International Conference on Science Communication for Scientific Temper

Organised by: CSIR-NISCAIR, Vigyan Prasar and NCSTC
Date : 10 - 12 January 2012
Venue : NASC Complex, Pusa, New Delhi

Machine Translation and Communication

10th January, 2012
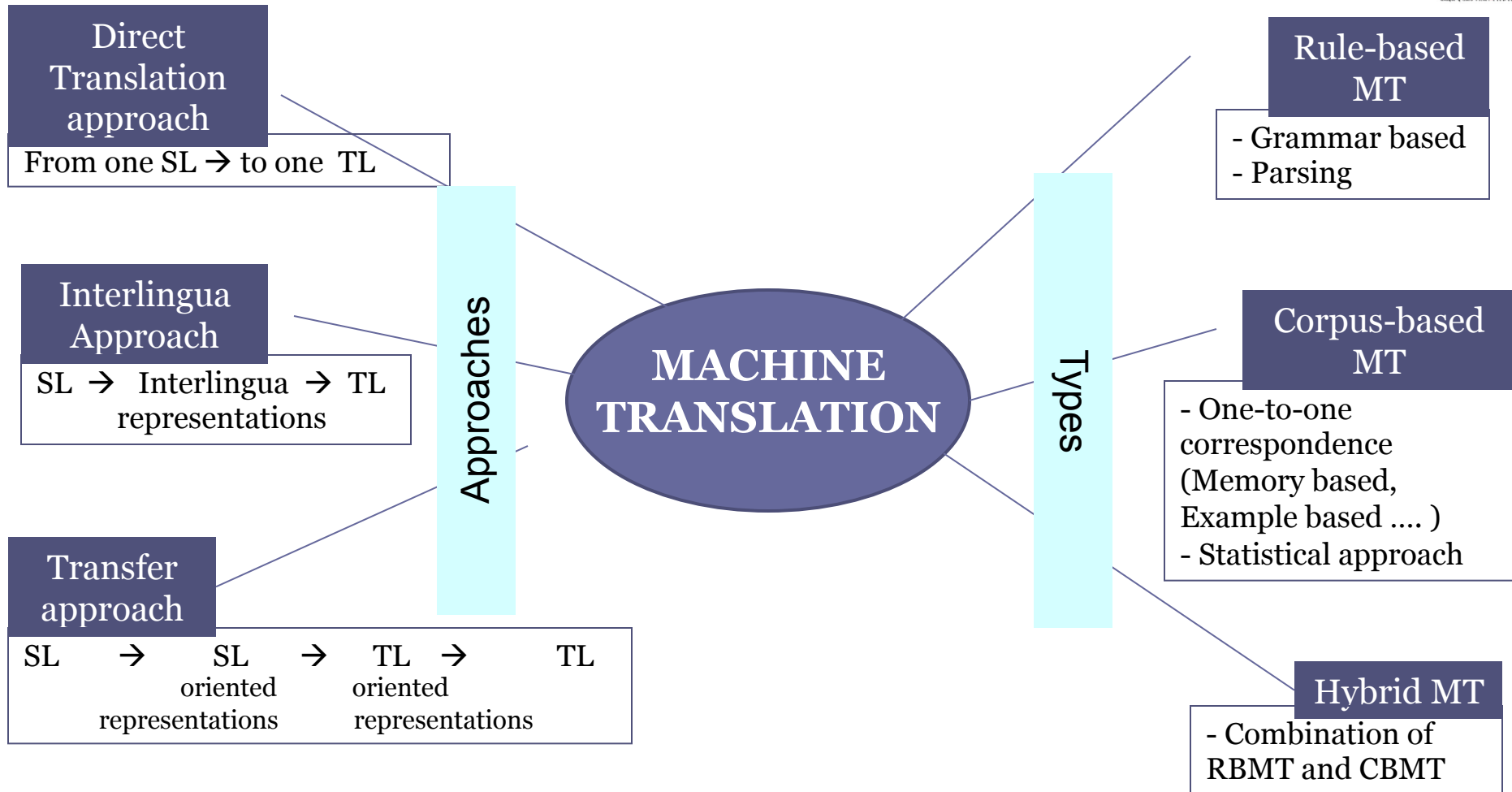**Ashok K Chawla,** Chief Scientist & Head, Translation Services Division, CSIR-NISCAIR, India
Rida Kumar, Trainee, On-going Translation Training Course, CSIR-NISCAIR, India
Puneet Jha, Trainee, On-going Translation Training Course, CSIR-NISCAIR & Practising Translator, NEC HCL, India

Machine Translation:

process by which computer software is used to translate a text
One natural language (SL) → another natural language (TL)

> Mechanical Dictionaries suggested → 17th century

> First attempt → two patents on mechanizing translation → early 30's
  - George Artsrouni's "Mechanical-Brain", 22nd July 1933;
  - Petr Smirnov-Troyanskii's "Translating Machine", 5th Sept. 1933

> Machine Translation noticed → 1949 through
  - Warren Weaver's memorandum, "Translation", published in
    Machine translation of languages: fourteen essays, 1955

> Introduction of mainframe computers in public R&D

> Huge money and efforts invested in Machine Translation
  - Many applications and online translation sites developed
    - o Online translation sites: Google, Bing Translate, Worldlingo, Systran, Babylon
    - o CAT tools: Atlas, Trados ....

Direct Translation approach

From one SL → to one TL

Interlingua Approach

SL → Interlingua → TL representations

Transfer approach

SL → SL oriented representations → TL oriented representations → TL

Approaches

MACHINE TRANSLATION

Types

Rule-based MT

- Grammar based
- Parsing

Corpus-based MT

- One-to-one correspondence (Memory based, Example based ....)
- Statistical approach

Hybrid MT

- Combination of RBMT and CBMT

Many Variant Approaches and Types Available, but
Practically all MT tools fall under three types
Rule-based, Corpus-based and Hybrid MT
Main Stream: Hybrid Approach (Currently)

Commonly Used Tools / Applications

Apertium	Free/open-source rule-based machine translation platform
OpenLogos	Free/open-source version of the historical Logos machine translation system
Anusaaraka	English-Hindi Machine translation system
Moses	Statistical machine translation
SDL Trados	Computer assisted translation software (CAT). Provides translation management software, content management and language services
Google Translate	Free statistical machine translation by Google Inc.
Bing Translation	Statistical machine translation technology, developed by Microsoft Research
Babylon translation	Computer dictionary and translation program for Microsoft Windows
Systran	Hybrid machine translation (SMT) technology; is one of the oldest machine translation companies
Worldlingo Translation	Hybrid MT technology, MT partner in Microsoft Windows and Microsoft Mac Office

Some tools have Beta versions available free of cost online and full versions are marketed commercially

Experiments

Exp.	SL	Intermediary	TL	MT tools used	Process	Comparison with	No. of expressions	Purpose
1	EN	-	FR MT	5	MT of given expression using 5 online MT tools in TL; selection of best translated sentence out of the 5	Reverse v/s original	37	Simple MT function (alone)
2	JP	-	EN MNL	5	MT of given expression using 5 online MT tools in TL; selection of best translated sentence out of the 5	with best MT	19	Manual v/s MT translation
3	JP	FR	EN MNL	5	MT of given expression using 5 online MT tools in intermidiary lang. by matching with MNL EN; selection of best translated sentence out of the 5; MT of this sentence into TL using same MT tool	with MNL EN translation	19	Transition through intermidiary language
4	EN	-	JP MNL	5	Manual translation by native into TL; reverse MT of given translation using 5 online MT tools in SL; selection of best translated sentence out of the 5	with original EN	22	Native v/s Non Native translation
5	EN	-	JP MNL	5	Manual translation by Non-native; reverse MT of given translation using 5 online MT tools in SL; selection of best translated sentence out of the 5	with original EN	22	

Abbreviations:

Exp. Experiment No. MT Machine Translation FR French
SL Source language MNL Manual translation JP Japanese
TL Target Language EN English

Grading

• Specially evolved to assess Machine Translation vis-à-vis Manual Translation
• Other criteria available so far did not cover communication
• Evaluation has been made slightly mild (shift from anti MT to pro MT)

| Points	Evaluation criterion
5	**Natural:** Translation communicates the message perfectly.
4	**Good but not perfect:** There are no mistakes in grammar and meaning is conveyed without any need for speculation. However, selection of words may be disputable.
3	**Understandable:** Grammatical errors and mistakes in selection of words exist. However, meaning desired to be conveyed can still be somehow speculated. Miscommunication is possible.
2	**Poor:** Numerous grammatical errors and meaning of the generated output can be speculated but only partially. Communication fails.
1	**X:** Translation does not make sense.

Sample Examples & Evaluation Process Used in Experiement
## Complicated non-standard / non-specification sentences usually miscommunicate

Original	Intermediary Language	Best Machine Translated	Compared with	Grading	MT Tool	Comments
First, it is necessary to sign this agreement.		D'abord, il est nécessaire de signer cet accord. REVERSE TRANSLATION: Initially, it is necessary to sign this agreement.	Original	5	Systran	Original sentence is grammatically structured and is as good as pre-edited one.
実りのある論議を交わした.		Had fruitful discussion.	MANUAL TRANSLATION : Had a fruitful discussion.	4	Bing	Sentence structure too simple. Almost word-to-word translation.
Our Summit takes place at a time when our relations are in excellent shape. Today we have discussed new vistas of our cooperation.	我々のサミットは両国の関係は絶好調であるときに行われ、今日、我々は協力の新しい展望について話し合をしました。	As for our summits as for relationship of both countries when being the best condition, it was done, today, we spoke concerning the view whose cooperation is new and did combination.	Original	3	Worldlingo	Focus in the first sentence has shifted from relations to conditions. Style of translation in the second sentence could lead to miscommunication.
I am delighted to extend a very warm welcome to India to Prime Minister Noda and Madam Noda on their first visit to India.	野田総理と令夫人のインド初訪問を心から歓迎いたします。	We sincerely welcome the visit of India's first Prime Minister Noda lady.	Original	2	Google	Miscommunication or communication failure.
私のプロフィールを見ていただくと、大学を卒業してから1年間に本田に働きました.		I look at the profile of the university, I graduated from Honda to work in 1 years.	MANUAL : As you look at my profile, I worked for one year in Honda after graduating from university.	1	Babylon	No communication

Grading of Machine Translated expressions
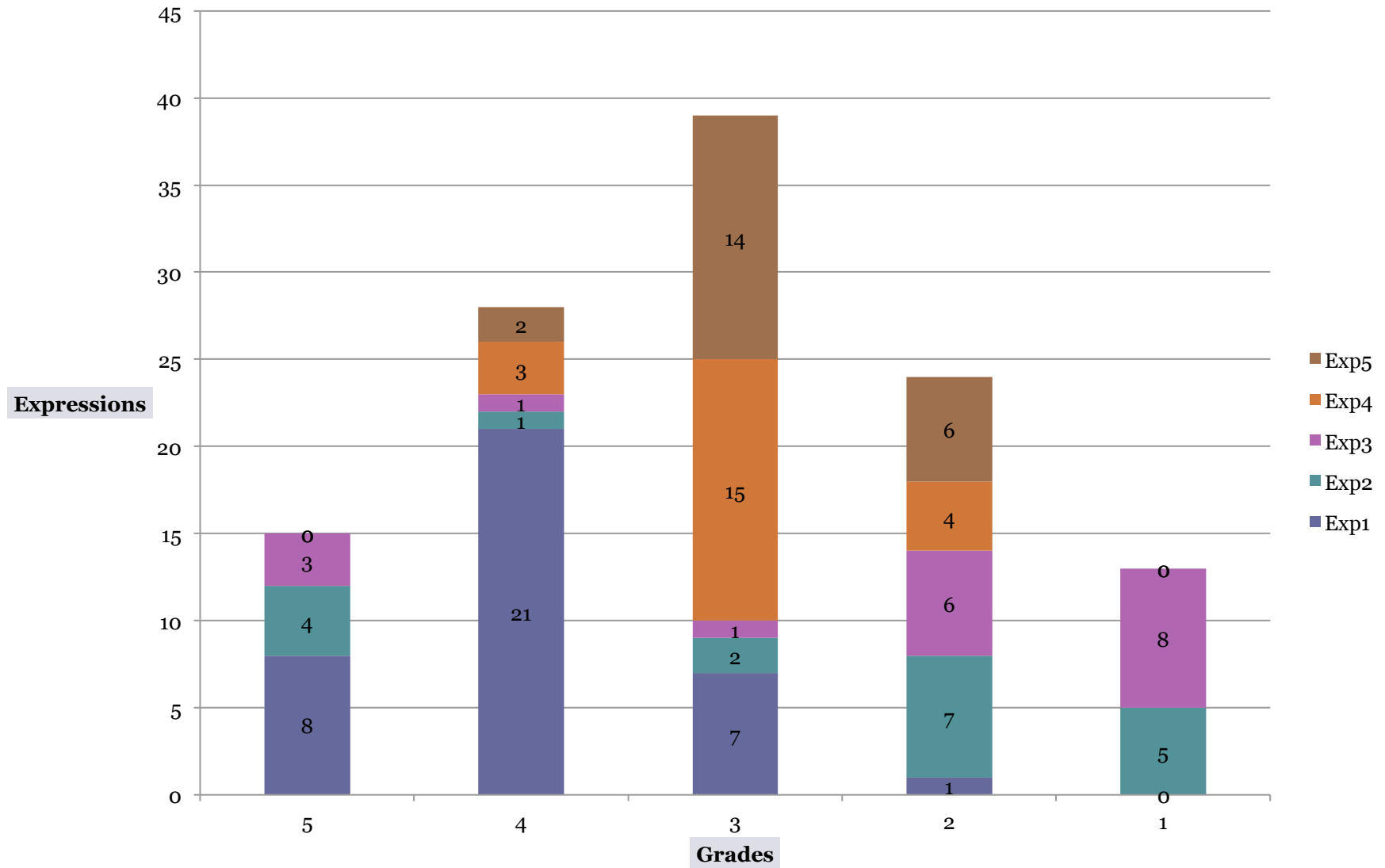
Performance of selected MT tools in experiments
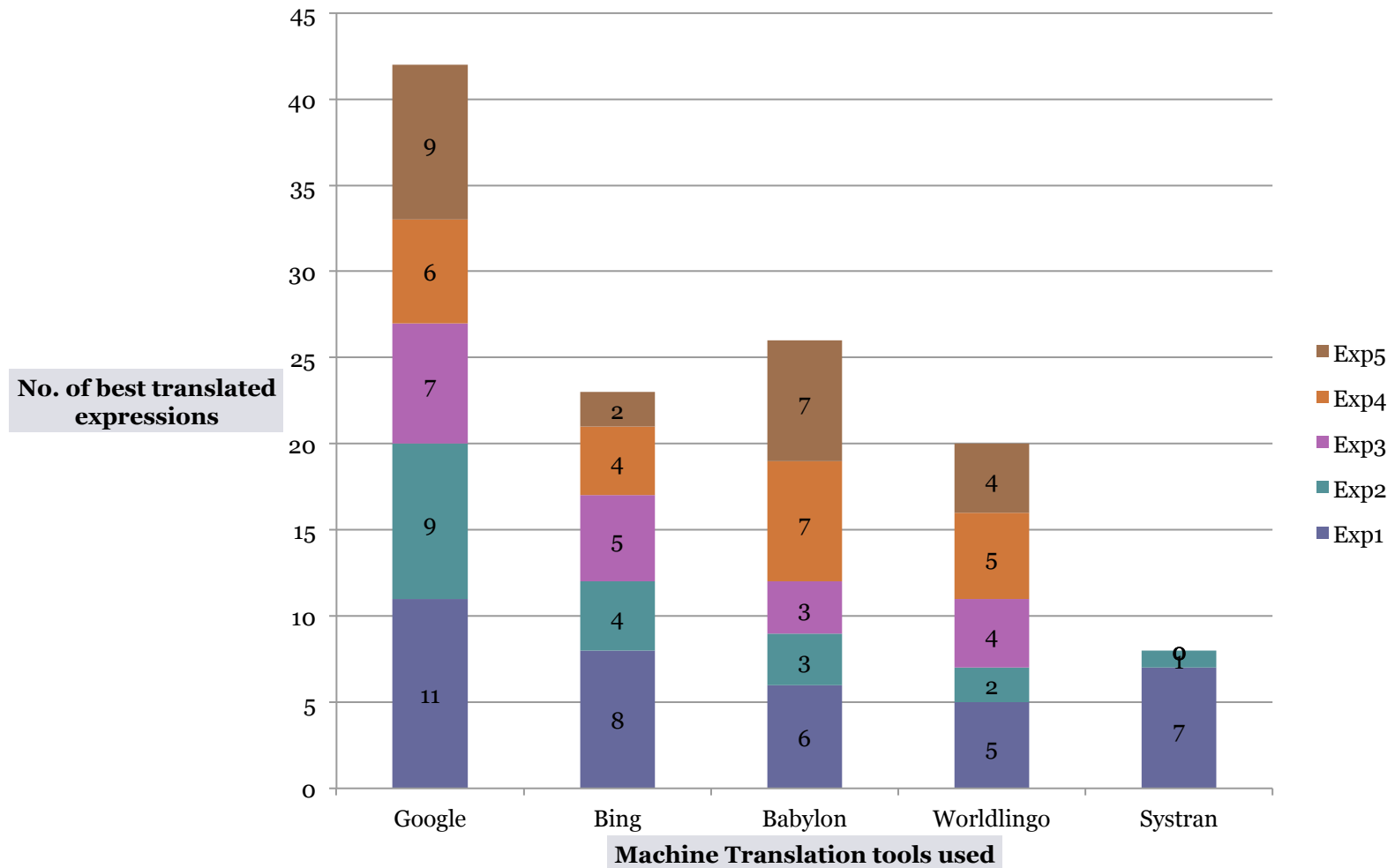
Findings

1. With these limited examples from limited domains
    - ✧ **Google appears to be performing best**

2. MT tools were totally non-productive & a burden till 1980's
    - ✧ **This study shows potential through selective use**
    - ✧ **Pre-editing can be a facilitator**

3. Productivity of Machine Translation should improve
    - ✧ **With the improvements in processing speed, memory, cognisance level, storage devices and capacities**
    - ✧ **With more usable terminology banks**

4. However, it may be out of proportion to say that:
    - ✧ **MT tools can replace Manual Translation in near future**

5. Communication
    - ✧ **Part-communication in few selective/domain specific cases may be possible**
    - ✧ **Stand-alone use can lead to miscommunication**

Limitations

1. Non-uniformity of tools
2. Limited number of examples
3. Difficulty in selection of uniform samples from different fields
4. Grading criteria standard not available

Proposal

1. Current study to continue with
   - More original examples of similar grades will be selected
   - Domain-wise or category-wise use of different tools
   - Grading criteria standard to be prepared

2. Machine translation tools may be used
   - In selective, intelligent, pragmatic and productive manner
   - To enhance quality, quantity and satisfaction level of manual translation and
   - For R&D purpose by practicing translators

Thank you

Please send comments / suggestions at
akchawla@niscair.res.in
kumar.rida@gmail.com
pjha21@gmail.com