Instant Annotation: Early Design Experiences in Supporting Cross-Cultural Group Chat

Na Li[1], Mary Beth Rosson[2]

College of Information Sciences and Technology
The Pennsylvania State University
329B IST Building, University Park, 16802
[1]nzl116@psu.edu, [2]mur13@psu.edu

ABSTRACT

Cross-cultural group chat is an important option for supporting communication in both industry and education settings. However, studies of such interactions have reported persistent communication problems that appear to be due to mismatches in non-native and native speakers' language proficiency. With this problem in mind, we have been exploring a conceptual design called Instant Annotation. Our design concept supports a kind of threading in chat using annotation, thus offering para-communication support in cross-cultural group chat. As part of this design investigation, we studied native and non-native speakers in a group chat activity, shared the new design concept, and interviewed users to gather their feedback about the Instant Annotation concept. The results pointed to three different design use cases and led us to envision four general design features that we will explore in our ongoing work. We discuss the cross-cultural communication problem, findings from the interview study, the current design and future directions.

Categories and Subject Descriptors

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces: User-centered design

H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces: Synchronous interaction

General Terms

Design, Human Factors.

Keywords

Cross-cultural, native speaker, non-native speaker, scenario-based design, CSCW, group discussion, online chat, annotation, tagging.

1. INTRODUCTION

Cross-cultural communication is taking place everywhere as the world is getting flat. It happens in working places, classrooms, international conferences, online chat boards, emails, and more. Research conducted across many disciplines (e.g. education, psycholinguistics, sociology, HCI and CSCW) shows that language proficiency plays an important role in cross-cultural communication. In particular, non-native speakers often suffer from communication problems caused by language proficiency issues both in face-to-face communication and distributed online communication [6,18].

Non-native speakers receive and understand information more slowly; they need more time to organize expressions; and they frequently make grammatical errors, due to an increase in cognitive load when processing a second language [8]. In this context, comparison studies have found that text-based computer-mediated communication (CMC) can be beneficial to non-native speakers [24]. Text communication typically requires less immediacy in response than face-to-face or oral communication, thus alleviating the cognitive burdens of non-native speakers. With CMC tools, non-native speakers have more time to read others' expressions and generate their own expressions. Nonetheless, studies show that there are communication problems caused by non-matching levels of language competence in text-based CMC [7]. Many such problems have been reported in classroom or group studies [4,6,7], perhaps indicating that they are particularly prominent in multi-person communication settings.

A common problem found in those studies is a disruption in turn-taking, a flaw in existing text chat systems that has been identified in the literature [22]. Because there are no social cues indicating start and end of speaking as in face-to-face communication, speakers may send messages at the same time. Therefore, any given message may not be a response to the most recently received message but rather refer to a point made several messages earlier. This disruption of the sequential turn-taking system may generate communication confusions, particularly when overlapping threads are semantically related. The disruption might arise more easily in cross-cultural group chat due to the unmatched levels of language proficiency and differences in general communication styles.

Although researchers from different areas have noted communication problems in cross-cultural group chats, there is a relatively little research that articulates users' needs and requirements in these settings; there are few examples of CMC tools that have been designed to support cross-cultural communication. Given these gaps in research and tool development, we have begun to explore the needs and requirements of mixed groups of non-native and native speakers, drawing from these studies to design new tools that can help them to communicate better with one another. Our work has been inspired by two streams of design research - threaded chat [22] and collaborative annotation [12]. In particular, we have designed an enhancement to group chat called Instant Annotation (IA) – a conceptual design in which participants use a parallel channel to annotate or otherwise comment on an ongoing discussion. We believe that IA might assist information retention and conversation management in cross-cultural communication.

SIGDOC'12, October 3–5, 2012, Seattle, Washington, USA. Copyright 2012 ACM 978-1-4503-1497-8/12/10...$15.00.

In this paper, we report our initial exploration of the design space with an eye on cultural differences in participants' reactions to the IA concept. We hope by doing this that we can tailor the design concept to the different needs of non-native and native speakers. We first used a warm-up task to immerse our participants' into the experience of cross-cultural group chat; we next presented the IA conceptual design and asked our participants to reflect on its usefulness. We show that by offering an early design concept for feedback, and asking users to weigh pros and cons of the conceptual design, we were able to learn about users' needs, enabling us to refine the design. In the balance of the paper we describe the communication problem we are addressing, the IA design concept, the methods we used to introduce this concept to users, our findings and the implications we drew for continued work.

2. BACKGROUND

Text-based instant communication tools have been studied for many years. Both their drawbacks and their benefits are well known. On one hand, text chatting affords a level of reviewability and revisability [2] that is difficult to match with audio or video channels. On the other hand, taking turns in text chat can easily be disrupted by overlapping threads of multiple topics. Several qualitative studies have reported participants' frustrations in online group discussion; in these cases the frustration seems to be directly linked to the frequent and flexible turn-taking common in text-based CMC [4,6,7]. These facts motivated our interest to explore new design features that could possibly solve the problem. In this section, we briefly review the disrupted turn-taking problem and two streams of designs in the literature that inspired our work.

2.1 Disrupted Turn-taking

Turn-taking has been studied for some time in oral conversations; it is seen as a vital component in the construction of any spoken interaction [20]. Turn-taking is accomplished through the dynamic collaboration of interlocutors, who exchange cues about whether they plan to hold a turn, start or end a turn, or interrupt a turn. Schegloff pointed out that turn-taking also occurs in text-based conversations but with slightly different dynamics than spoken conversations [21]. Due to the common lack of nonverbal social cues (e.g., seeing that someone is preparing to "speak"), several interlocutors may enter a conversation at the same time. As a result, even though utterances appear one at a time on the screen, and thus may seem to be sequential [25], there is no necessary logical linkage between each turn. In addition, in contrast to spoken conversations where utterances are produced and heard at the same time, an expression in a text chat is often seen by others only after the user finishes typing all the words and sends it to the system. Therefore, a given conversational turn may actually respond to a turn several turns before it. This difference disrupts the sequential nature inherent in face-to-face conversation [9].

The problems emanating from disruptions in turn-taking are exacerbated in group chat, because there are more interlocutors competing for turns. Gonzalez [6] found that people often introduce new topics without finishing previous ones and they only selectively attend to the topic that is of most interest to them. Therefore, some participants may feel overwhelmed and even lost in parallel and fast-flowing discussions, especially speakers who have slow keyboarding skills, slow reading/writing skills, or different cultural backgrounds [17].

Asymmetrical relationships are often formed on the basis of language competency, similar to the asymmetrical relationships built upon expertise or authority [26]. When a native speaker perceives that a non-native speaker has language deficiencies that interfere with communication, the native speaker may then assume control of the conversation [5]. This language competency imbalance may then affect turn-taking: Native speakers are likely to take over most of the turns, whereas non-native speakers have fewer opportunities. When this imbalance is severe, non-native speakers may simply refrain altogether from participating in the conversation, even if they have great ideas and are eager to participate at the beginning.

2.2 Threaded Chat

Threaded chat helps people to organize their chat logs into threads. This design feature can be integrated into chat systems or other tools for different purposes, such as managing conversation structure or supporting side chat. Smith et al. [28] proposed an IM tool that supports threaded chat, in which the threaded structure is integrated in the main chatting window thus making it a primary element in such communication. To initiate a new topic, a speaker clicks the "root" of the conversation tree to create a new thread. To respond to an existing thread instead, the speaker clicks on that thread to create a turn placeholder for editing. Although this method makes the relationships between turns clear, it interrupts the flow of the conversation quite a bit, such that it "hops around" and feels unnatural. As a result, users' satisfaction rates are low compared to standard chat.

Threaded chat can also be found in other tools but integrated as a more peripheral element of the design. The primary activity those tools support is not chatting; however chatting is included to assist the primary activity. For example, some anchored discussion boards designed to support students' collaborative learning have offered threaded chat on a sidebar next to the discussion window.

A representative example is WebAnn [1], used by students to discuss specific points in a digital document. To create a discussion point in WebAnn, a user selects the text snippet in the document to be annotated, causing a new thread to be created in the sidebar for editing. Other users can reply by clicking on that thread. Conceptually, the selected text snippet becomes the thread's anchor and the central point of the on-the-side discussion thread. A study of WebAnn showed that positioning threaded discussions next to the anchoring text raises people's awareness of others' interests and activities [16]. Two other studies with similar designs showed that people are motivated to participate after seeing others joining the activity [11,12].

2.3 Collaborative Annotation

The idea of collaborative annotation is similar to the anchored discussion boards in that it also provides peripheral spaces for anchored points. In fact, researchers at times have referred to designs like WebAnn as collaborative annotation tools. The reason for this is that the sidebar is not exclusively used for chatting; sometimes it is used for annotating the anchored content. This means that the thread does not represent a conversation but rather one or more notes or tags about the anchor. In this case, the structure of the "conversation" space is flat compared to the tree-like structure of a threaded chat.

Studies of collaborative annotation tools suggest that serious challenges arise for designers when the annotation features must be integrated with real-time communication and collaboration [12]. Because synchronous interaction is usually attention demanding, adding a subtask such as annotating points could be

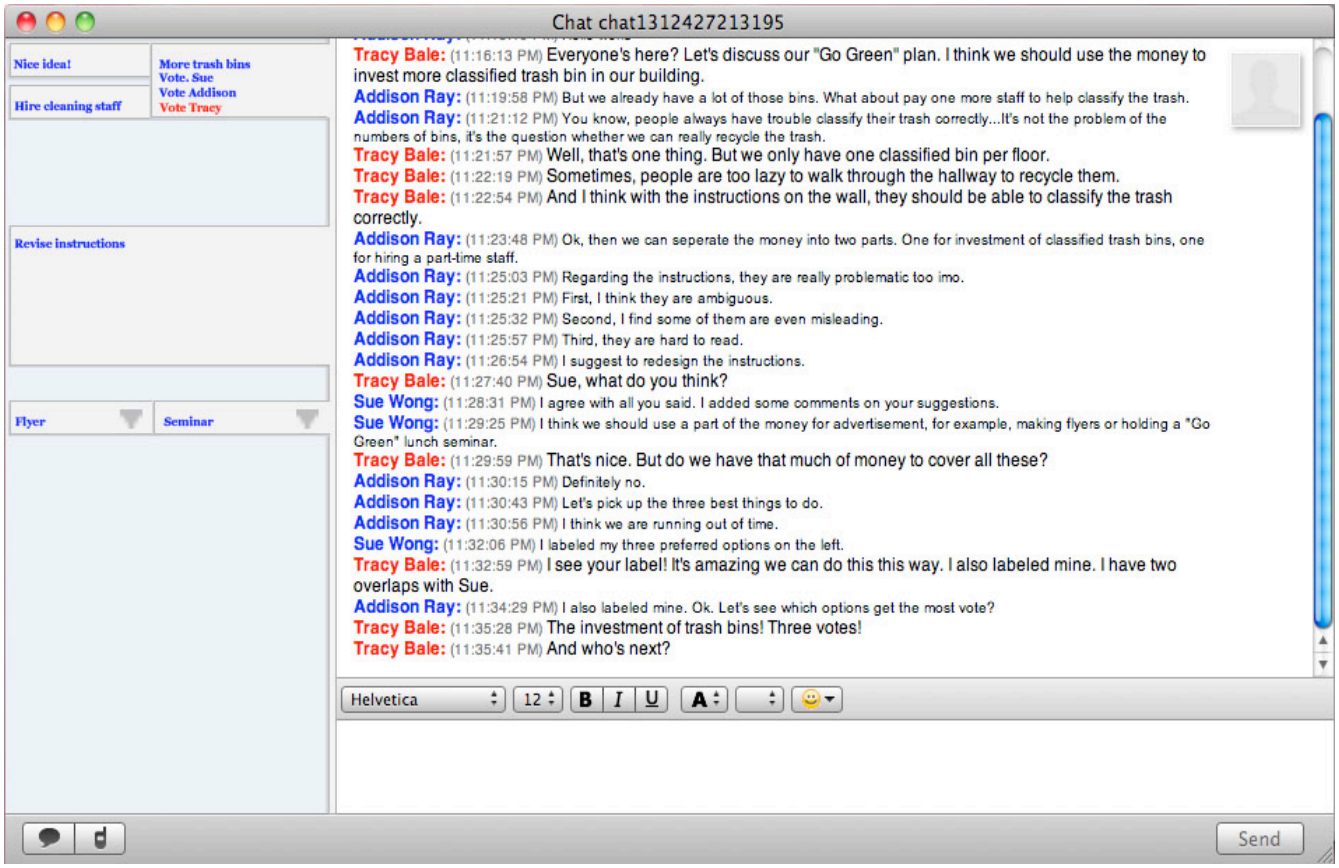

Figure 1. A mock-up of the Instant Annotation design concept.

difficult. Kelkar et al. [12] described a "live" collaborative tagging system for real-time audio meetings. Their system allows users to annotate utterances that are indexed by a timeline. However, their system did not support replying or otherwise elaborating an existing annotation. They found that users could not multitask well between tagging and active participation in the meetings.

3. INSTANT ANNOTATION

Inspired by threaded chat design and collaborative annotation design, we developed a conceptual design – Instant Annotation (IA). It supports conversation management and information summarization through an annotation space that is positioned near the main chat area (see Fig. 1 the light gray rectangular "tabs" in the left subpane hold annotations related to the text chat in the right). For convenience, in the paper we will refer to the right subpane as the "IM" window, and the left as the "IA" window).

To start a new annotation, a user clicks on a line of text in the IM window; this will generate an annotation tab in the IA window. The user can then add text to that tab. Other interlocutors will see the annotation as soon as the user clicks the comment button at the bottom left. The annotation tabs are tied to their anchoring line(s) of the chat log and will scroll up with that content. If more than one annotation is created for the same line(s) of text chat, they will appear side by side as shown in the top and bottom annotations in Fig. 1. To reply to an annotation, a user double clicks the annotation; when this happens a reply tab is created under the original annotation. In this way we are able to distinguish threaded chat and unrelated annotations of the same chat content.

Our design concept differs from prior work in several ways. We are the first to apply annotation tools to real-time text chat. Our design also supports a mix of threaded chat (i.e., comments and replies) and general annotation (i.e., multiple unrelated notes). As a result, users can seamlessly transition between threaded side conversations and simple annotations. Finally we have as a primary design goal to display and support the sidebar communication in as non-invasive a fashion as possible, so as to minimally disrupt the flow of the main conversation. We expect that users will only use it when they have a particular need in managing their conversations. We turn now to the methods we used to introduce and to gather feedback and design directions for the IA conceptual design.

4. METHOD

Early in design, when designers are still exploring the problem space, the emerging design concept may not have been specified well enough to create an interactive prototype for user evaluation. Nonetheless, it can be very beneficial to gather feedback and suggestions at this time, particularly if the concept is novel; users' input can help in formulating more detailed requirements as the design work continues. Often such input is gathered through field studies, with the goal of learning about users' needs in the real world. However, because the scope of the IA project is small and focused on a particular set of concerns related to language proficiencies, fieldwork was not a practical option. Instead we

sought to engage participants in a familiar usage experience in a lab setting (a group chat) where we could closely observe their behavior; with the group chat experience fresh in their minds, we then injected the IA design concept as an object of inspiration and reflection.

The lab-based method we used can be viewed as an adaptation and merging of scenario-based design (SBD) [19] and technology design probes [10]. In SBD, hypothetical users are envisioned as they interact with novel technology features in a familiar activity context and the ideas evoked by this exploration are used to transform the current activities. When using design probes, a novel technology concept is introduced into an ongoing activity or situation, usually for some period of time, so that the designers can observe what sorts of reactions or new behaviors it evokes in users. Our approach was to ask users to first enact a familiar activity (providing a personal usage scenario), and leverage that usage context to gather reactions to a novel technology concept.

More specifically, we engaged native and non-native speakers in an online group chat using technology familiar from their everyday lives – AOL Instant Messenger (AIM). Following this experience, we presented the IA design concept (Fig. 1), and gathered feedback using one-on-one semi-structured interviews. In the interviews we asked participants to reflect back to the group chat experience and to consider whether and how they might incorporate the IA concept into this cross-cultural communication process. We had intensive discussions with our participants, encouraging them to voice both pros and cons about their reactions to the design concept.

4.1 Participants

Over a two-month period we assembled five groups for study, each with two native speakers (Americans) and two non-native speakers (Chinese). After each session and set of interviews we did an initial coding of the participants' reflections and concluded the study after reaching the point of theoretical saturation [23], when themes were repeated in the data and no new themes were emerging.

We chose to study Chinese as the representative samples of non-native speakers because both their language and culture are very different from the western world; therefore they are likely to experience communication problems when communicating with people from the western world. We do not expect that our findings will be specific to Chinese non-native speakers, but future work will be needed to determine whether they apply to other non-native speaker populations.

The participants were undergraduate and graduate students majoring in information science or psychology from a large university in the Northeast United States; students' ages ranged from 20 to 43. There were 8 females and 12 males. Because we are not interested in gender effect in our study, and to avoid gender-related social and communication dynamics, we ensured that all groups were of the same gender. Most of the native speakers had some experience in collaborating with non-native speakers as part of group work (two Americans had not had this experience). Some of them had studied with non-native speakers in course projects in the past, and some of them worked closely with non-native speakers every day.

For the Chinese participants in our study, all had been living in the U.S. for less than four years, and all reported an advanced English proficiency (indicated in their pre-task survey). Advanced English proficiency was described as "I can carry on a conversation with a native speaker of the language, although it is highly evident that I am not a native speaker of the language." Thus any difficulties observed should be seen as persisting even once a non-native feels relatively comfortable conversing in a new language.

4.2 Task

We set a simulated group task that could stimulate participants' thoughts about cross-cultural group chat. The task was adapted from a study of second language learners' communication media preferences [4]. Each participant was asked to assume the role of a "Go Green" team member and to discuss with their group members how to spend $5000 to support environmental sustainability.

Four participants chatted about the sustainability topic in AIM for 15 minutes. They were asked to generate at least eight ideas and to decide on the best three. This combination of brainstorming and decision-making within the task make it similar to the real world situations where formal discussions occur (e.g. in business or academic settings).

Before entering the group chat, participants completed a pre-task survey; they also completed a post-task survey. Our analysis and discussion of these survey results is reported elsewhere [14].

4.3 Interviews

We conducted a semi-structured interview with each participant, spending approximately 30 minutes in discussion. During the interview, we presented the IA design concept: We showed the mock-up images to participants. Using the mock-up image we did a walkthrough to "demo" basic functions the IA could provide. We then asked several open-ended questions as described below; in this paper we focus particularly on answers to the second and third question. The open-ended question format allowed us to pursue other points that came up in each participant's response.

1. What was the participant's experience in this cross-cultural group discussion, especially their experience communicating with people from other countries?
2. Will a chat client featuring Instant Annotation help him/her in cross-cultural group discussions and if so how? Providing use cases if possible.
3. What are the limitations of Instant Annotation? Or, what other features can he/she imagine to assist cross-cultural group discussions?

Each interview was recorded and transcribed to text. Participants were interviewed in their first language. For interviews with Chinese participants, the transcriptions were translated back to English by the first author.

The transcripts were analyzed informally to discover themes related to cultural differences and communication difficulties. Specifically, we informally coded the transcripts with descriptive words, such as "tagging topics", "taking notes of important contents", and "side chat", etc. And then we sort through similar codes and merge them into higher-level concepts. We searched for themes that could cover the full range of concepts.

5. DATA ANALYSIS

Although we also collected and analyzed the chat logs of the five groups, the analysis of those data are reported elsewhere [14]; our focus in this paper is the participants' reactions to the IA design concept *after* experiencing the group chat. In general, we found

that most participants (18 out of 20) expressed an interest in using an IM tool that included an IA space, and they voiced many ideas about how they would use it based on the experiences they had just has as well as their more general experiences using group chat tools and interacting with speakers from different cultures. Collapsing across these various ideas, we have organized their ideas into three high-level categories: tagging, side chatting, and other concepts.

5.1 Tagging

Many participants envisioned that they could use an IA space for tagging; that is, to provide a conceptual level description of a piece of text. They also indicated what they would use the tags for; these more specific ideas led to three sub-themes in our discussion of tagging as a design direction. We refer to these subthemes as *use cases* by analogy to the hypothetical uses that are often generated during requirements engineering [13]. Of interest to our general research project, these use cases suggested that the non-native and native speakers differed in how and why they would incorporate tagging into their cross-cultural communication.

One common use case for tagging in the IA space was to quickly retrieve earlier discussion points. Participants indicated a need to remember these discussion points when a discussion went for a long time or when it became complex with a number of rich ideas. We noticed that one participant had noted his group's ideas on a piece of paper, and all five groups had at least one member who summarized the emerging ideas in the chat window during the chats.

Going beyond the general need to manage a long and complicated discussion, two of the non-native speakers envisioned a more specific scenario: They would use the tags to help reveal the structure of a conversation. In these cases, they said they needed a clearer view of the flow of a conversation because of disruptions in turn-taking. As one interviewee said:

"The discussion of a topic may be disrupted by other discussions, the tags will help you to follow a topic more easily. When a conversation goes really long, discussions of several topics may be mingled together, the tags will make the structure of the conversation clearer. " (Interviewee 1, Chinese)

Only non-native speakers suggested the use of tags for capturing conversation structure. In this sample at least, native speakers seemed not to have comprehension problems caused by disruptions in turn-taking, perhaps because they are quite familiar with this from their everyday online chatting activities.

A second use case for tagging was to make a note of points that have not yet been shared, so that they can be discussed later. One surprise was that it was native speakers who repeatedly mentioned this possible use, not the non-native speakers. Some native speakers tended to speak less, giving the non-native speakers more chances to talk. Such a strategy might reduce their chances of introducing new ideas while they promote the non-native speakers' opportunities to contribute:

"I had a couple ideas but I tried to only kind of go in turn as much as possible so that everybody had a chance." (Interviewee 4, American)

"Sometimes I waited to give everyone else time to type, because I typed faster than some people, I was trying to not say anything." (Interviewee 12, American)

Although not explicitly calling out the cross-cultural demands of the conversation, we speculate that it was at least partly the native speakers' awareness of the culture difference – and the accommodations they made in response – that led to production loss [3] in these groups. It is interesting to consider whether and how often such accommodation happens in cross-cultural communication. It suggests a social sensitivity along with finely tuned cross-cultural communication skills, perhaps a function of the prior experience many of the native speakers had in working with student peers from other countries. Interestingly for our design project, this use case could be supported by the same IA affordance as the more general tagging goal (i.e., a place to hold "extra" ideas). In this case though, we can hypothesize that a side conversation space might be at least as important to native speakers who are being considerate as to non-native speakers who are feeling stressed by the pace of the conversation.

Two native speakers also suggested the use of tags as a temporary reminder for ideas that they wanted to express but that would not have let them "keep up with the flow of the current chat" (Interviewee 13, American). In this case, it was not so much that they wanted to give their non-native counterparts a chance to contribute but rather than they judged that this was not the right time to change a topic. Thus they elected to wait for the next turn because they did not want to "completely ruin someone's thought" (Interviewee 4, American). When asked to compare using tags as a reminder versus paper and pencil, they said the tags would be used as "visual markers" that were easy and convenient to access, whereas paper and pencil might not be always available.

The third use case of tagging was to promote awareness of a discussion point. In two different groups, the two Chinese participants' ideas were at times ignored, apparently due to the overwhelming and parallel discussions taking place. These two Chinese speakers were significantly less talkative than the other three participants in their groups and their expressions tended to be short and simple, making them easy to miss.

One of the Chinese participants made several efforts to re-raise or re-address the ignored ideas from her Chinese group member. This caused her to be seen as a coordinator who organized and shifted topics back and forth in her group. She later provided a compelling scenario of using collaborative tagging to replace her role: *"If we missed an idea and we all wanted to catch it up, we would all annotate that idea. Then we would easily shift back to that idea and further discuss on it. " (Interviewee 15, Chinese)*

Another Chinese participant expressed similar ideas about collaborative tagging as a way to raise awareness of others' contributions. On the other hand, the native speakers voiced no concerns about the highly parallel and disrupted communication style, taking it as "a common feature that all chatting systems have".

5.2 Side Chatting

Another general design scenario that many participants envisioned was using the IA space for a side conversation. A common need recognized by both native and non-native speakers was to easily comment about earlier discussion points. As stated by the interviewees below,

"Because each tab corresponds to a topic, if I comment there, they can easily see it. If I respond to an earlier point in the main window, then you have to search above to anchor the point. With the IA window and maybe a bright alert sign when new comments are added, it's easily for people to see what they say, what others

respond, and what others respond to those response." (Interviewee 16, Chinese)

"Sometimes I don't know how to say it in English, especially when many people are discussing, I miss the chance to speak out, after a while, I forget it myself. … With the IA space, I can make up the discussion points I missed." (Interviewee 1, Chinese)

The side chat feature provides users an alternative way to communicate, one that requires little extra effort to foreground an earlier discussion point before commenting on it. Furthermore, other users do not need to search through the complex chat record to access the point. It is as easy as going right to the point, clicking to expand the comments and reading them. Non-native speakers saw this as a chance for them to jump into a conversation at a later point, for example after they had a chance to process and reflect on what others were saying. Interestingly, they felt that they would only respond to ideas in the IA space when a discussion of the idea was over, because responding while a topic was still in discussion "would interrupt the flow of the main chat too much" (Interviewee 8, Chinese).

The only time that the interviewees predicted that they might respond to currently active ideas in the IA space is when several parallel discussions were taking place. In this case, the rationale for responding next to a piece of text was similar to the one above, in that it would save effort to foreground which point was the target of the new response.

As an example, this functionality was achieved in a different way in group 5's discussion. The conversation of Group 5 had many cases of disruption, for example one parallel discussion that involved eight out of the ten ideas proposed through the whole discussion process. However, group 5's members thought their conversation was clear, because they successfully developed a communication protocol within the group: Adopting a practice often seen in microblogs, they used @groupmember's name to direct responses to the right person. The following snippet from their chat log showed how this worked for them (participants' screen names have been replaced by letters).

A: I have one idea.

B: @A: yes?

C: We can set up a group to collect student, staff, and faculty's address and then make the carpool assignment.

D: Note to tall: we are at 12 minutes of discussion so far.

B: @C: We kind of already have that.

A: We can purchase some cloth bags and distribute them among the students.

This spontaneous practice is similar to making a comment to the side of a piece of text, which is also able to direct attention of the right person to the right place. One of the participants of group 5 also offered a compelling scenario that uses the IA space to organize a major discussion and a sub-discussion in his lab meeting: *"If you are on a task where some people are worrying about when we are gonna schedule things versus how do we organize logistics for something else, people who are scheduling things can have their own discussion about something versus this general conversation about logistics." (Interviewee 17, American)*

This scenario also applies to the case of native speakers placing tags as reminders for ideas to discuss later. They could chat on these ideas while waiting for the non-native speakers to input their ideas in the main chatting window.

Some interviewees suggested that the IA space could serve as a convenient mechanism to express agreement/disagreement (e.g. Nice idea!) and other quick opinions for that matter. For example, one interviewee said, *"Because the normal chat does not hold a lot of ways for you to kind of make different expressions unless you use the happy faces or those kind of things. But I think this is a much better way to do it." (Interviewee 3, American)*

5.3 Other Use Cases

Besides tagging and side chatting, interviewees offered several other use cases of the IA space. They could use it to note down important facts, such as phone numbers, people's names, and addresses, etc. This is similar to Micronote [15], which provides a temporary note for fast retrieval. Inspired by the decision making task, participants also mentioned that they might use the IA space to vote for the three best ideas. Although this use case is specific to the task they were given, it might generalize to other decision-making tasks, especially those tasks requiring majority vote of ideas for the final decision.

One participant suggested adding the IA feature to chat tools used for larger groups. He provided an example of annotating in a chat room: Many online live sportscasts provide a chat channel next to the video, so that fans can express opinions, reactions or converse with other fans. Fans from the two sides often debate for their team, which generates some discussion points. Because there are hundreds of fans in the room, the screen updates very rapidly, which makes these conversations severely disrupted.

In this situation the IA space could help to organize discussions by allowing responses next to a piece of text. It also helps to remind people of important contents. In fact, this scenario also incorporates tagging and side chatting. However, an interesting point raised by this scenario was the scale of the chatting activity. The IA feature may be even more useful in chatting contexts that involve many people as opposed to the small groups we studied.

6. DESIGN CHOICES

We have shared users' reactions to and reflections about uses of the IA design concept, after first being primed with an experience in online chat among native and non-native speakers. We turn now to a set of design implications that are entrained by the range of usage ideas and that we are now considering in the elaboration and realization of the IA design concept.

6.1 Tag Access Control

The interviewees were sensitive to the tags' access rights, though they differed in whether they believed tags should be used for private or public purposes. Although their interview comments did not elaborate this issue enough to clarify the distinction, they asked questions like "Who will see my tags?" and "What does others' IA space look like?" This suggests that tag ownership and control will be an important issue as we elaborate our design.

When talking about use cases for the IA space, at times participants distinguished between personal versus others' chat content. For example*, "I will definitely tag others' ideas, so that I can review them easily." (Interviewee 2, Chinese)* This participant also said that he would not tag his own ideas. Another participant also saw tagging as a communication tool with public access*, "I will tag something when I want to emphasize it to others." (Interviewee 10, Chinese)* In contrast, interviewee 3 only thought about tagging his own content*, "If you don't want to express yourself in this group chat, so if you want to hold something and send it as a sidebar to somebody outside the chat. You could take that idea offline and hold it there for yourself for later." (Interviewee 3, American)*

These different views of the tags lead to the general design question, "Who will see what in the IA space?" The simplest option is to make the side pane entirely public or entirely private. However, such a design might only satisfy one portion of users' needs. A second option is to give users the right to decide what access mode they want for each tag. However, this extra operational cost may add considerable burden to users who do not want to worry about access for each comment. A better compromise might be to make the IA window entirely public but give users the opportunity to "hide" any tags they wish. Yet another option is to have two side panes, one private and one public. The worry there is that users might be more distracted by a two-part IA display, or by navigating between them if they were layered using tabs.

6.2 Notifications

While sharing their thoughts about using a side chat to raise awareness of a discussion point, several interviewees asked how other users would be notified of new contents. Because the annotations are anchored by a piece of text, a natural design is to display annotations next to the text as the main chat proceeds. But will users notice a new annotation or a response to an annotation if it is not shown in the current window to which they are attending? As the chat continues, earlier comments scroll up out of sight; this means that a new annotation intended for public viewing might be missed by other users.

With this concern in mind, one participant proposed to include a dashboard at the top of the IA pane. When new content is added, the dashboard could display a hyperlink that other users can click on to go to the content, somewhat like the function of an anchor link in an html page. The dashboard would be updated whenever new annotations are posted. An alternative design would be to have a pop up window at the right bottom corner (i.e., in the IM window). When new content is added, it could pop up for a few seconds and then provide a hyperlink at the corner until the user clicks it. We think that both of these designs could have a positive effect on raising awareness. If new contents about the same discussion point occur often, it indicates a general interest of that discussion point; people may be attracted to see what others say in this hot topic. A history pane containing the recent annotations the user visited may also be useful for the user to quickly revisit the tags or side conversations that are emerging.

6.3 Annotation Ownership

In the side chatting scenarios described earlier, users communicate under an annotation tab. One issue that arose was how to distinguish among different annotation contributors, so that the participants can communicate unambiguously. One simple design is to color-code each annotation, so that they represent the color of the user. However, when many users are participating in a chat, there may be too few colors to be distinctive. Another design is to automatically add users' names or icons as a prefix to an annotation. However, because the IA space may be small relative to the main chat window (implying that it should be used for concise expressions), the addition of user information may overfill the space. This may be an issue that is best left to a group, for example depending on how many participants it includes.

6.4 Operational Cost

Finally, two interviewees expressed concerns about the extra cognitive costs of the IA features. Reading and contributing to an online group chat is already demanding; they wondered about the effort required to also attend to and contribute to an IA pane. We also have had this concern from the beginning of this research program. While we have proposed a mechanism for annotation creation in our design scenario, we recognize that we may not have yet found a good solution. For example, we may replace the action of right clicking on a text line with a hot key, to see whether it reduces the interaction costs, and there may be other keyboard shortcuts that could simplify navigation among lines of text in the chat log as well as navigation to and from the IA window. These will be an important focus in our iterative design process as we prototype, evaluate and refine the IA concept.

7. REFLECTIONS ON THE STUDY METHODS

This ongoing project is an exploratory project stemming from the first author's doctoral research. We found the design method we have described here to be very useful for exploring the design space from the users' perspective very early on in the design process. Thus one goal of this paper is to share our experience of using the method with others in the HCI and CMC design community, so that projects with similar characteristics could benefit from our experience.

From the start, our over-arching goal has been to design a new technology that could assist cross-cultural group communication. A traditional design approach would be to carry out fieldwork to understand target users, develop the technology and then evaluate it to see what the users like or dislike [19]. At that point, the design may be revised and reevaluated. However, in our research project, time and labor are limited. We wanted to use a lightweight method that could quickly reveal some specific and useful implications about our design ideas. So we started with a lab task instead of fieldwork. The task in our study was carefully plotted, so that it could mimic situations that might happen in the real world. This allowed us to investigate users' problems and needs in-situ, even though the "in-situ" was a simulated communication situation. In fact, we observed many expected user behaviors, such as non-native speakers whose comments were ignored, parallel discussions in the chat, listing things to remember, and so on. These observations confirmed that the design scenario we had developed was realistic. We found that our participants were fully engaged in this process, as revealed both by the active discussions during the task, and the many creative reactions and reflections they provided in their interviews.

Because we only have one developer in the project, we cannot afford to spend significant time in iterative tool development. To accelerate the work of this one developer, we emphasized the development and evaluation of a conceptual design at the start of the project. After a recent experience with online group discussion, in combination with any relevant past experience, our participants were able to provide many inspiring but also feasible suggestions for how the conceptual design might be integrated into their online group discussion behaviors. We believe that this lightweight approach may be useful to other design research projects that are limited in time and labor. The combination of a realistic task followed by an interview about a design concept has provided a way to gather rich and useful information from participants.

Thinking more broadly, the approach could be used in many rather different projects. For example, suppose a designer envisions a bike-route application for a town. Residents who ride bikes could be asked to first finish a route-planning task using Google Maps; with that in mind the designer could seek feedback

for an early idea about the route-mapping project. Because participants would have just practiced route-planning, their experiences of the problems and processes of this task are fresh, and they are motivated to think seriously about the novel design concept. Because they were recently "contextualized" by a concrete route-planning experience, they are tuned to think as bike-riders who need to plan routes.

The benefit of providing a conceptual design for evaluation (i.e. versus thinking more abstractly about design features) is that it creates a specific base from which to form reactions and new ideas. At the same time, it may limit the ideas possible because participants will focus on what has been presented rather than totally new concepts.

One limitation of design concept interviews is that problems that users foresee may arise from lack of experience with the task or the technology rather than the design itself. Therefore, it is important to choose participants who have the right kinds of technology and task experience (in our case, text chat, including groups and including cross-cultural groups). If we had recruited less experienced chatters or chatters without cross-cultural experience, the chats and subsequent interviews would likely have been much less rich. Similarly, in the bike-route example, if participants are not familiar with Google Maps or riding bikes, their experiences of the task and design ideas may be dominated by frustrations of learning these things, and the post-task interviews would be less valuable. We suggest screening participants to select those experienced with the task and the technology.

A more general limitation of the work reported here is the nature of the study and data analysis. We observed and interviewed five groups, each composed of two native and two non-native speakers; all were university students. Our focus throughout has been on understanding users' ideas; in this we have relied very much on the comments offered by individual participants rather than attempting to draw conclusions about general populations. This is appropriate for the phase of design research currently in focus – our current goal is to acquire as rich as possible an understanding of design opportunities and challenges for the IA concept, not to formally test hypotheses or to generalize our findings to other groups or communication tasks.

In the next design iteration, we plan to explore some of the more unexpected communication patterns we observed, such as the tendency for some native speakers to hold back in deference to non-native speakers. We will also revise our conceptual design in consideration of the design implications we have presented. Following this iterative process, we will prototype and evaluate an operational tool (i.e., once we no longer find new useful suggestions from users about the conceptual design). In summary, the point of our method is to closely integrate our ideas about IA design and possible use scenarios, so that representative users can participate in each iteration of the design process.

8. CONCLUSIONS

Online cross-cultural group chat has many communication problems. Prior work has shown that non-native speakers are frustrated by their lower language competencies and disruptions in turn-taking. Inspired by the literature and related work of threaded chat tools and collaborative annotation tools, we proposed a novel conceptual design, Instant Annotation. We conducted a qualitative study to understand the communication problems between native and non-native speakers in a group chat setting that used a text-based communication tool, and solicited feedback and ideas about the IA concept. A data analysis of the chat log and the interviews showed several communication problems in this setting. Some of them were predicted by our theoretical analysis, while others were unexpected. One interesting phenomenon we found was that being aware of the language differential, native speakers at times hold back their ideas in order to give non-native speakers more chance to talk.

Our analysis of users' suggestions led us to three classes of design scenarios for IA. The users thought they might use IA for tagging, side chatting, and several other more specific use cases. However, they differed in their intentions of using these features. Some of the differences seem to be due to culture or language proficiency differences; these point to interesting research directions for our future work. For example, we plan to implement an IA prototype and run the experiment again. We will log users' behavior using the IA pane during the chats, and compare their behavior to confirm our findings in this paper. We generated four design choices from users' feedbacks to revise our design. We also reflected on our experience of using a "naturalistic" warm-up task as a method for engaging and contextualizing users' reflections and ideas about a novel design concept. We offer this as a general method that can be used in a variety of exploratory design efforts.

9. REFERENCES

[1] Brush, A. J. B., Bargeron, D., Grudin, J., Borning, A. and Gupta, A. Supporting interaction outside of class: anchored discussions vs. discussion boards. In Proc. Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community, International Society of the Learning Sciences.

[2] Clark, H. H. Using language. Cambridge University Press, 1996.

[3] Diehl, M. and Stroebe, W. Productivity loss in brainstorming groups: Toward the solution of a riddle. Journal of Personality and Social Psychology, 53, 3 (1987), 497-509.

[4] Freiermuth, M. R. Native speakers or non-native speakers: Who has the floor? Online and face-to-face interaction in culturally mixed small groups. Computer Assisted Language Learning, 14, 2 (2001), 169-199.

[5] Gaies, S. J. Modification of discourse between native and nonnative speaker peers. In Proc. TESOL '82.

[6] Gonzalez, D. Teaching and learning through chat: a taxonomy of educational chat for EFL/ESL. Teaching English with Technology, 3, 4 (2003), 57-69.

[7] Hara, N. and Kling, R. Students' frustrations with a Web-based distance education course. First Monday, 4, 12 (1999).

[8] Harrington, M. and Sawyer, M. L2 working memory capacity and L2 reading skill. Studies in Second Language Acquisition, 14(1992), 25-38.

[9] Herring, S. C. The rhetorical dynamics of gender harassment on-line. Information Society, 15(1999), 151-167.

[10] Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderback, B., Lindquist, S. and Sundblad, Y. Technology probes: inspiring design for and with families. In Proc. CHI '03, ACM.

[11] Kalnikait, V. and Whittaker, S. Social summarization: does social feedback improve access to speech data? In Proc. CSCW 2008, ACM, 9-12.

[12] Kelkar, S., John, A. and Seligmann, D. D. Some observations on the "live" collaborative tagging of audio conferences in the enterprise. In Proc. CHI '11, ACM, 995-998.

[13] Kulak, d. and Guiney, E. Use Cases: Requirements in Context. Addison-Wesley Professional, 2000.

[14] Li, N. and Rosson, M. B. At a Different Tempo: What Goes Wrong in Online Cross-Cultural Group Chat? In Proc. GROUP '12, ACM.

[15] Lin, M., Lutters, W. G. and Kim, T. S. Understanding the micronote lifecycle: improving mobile support for informal note taking. In Proc. CHI '04, ACM.

[16] Marshall, C. C. and Brush, A. J. B. Exploring the relationship between personal and public annotations. In Proc. JCDL '04, ACM, 349-357.

[17] Mynard, J. Introducing EFL students to chat rooms. The Internet TESL Journal, 8, 2 (2002).

[18] Neeley, T., Hinds, P. J. and Cramton, C. Walking through jelly: Language proficiency, emotions, and disrupted collaboration in global work. SSRN eLibrary(2009).

[19] Rosson, M. B. and Carroll, J. M. Usability engineering: Scenario-based development of human-computer interaction. Morgan Kaufmann, 2001.

[20] Sacks, H., Schegloff, E. and Jefferson, G. A simplest systematics for the organization of turn-taking for conversation. Language, 50, 4 (1974), 696-735.

[21] Schegloff, E. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. Georgetown University Press, 1982.

[22] Smith, M., Cadiz, J. J. and Burkhalter, B. Conversation trees and threaded chats. In Proc. CSCW '00, ACM.

[23] Strauss, A. L. and Corbin, J. M. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications, 2008.

[24] Wang, H.-C., Fussell, S. F. and Setlock, L. D. Cultural difference and adaptation of communication styles in computer-mediated group brainstorming. In Proc. CHI '09, ACM.

[25] Werry, C. C. Linguistic and international features of internet relay chat. John Benjamins Pub Co, 1996.

[26] Woken, M. D. and Swales, J. Expertise and authority in native-nonnative conversations: The need for a variable account. Multilingual Matters, 1989.