

---

# Collaborative Translation by Monolingual Users

**Chang Hu**

Human-Computer Interaction Lab  
Computer Science Department,  
University of Maryland  
College Park  
USA  
changhu@cs.umd.edu

**Abstract**

This paper describes a research effort to support collaborative translation by monolingual speakers, or people that speak only the source or target language. I hypothesize that sharing knowledge across the language barrier is possible with a combination of automated (but poor quality) machine translation, language-independent communication, and existing background knowledge. I demonstrate this possibility with proof-of-concept experiments.

**Keywords**

Collaborative translation, monolingual, computer-aided translation, machine translation

**ACM Classification Keywords**

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

**Introduction**

English is losing its leading position to other languages on the Internet [3,5], but the translation among those languages and English is falling behind. A high-quality machine translation (MT) engine requires so much training data that it has not been possible to build one to date. On the other hand, human translation often

---

Copyright is held by the author/owner(s).

*CHI 2009*, April 4 – April 9, 2009, Boston, MA, USA

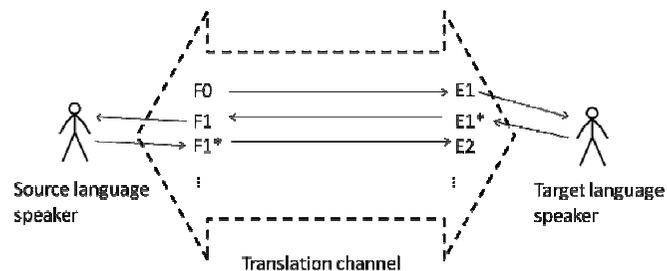
ACM 978-1-60558-247-4/08/04.

relies on highly-trained bilingual individuals, so it involves high costs in both labor and time [7].

Humans master languages with much less training data. No individual has read the whole Web, yet just about everyone speaks at least one language fluently. Besides, there are far more fluent monolingual people than professional translators. Therefore I propose an approach that combines low quality, but readily available and inexpensive machine translation with single language expertise among humans to create high-quality translation.

### System

I propose a system that improves translation quality over a poor translation channel by negotiation between two participants with imbalanced language skills (Figure 1). The system consists of two people on either end of a translation channel that don't speak the other's language. One user enters text in the source language (foreign language, or F) at one end of the channel, and the user at the other end tries to produce translated text in the target language (English, or E). I make the



**figure 1.** Round-trip translation system, solid arrows represent translating/editing

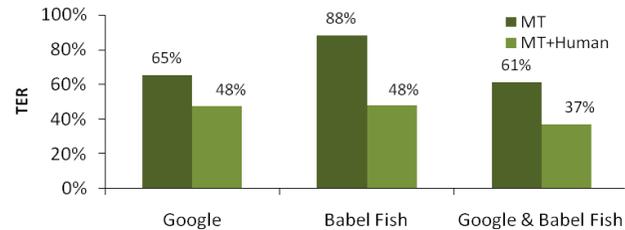
following assumptions about the users and translation channel:

- **The translation channel is poor**, whether it is a single or multiple machine translation engines, one or many amateur human translators, or any combination. However it does provide better-than-random translation.
- **The translation channel's internal details are not exposed to the users.** Unlike human-enhanced MT systems in the Linear B experiment [1], users only see the input/output of the channel.
- **The users' language skills are imbalanced.** Users on either side are experts in one language but have poor or no skills in the other.

Translation using this system often takes a few round-trips. In each round, the source language speaker sends the source text through the translation channel, whose output is edited by the target language speaker and sent back. Informal questions and answers for clarification can be sent through the channel, too. Alternatively, either participant can point to public context e.g. related online pictures, Wikipedia pages, etc. There is also support for a fixed set of communications and actions. For example, participants can agree on the next step in a language-independent way. They can either take one more round-trip or stop with success/failure. Any message sent across the translation channel other than the text in question is called a comment. This system is in line with Translingual Instant Messaging (TrIM) [4], which enables users to chat across the language barrier, except that my approach is aiming at translating a particular piece of text through round-trips.

Protocol for multiple engine experiment:

1. Every Chinese sentence (**F0**) in the data set was sent through the machine translation channel (one or two MT engines) and translated into an English sentence **E1**.
2. **E1** was edited by the English speaker, producing the edited sentence **E1\***. If two translation engines were used, the user was shown outputs from both engines and edited them into one sentence.
3. A bilingual user generated the corresponding ground truth sentence (**E0**) by looking at both **E1\*** and the Chinese sentence.



**figure 2.** Average TER between sentences **E1** and ground truth **E0** (lower is better).

### Hypotheses

I have the following hypotheses:

- **Editing by monolingual users improves translation quality.** Even monolingual target users can reason out much of the intended meaning of imperfect automatically-translated sentences (as proved by Callison-Burch [1] and TrIM [4]).
- **Redundancy improves translation quality.** Natural language includes natural redundancy, which is one of the factors that make this overall approach possible. More redundancy can be added by using redundant translation engines since they will make different kinds of errors. It is also possible to add further redundancy through comments.

### Experiments

I conducted two formative experiments to test these hypotheses. In the first experiment, two translation engines (Google and Babel Fish) were used to translate from Chinese to English, followed by human editing. In the second experiment, two monolingual users collaborated using an MT engine and comments to

translate from Chinese to English. Both experiments showed encouraging results.

I used 25 randomly sampled sentences from Chinese Wikipedia articles for either experiment respectively. None of these sentences contained non-Chinese words or characters. There were two monolingual participants in the experiment, a Chinese speaker and an English speaker. Similarity was measured between ground truth and all the English sentences with Translation Edit Rate (TER) [6]. The more similar two sentences are, the smaller TER is.

The results for experiment #1 (Figure 2) showed that monolingual human editing improved the similarity to ground truth. It also showed that using both translation engines together improved translation quality, compared to using any one of the engines alone.

The results for experiment #2 (Figure 3) showed improvement between each of the round-trips. The difference between human-editing improvements in the two experiments is due to different sentences used.

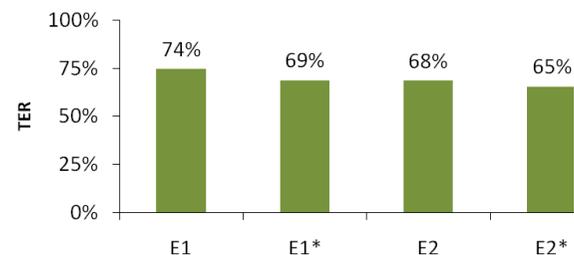
### Discussion and Future Work

These early experiments give evidence to support my hypotheses about human editing, redundancy and comments. They shed some light on the feasibility of the round-trip translation system.

The translation system can be seen as two participants negotiating an invariant (i.e. the meaning of the sentence) via a translation channel. Each participant has a perfect language model; their knowledge of the translation model varies. At the extreme, neither participant knows the other language, thus neither has

Protocol for round-trip experiment:

1. Every Chinese sentence (**F0**) was translated by the MT engine into an English sentence (**E1**).
2. The English speaker edited the machine translated version (**E1**), asked questions and attached other comments, producing **E1\***.
3. The edited English version (**E1\***) was translated back into Chinese along with comments (**F1**).
4. The Chinese speaker edited the back-translation responded to comments, and/or added comments (**F1\***).
5. The edited Chinese sentence and comments (**F1\***) were translated by the MT engine, producing **E2**.
6. The English speaker edited the translation, producing **E2\***.
7. A bilingual translator generates the ground truth **E0**.



**figure 3.** Average TER between English sentences and ground truth (lower is better). See the protocol and definition of each result. Briefly, **E1** shows MT; **E1\*** shows (**E1**) + human; **E2** shows (**E1\***) after another round-trip with MT; and **E2\*** shows (**E2**) + human.

any knowledge of the translation model. It has been proved that improving fluency on the target side alone can improve translation quality [1]. Since monolingual speakers have perfect language models, they can improve translation quality without any knowledge of the translation channel.

Language-independent comments may bring greater improvement since they are not affected by the translation channel. In future experiments, I am going to include such comments through a dedicated user interface. For example, if word correspondence between source and target sentences can be established, users can then point to specific parts of a sentence. The UI will also include language-independent action and communication templates. Such interfaces can not only support monolingual users, but are also beneficial to users with various language skills. I am also working on a theoretical model for collaborative monolingual translation.

## Acknowledgements

I would like to thank my advisor Dr. Benjamin B. Bederson. This work is a close collaboration with Dr. Philip Resnik and Dr. Jimmy Lin, researchers on computational linguistics. I would also like to thank Alexander J. Quinn for his comments.

## References

- [1] Callison-Burch, C., Linear B System Description for the 2005 NIST MT Evaluation Exercise, <http://cs.jhu.edu/~ccb//publications/linear-b-system-description-for-nist-mt-eval-2005.pdf>
- [2] Category: Available translators in Wikipedia, [http://en.wikipedia.org/wiki/Category:Available\\_translators\\_in\\_Wikipedia](http://en.wikipedia.org/wiki/Category:Available_translators_in_Wikipedia)
- [3] Internet World Users by Language, Internet World Stats, <http://www.internetworldstats.com/stats7.htm>
- [4] Maybury, M., Griffith, J., Holland, R., Damianos, L., Hu, Q. and Fish, R., Virtually Integrated Visionary Intelligence Demonstration, MITRE technical papers, [http://www.mitre.org/work/tech\\_papers/tech\\_papers\\_05/05\\_0140/05\\_0140.pdf](http://www.mitre.org/work/tech_papers/tech_papers_05/05_0140/05_0140.pdf)
- [5] Sifry, D., State of the Blogosphere, April 2006 Part 2: On Language and Tagging, blog entry, <http://www.sifry.com/alerts/archives/000433.html>
- [6] Snover, M., Dorr, B., Schwartz, R., Micciulla, L., and Makhoul, J., A Study of Translation Edit Rate with Targeted Human Annotation, Proc. of Association for Machine Translation in the Americas, 2006.
- [7] Wikipedia entry "translation", <http://en.wikipedia.org/wiki/Translation>