ICIC Express Letters ICIC International ©2017 ISSN 1881-803X
Volume **11**, Number **8**, August **2017** pp. 1331–1337

CAN WE SUPPORT PERIPHERAL PARTICIPANTS BY SHARING SCREEN-CAPTURE IN REAL TIME? – THE STUDY OF IDIOSYNCRATIC NETWORK SERVICE IN HANDHELD MEDIATED COLLABORATION

Kibum Kim

Department of Game & Mobile Engineering
College of Engineering
Keimyung University
1095 Dalgubeol-daero, Dalseo-Gu, Daegu 42601, Korea
kibumkim@kmu.ac.kr

Received January 2017; accepted March 2017

Abstract. *Ubiquitous and pervasive computing provides that people are sometimes working together in the presence of communicating artifacts. Often, they have shared visual context, that is, they are looking at the same devices and displays at the same time. This allows them to engage in the processes that create and maintain common ground. However, sometimes they have partial, intermittent, or unpredictable access to shared information. A substantial body of research in the past has focused on the role of shared visual context in focused interaction, especially in dyads and usually at a distance. However, pervasive and ubiquitous computing raises the importance of different kinds of interaction, which is peripheral participation. A new handheld network service, "Look", was devised to support peripheral participants' acquisition of conversation meaning by sharing visual co-presence of the task. The experiment results showed in general that Look facilitated the acquisition of sufficient understanding for effective action by peripheral participants.*

Keywords: Computer mediated communication, CSCL (Computer Supported Collaborative Learning), CSCW (Computer Supported Cooperative Work), Visual communication

1. **Introduction.** Over the past decades, personal mathematical calculators have been exceedingly successful in the classroom because they can be tailored to satisfy the learning needs of individual students. In much the same way, handhelds and smartphones promise to provide the educational benefits of their technological kin, traditional desktop computer models, but in an inexpensive and portable manner. This cheaper but effective technology could be used to reach every student in every school, thus efficiently bridging the digital divide. However, a challenge presented by the small screens is the need for others who are not direct participants in the activity. In the classroom, this problem happens when a teacher attempts to walk around students' groups for checking whether they are following directions correctly or paying attention to. Another situation involves the difficulty inherent when a distracted student comes back to learn what is happening and join the on-going activity of peer students. These peripheral participants have different access to information and different cognitive burdens versus central participants, such as a speaker or an addressee.

Clark investigated and found a theoretically predicted advantage that main participants have over peripheral participants in building and maintaining common ground [1]. However, very little is known about how to support peripheral participants' needs. Our study reported here is distinctive with prior work in that our investigation has implications to the sharing of workspace for peripheral participation in use of handhelds. In particular, it highlights a gap in predictive theory about the nature of joint action. Current theory does not accommodate the range of situations that we find in ubiquitous and pervasive computing. For example, although Kraut et al. [2,3] examined the cost of delay in sharing visual workspaces, they have not examined punctuated, or non-continuous, sharing, such as the kind obtained by our system, "Look".

Prior empirical evidence suggests that, when people have a shared visual workspace, it is easier for them to establish common ground. For example, incorporation of the shared video for a workspace proves valuable when partners are expected to collaborate on task-oriented discourse [2]. Look was implemented to provide the ability of unobtrusive synchronous visual capturing of activities and attentive artifacts, which allows peripheral participants to explicitly share focus on a given task without interrupting central participants. In this study, we explored the challenges faced by peripheral participants, introduced a possible solution to overcome those challenges, and reported experimental results showing the proof-of-concept of the suggested solution, "Look".

2. **Related Work.** We surveyed previous work showing the importance of sharing visual workspaces for cooperative work and challenges faced by peripheral participants.

Previous study showed when pairs work together on a physical task, seeing a common workspace facilitates communication and benefits performance [3]. However, other study indicated that it does not matter what communication mode is used for an initial short meeting of a group that is to subsequently work together via asynchronous text communication [4]. As Clark and Brennan long ago noted [1,5], different communication media put different constraints on the grounding of information. The underlying theory describes the need that participants have to create and maintain joint focus on the mental and physical objects of collaboration [1]. Clark and Krych showed that monitoring an addressee's workspace during a task was associated with an eighth of the errors, and half the time needed for the work as opposed to no monitoring the workspace [6].

According to psycholinguistics, in a collaborative view of language usage, a peripheral participant faces several disadvantages in understanding what is said. First, a peripheral participant has limited resources in grounding the mutual beliefs, knowledge, and assumptions required for current purpose of understanding the conversation [7]. Grounding refers to the interactive exchange of evidences by discourse participants regarding what is understood. A peripheral participant cannot actively join the process of such coordination between a speaker and an addressee. Instead, s/he receives only what is given by central participants (i.e., speaker and addressee). Second, a peripheral participant cannot control the pace of the conversation, and once s/he loses track of the content, misunderstandings can accumulate easily [7]. A peripheral participant does not have an opportunity to keep the speaker informed of the state of his confusions or to clarify misunderstandings. Third, although the addressee can determine what the speaker means from conclusive evidence of their common ground, a peripheral participant can only conjecture about what the speaker means using inconclusive evidence [1]. Without knowing what constitutes both the speaker and addressee's common ground, the peripheral participant finds it difficult to determine exactly what their discussion means.

These considerations were reflected in the design of our experiments.

3. **Experiment.** In this experiment, we investigated whether a new handheld network service Look, which allowed sharing visual co-presence of the task among members of a social group, could facilitate the effective communication for a peripheral participant in handheld-mediated face-to-face collaboration. In particular, our study examined whether peripheral participants could benefit from a shared visual workspace. The performance of a task was indicated by task correctness and conversation efficiency. Peripheral participants are late-arrivers into conversation. They are ratified but not central participants

such as students trying to enter the conversation. They may talk and ask questions but are not supposed to interrupt the on-going activity. Despite the ability to talk, the peripheral participant is still at a disadvantage compared to the central participant. Although a speaker may adjust his or her utterance in a way that acknowledges the peripheral participant, the speaker's contributions are often completed without waiting for acknowledgement or validation from the peripheral participant [8].

To help with the problem of allowing the peripheral participant to enable meaningful monitoring of ongoing conversation and to gain adequate acquisition of a conversation meaning, we conducted a controlled laboratory experiment with Look and No-Look conditions.

3.1. **Participants.** Total 96 students were recruited to participate in the experiment. 32 groups of three people were randomly assigned to one of two conditions, Look and No-Look conditions, 16 per condition. Students were given extra credit for participating in the experiment. Participants' ethnic backgrounds were multi-cultural: African-American, Mexican-American, Asian-American, Chinese, Indian, Japanese-Brazilian, Jewish, Puerto Rican, and Caucasian (84.4%). None of the participants had prior knowledge of how to read Korean characters, were familiar with Korean culture, or had traveled to Korea. Participants' ages ranged from 18 to 22, with a mean age of 19 (standard deviation: 1.1). Fifty-one percent of the group were female. All participants were undergraduates, with half of them freshmen (50%). Psychology majors provided the largest group (15.6%) but other majors also took part.

3.2. **Apparatus.** We explored the various design options of proper Wireless Local Area Network (WLAN) architecture for handheld connection. Among the large variety of networking options, widely deployed standards for short-range wireless technologies are Infrared (IR) connection and wireless LAN networks, such as Bluetooth or IEEE 802.11 (Wi-Fi).

In our previous study of handheld mediated activity, we used IR connection as the network infrastructure [9]. Applications for which IR is well suited include those that require fast communication or security. Because IR uses a direct point-to-point communication with the beam of light, which is more focused than the wider-ranging radio signals, such communication is less susceptible to electronic eavesdropping. In beaming, the "seeking" or "controlling" device can clearly identify the target device simply by pointing toward it.

However, a situation that involves multiple people simultaneously using devices does not work effectively with IR, because at any given time data should be transferred between two devices without disrupting the line-of-sight connection. Such a connection must remain relatively stationary for the duration of the data transmission session.

The Bluetooth wireless technology is a low-power, low-cost, short-range RF technology that replaces cables between electronic devices. Bluetooth communication technologies can complement infrared's narrow angle of sight (30 degrees or less), short range (three feet or less) signal, and point-and-shoot overt use, with its omni-directional signaling, longer distance communications, and capacity to covert transmission. In this study, we implemented Look, based on the Bluetooth network. Additionally, we embedded and tested the Look function in the context of our handheld game. Participants use the stylus to play the game through the drag-and-drop interaction and to select soft buttons on the screen for initiating task actions, such as shuffling the image order as well as looking others' screen by tapping with the stylus on the button (see Figure 1).

3.3. **Procedure.** During a roughly two-hour period, participants were asked to play a game called Korean Characters Tangram in which one person was the director, another person was the matcher, and the other person was the peripheral participant.

Figure 1. Korean characters tangram game

In this game, everybody had a handheld device with twenty-five Korean characters on it. They played four rounds of the game. At the beginning of every round, the director hit the “Shuffle” button to put his characters in a specific order. The game asked the matcher and director to put the matcher’s characters into the same order as the director’s without one person ever looking at the other person’s screen. The matcher used the stylus to drag and drop characters on her screen to rearrange them. The director could use words to describe the characters, but he also had their Korean names and was encouraged to use the name of each character at least once in every round, maybe before he moved to the next character. The round ended when the director and matcher agreed that the characters were all in the same order. At the end of each round, their answers were compared to each other. Also, they individually responded quizzes about naming and identifying the Korean characters.

The director and matcher played four rounds of the game. During the first two rounds, the peripheral participant was out of the room. Then, the peripheral participant was brought back into the room from the third trial. The job of the peripheral participant was described as being to “catch on without being too much of a burden”. Participants could talk with one another, but the director’s and matcher’s “main job” was to work with each other. Their job was not to tutor the peripheral participant. The peripheral participant also put his characters into the same order that the director had, but only half of the peripheral participants were allowed to use the “Look Others” button. By clicking the “Look Others” button, the peripheral participant could capture a screenshot of the matcher’s handheld. The peripheral participant could use this screen capture in any way that helped himself. In the No-Look condition, the “Look Others” button option was not available to the peripheral participant and he could not capture a screenshot of the matcher’s handheld. The peripheral participant would do two rounds altogether. All the experiment sessions were video- and audio-taped and discussions between participants during the game were transcribed.

4. **Results.** We collected two sets of dependent measures: task performance and conversation efficiency. Task performance was measured by the accuracy of execution. For

conversation efficiency, we reviewed all of the video recordings and noted all events of turn-taking in the conversation.

4.1. **Task correctness.** Task correctness was evaluated by comparing the order of Korean character figures arranged by the peripheral participant with the target order described by the director. The numbers of matching figures between the peripheral participant and the director were counted [1]. These scores were then compared across two conditions, Look and No-Look conditions. Because Look allowed to capture a screenshot of the matcher's handheld, it was expected that peripheral participants whose handhelds were equipped with the Look functionality should better understand the conversation and thus should be able to rearrange the figures more correctly. As expected, performance did differ across two conditions (see Figure 2).

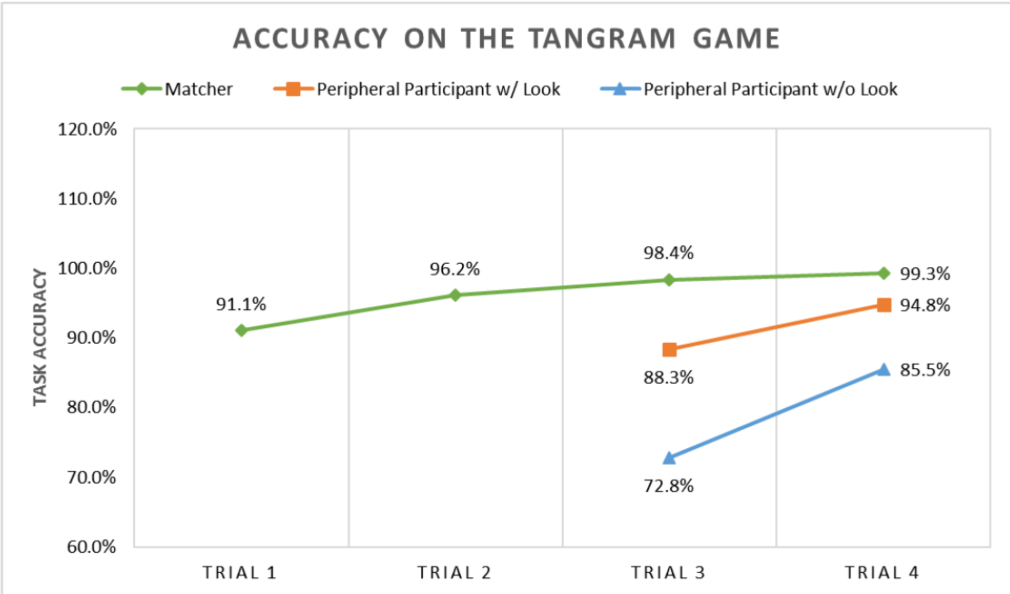

FIGURE 2. Percentage correctness of matching Korean characters tangrams on each trial by the matcher, peripheral participants with and without *Look*

During trial 3, in which the peripheral participant entered the discourse at the first time, experiment results indicated an average of over 88 percent of correctness (Standard Deviation (SD): 9.96) among those peripheral participants who used Look, compared to an average of about 73 percent of correctness (SD: 26.94) for peripheral participants without the Look functionality. Given the heterogeneity of variance, we conducted Welch's T-test. The data analysis showed the significant difference between groups Look and No-Look at trial 3 ($T(19.0) = 4.66$, $p < .044$).

As the peripheral participants repeated tasks through trial 4, the average accuracy per trial increased in both groups: Look (95% (SD: 7.83)) and No-Look (86% (SD:19.31)) but no significant difference was found ($T(19.8) = 3.15$, $p = .091$ by Welch's T-test). The absence of significant difference in trail 4 was possibly due to the effect of the accumulated common ground by repeated trials.

4.2. **Conversation efficiency.** During the experiments, all sections were recorded on digital video- and audiotapes. In the video analysis, we focused on the number of turn-takings in the conversation. A turn was defined as a stretch of talk contributed by a single speaker.

Previous literature shows that the effort of achieving common ground is indicated by more turns of talk [6,10]. The act of grounding between participants in a conversation requires that A presents an action and/or a signal *s* for B to understand, and B in turn eventually validates that action and/or that signal as having been recognized or understood. When these two phases are accomplished properly, they constitute the shared basis for the mutual belief that B understands what A means by signal s [1].

Displaying understanding gives partners the opportunity for such validation or correction. Using the Look network service, participants make displays and exemplifications of understanding practicable for the purpose of validation. Therefore, Look can be used to allow that the peripheral participant's presentations overlap the central participants' verbal descriptions, and thus they continue the conversation without separate turns: When the workspace is made visible by the Look, the peripheral participant will be able to continuously reformulate his/her tryouts without forming turn-taking. However, without Look, the workspace is not visible, so the peripheral participant will seek validation from central participants, which requires both parties to take more turns.

This difference was reflected in the mean number of turns by the peripheral participant, as shown in Figure 3. In trial 3, the peripheral participant without Look took over five times as many turns as the peripheral participant with Look (67 turns vs. 13 turns, $T(17.2) = 13.66$, $p < .002$ by Welch's T-test). A similar pattern of result was shown in trial 4. Without Look, there was an average of 37 turns but with Look, an average of 10 turns occurred ($T(23.0)=12.35$, $p < .002$ by Welch's T-test).

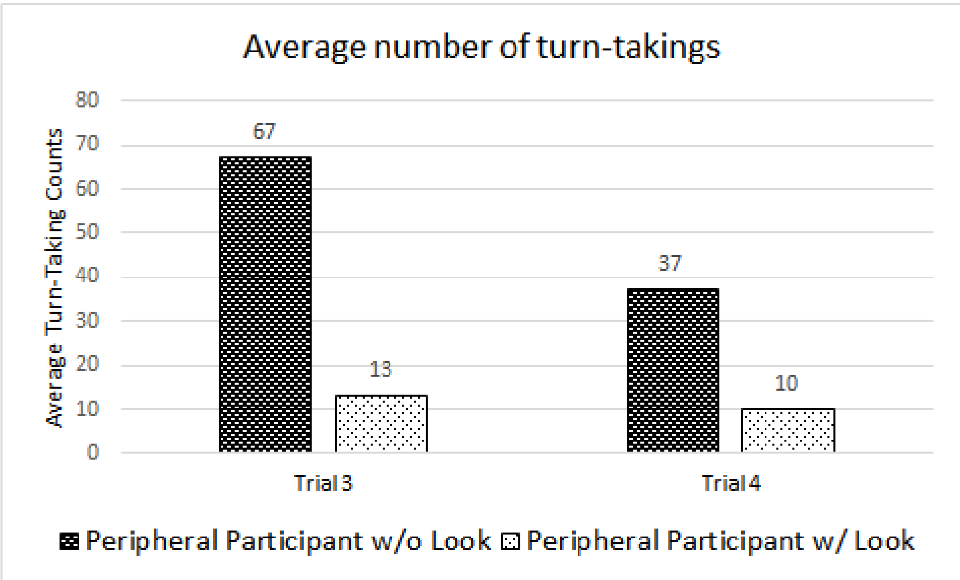

FIGURE 3. Average number of turn-takings by peripheral participants under *Look* and *No-Look* conditions

5. **Discussions and Conclusion.** Three factors differentiated our study from other related work on shared visual spaces and peripheral participants. First, we manipulated at least two layers of difficulty in creating common ground of a participant (i.e., to become a latecomer and a peripheral participant). This setting emphasizes the peripheral participant's role in the classroom (e.g., teacher or peer commentator) and the experiment here is abstracted away from a situation found in the classroom use of handhelds to aid learning. Compared to normal classroom practice, it increased the need for deictic resolution, and therefore the potential for errors. If Look reduces errors in a stressful situation, it is highly likely to do so in a daily classroom usage.

Second, this study investigated whether implementing a minimal shared visual workspace for peripheral participants using wirelessly connected handheld was worthwhile. Look involved a kind of sharing that was punctuated rather than smooth. One way of thinking about this situation is that the peripheral participant has less information about what is happening than if s/he had continuous real-time video monitoring. However, another point of view is that the peripheral participant has an amount and kind of information that is controlled by him- or herself. The peripheral participant's control over viewing may arguably prove more important to understanding than continuous visual information. Despite the fact that Look provided only a snapshot of the referents, the peripheral participants appeared to utilize it to acquire sufficient comprehension to engage in the nuanced negotiation about their participatory status in the conversation.

Third, a subtle social role of the peripheral participant in an interaction was successfully managed in our experiments. When participants are allowed to speak to each other, the exact status of peripheral participants is hard to manage; people are polite to and cannot ignore each other. In previous work on overhearers, the absolute separation between an overhearer and discourse participants was enforced by physical separation between rooms. Indeed, sometimes the overhearer was a witness only to a videotaped interaction (c.f. [7]). In our case, controlled interaction was made inevitable by the constraints of the handheld device and we addressed the situations found in ubiquitous and pervasive computing.

Although the demonstrations in our experiments are limited, we feel that there is enough evidence to go through the effort to incorporate the Look functionality in a more contextualized co-located synchronous collaboration environment.

Acknowledgment. This work is partially supported by the Keimyung University Research Grant. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] H. Clark, *Using Language*, Cambridge University Press, Cambridge, MA, USA, 1996.
[2] R. E. Kraut, D. Gergle and S. R. Fussell, The use of visual information in shared visual spaces: Informing the development of virtual co-presence, *Proc. of the Conference on Computer Supported Cooperative Work*, pp.31-40, 2002.
[3] D. Gergle, R. E. Kraut and S. R. Fussell, Using visual information for grounding and awareness in collaborative tasks, *Human-Computer Interaction*, vol.28, no.1, pp.1-39, 2013.
[4] H. Han, S. R. Hiltz, J. Fjermestad and Y. Wang, Does medium matter? A comparison of initial meeting modes for virtual teams, *IEEE Transactions on Professional Communication*, vol.54, no.4, pp.376-391, 2011.
[5] S. E. Brennan, *Seeking and Providing Evidence for Mutual Understanding*, Ph.D. Thesis, Stanford University, Stanford, CA, 1990.
[6] H. Clark and A. Krych, Speaking while monitoring addressees for understanding, *Journal of Memory and Language*, vol.50, pp.62-81, 2004.
[7] M. Schober and H. Clark, Understanding by addressees and overhearers, *Cognitive Psychology*, vol.21, no.2, pp.211-232, 1989.
[8] A. F. Monk and L. Watts, Peripheral participation in video-mediated communication, *International Journal of Human-Computer Studies*, vol.52, no.5, pp.933-958, 2000.
[9] K. Kim, *The Effects of Handheld Network Service, Look, on the Acquisition of Common Ground*, Ph.D. Thesis, Virginia Tech, Blacksburg, VA, 2007.
[10] A. Anderson, Achieving understanding in face-to-face and video-mediated multiparty interactions, *Discourse Processes*, vol.41, no.3, pp.251-287, 2006.