Design for Computer-Mediated Multilingual Communication with AI Support

Hajin Lim
Department of Information Science
Cornell University
Ithaca, NY 14853, USA
hl934@cornell.edu

Abstract
Machine translation is often not enough for people to engage across languages due to translation errors and lack of cultural background. In addressing these challenges, my dissertation explores how AI-augmented analytics can improve computer-mediated communication between speakers of different native languages. First, to support better sense making of foreign language posts in social media, I designed SenseTrans, a tool that adds contextual information using AI-analytics such as sentiment analysis. In my future work, I intend to explore 1) how people perceive, interpret and make use of AI-generated information, and 2) how AI-augmented analytics could be applied to other settings.

CSCW '18 Companion, November 3–7, 2018, Jersey City, NJ, USA

ACM ISBN 978-1-4503-6018-0/18/11.
https://doi.org/10.1145/3272973.3272982

Author Keywords
Multilingual Communication; Cross-lingual Communication; Machine Translation; AI-augmented Communication; AI-assisted Communication; social network sites; social media; Sense making

ACM Classification Keywords
Human-centered computing → Collaborative and social computing

Introduction
Advances in machine translation (MT) and computer-mediated communication (CMC) tools now allow people to interact across linguistic and cultural boundaries. However, these technologies often cannot suffice to support successful cross-lingual communication. Because of the imperfect quality of MT and people's lack of cultural and contextual knowledge, users still experience difficulties in communication with speakers of different native languages.

Through my doctoral research, I aim to explore design possibilities to improve multilingual communication based on grounded understandings of how people consume and make sense of foreign language messages and contents. I am particularly interested in incorporating artificial intelligence (AI) analytics such as natural language processing (NLP) into CMC tools to facilitate better multilingual communication. Such techniques have been mostly used to analyze large scale data (e.g., opinion mining) rather than individual use. My approach explores AI-based analytics as something that can also benefit our daily conversations especially in multilingual communication contexts.

Project Overview

As a starting point, I focused on multilingual communication in social network sites such as Facebook. I chose this focus because 1) it is relatively common to have some foreign connections in one's network and 2) social network sites often provide MT features that allow users to translate other's foreign language posts.

In the first phase of the research, I aimed to gain a holistic understanding of how people consumed and made sense of foreign language posts in social media streams using both quantitative and qualitative research methodologies.

Study 1: Understanding how people consume and make sense of foreign language messages (Completed)
To investigate how language affects people's attention to and engagement with social media posts, I conducted an eye-tracking lab study [3] in which monolingual English speaking participants (N=22) viewed a mock newsfeed containing English and foreign language posts. The result showed that participants paid less attention to, and showed less willingness to engage with, foreign language posts compared to English posts.

I also conducted a qualitative interview study [2] examining participants' actual Facebook timelines (N=23) and probing the reasons that people were reluctant to pay attention to and engage with foreign language posts. Interviewees reported that they felt a lack of relevance and cultural knowledge when they tried to understand foreign language posts. Even with MT results, literal translation often could not provide sufficient understanding of posts due to the imperfect quality of MT and people's lack of contextual knowledge associated with each post. As a strategy to make sense of foreign language posts, some focused on simply extracting the emotional components of a post (e.g., emoji) while others tried to gain a fuller understanding of a post by searching the information about keywords.

Study 2: Design and Evaluation of AI-Augmented Interface to support better sensemaking of foreign language posts (Completed)
Based on the previous findings, I designed and developed a browser extension, SenseTrans [1]. that provides emotional and contextual knowledge in addition to MT to support better sensemaking processes surrounding social media posts. Aligned with the sensemaking strategies that interview participants reported (i.e., emotion vs. context focus), SenseTrans provided *emotional knowledge* about a poster's emotional states and *contextual knowledge* about the social and physical context of a post. By making use of available NLP techniques such as sentiment mining and named entity extraction, SenseTrans provides auto-

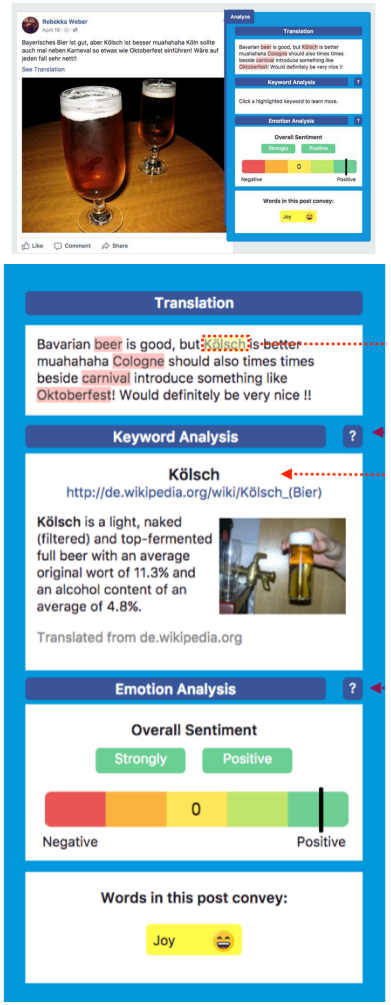

Figure 1 (Upper) SensTrans on a German Facebook Post (Lower) SeseTrans interface components

- Top: MT results and keyword highlights
- Middle: Retrieval of relevant information from Wikipedia
- Bottom: Sentiment and Emotion analysis

generated emotional and contextual annotations based on the text content of a status update.

The SensTrans interface consists of three parts (**Figure 1**). First, in the *Translation area*, users can see the machine translation result for the selected post. Next, the *Keyword analysis* provides contextual information about a post by identifying named entity words (e.g., place, important figure) in the post and highlighting those words on the MT result. When users click highlighted words, it displays relevant information about each word retrieved from Wikipedia. Lastly, the *Emotion analysis* provides information about the main sentiment and prevalent specific emotions in a post.

To examine whether providing emotional and contextual information about a post could improve people's sensemaking outcome, I conducted a between-subject laboratory experiment. In the experiment, native English speakers (N=24) were asked to browse five fictitious Facebook profile pages written in non-English languages using one of two interface conditions: the full version of SenseTrans, or a machine translation(MT) only version that presented the translated text. After browsing each profile, participants answered a survey about their perceived comprehension and willingness to engage with each post and overall cognitive loads in sensemaking.

The results were promising. Participants using SenseTrans reported significantly greater perceived comprehension of the posts, and greater willingness to engage with the posts by liking or commenting. Also, no additional cognitive load was associated with using an interface that provided more information. These results provide support for the idea of incorporating AI-analytics into multilingual communication settings.

Study 3: In the Wild Deployment Study of SenseTrans (Ongoing)
Study 4: Design and Evaluation of AI-augmented Interface for Supporting Multilingual IM Communication (Planned)
To explore the expandability of my approach, I plan to design and evaluate an interface that can support real-time multilingual conversation using Instant Messenger (IM). Comparing to previous settings (i.e., social media posts sensemaking), IM conversation is more synchronous and interactive, which requires more rapid and frequent message processing and creation. I am especially interested in exploring how to minimize the possibilities of misunderstandings between speakers. As demonstrated in SenseTrans, augmenting emotional and contextual information might be useful in this context too. In addition to that, I am investigating 1) what other types of information would be useful, and 2) how this information could be presented in IM communication contexts. After iterations of the design, I will conduct a user study to evaluate whether this new interface will help users to understand other's message better and to compose messages clearly to minimize potential misunderstandings in multilingual IM conversation.

Study 5: Building Framework in Supporting Computer-mediated Multilingual Communication with AI Support (Planned)
In this stage, I will incorporate the findings from my previous design processes and evaluation studies to build a framework about how we can support computer-mediated multilingual communication using

AI analytics. This framework will highlight design processes for incorporating AI analytics into CMC tools in multilingual communication settings. Also, it will examine how users perceive, interpret and make use of AI-generated information. Based on this framework, I hope to expand my research focus to other types of inter-group miscommunication (e.g., communication between different political views, different age groups, or different religions), by applying the design approach that aims to facilitate better communication process by augmenting AI into CMC tools.

Expected Contribution

My dissertation work will contribute to CMC, online multilingual communication, and AI-assisted communication research. With the advancements in MT technologies, my research will provide useful design implications for future CMC tools to support online multilingual communication effectively. Also, my design approach, augmenting online communication with AI analytics, can be expanded to other inter-group domains in facilitating the understanding gaps and minimizing potential miscommunications.

Objective for the CSCW Doctoral Colloquium

As a fourth-year Ph.D. student, attending the Doctoral Colloquium in CSCW will be one of the valuable milestones for me. As my research is deeply embedded to CSCW, I expect the feedback and advice from mentors and peers in CSCW communities will be helpful. Primarily, I hope I could gain insights into the unforeseen aspects and new possibilities of my research. I am also excited to contribute to other participants' ongoing projects.

References

1. Hajin Lim, Dan Cosley, and Susan R Fussell. Beyond Translation: Design and Evaluation of an Emotional and Contextual Knowledge Interface for Foreign Language Social Media Posts. In *Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems* (CHI '18).

2. Hajin Lim and Susan R Fussell. 2017. Making Sense of Foreign Language Posts in Social Media. *PACM on Human-Computer Interaction* 1, 2.

3. Hajin Lim and Susan R. Fussell. 2017. Understanding How People Attend to and Engage with Foreign Language Posts in Multilingual Newsfeeds. In *Eleventh International AAAI Conference on Web and Social Media*.