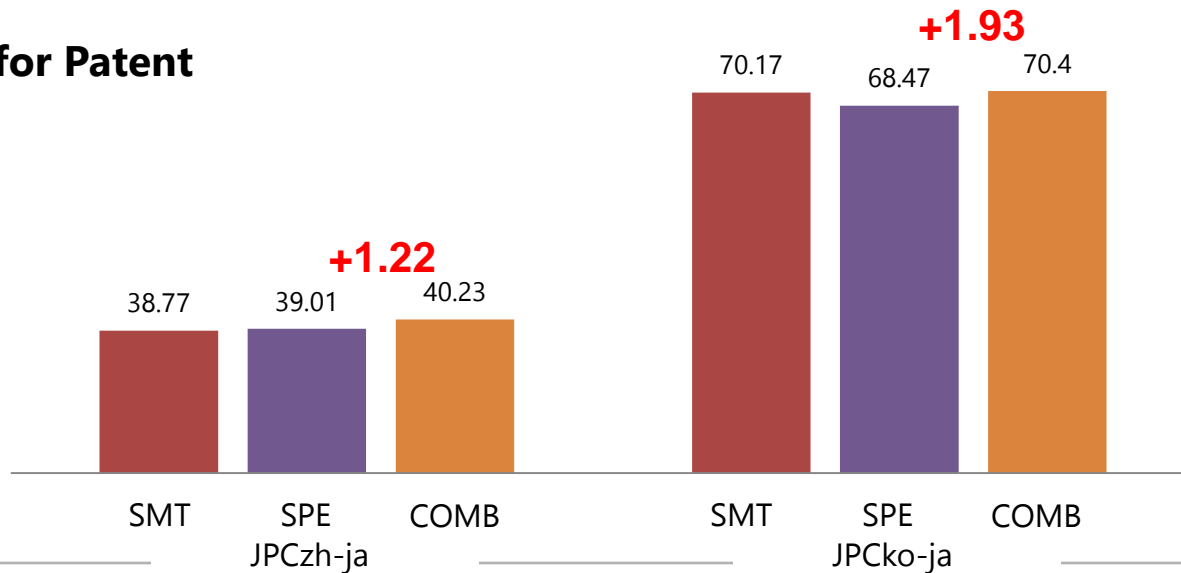
Experimental Results

Systems	Rerank	ja-en		en-ja		ja-zh		zh-ja	
		BLEU	RIBES	BLEU	RIBES	BLEU	RIBES	BLEU	RIBES
RBMT	No	15.31	0.677	14.78	0.685	19.51	0.767	15.39	0.767
SMT	No	17.41	0.620	25.17	0.642	28.20	**0.810**	36.34	**0.810**
	Yes	17.85	0.619	25.37	0.643	28.46	0.809	36.69	0.809
SPE	No	22.65	0.717	31.10	0.767	29.48	0.809	35.76	0.809
	Yes	22.92	**0.718**	31.73	**0.770**	29.49	0.809	36.06	0.809
COMB	Yes	**23.00**	0.716	**31.82**	**0.770**	**29.60**	**0.810**	**37.47**	**0.810**

System Combination (COMB) achieved improvements of BLEU and RIBES score than SPE.

COMB is the best system except JPCko-ja task.

Systems	Rerank	JPCzh-ja		JPCko-ja	
		BLEU	RIBES	BLEU	RIBES
RBMT	No	25.81	0.764	51.28	0.902
SMT	No	38.77	0.802	70.17	0.943
	Yes	39.18	0.805	**70.89**	**0.944**
SPE	No	39.01	**0.813**	68.47	0.940
	Yes	39.30	0.811	68.76	0.940
COMB	Yes	**40.23**	**0.813**	70.40	0.942

Which systems did the combination selected?

SAME 3% — **SMT 14%** — **SPE 83%**
ja-en

SAME 2% — **SMT 9%** — **SPE 89%**
en-ja

SAME 5% — **SMT 40%** — **SPE 55%**
ja-zh
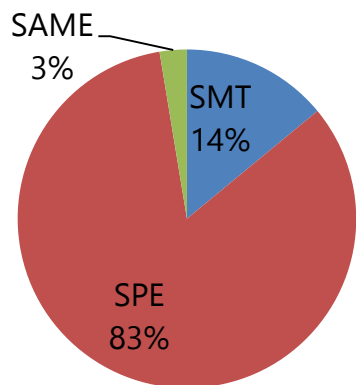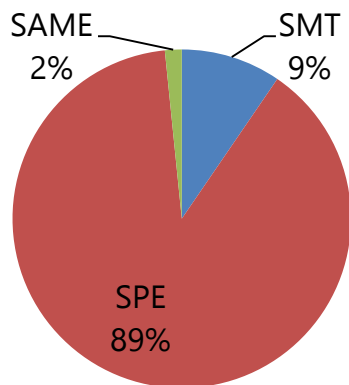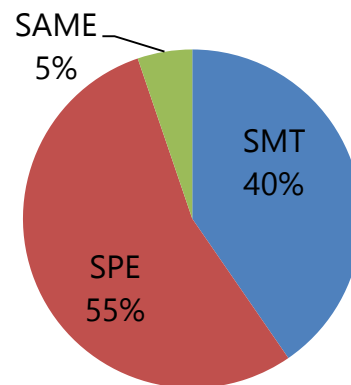
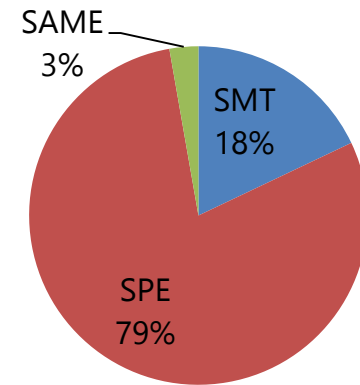**SAME 3%** — **SMT 18%** — **SPE 79%**
zh-ja

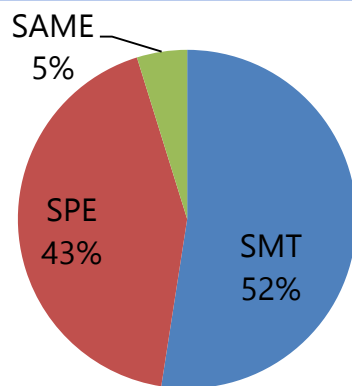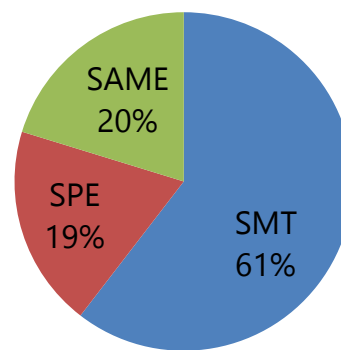> **ja-en/en-ja/zh-ja:** about 80% translations come from SPE.
>
> **ja-zh and JPCzh-ja:** COMB selected SPE and SMT, equivalently.
> (Because RBMT couldn't translate well, % of SMT increased.)

SAME 5% — **SMT 52%** — **SPE 43%**
JPCzh-ja

SAME 20% — **SMT 61%** — **SPE 19%**
JPCko-ja

"same" means that COMB results were included both SMT and SPE.

TOSHIBA
Leading Innovation >>>

Translation Examples

zh-ja: COMB was selected from n-best of SPE.

SRC	还对在完成来所登记之前的各个环节进行了介绍。
REF	来所登録が完了するまでの流れ等も紹介した。
RBMT	さらに完成が登記する前でのそれぞれの段階に対して紹介を行った。
SMT	通所を完了しても登録までの各環節について紹介した。
SPE	登録の完了までの各段階について紹介した。
COMB	登録が完了するまでの各段階について紹介した。

JPCzh-ja: COMB was selected from n-best of SMT.

SRC	图6是用于说明本发明第一实施例中的移动台站的活动水平控制的图;
REF	図6は、本発明の第1の実施例における移動局の活動レベルの制御を説明するための図である。
RBMT	図6は、本発明の第一実施例中を説明するのに用いる移動台ステーションの活動レベリングコントロールの図である;
SMT	図6は、本発明の第1の実施例における移動局のアクティブレベル制御を示す図である。
SPE	図6は、本発明の第1の実施形態を説明するための移動局の活動水平コントロールの図である。
COMB	図6は、本発明の第1の実施例における移動局のアクティブレベル制御を示す図である。

Toshiba MT system of WAT2015

- **We additionally applied some pre/post processing.**

Technical Term Dictionaries

Selecting RBMT dictionaries by devset.

+ JPO patent dictionary (2.2M words for JPCzh-ja)

English Word Correction

Edited-distance based correction.

continous -> continuous
behvior -> behavior
resolutin -> resolution

KATAKANA Normalization

Normalize to highly-frequent notations for "一".

スクリュ -> スクリュー
サーバー -> サーバ

Post-translation

Translate remaining unknown words by RBMT.

アルキメデス数 ->阿基米德数
流入마하수 -> 流入マッハ数

Official Results

- **SPE and SMT ranked in the top 3 HUMAN in ja-en/ja-zh/JPCzh-ja.**

System	ja-en			en-ja			ja-zh			zh-ja		
	BLEU	RIBES	HUMAN	BLEU	RIBES	HUMAN	BLEU	RIBES	HUMAN	BLEU	RIBES	HUMAN
SPE	22.89	0.719	**25.00**	32.06	0.771	40.25	30.17	0.813	2.50	35.85	0.825	-1.00
COMB	23.00	0.716	21.25	31.82	0.770	-	30.07	0.817	**17.00**	37.47	0.827	18.00

System	JPCzh-ja			JPCko-ja		
	BLEU	RIBES	HUMAN	BLEU	RIBES	HUMAN
SMT	-	-	-	71.01	0.944	4.50
SPE	41.12	0.822	**24.25**	-	-	-
COMB	41.82	0.821	14.50	70.51	0.942	3.00

- **The correlation between BLEU/RIEBES and HUMAN is not clear in our system.**

BLEU-HUMAN $R^2 = 0.2338$

RIBES-HUMAN $R^2 = 0.3813$
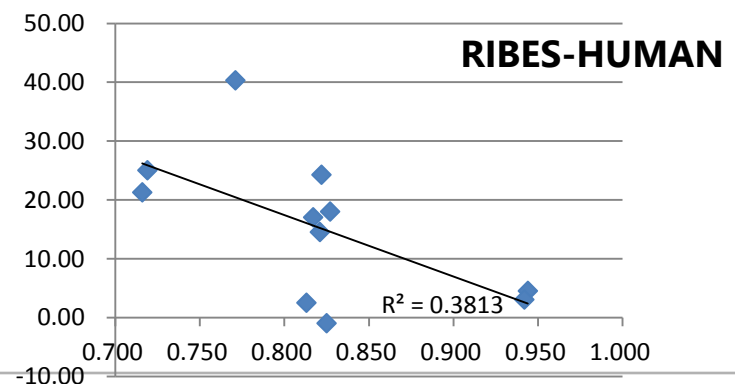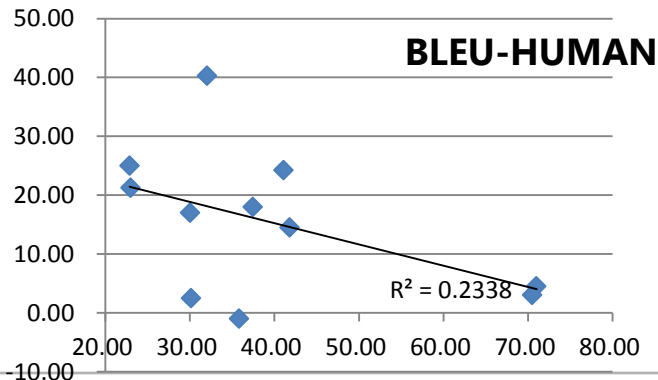
Crowdsourcing Evaluation

- **Analysis of JPCko-ja result (COMB vs Online A)**
  - In in-house evaluation, COMB is better than Online A.

	BLEU	RIBES	HUMAN Baseline	COMB	Online A
COMB	70.51	0.94	3.00	-	10.75
Online A	55.05	0.91	38.75	-10.75	-

  Official (Crowdsourcing)   **In-house evaluation results**

  - Effected by differences in number expressions !?

  SRC : 시스템(100) ⇒ Online A: システム(100)
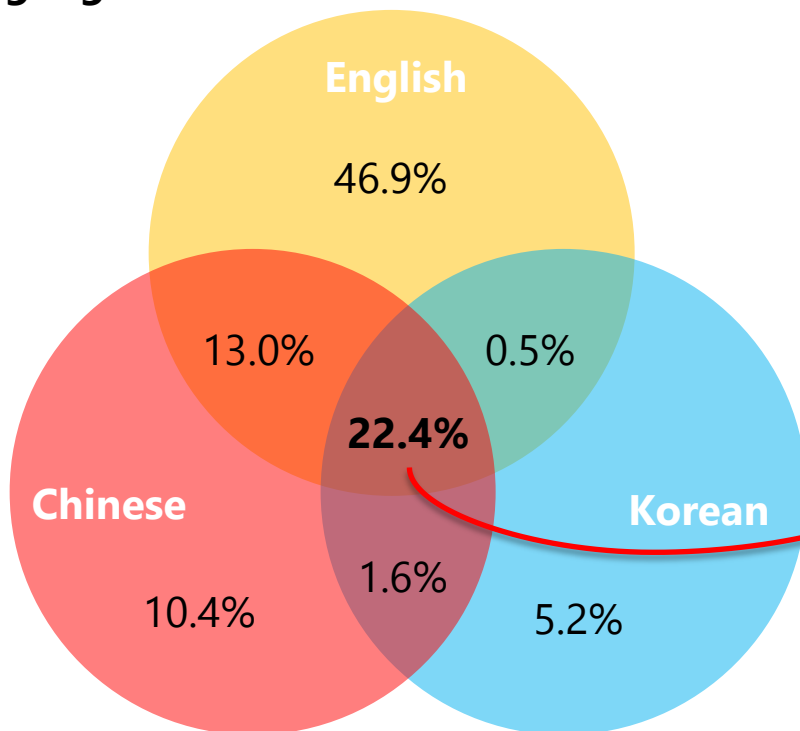
  COMB(SMT): システム100

  ⇒ Equally evaluated in-house evaluation.

  - **Crowd-workers should be provided an evaluation guideline by which such a difference is considered.**
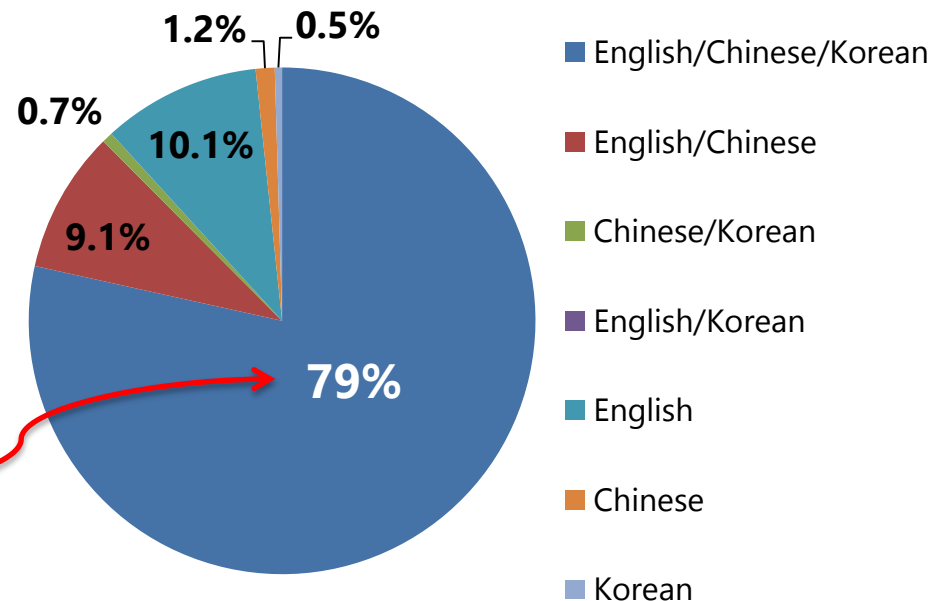
Findings of Crowdsourcing Evaluation (1)

- **Many workers evaluated more than one language pairs.**
  - 22.4% workers evaluated all languages.
  - 22.4% workers evaluated 78.5% sentences.

Languages of workers

English 46.9%

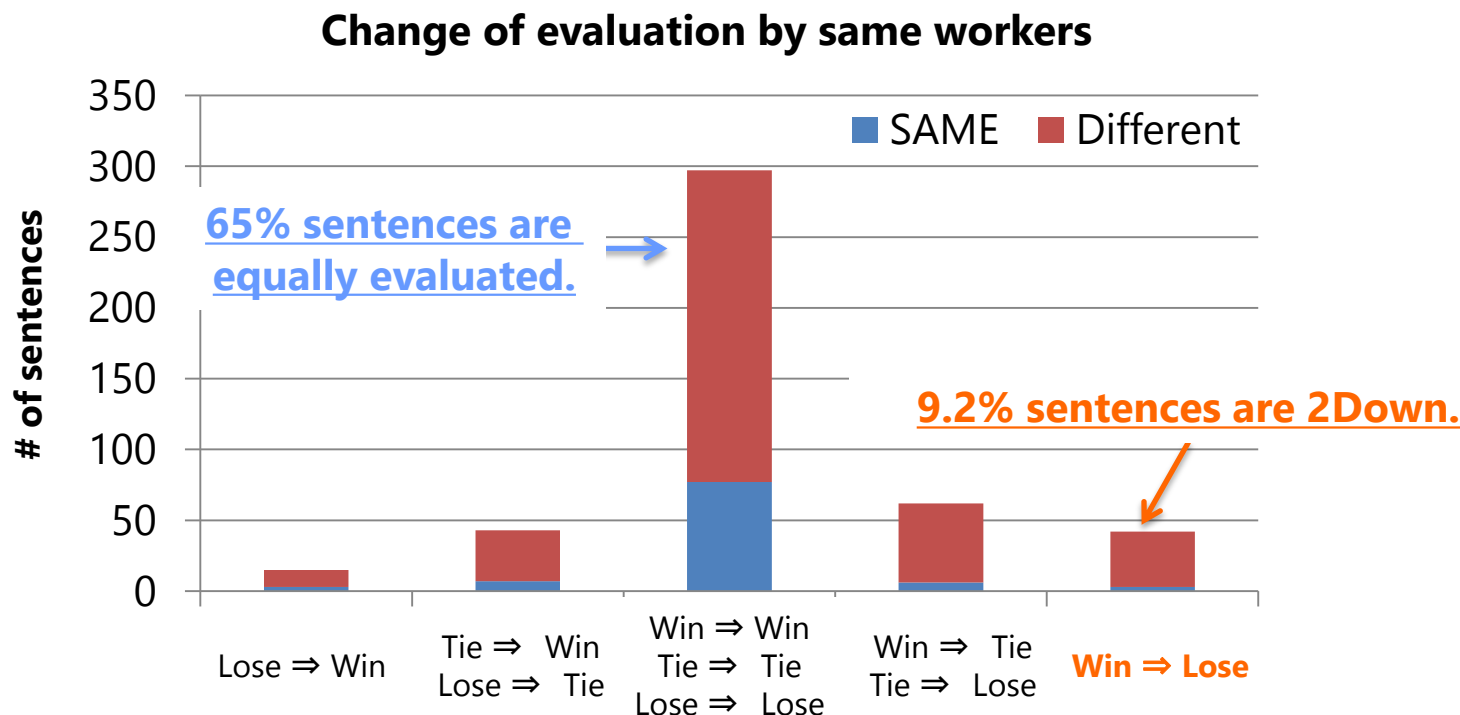13.0%   0.5%

22.4%

Chinese   1.6%

10.4%   5.2%

Korean

Ratio of sentences evaluated by workers

1.2%   0.5%

0.7%

10.1%

9.1%

79%

- English/Chinese/Korean
- English/Chinese
- Chinese/Korean
- English/Korean
- English
- Chinese
- Korean

Findings of Crowdsourcing Evaluation (2)

- **In JPCko-ja, we got two evaluation results.**
  - Unofficial results: evaluation of original translation (HUMAN=29.75)
  - Official results: evaluation of normalized translation (to full-width)

    Ex.) T溶融(DSC)=89.9℃;T結晶化(DSC)=72℃(5℃/分でDSCで測定)。

    ⇒ T溶融（DSC）＝89．9℃；T結晶化（DSC）＝72℃（5℃／分でDSCで測定）。

Change of evaluation by same workers

65% sentences are equally evaluated.

9.2% sentences are 2Down.

TOSHIBA
Leading Innovation >>>

Summary

- **Toshiba MT system achieved a combination method between SMT and SPE by RNNLM reranking.**

- **Our system ranked the top 3 HUMAN score in ja-en/ja-zh/JPCzh-ja.**

- **We will aim for practical MT system by more effective combination systems (SMT, SPE , RBMT and more...)**

TOSHIBA

Leading Innovation >>>