

# **TOSHIBA**

Leading Innovation >>>

## **Toshiba MT System Description for the WAT2015 Workshop**

Satoshi SONOH  
Satoshi KINOSHITA

Knowledge Media Laboratory,  
Corporate Research & Development Center,  
Toshiba Corporation.

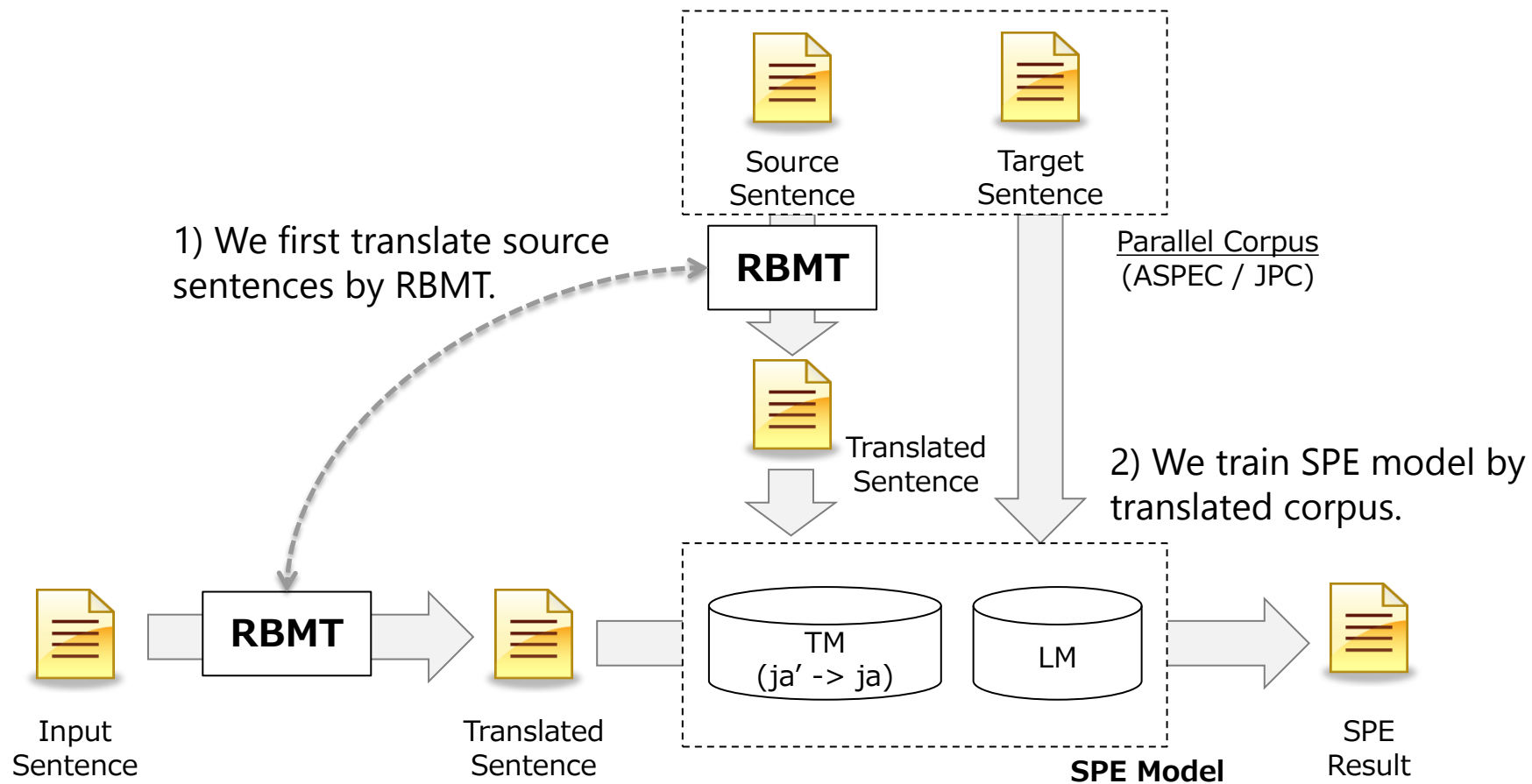
# Motivations

---

- **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.



本発明具有以下效果。

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**

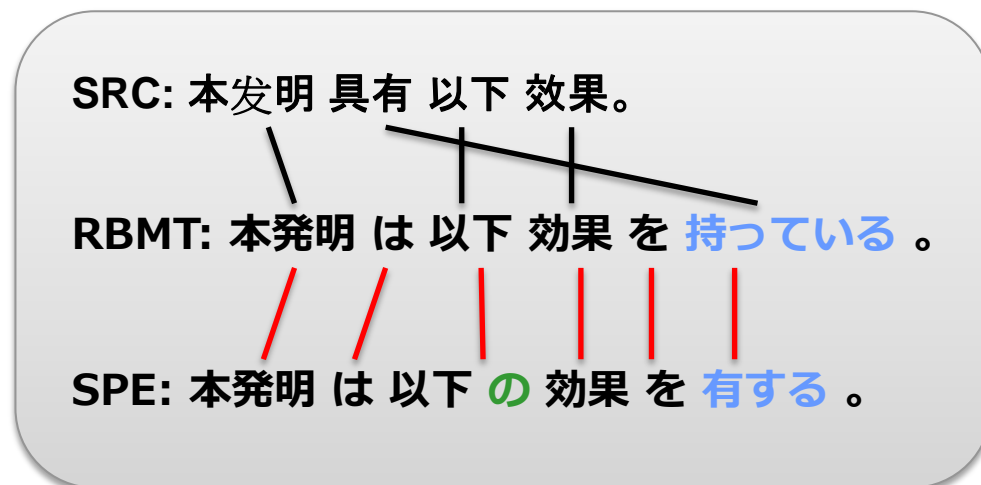
- Use of more fluent expressions
- Insertion of particles

- Recover translation failure

- **From SMT's standpoint**

- **Pre-ordering by RBMT**

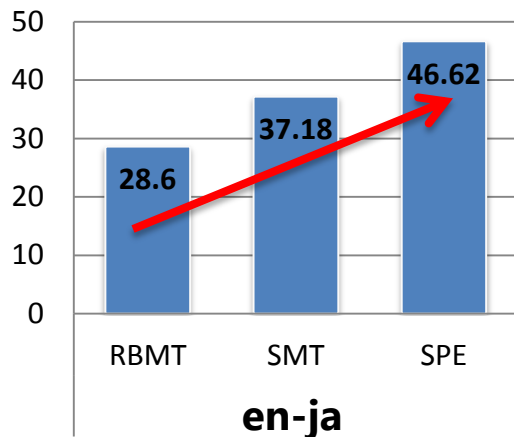
- Reduction of NULL alignment (subject/particle)
- Use of syntax information (polarity/aspect)
- Enhancement of lexicon



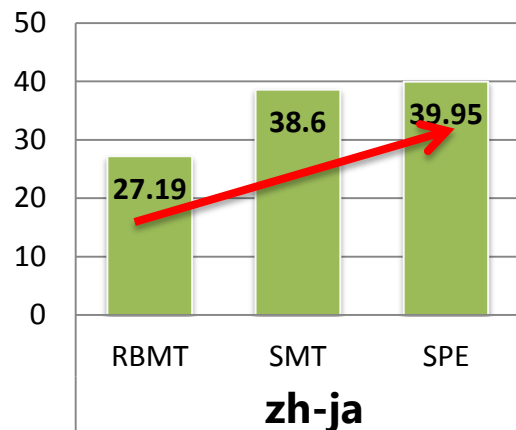
# SPE for Patent Translation

Corpus: JPO-NICT patent corpus  
 # of training data: 2M(en-ja), 1M(zh-ja/ko-ja)  
 # of automatic evaluation: 2,000  
 # of human evaluation: 200

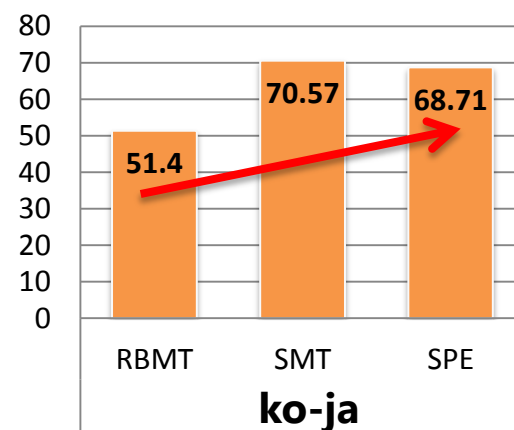
BLEU



BLEU



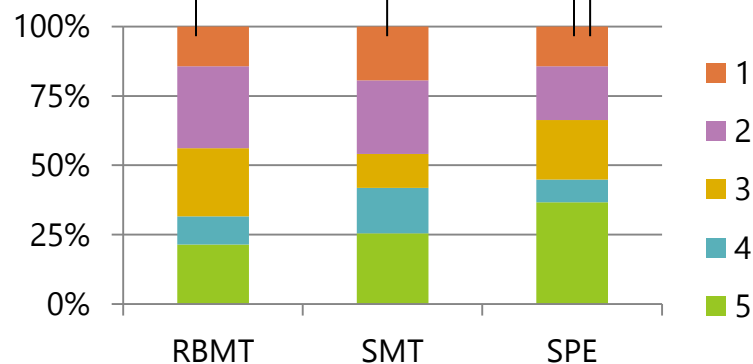
BLEU



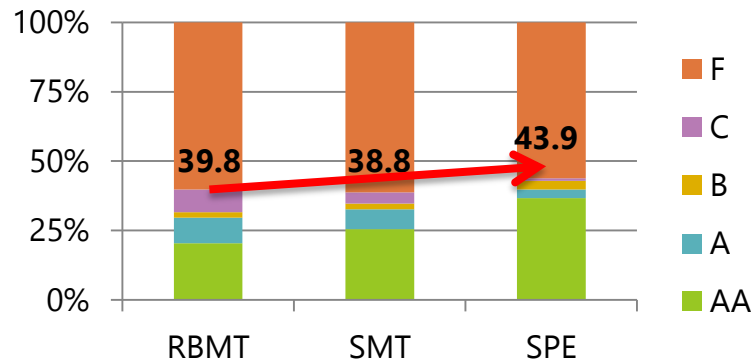
## SPE shows:

- Better scores than PB-SMT in automatic evaluation
- Improvements of understandable level ( $\geq C$  in acceptability)

Adequacy



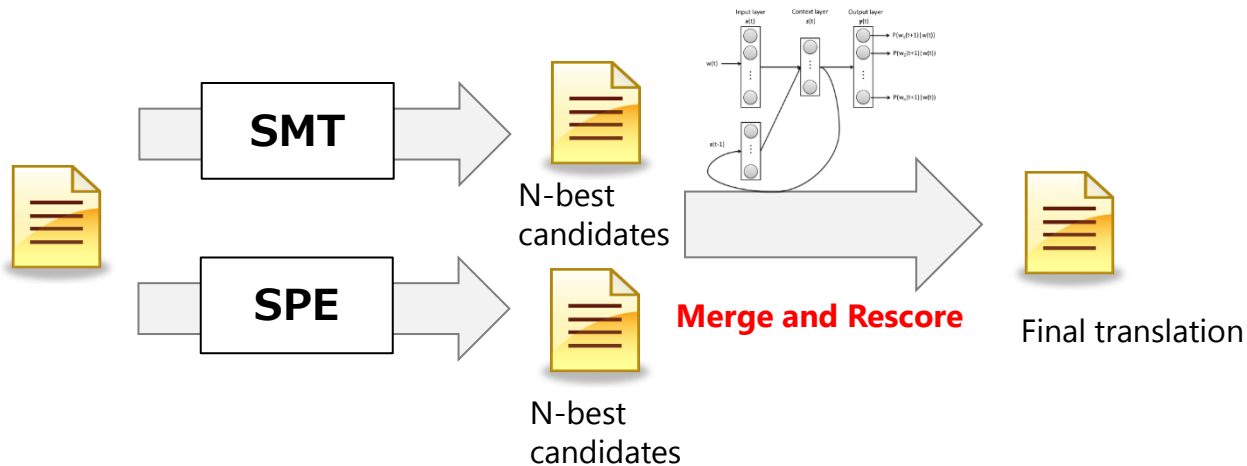
Acceptability



Human evaluation for zh-ja

# System Combination

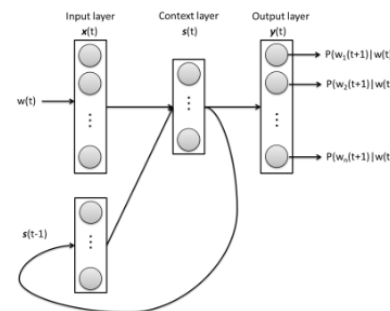
- **How combine systems?**
  - Selection based on **SMT scores and/or other features.**
  - Selection based on estimated score (Adequacy? Fluency? ...)
    - Need data to learn the relationship...
- **Our approach in WAT2015:**
  - **Merge n-best candidates and rescore them.**
  - We used **RNNLM** for reranking.



# RNNLM reranking and Tuning

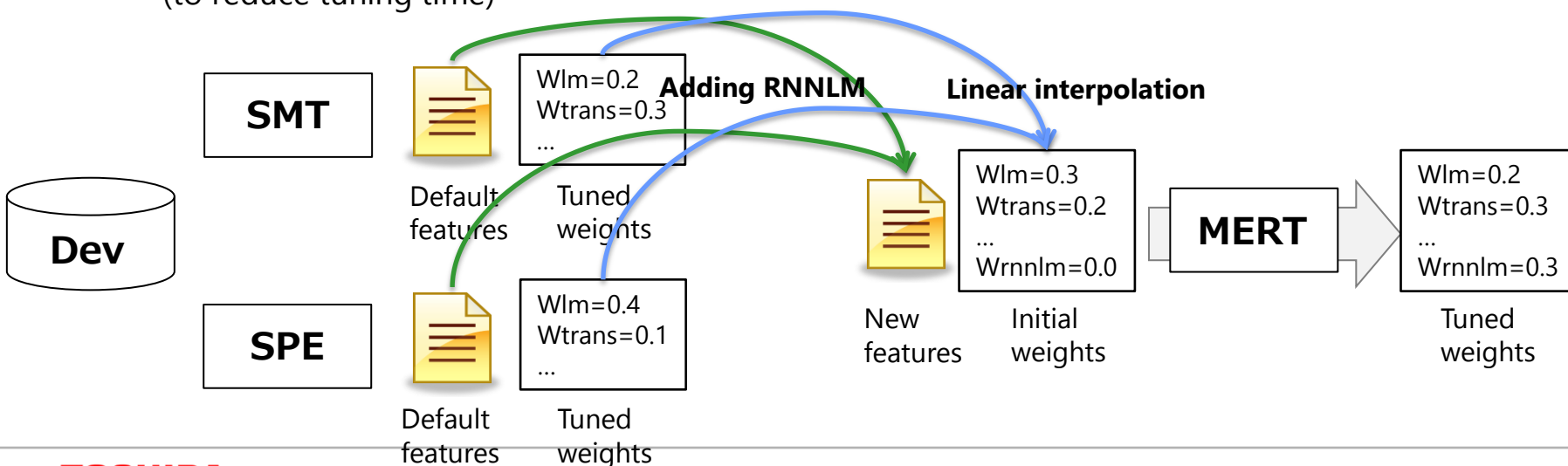
- **Reranking on the log-linear model**

- 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



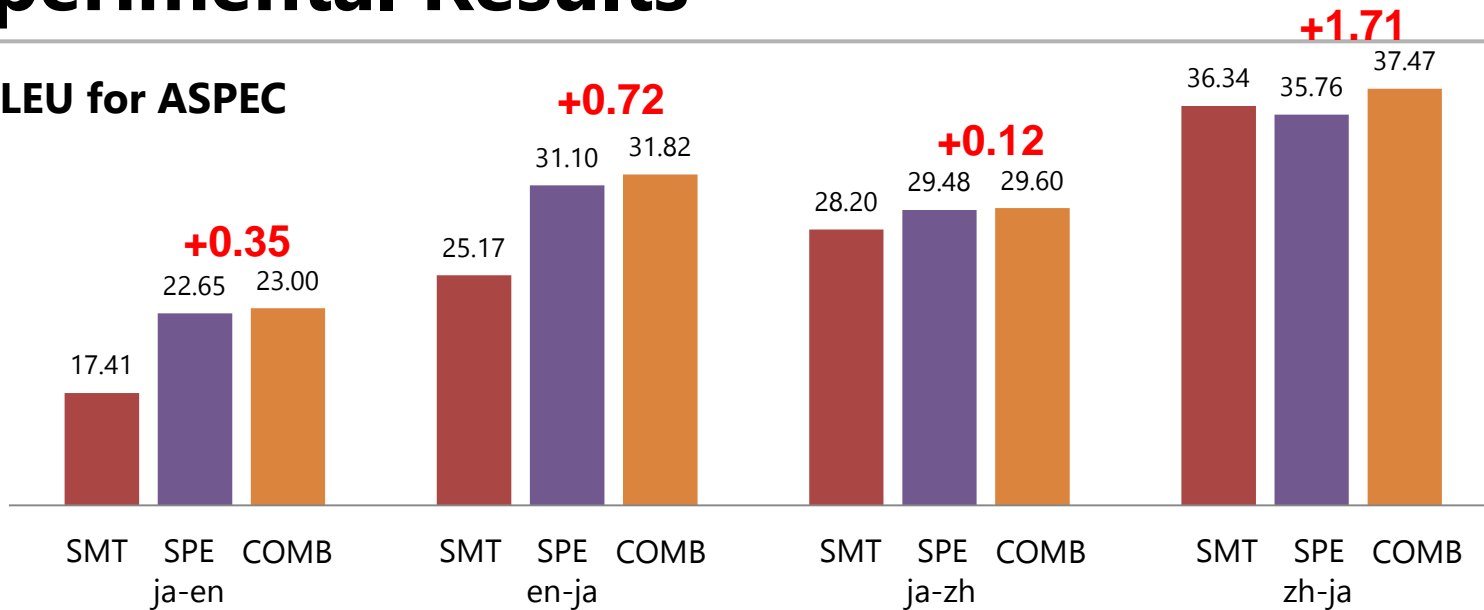
- **Tuning**

- Using tuned weights without RNNLM, we ran only 1 iteration.  
(to reduce tuning time)

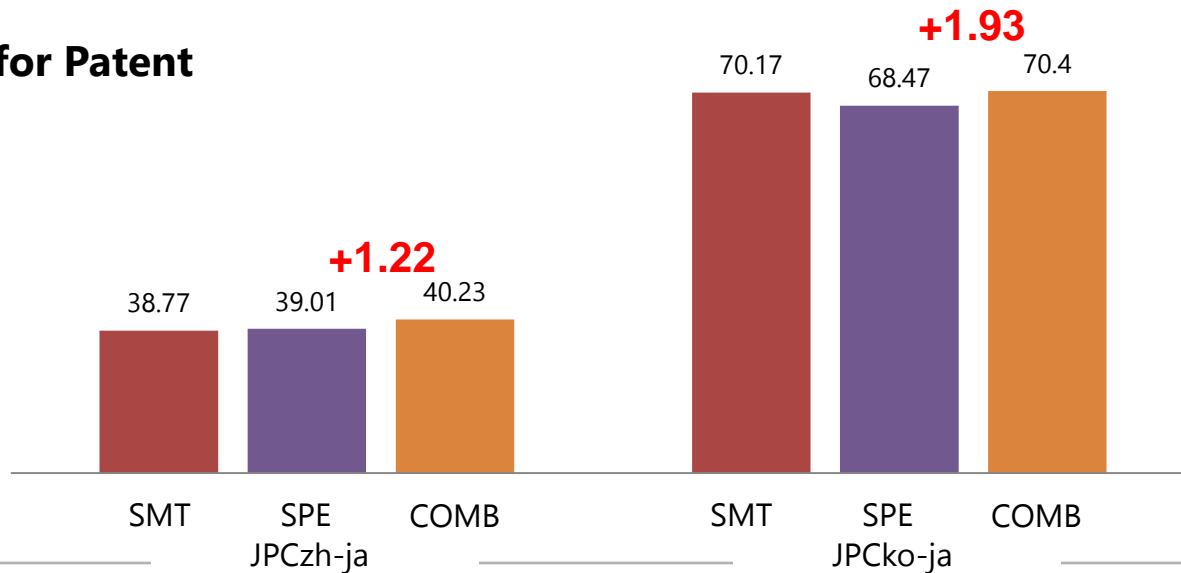


# Experimental Results

## BLEU for ASPEC



## BLEU for Patent





# Experimental Results

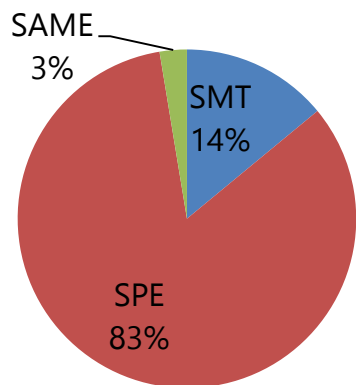
Systems	Rerank	ja-en		en-ja		ja-zh		zh-ja	
		BLEU	RIBES	BLEU	RIBES	BLEU	RIBES	BLEU	RIBES
RBMT	No	15.31	0.677	14.78	0.685	19.51	0.767	15.39	0.767
SMT	No	17.41	0.620	25.17	0.642	28.20	<b>0.810</b>	36.34	<b>0.810</b>
	Yes	17.85	0.619	25.37	0.643	28.46	0.809	36.69	0.809
SPE	No	22.65	0.717	31.10	0.767	29.48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

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

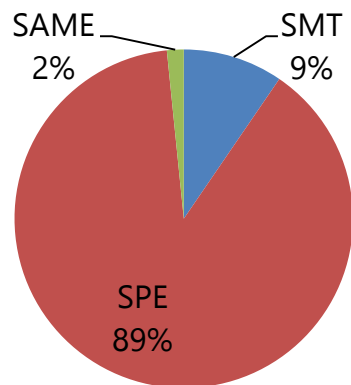
**COMB is the best system except JPCko-ja task.**

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	<b>70.89</b>	<b>0.944</b>
SPE	No	39.01	<b>0.813</b>	68.47	0.940
	Yes	39.30	0.811	68.76	0.940
COMB	Yes	<b>40.23</b>	<b>0.813</b>	70.40	0.942

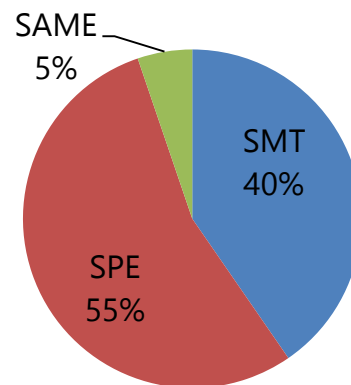
# Which systems did the combination selected?



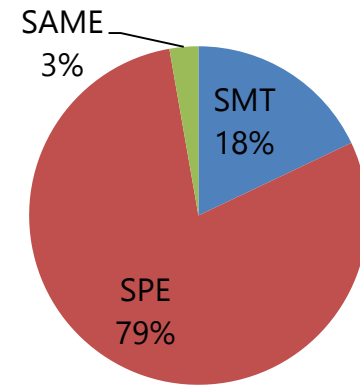
ja-en



en-ja



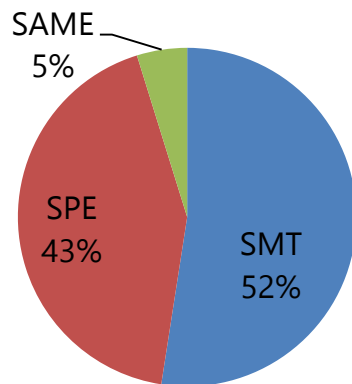
ja-zh



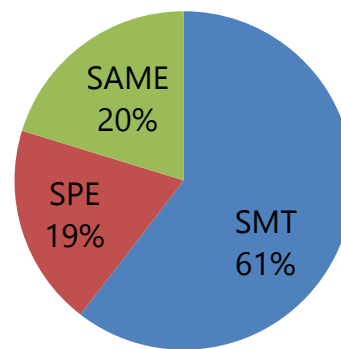
zh-ja

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.)



JPCzh-ja



JPCko-ja

# 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  
behavior -> behavior  
resolutin -> resolution

## KATAKANA Normalization

Normalize to highly-frequent notations for “—”.

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 -> 阿基米徳数  
流入마하수 -> 流入マッハ数

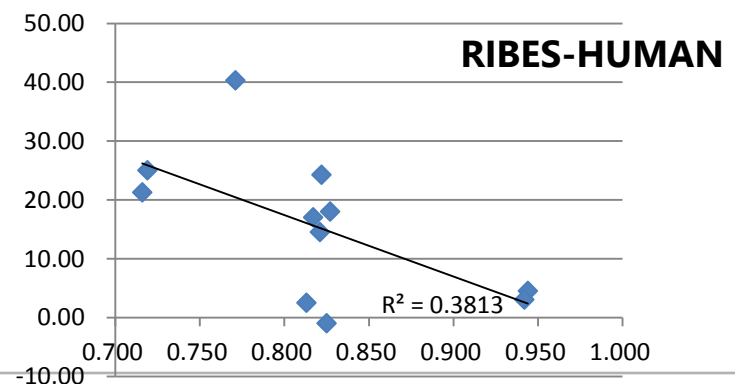
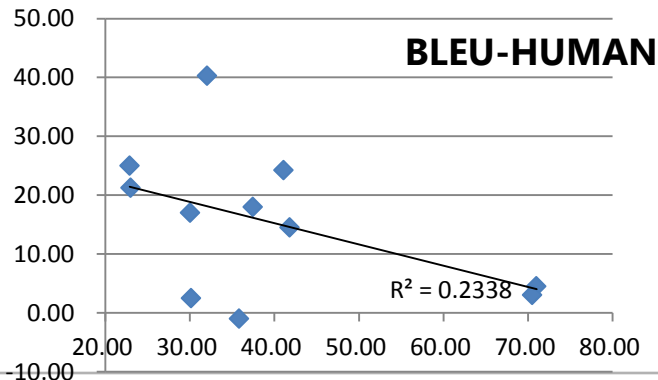
# Official Results

- 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	<b>25.00</b>	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	<b>17.00</b>	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	<b>24.25</b>	-	-	-
COMB	41.82	0.821	14.50	70.51	0.942	3.00

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



# 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.**

# 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 >>>**