Supporting Local Jargon in Multilingual Collaboration
Extended Abstract for Demo

Mika Yasuoka\textsuperscript{1}, Toru Ishida\textsuperscript{1,2}, Yohei Murakami\textsuperscript{2}, Donghui Lin\textsuperscript{2}, Masahiro Tanaka\textsuperscript{2}, Rieko Inaba\textsuperscript{1}

\textsuperscript{1}Department of Social Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto
+81-75-753-4821

\textsuperscript{2}Language Grid Project, National Institute of Information and Communications technology (NICT)
3-5 Hikaridai, Seika-cho, Kyoto
+81-77-498-6907

SUMMARY
In this paper, we describe our demo system the Language Grid Toolbox. The Language Grid Toolbox facilitates multi-language-mediated collaboration by extending a normative machine translation functionality with a local dictionary function that enable users to create and maintain a dictionary for local jargon iteratively. We present the theoretical background of the system design and briefly describe our Language Grid Toolbox system.

Categories and Subject Descriptors
H.5.3 [Group and Organization Interfaces]: Computer supported cooperative work, Web-based interaction.

General Terms: Design, Human Factors, Languages.

Keywords: Multi-language mediated collaboration, local dictionary, local language, community jargon, common ground.

1. BACKGROUND AND MOTIVATION
Collaboration across national borders has drastically increased in response to the development of communication technology and the global economy. In a large number of such international projects, collaborators often gather with different native languages. Projects, however, are often conducted utilizing a single standard language, in many cases, English. In spite of such a predominant position in the use of a single language, many people have difficulty communicating and thinking in non-native languages which degrades their work efficiency and quality.

Lately, an increasing number of international organizations and multilingual communities have offered computational communication support with machine translation functionalities. Our multilingual translation support framework, the Language Grid\textsuperscript{[2]}, contributes in promoting machine translation usage in a wide variety of increasingly internationalized fields such as education, medical service, public city service and NPO activities. As easily imagined, language support in collaboration appears to be attractive to those who are required to use non-native languages for collaboration. Thus, the demand for language supports with computational artifacts has continuously grown.

In spite of such growing demand for the use of cutting edge technology to facilitate international projects, current machine translation has limitations in practical use. In addition to normative challenges of machine translation such as quality and reproducibility, communication mediated by machine translation has extra communicational challenges caused by inconsistencies and asymmetries of references in pre- and post-translation\textsuperscript{[6]} as well as fundamental technical challenges such as typographical mistakes. With accumulation of such mistranslations, collaborators face more critical difficulties in constructing and maintaining common ground among collaborators\textsuperscript{[7]}.

In order to support construction and maintenance of common ground in multilingual collaboration, we focus on local language or jargon known in communities of practice\textsuperscript{[5]} or semantic communities\textsuperscript{[4]}, respectively. Local jargon such as words, expressions and terms used in groups of people is socially constructed during the collaboration period or in a specific community. Local jargon can also be professional expressions acquired through professional education or fostered through professional experience in practice. Such local jargon contributes substantially to the construction and the maintenance of common ground. On the other hand, collaboration without sharing local jargon prevents collaborators largely from having common ground even if they speak the common language (e.g., English)\textsuperscript{[e.g.,4,5]}.

The importance of project jargon in multilingual collaboration is already reported in several cases. For example, the ICE project\textsuperscript{[3]} shows that machine translation mediated collaboration can be successful in spite of all language related difficulties if collaborators have a certain level of local jargon based on work culture such as computer science expressions. This and related experiments also reported collaborative repair activities as well as self-repair activities, which contributed to create local jargon. In such self-repair activities, participants improved the translation quality through iterative try-and-error interactions with the machine translation result. Many participants who conducted self-repair activities reported their findings to the community which improved general translation results\textsuperscript{[3]}. In collaborative repair activities, collaborators refined expressions through interaction with other members and confirmed meanings of expressions through the interaction process\textsuperscript{[2, 8]}.

This research indicates that consistent translation through the entire project period will play a beneficial facilitation role in...
developing and maintaining common ground. In addition, facilitating collective intelligence seen in conscious creation, maintenance and development of local jargon in collaboration will accelerate this process. All in all, offering such facilitation functionalities will contribute to multilingual collaboration.

2. SYSTEM
The Language Grid Toolbox (see Figure 1) is developed as one of the collaboration tools of the NICT Language Grid project. Originally, the Language Grid project was initiated as a reaction to the current situation in language services. Briefly described, currently quite a few languages services exist, however, they often have less usability and no flexibility to add new words, customize domain specific expressions or sentences, and combine several services.

Figure 1. The Toolbox and its local dictionary interface.
The Language Grid is developed as a multilingual service infrastructure, which connect several existing standards as well as local language services as atomic components and offer cross-language services with better accessibility and usability. Among several collaboration tools developed based on the Language Grid infrastructure, the Language Grid Toolbox offers a unique interface and further improved usability. The Language Grid Toolbox facilitates multilingual collaboration by offering a distinctive local dictionary function where users can create, develop and maintain a dictionary for local jargon iteratively. By doing so, the user can combine existing normative machine translation functionalities and their local jargon dictionaries.

3. CONTRIBUTIONS
It is only within the last decade that the number of global collaborations facing multilingual challenges has increased drastically. For that reason, understanding communication in multilingual collaboration is still quite limited compared to profound investigations and understanding communication and communicative processes in normative cooperative work accumulated in the Computer Supported Cooperative Work (CSCW) community.

Several critical multi-language mediated communication processes can be investigated. First of all, it is beneficial for the community to understand this emerging communication style, machine translation mediated communication. Second, it is a great contribution for the CSCW community to understand unknown aspects of grounding process and common ground in multilingual collaborations. For example, it appears useful to be able to translate local jargon such as identical key terms consistently through the entire multilingual project. However, to our knowledge, no investigation has been made with this perspective and thus our understanding is limited as for how such consistent translation of words, expressions and terms can contribute to the overall translation quality and how such consistent translation contributes to common ground. It is very fruitful to understand these key components deeper to achieve better multilingual communication. Lastly, it would be very insightful to investigate how local jargon multilingual dictionaries could be created and maintained by collaborators as seen in previously mentioned multilingual cases. This will contribute to understanding the nature of unique self-evolving community styles such as Wiki and open source communities.

4. ACKNOWLEDGEMENT
This research is partially supported by Kyoto University Global COE Program: Informatics Education and Research Center for Knowledge-Circulating Society.

5. REFERENCE