



Open Source Software Development with Your First Language: Intercultural Collaboration Experiment 2002

Saeko Nomura, Toru Ishida, Mika Yasuoka

Naomi Yamashita, Kaname Funakoshi

Department of Informatics, Kyoto University

Japan Science Technology Corporation

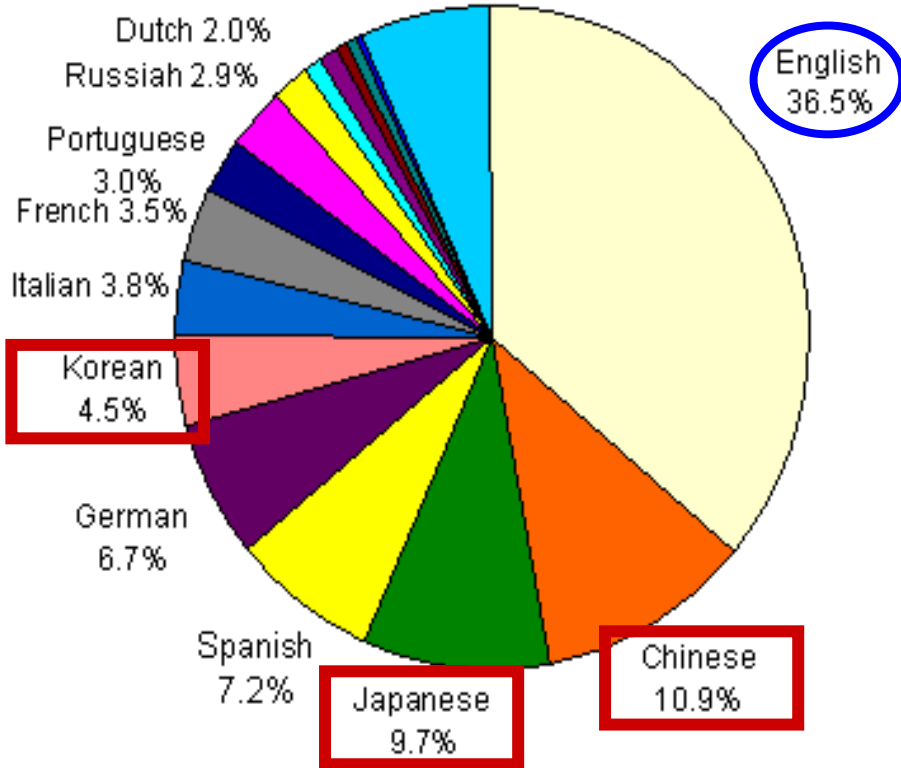
NTT Communication Science Laboratories

Presentation Outline

- Research background
- Brief introduction of **ICE2002**
 - Open source software development using machine translation
- Multilingual communication tools using machine translations
 - **TransBBS**, **TransWeb**
- Multilingual communication features on TransBBS
- Future research plans

Research Background

Online Language Populations
Total: 619 Million
(Sept. 2002)



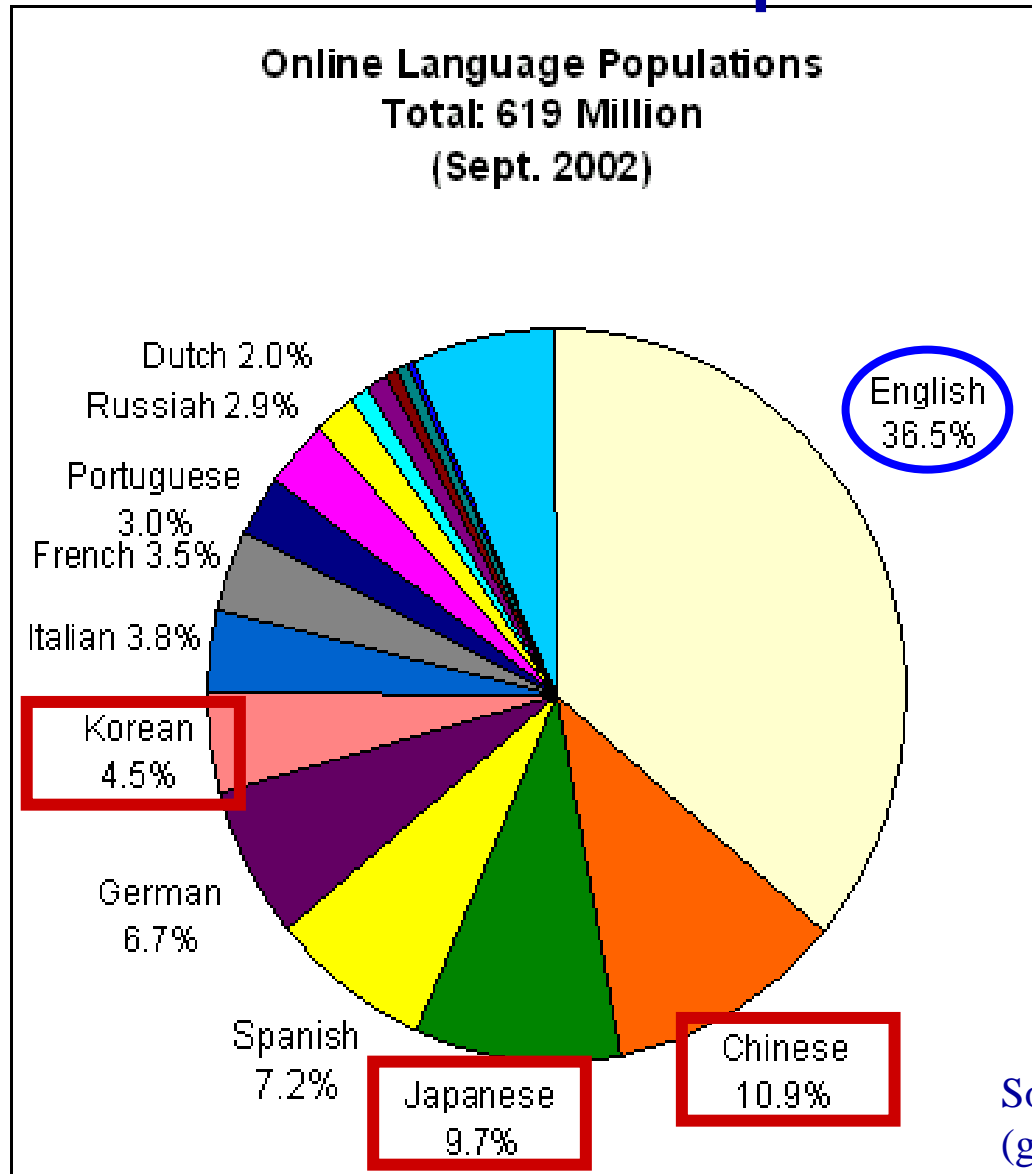
Source: Global Reach (global-reach.biz/globstats)

However...

Language Barriers Hinders Asia-Wide Collaboration

- Asian people do not learn neighboring languages.
- For most Asian people, it is hard to participate to multinational discussions that require high English skills.
- In Asia-wide collaboration projects, communication problems are so serious.

The upsurge of Asian Internet Populations



Source: Global Reach
(global-reach.biz/globstats)

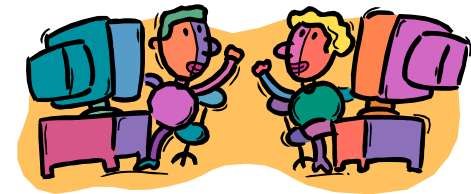
Machine Translation Can Be a Bridgebuilder?

- Current MT quality is considered as hardly adequate to be applied to practical human-to-human communications.
 - MT has been applied to and evaluated on *written documents*, such as news papers.
 - No “interactive” factors.



- It should be applicable to *human collaboration context* of *multinational project teams*.

- Activities which achieve a common goal.
- Communication context is pretty narrow.
- Collaborators share some technical terms among them.



Intercultural Collaboration Experiment 2002

- Collaborative Experiment among Universities in Asia

- **China:** Shanghai Jiao Tong University
- **Japan:** Kyoto University
- **Korea:** Seoul National University, Handong University
- **Malaysia:** Malaya University

- Task in the Experiment:

- **Open source software development** with communicating through machine translations.

- Assignment: “Develop software which supports intercultural collaborations”



Multilingual Collaboration Tools (1)

Copyright (C) 2002 Japan Science and Technology Corporation.

The screenshot shows the TransBBS web interface. At the top, the title "TransBBS" is displayed in large red letters. Below it, there is a navigation menu and a language selection table. The table has columns for "Interface Language" and "Message Language" with options for Chinese, Japanese, Korean, Malay, and English. Below the table is a form for composing a message, with fields for "Discussion", "Name", "Action", "Object", and "Where". A red box highlights the "Discussion" field, with the text "1. Input message in your first language" overlaid. Another red box highlights the "Post" button, with the text "2. Translate, Repair, then Post the message" overlaid. At the bottom, a red box highlights a list of posted messages, with the text "3. Posted Messages are Displayed" overlaid. The messages are displayed in a table with columns for language (C, J, K, M, E) and content.

Bulletin board system
(discussion space among
software development
team members)

Translation among
Chinese, Japanese, Korean,
Malay, and English.

Translation services are provided by



NTT MSC Sdn. Bhd.

NTT Cyber Space
Laboratories

powered by
KODENSHA

Multilingual Collaboration Tools (2)

Copyright (C) 2002 Japan Science and Technology Corporation.

TransWeb

Knowledge Sharing Tool
(Software development documents browsing tool)



Translation among
Chinese, Japanese, Korean,
Malay, and English.

Translation services are provided by

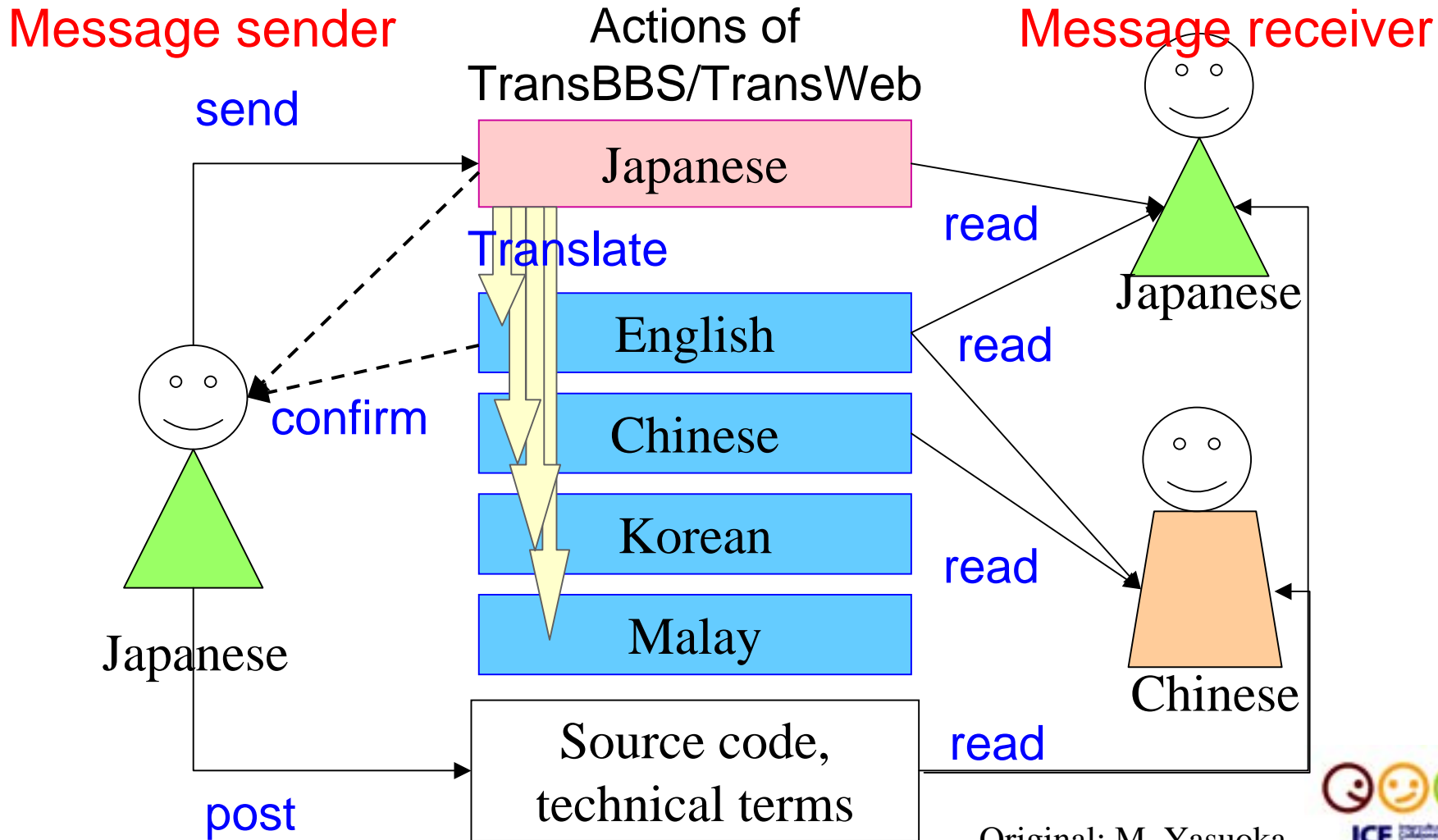


NTT MSC Sdn.Bhd.

NTT Cyber Space
Laboratories

powered by
KODENSHA

User Actions on TransBBS/TransWeb



Poor Translation Quality

“TransBBS was a *noisy* channel”

japana-82 || 投稿者: calvin sheng || [document](#)

Subject: 翻译效果好象不太好啊

投稿日時: 2002年05月21日 火曜日 12時54分08秒



日本語(retranslate)	英語(retranslate)	中国語	韓国語(retranslate)
翻訳した内容は私 はあまりわからない ことを見る	The content translation comes out I see to don't quite know	翻译出来 的内容我 看不太懂	번역한 내용은 는 그다지 모르 것을 본다

Translated Messages

Original Message

Human Adaptation to Machine Emerged

- Repair translation (*self-initiated repair*) again and again before posting
 - Create translatable sentences, even if these sentences are *not* smooth in natural usages

japana-149 || 投稿者: FUJISHIRO Yoshiyuki || [document](#)

Subject: We want to know whether you got my message(japana-148)

投稿日時: 2002年05月23日 木曜日 13時58分42秒



日本語	英語(retranslate)	中国語(retranslate)	韓国語(retranslate)	Malay(retranslate)
<p>あなたが私の投稿(japana-148)を理解できたかどうかを私たちは知りたい。あなたが理解した部分をあなたの言語を使って投稿してください。また、</p>	<p>We want to know whether you could understand the my contribution japana-148 or not. We, please contribute using your language the part that you understood. I want us</p>	<p>你能理解与否 我的投稿 (japana-148)我们想知道。你理解的部分使用你的语言请投稿。同时，对(japana-148)的你的意见也</p>	<p>당신이 나의 투고(japana-148)을 이해할 수 있었던 것인가 아닌가를 우리들은 알고 싶다. 당신이 이해한 부분을 당신의 언어를 써서 투고해 주십시오. 또한, (japana-</p>	<p>Kami hen whether a dapat tela japana-14 saya, or n contribute anda, kam bahagian y awak tela</p>

Human-to-human Collaboration Emerged

- To improve translation quality
 - Collaborative translation repair
- Even though translation quality is pretty low,
 - Active discussions on technical topics
 - *Program source* performs as the achieved consensus checker.

malay-252 || Posted by: Puja Lakshmi || [document](#)
 Subject:
 Date Posted: Monday, June 3, 2002 13:39:33

Japanese(retranslate)	English	Chinese(retranslate)	Korean(retranslate)	Malay(retranslate)
メール横断は、attachment..will that 翻訳 また で、もしメールの持ち意志、内容翻訳のについて、および持たせているよいidea.Butですか？	Transmail is a good idea. But how will the content be translated and if the mail has an attachment..will that be translated too?	Transmail是一个好主意。但是如何将内容被翻译和如果邮件有一份附件。。也被翻译的意愿？	메일 횡단은, attachment.. will that 번역 포이어서, 만약 메일의 가정 의 지내용번역의에 대해서, 미처 갖게 하고 있는 좋은 idea. But 입니까?	Transmail adalah suatu cadangan baik. Tetapi diterjemahkan bagaimana akan kandungan dan jika surat mempunyai suatu lampiran .. Akan yang diterjemahkan juga?

japana-254 || Posted by: FUJISHIRO Yoshiyuki || [document](#)
 Subject:
 Date Posted: Monday, June 3, 2002 19:53:43

Japanese	English(retranslate)	Chinese(retranslate)	Korean
Re: malay-252. 添付ファイルは翻訳されないでしょう。なぜなら、それは非常に難しいからです。	Re: malay-252. The appending file will not be translated. Because it is very difficult.	Re: malay-252. 附加文件不被翻译吧。要是为何，那非常难。	Re: 화일 요, 오 대단 니다

japana-272 || Posted by: Won Je Yang || [document](#)
 Subject:
 Date Posted: Wednesday, June 5, 2002 16:02:57

Japanese(retranslate)	English(retranslate)	Chinese(retranslate)	Korean
TransMailの場合 貼付ファイルの問題は そのように(そんなに)重要に見られる かもしれませんが。SMTPを経る前に Wrapper役割をして くれるserverを経たら 十分に可能だと 思うと見ます。	You can be seen or there is not a problem of the sticking file in the case of TransMail importantly like this (so). It sees that will be possible sufficiently if passing through server to raise by playing Wrapper role to pass through SMTP.	TransMail机动警 察右粘贴文件疑 制那样(那样)重 要河事是波特不 有。经过SMTP前 相暴的多面充分 地可能下吏la短 发啦 Wrapper作用乙 亥焦耳server.	TransMail의 경 찰부 파일의 문제 는 그렇게 중요하 게 보이지 않는군 요. SMTP를 거치 기 전에 Wrapper 역할을 해 줄 server를 거친다 면 충분히 가능하 라라 합니다.



Software Submitted in ICE2002

Modules

TransSearch (Group China)

- Translate query terms and send it to “google.” Search results are browsed in TransWEB.

TransMail (Group Japan)

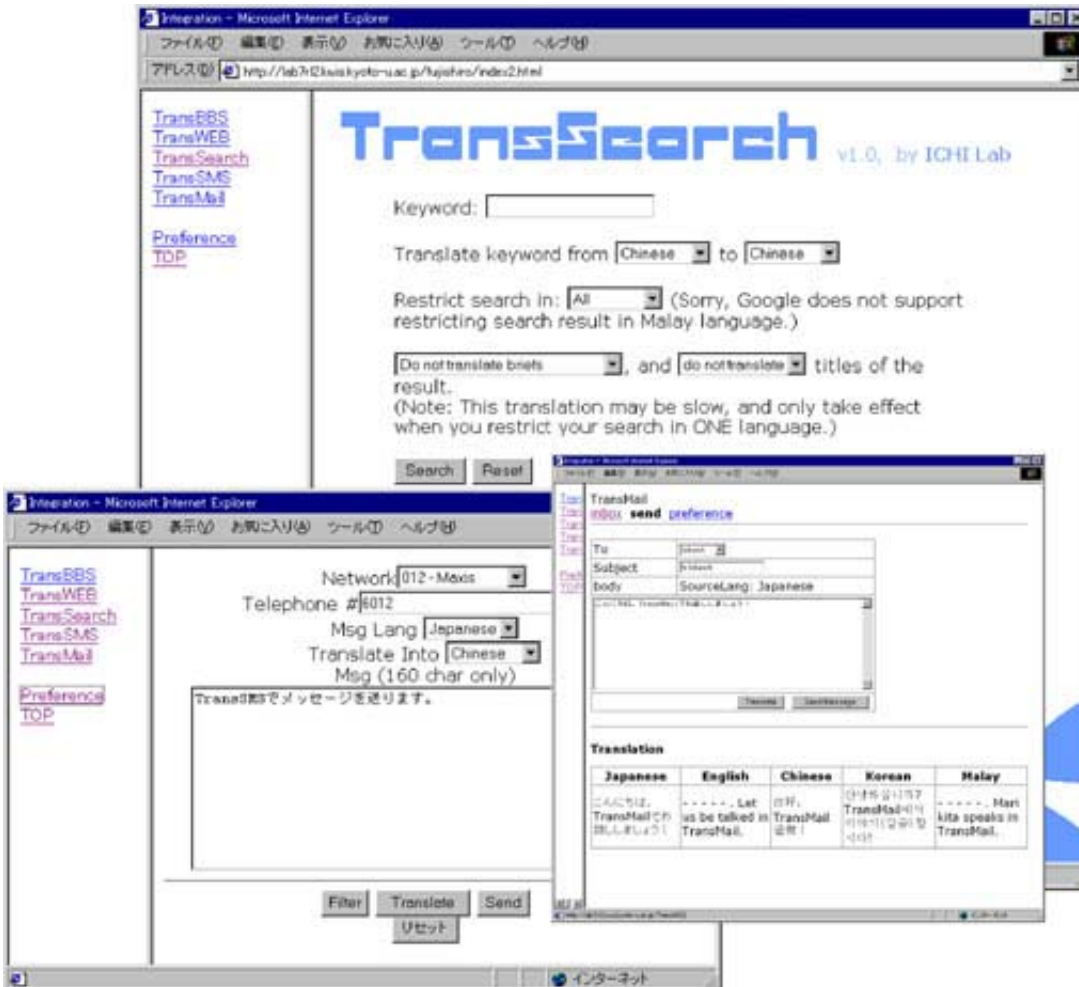
- Web base email service with translation function.

TransSMS (Group Malaysia)

- HTML-interfaced short message service for PDA users with translation function.

TransChat (Group Korea)

- Multilingual Chat Space



Research Issues (1)

Support *Self-Initiated Repairs*

Analysis on user mental model toward MT

- Little satisfaction with the translation results even after repeated self-initiated repairs...
 - ➔ The mental model of MT is hard to be built by users.
- Toward a user-centered MT systems in the context of multilingual communications...
 - Joint research with natural language processing researchers to design user-centered MT systems
 - Controlled experiment and cognitive walkthrough to understand mental models toward MT

Research Issues (2)

Support *Collaborative Repairs*

Development of *nonverbal* collaborative repair function

japana-7 || Posted by: Tomoko
Koda || [document](#)
Subject: Howdy
Date Posted: Thursday, May 1, 2003 13:59:09 (+0900)

[Cool Opinion!](#) [Translation is OK](#) [Terrible Translation](#)

[Need Evaluation](#) 0 / 0 / 0 ([detail/edit](#))

Chinese(retranslate)	Japanese	English(retranslate)
我从4月成为了ICE项目的成员。我的名字是Tomoko Koda。	私は4月からICEプロジェクトのメンバーになりました。私の名前はTomoko Kodaです。	I became a member of an ICE project from April. My name is TomokoKoda.

[Reply](#)

Message receivers: drag the mouse over the problem part.

System: drags the problem part in an original message.

A message sender: repairs it until getting “OK” by receivers.

The system: learns the protocol and gives suggestions from next time.

MT as a “Social Agent”

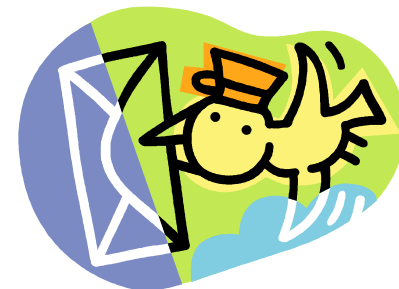
“**Interactivity**” as a new evaluation index of MT!

c.f.) conventional index: “**Accuracy**”

- How correctly MT can convert original language to target language
- ICE applies MT to a *collaboration context*, not only to information reception.
- ➔ MT should not be a *transparent* noisy channel, but a *social entity* which enables multilingual communication

- **Interactivity**

- Easy to conduct repair protocols
- Reliable message transmission to receivers
- Easy to build right MT mental model etc. etc.



Thank you!



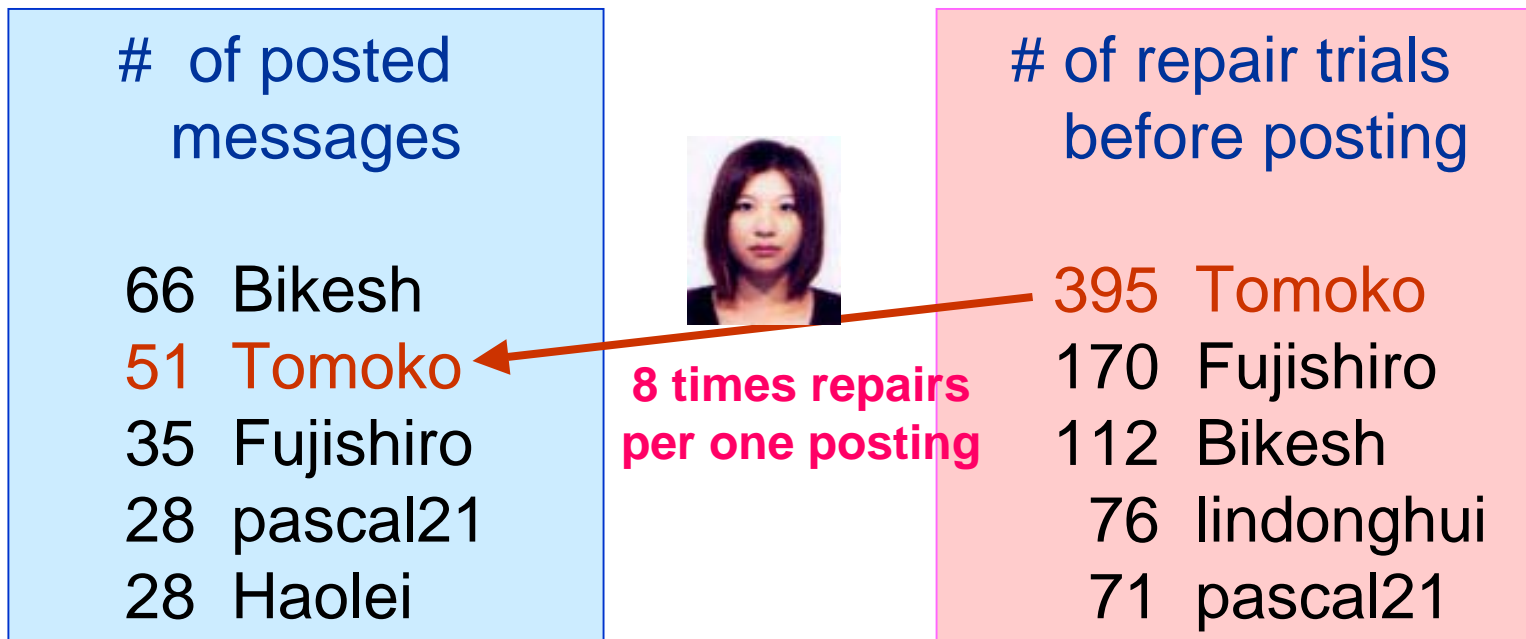
ICE Intercultural
Collaboration
Experiments

<http://ice.kuis.kyoto-u.ac.jp/ice/>

saeko@kuis.kyoto-u.ac.jp

Translation Repair (Self-initiated repair) Trials Before Posting to TransBBS

In a collaborative context, participants want to convey their messages correctly to other team members.



Data on TransBBS from Oct 10th to 30th

Intercultural Collaboration Experiment 2002

- Collaborative Experiment among Universities in Asia
 - China: Shanghai Jiao Tong University
 - Japan: Kyoto University
 - Korea: Seoul National University, Handong University
 - Malaysia: Malaya University
- Task in the Experiment:
 - **Open source software development**
 - Assignment: “Develop software which supports intercultural collaborations”
- Schedule
 - May-July, October-December 2002

Japan



ICE2002 Organizers



Korea



ICE Intercultural Collaboration Experiments



Malaysia



China