



Translation Paraphrases In Phrase- Based Statistical Machine Translation

Statistical Machine Translation

- Statistical Analysis of bilingual texts.
- Started in 80's (IBM Candide).
- Phrase-based concept introduced (Och)
- Koehn et al. (2003) – Introduced the concept of Phrase based Statistical Machine Translation

$$\arg \max_e p(e|f) = \arg \max_e p(f|e) p(e)$$

This works well

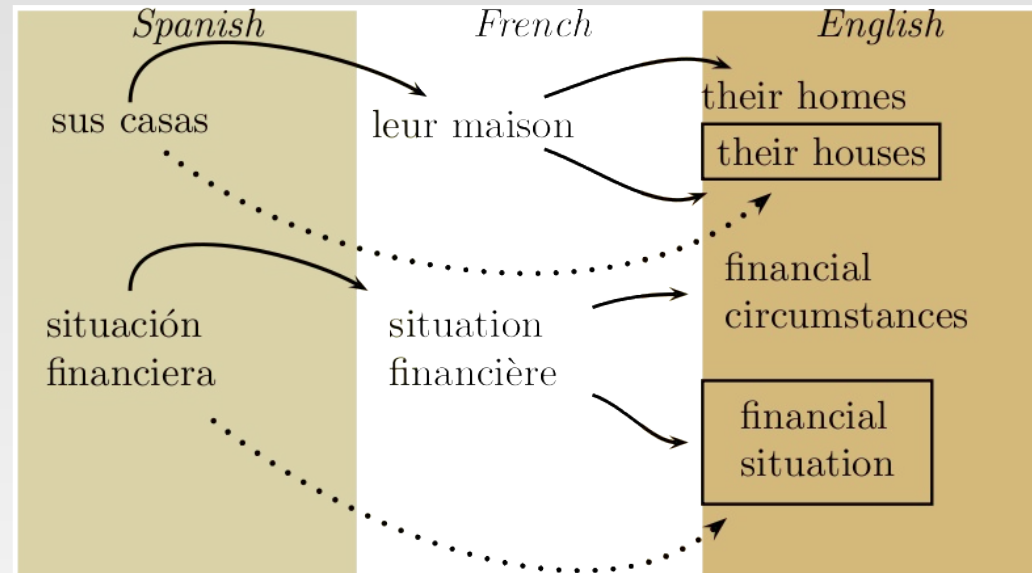
- STMT have proven to be very effective.
- Not bound to source/target language.
- Outstanding results compared to other MT systems.

but&

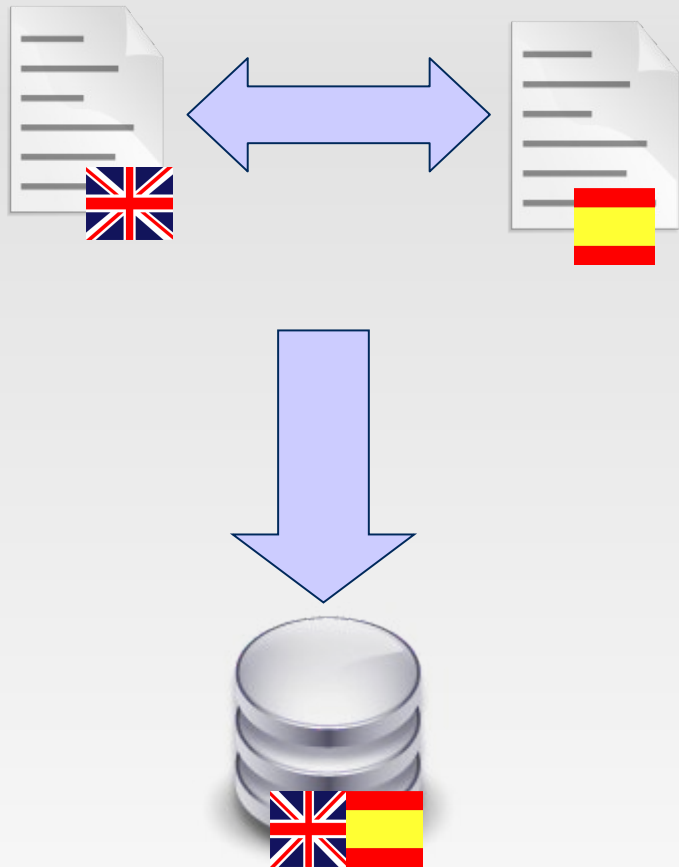
- Limited by **training** and availability of resources(corpora).

Translation Paraphrases

- **Paraphrases:** different phrases carrying similar meaning
- **Translation Paraphrases:** Mechanism of preserving meaning through translation
- Serve to give more **flexible interpretations** of source texts, as well as to reinforce good translations; regardless of the translation process.

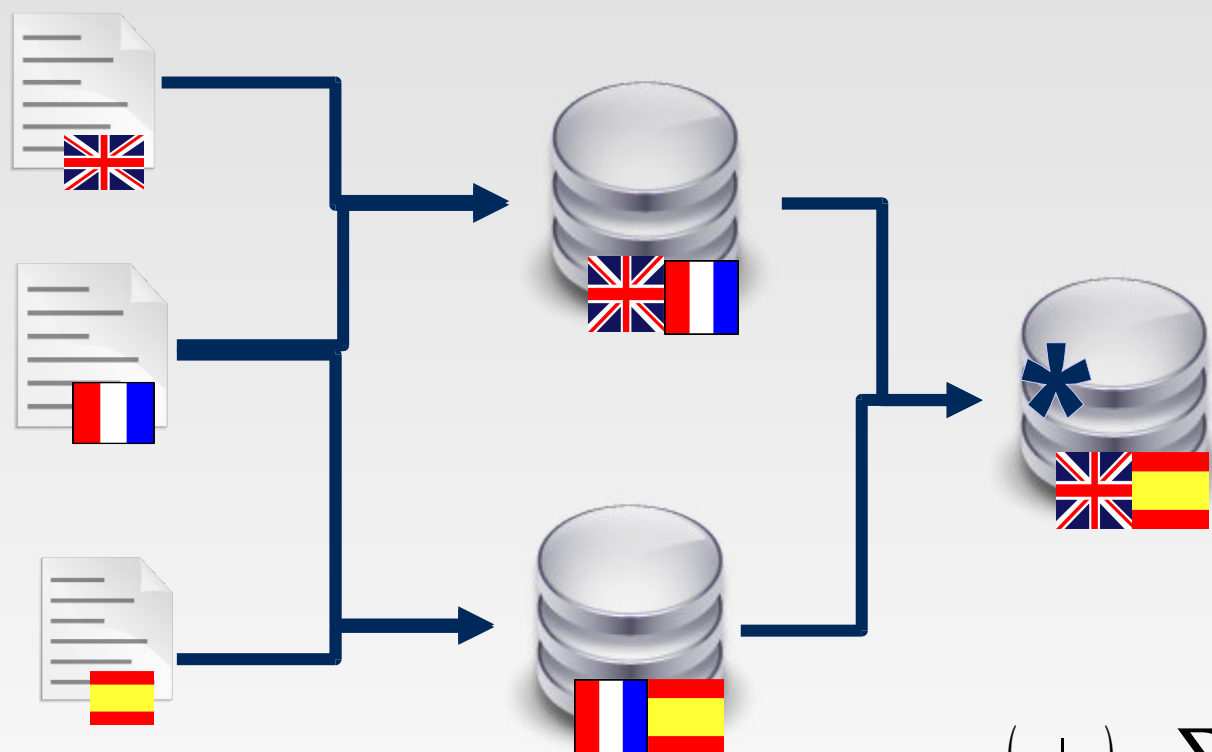


Training



- " Patterns are discovered through alignments.
- " Probabilities are assigned
- " Learned data is stored in phrase tables

Training pt. 2



$$p_{tp}(e|s) = \sum_f p_o(e|f)p_o(f|s)$$

How much extra information?

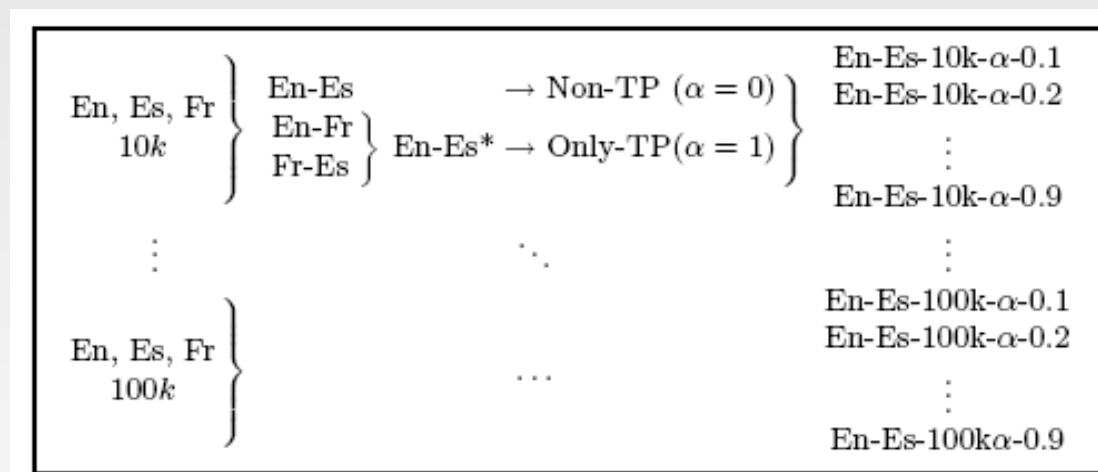
- Comparative Experiments
- Different Training sizes
- Alpha variation



$$p_{mix}(e | s) = (1 - \alpha) p_o(e | s) + p_{tp}(e | s)$$

Experimental Setup

- Training Corpus:
 - Europarl
 - Size (# lines) 10k, 20k, 40k, 80k, 100k
- MERT Training:
 - Europarl 2002
- Software Used:
 - Giza++
 - Moses Decoder
 - SRILM



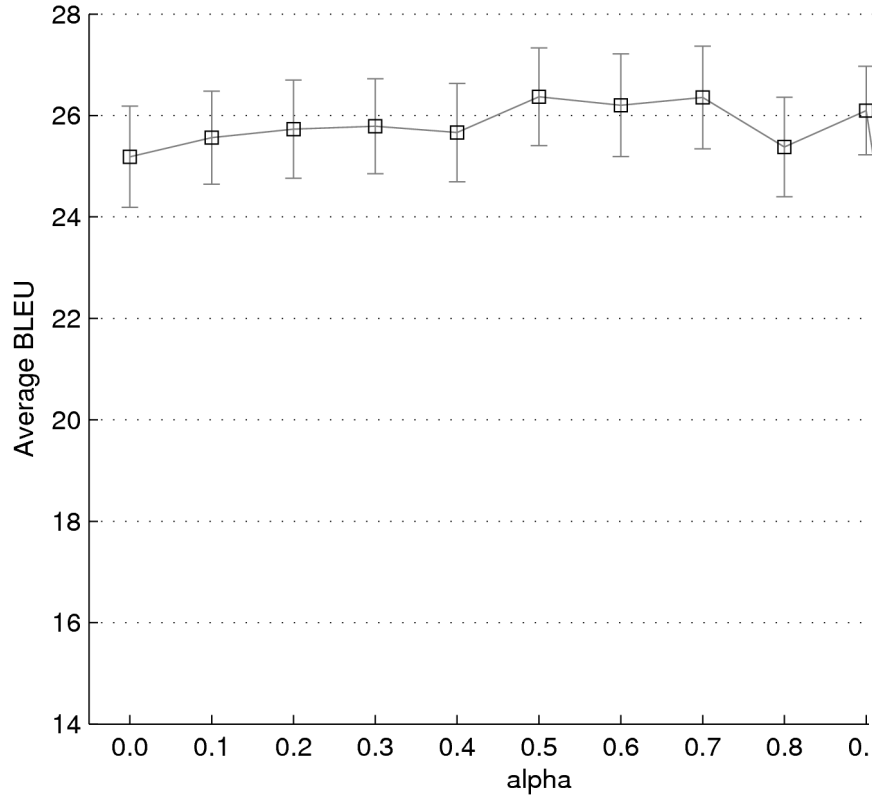
Results

Table 1. Experimental results presented by alpha and number of training sentence pairs. For each registry we have the average BLEU of the 30 translation problems (\bar{x}) and their standard error ($\sigma_{\bar{x}}$).

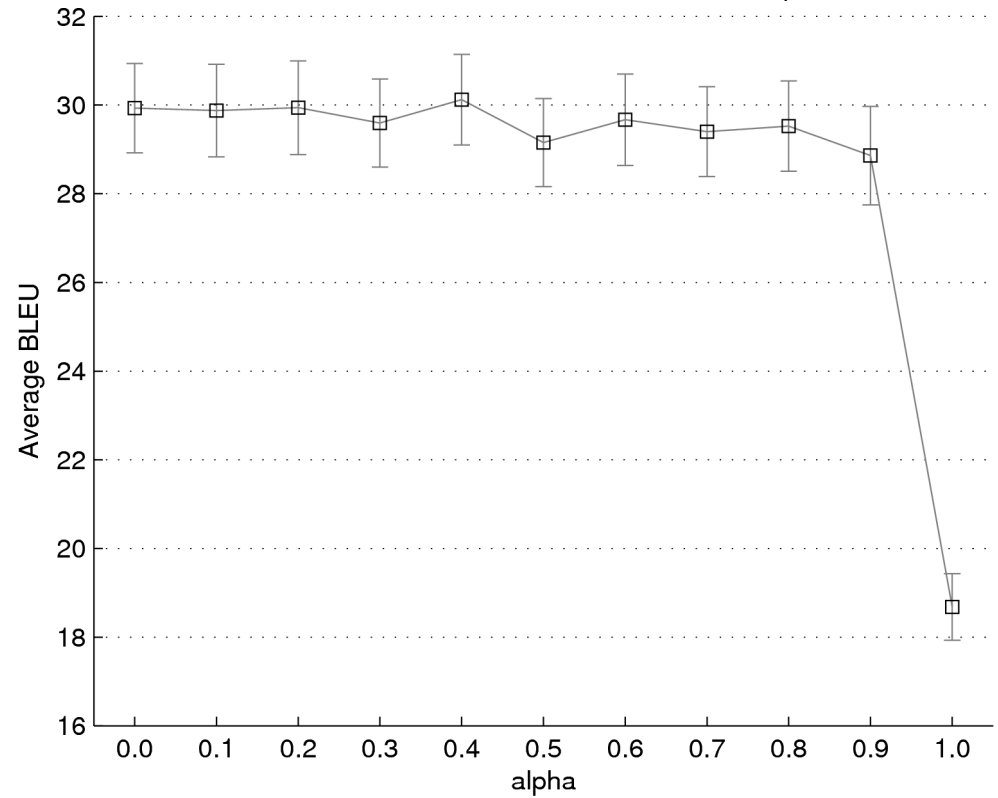
α	10k		20k		40k		80k		100k	
	\bar{x}	$\sigma_{\bar{x}}$	\bar{x}	$\sigma_{\bar{x}}$	\bar{x}	$\sigma_{\bar{x}}$	\bar{x}	$\sigma_{\bar{x}}$	\bar{x}	$\sigma_{\bar{x}}$
0.0	24.01	0.43	25.19	0.49	27.58	0.44	29.93	0.49	30.60	0.48
0.1	24.22	0.46	25.56	0.45	28.08	0.48	29.88	0.51	30.51	0.51
0.2	24.39	0.46	25.73	0.47	28.17	0.47	29.94	0.52	30.40	0.48
0.3	24.36	0.46	25.79	0.46	28.04	0.46	29.60	0.49	30.69	0.49
0.4	24.24	0.46	25.66	0.47	28.26	0.46	30.12	0.50	30.72	0.48
0.5	24.23	0.44	26.37	0.47	28.69	0.45	29.15	0.49	30.43	0.50
0.6	24.61	0.46	26.20	0.49	28.23	0.45	29.67	0.50	30.31	0.48
0.7	24.10	0.47	26.36	0.50	28.34	0.46	29.40	0.50	30.63	0.47
0.8	24.18	0.46	25.38	0.48	28.23	0.44	29.52	0.50	30.29	0.47
0.9	23.94	0.43	26.10	0.43	27.44	0.45	28.86	0.54	29.51	0.44
1.0	13.74	0.35	15.66	0.34	17.31	0.37	18.68	0.37	19.39	0.37

Confidence Intervals

Means and Confidence Intervals for Each alpha



Means and Confidence Intervals for Each alpha



Anova

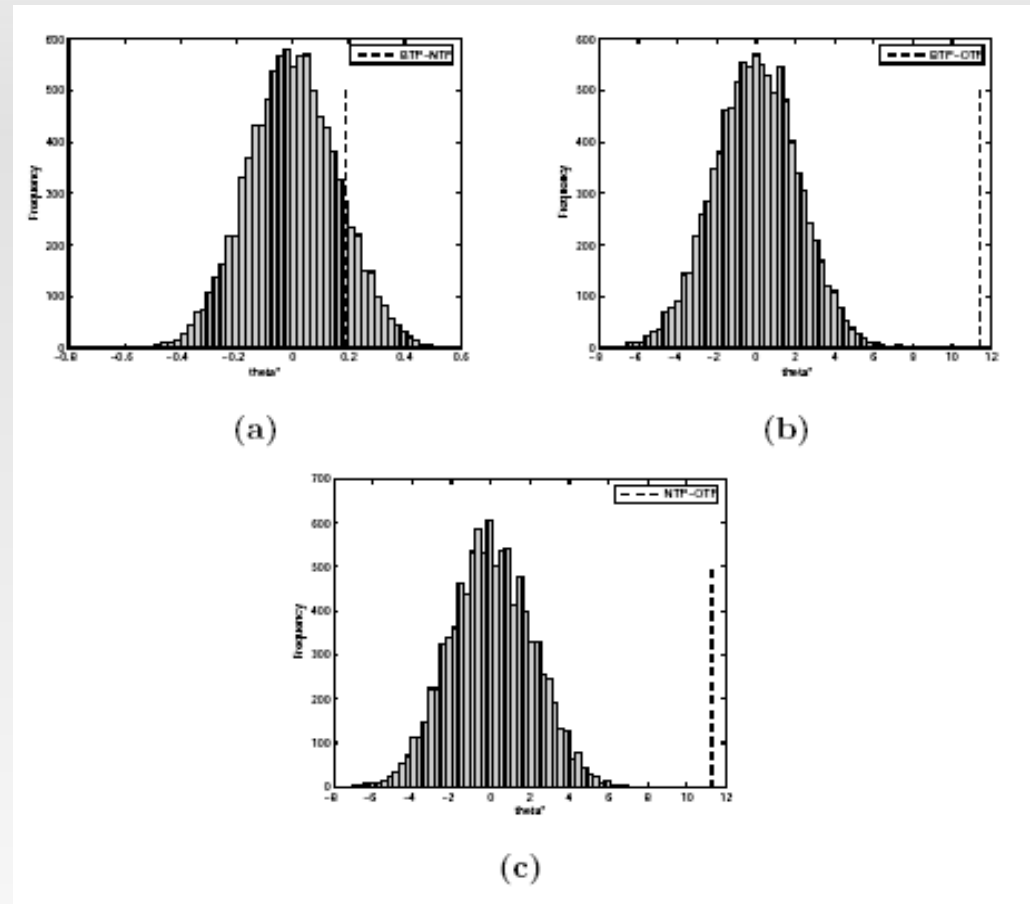
Table 2. ANOVA results for each k-group

k-size	Source	SS	df	MS	F	pvalue
10k	Groups	3011	10	301.1	51	0.00
	Error	1883.3	319	5.9		
	Total	4894.3	329			
20k	Groups	2866	10	286.6	44.87	0.00
	Error	2037.7	319	6.39		
	Total	4903.7	329			
40k	Groups	3213.4	10	321.34	53.13	0.00
	Error	1929.3	319	6.05		
	Total	5142.7	329			
80k	Groups	3297.9	10	329.79	45.32	0.00
	Error	2321.1	319	7.28		
	Total	5619	329			
100k	Groups	3343.5	10	334.35	50.05	0.00
	Error	2131	319	6.68		
	Total	5474.5	329			

Running pairwise comparisons

- Selected groups
 - Non-TP
 - Best-TP
 - Only-TP
- Algorithm

$$\theta = \sum_{i=1}^N (a_i - b_i) / N \text{ for } a_i \in S_A \text{ and } b_i \in S_B$$



Practical Comparisons

ORIGINAL	alpha 0	best alpha	alpha 1
that is why the commission has lent substantial and tangible support to the process of transition to democracy in chile .	for this reason the commission has brindado a process with considerable support and the transition to democracy in chile .	for this reason the commission has given a considerable support and with the process of transition to democracy in chile .	for this reason the commission has given a process with considerable support and the transition to democracy in chile .
mr president , i am very pleased to speak on behalf of the group of the greens / european	mr president , i am very pleased to be able to speak on behalf of the group of the greens	mr president , i am very pleased to act on behalf of the group of the greens	mr president , i am mucho am pleased to act on behalf the group of the greens
can assure you that the commission will honour this agreement in full .	can assure you that the commission will fully atener to this agreement .	i can assure you that the commission will fully atener to this agreement .	les i can assure you that the commission is going to completely atener this agreement

Pairwise Results

Table 4. Summarized results for pairwise comparisons presented by comparison and k-group

size-k	Non-TP vs Best-TP		Best-TP vs Only-TP		Non-TP vs Only-TP	
	θ	pvalue	θ	pvalue	θ	pvalue
10k	0.6003	0.0001	10.87	0.0001	10.268	0.0001
20k	1.1817	0.0000	10.71	0.0001	9.52	0.0001
40k	1.1	0.0000	11.38	0.0001	10.27	0.0001
80k	0.1910	0.1112	11.443	0.0001	11.252	0.0001
100k	0.121	0.1116	11.325	0.0001	11.204	0.0001

Conclusions

- As we increase training size TPs have a lower impact.
- For small training sizes, there is a significant improvement in translation quality by the utilization of TPs.
- TPs by themselves produce poor translations.

Future work

- Test with Scarce data.
- Test with out-of domain translations
- Assess intermediate language dependency of TP using information of language typology.