[*Translating and the Computer 22*: Proceedings of the Twenty-second international conference…16-17 November 2000. (London: Aslib, 2000)]

Reconciling User Expectations and Translation Technology to Create a Useful Real-World Application

Raymond S. Flournoy
Amikai Inc.
635 Tennessee Street, #405,
San Francisco, CA 94107 USA
Email: ray.flournoy@amikai.com

Christopher Callison-Burch
Amikai Inc.
635 Tennessee Street, #405,
San Francisco, CA 94107 USA
Email: ccb@amikai.com

Abstract

In this paper we discuss the motivations for the development of Amikai's web-based translated chat room application. Like other successful machine translation (MT) systems, Amikai's translated chat attempts to reconcile overly-optimistic user expectations with the limited capabilities of current MT technology by adjusting user expectations and limiting the scope and domain of the translation task. We explain why chat is a natural application for MT technology and briefly describe aspects of the user experience within the Amikai system.

Introduction: Making Useful Machine Translation Products

The history of natural language processing in general, and machine translation research in particular, has been marked by cycles of heightened interest and expectations within the general public, followed by periods of disillusionment when people's grand dreams for the technology are not realized. One of the most dramatic of these was probably the ALPAC report in 1966[1], which effectively killed MT research in the US for years to come. It is undoubtedly this process of repeated unfulfilled expectations which has caused many people - linguists, AI researchers, and laypersons alike - to have a jaundiced view of machine translation's potential usability in everyday life.

However, MT does in reality have some successful and commercially sustainable applications, most notably as translation helpers for human translators, such as the TRADOS system [2] or IBM's TranslationManager [4], and in the context of highly constrained subject matter and language usage, such as the Météo system for weather reports [3] or NTT's ALTFLASH business wire translator [7]. These types of applications share a common trait: they represent tasks where users' expectations have been modified to match the specific strengths and limitations of the MT technology, and the MT systems themselves have been limited in scope, specifically tailored to a narrow set of user needs. The mismatch between exaggerated expectations and limited technology is resolved, and useful systems result.

Amikai Inc. has built upon this idea of focusing the technology and adjusting user expectations to produce a multilingual, translated chat application. The system

provides a translation back-end to a classic internet chat room, allowing speakers of different languages to converse freely in real time over the web.[1] Through initial user tests, the system has proven highly usable, enabling satisfying, productive interactions between people around the globe, in spite of often inexact or completely incorrect translations. The success of the translated chat application can be attributed to three factors, all of which involve bridging the gap between user expectations and the state of current MT technology.

First, Amikai's system concentrates on chat, a forum in which content is cheaply produced and has a very short shelf life. In this context, exact and complete translations are not required; slight mistranslations or awkward phrasing are easily excused in the course of the dialogue. As long as meaning sufficient to maintain the flow of conversation is communicated, the MT system has successfully done its job. In addition, the Amikai MT system is tuned for the specific domain of chat communication, taking advantage of the large amount of repetition and the highly stereotyped interactions of chat language to improve the usefulness of translation in the chat rooms.

Second, the system trains users to understand better the strengths and limitations of the MT engines. The training mechanisms include a tutorial on translation-friendly language, as well as an interactive tutorial daemon (currently in development) which critiques a user's input. In addition, the chat interface to the MT engines introduces users to machine translation in a playful, hands-on environment where experimentation and exploration are encouraged. By enabling users to gain experience rapidly with an MT system, Amikai's translated chat quickly gives users a more concrete sense of the usefulness of the machine translation technology.

Finally, Amikai's translated chat improves the level of translation through user-to-user feedback in the form of a "Huh?" button, which a user can click when he or she doesn't understand another user's translated output. This immediately lets a user know when a translation was not of sufficient quality, prompting the user to try to find a better phrasing which will be more easily translated by the engine. This feedback loop is especially important for monolingual users who otherwise would have no way to measure the success of the translations of their own statements. In this way, the MT technology is made less mysterious and more personal, and users are better equipped to find the best ways to use the technology for their own communication needs.

---

[1] Amikai's translated chat is available through the Amikai website: www.amikai.com. Chat between English, French, German, Japanese, Italian, Portuguese, and Spanish are currently supported, with Korean and Chinese debuting shortly.

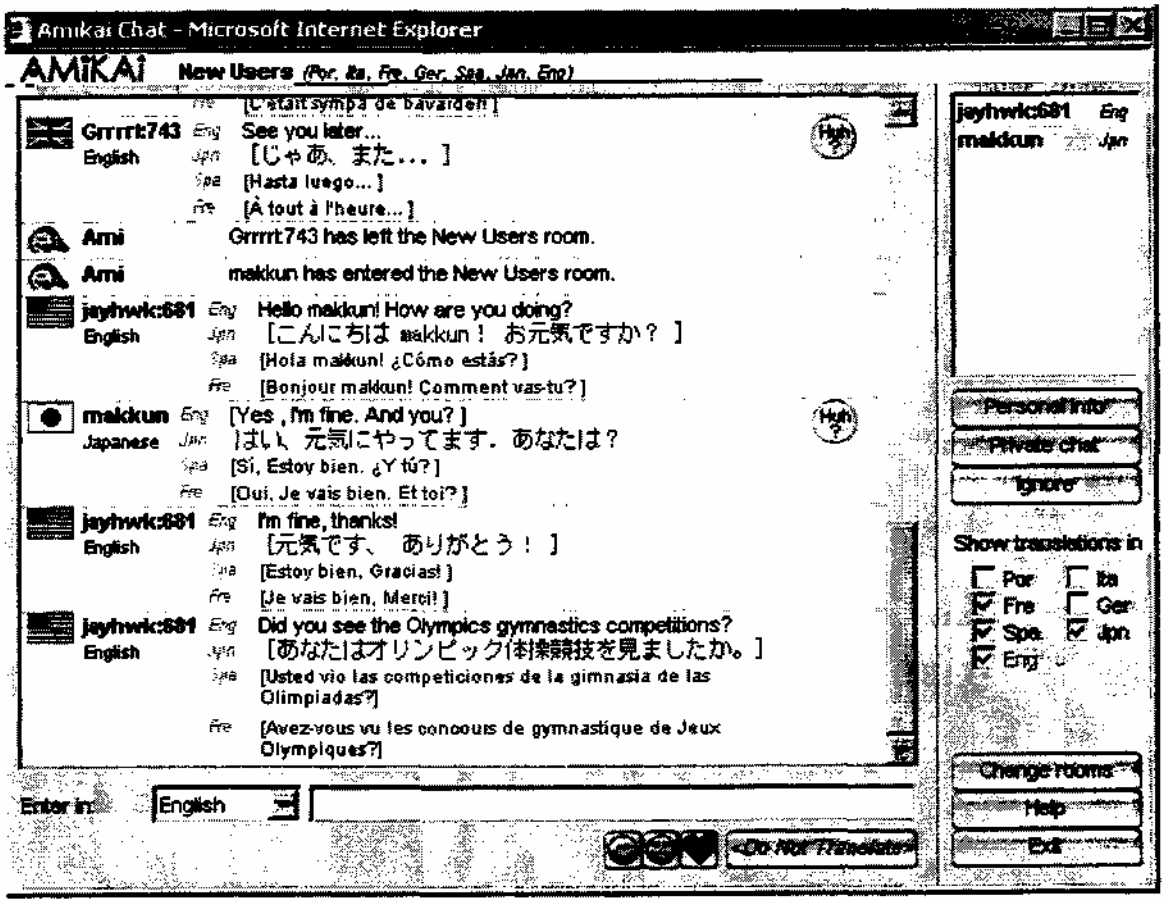

Figure 1. Amikai's Translated Chat Application.

The Chat Domain

Internet chat is an ideal application for machine translation for two main reasons: chat does not require a high level of precision, and the domain features a high-level of repetition and highly stereotyped interactions which allow the system to use a variety of heuristics to improve performance.

Chat can be called a "short shelf-life, high turnover" product. Chat inputs are cheaply produced and quickly discarded; furthermore, chat is inherently an interactive process in which information is repeatedly requested, transferred, clarified, and acknowledged. Thus when mistranslations happen, rephrasing and repetition can occur without great cost or inconvenience, and the imprecision and awkwardness of machine-produced translations do not critically affect the understanding and maintenance of the dialogue flow. For these reasons, chat does require the high level of translation quality that most other MT applications require.

Added to this, chat language is characterized by a high level of repetition and stereotyped interactions which allow a machine translation system to be focused and biased for more satisfactory performance. General purpose MT systems are expected to be able to interpret the breadth of language phenomena, a daunting task, so traditionally MT systems have sought to improve performance primarily through constraining the language they cover to grammatically well-formed, formal language. Colloquial language has conventionally been considered too intractable and unpredictable for processing. However much of the variation in chat language is

basically typographic in nature - nonstandard capitalization, spelling, abbreviations, accenting, and punctuation. In three sets of chat data across three different languages (English, Spanish, and Japanese) we found that after allowing for this type of surface variation, the top 300 most frequent phrases covered approximately 15-19% of the data. (Table 1) In comparison, a similar analysis of English newspaper articles showed only 1.3% of the sentences covered by the top 300 phrases. (Table 2)

Language	Lines of input	% in top 300 phrases
English	135250	18.4%
Japanese	153799	17.9%
Spanish	50733	15.3%

Table 1. Percentage of chat data in top 300 phrases, across languages.

Language	Lines of input	% in top 300 phrases
English	88613	13%

Table 2. Percentage of newspaper data in top 300 phrases.

On top of the repetition, the conversations in chat tend to follow fairly scripted interactions, such as greetings, questioning and answering, and summarizing and signing off. This often provides the MT system with more context to interpret the input more accurately. While theories about discourse analysis still have not been completely realized in working MT systems[2], the chat setting allows the Amikai system to apply certain heuristics to aid interpretation. For example, certain questions without subjects in Japanese are assumed to require a "you" subject, and ambiguous words such as "nice" or "great" are interpreted with their conversational senses.

Because of these two characteristics of chat, users are more willing and able to work with imperfect translations while the system itself is more focused to the specific chat domain, producing a more satisfactory and useful system.

User Training

A large factor in the usability of the Amikai translated chat comes from the fact that users gradually learn how to phrase their inputs to maximize the accuracy of the translations. This training-oriented approach is similar to that employed by Palm Computing's *Graffiti* system [6]. The success of *Graffiti* over other handwriting recognition system is due to the fact that it requires that user input be constrained. The Palm application circumvents the problem of recognizing fluid, natural handwriting by requiring the user to input characters as discrete units rather than as continuous text, and to learn a specialized alphabet rather than using an individual writing style. These constraints increase the accuracy of the recognition dramatically, and therefore create an overall user experience that is positive enough to make the

[2] See [5] for examples of approaches to discourse analysis in the JANUS system.

constraints seem negligible. Because older handwriting recognition systems failed to fulfill user expectations of high accuracy on freeform writing, they were by contrast much less usable.

Amikai's translated chat application similarly attempts to focus the application of MT technology and adjust user expectations. By training users to constrain their input, we produce a marked improvement in the quality of translations over those produced for standard chat room dialogue. By giving users reasonable expectations about the quality of machine translation and reasonable constraints to maximize the quality, Amikai's translated chat becomes a useful tool.

The user training involves a number of components. As an introduction for chat room users who are initially unfamiliar with machine translation, we provide a set of translation tips for each language that give examples of language usage that does not translate well. For example, the English translation tips section instructs users to form complete sentences, avoid slang and idioms, beware of the word "got", because it has too many different meanings, and keep sentences simple and short, among other recommendations. Each tip is accompanied by examples of translation-friendly and translation-unfriendly sentences. These tips are a passive means of training users to better understand the limitations of machine translation, and to constrain their input accordingly.

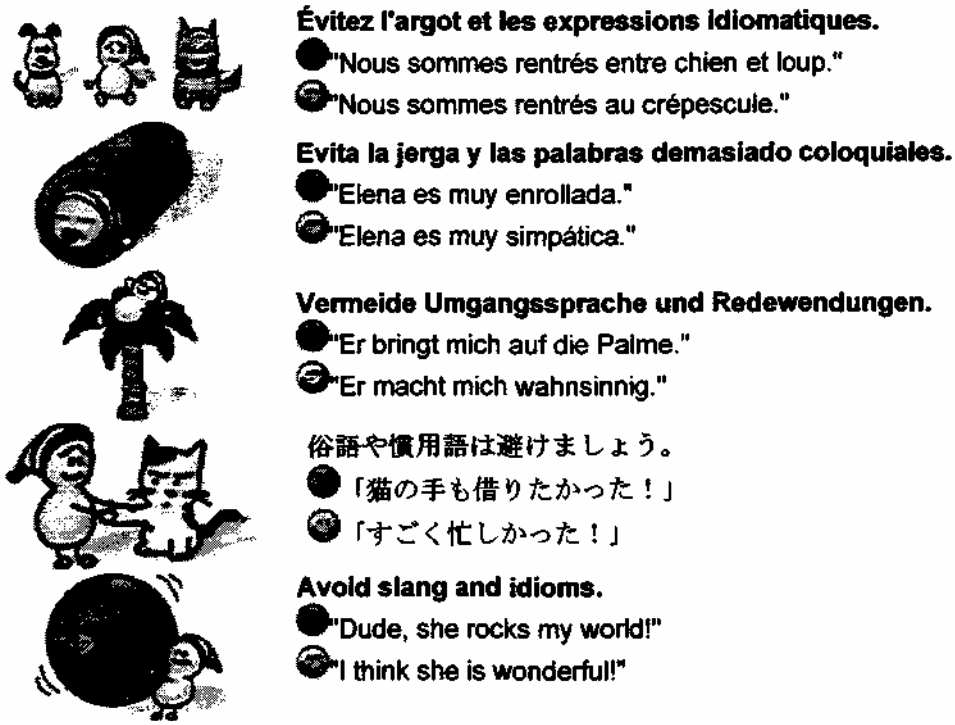

Figure 2. Example translation tips from French, Spanish, German, Japanese, and English.

The chat interface itself also fulfills a training role, as it provides a hands-on environment for users to experiment and actively explore the limitations of MT. In using the chat rooms, bilingual users are able to observe and correct bad translations. Because of this, bilingual users are the most quick to familiarize themselves with the type of input that translates well.

In addition, we are currently developing an interactive instructional component for Amikai's translated chat, which will give users feedback about how translation-friendly their input is, and will allow monolingual users to train in the same sort of active fashion that bilingual users do. The interactive tutorial component scans a user's input for trouble spots, such as misspelled words, zero anaphora, and sentence fragments while he or she is chatting. If it detects that an input contains one of these difficulties, then the user is warned that his or her input may not translate well, and is shown the translation tip that illustrates the suggested type of usage. Through this interactive instructional component as well as through the passive translation tips, users are encouraged to constrain their input to be more translatable. At the same time, their expectations about how well machine translation works and about when it is effective are adjusted to be in-line with the current state of the technology.

User-User Feedback

As mentioned earlier, chat is inherently an interactive process, so requesting rephrasing and clarification can be a natural part of the conversational flow. In order to encourage this, Amikai supplies a simple, clear mechanism for users to give each other feedback when a translation is not understandable. As with the previous two points, this user-user feedback serves the dual purpose of adjusting user expectations while making allowances in shortcomings in the MT technology, and therefore making the overall translated chat system more usable.

The mechanism for user-user feedback is a button - dubbed the "Huh?" button - placed next to each line of translated input, which when clicked sends a message to the line's author informing him or her that the line needs to be rephrased for retranslation. From the user's perspective, this serves as a reminder that the MT engines are not always perfect, while giving the user a mechanism for trying to remedy these imperfections. Especially for monolingual users this is an important tool for feeling more comfortable with the technology - you are not simply sending your translations out blindly, unsure whether the messages are being understood. Feedback from your conversational partner is quick and clear.

Ideally the MT engine would identify improperly formed input or ambiguous sentences and warn the user appropriately. But natural language processing technology is not at the state that it can identify these problem inputs with certainty, and arguably the technology might never be at that state. The "Huh?" button helps to make up for this shortcoming by helping users to identify these problem inputs for each other. What the system cannot do automatically, users are encouraged and enabled to do for each other. The "Huh?" button serves the additional purpose of identifying problem inputs for the system designers, helping to point out problems that need to be addressed in the next version of the system.

The "Huh?" button helps to give users more control over the system thereby increasing the utility of the overall MT application.

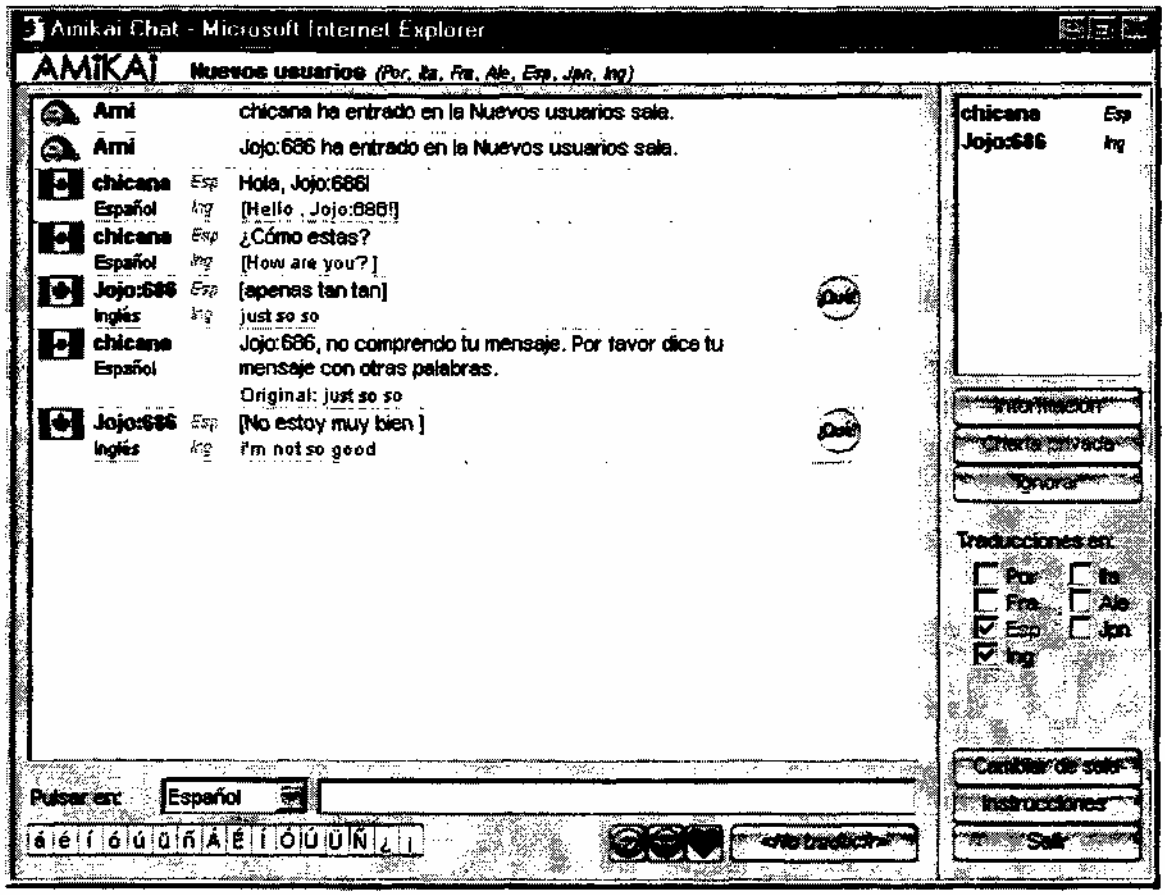

Figure 3. A demonstration of the "Huh?" button.
On the Spanish screen it is called a "¿Que?" button.

Summary

Amikai is attempting to follow the strategy used in other useful MT applications by choosing a manageable task for which high-quality translation is not required, tailoring the system for the task, and educating users about the strengths and weakness of the technology. Internet chat is such a task. By optimizing for the chat domain while reconciling the mismatch between users' high expectations and MT's current state, Amikai is building a tool which makes MT technology usable and useful for a large audience of users

Bibliography

[1] ALPAC. 1966. *Language and machines: computers in translation and linguistics.* A report of the Automatic Language Processing Advisory Council, Division of Behavioral Sciences, National Research Council. National Academy of Sciences, Washington, DC.

[2] Berry, M. 1992. "The Trados Translator's Workbench II", in *Proceedings of the 33rd Annual Conference of the American Translators Association,* San Diego, California.

[3] Chandioux, J. 1989. "Météo: 100 million words later", in *American Translator's Association Conference 1989: Coming of Age,* Medford, New Jersey.

[4] IBM. "TranslationManager", http://www-4.ibm.com/software/ad/translat/tm/tmfeat20.htm

[5] Lavie, A., Levin, L., Qu, Y., Waibel, A., Gates, D., Gavalda, M., Mayfield, L., and Taboada, M. 1996. "Dialogue Processing in a Conversational Speech Translation System", in *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96.* Fourth International Conference on Spoken Language Processing, Philadelphia, PA.

[6] MacKenzie, I. S., and Zhang, S. 1997. "The immediate usability of Graffiti". In *Proceedings of Graphics Interface '97,* pp. 129-137. Canadian Information Processing Society, Toronto.

[7] Uchino, H., Ooyama, Y., and Furuse, O. 1999. "ALTFLASH: A Japanese-to-English Machine Translation System for Market Flash Reports", in *Proceedings of MT Summit VII,* Kent Ridge Digital Labs, Singapore.