Supporting Cooperative Language Learning: Issues in Interface Design for an Interactive Table

Meredith Ringel Morris*, Anne Marie Piper, Anthony Cassanego*, Terry Winograd***
Stanford University
*Department of Computer Science, **School of Education
merrie@cs.stanford.edu, {ampiper, tonyc}@stanford.edu, winograd@cs.stanford.edu

ABSTRACT
The recent introduction of computationally-enhanced tables that support simultaneous, multi-user input has important implications for co-located, face-to-face activity. Educational applications particularly stand to benefit from this new technology, which can combine the benefits of small group work with the enhancements offered by digital media. In this paper, we explore how the unique affordances of interactive tables provide a match for the needs of foreign language education, and how the design of tabletop software can be subtly altered to encourage desired educational outcomes. We present three prototype applications, and explore four design variations (feedback modality, feedback privacy, spatial configuration, and interaction visualizations) to assess their impact on student participation and self-assessment. We present observations of the use of our prototypes in two settings: (1) a controlled laboratory study and (2) authentic use by students as part of a language course at our university, and discuss our preliminary findings and avenues for future exploration.

Author Keywords
Educational interfaces, tabletop interfaces, computer-supported cooperative learning, computer-supported cooperative work, co-located groupware.

ACM Classification Keywords
H5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces – computer-supported cooperative work.

INTRODUCTION

Recent advances in computing hardware have opened new possibilities for designers of co-located groupware. For instance, multi-user technology like DiamondTouch [6] supports simultaneous input by four co-present users (Figure 1).

Figure 1. Four students sit around a touch-sensitive DiamondTouch table, working together to match pictures with words in a foreign language, using the MatchingTable software.

The importance of collaboration in small-group work and methods for facilitating effective group work, specifically through group problem-solving tasks, is a prominent research topic in the field of education [25]. Small group work is also particularly valuable in the domain of foreign language learning, where peer-to-peer interaction provides important opportunities to practice conversational skills [8]. Educational activities may benefit significantly from interactive table technology because it combines the face-to-face interaction style of traditional small-group work with the enhancements of digital media. Digital technology offers many benefits for educational activities – in particular, digital technology can help address the problem of having one teacher for many students. When students are working on a small group activity, the teacher can only assist one group at a time. With digital technology, however, groups can still receive feedback regarding their progress even when the teacher is busy helping other students. Allowing students to immediately know they have found the correct answer has pedagogical benefits and increases group efficiency, as shown in a study on immediate and delayed feedback in the context of a LISP tutor [5]. The reciprocal benefit applies to the instructor – even though she cannot monitor the behavior of all students simultaneously, the digital technology can keep interaction records that can be reviewed after class, so that she can discover which students need extra help or which topics need more explanation. Interactive tables combine the benefits of face-to-face small group work with the ability of digital technology to provide feedback to students and interaction records to teachers, making them an excellent platform for educational groupware.

RELATED WORK

The benefits of facilitating effective small-group work with problem-solving tasks are widely researched in the educational community. Small-group work presents opportunities for learners to share insights [2], explain their thinking [10], observe the strategies of others [1], and listen to explanations [4].

Single Display Groupware (SDG) [23], supports co-located cooperative work by multiple users sharing a single, typically large, display. SDG facilitates cooperative work by providing a shared focus and context for the group. Until recently, such displays typically had a vertical, white-board-style form-factor, but the recent introduction of technologies such as DiamondTouch [6] and Lumisight Table [11] have allowed the development of interactive tables that allow single display groupware to support a face-to-face (rather than shoulder-to-shoulder) work style. Face-to-face work has many benefits compared to whiteboard-style interactions [18], including encouraging more group members to participate in interactions [13].

Most prior work on tabletop user interfaces focuses on applications of interactive tables for entertainment, such as playing games [11, 16, 22] and viewing digital photographs [14, 21]. A few projects, such as the InteracTable [24], UbiTable [20], and RoomPlanner [26] have explored the use of tables for productivity-based tasks such as document annotation, document sharing, and furniture layout, respectively. Our work on educational tabletop interfaces explores the potential utility of this new technology platform for computer-supported collaborative learning (CSCL) tasks. Recent work on the StoryTable [3] and Read-It [14] describe educational tabletop applications for very young (kindergarten-aged) children. Our work focuses on designing tabletop applications for teenage through college-aged students learning a foreign language.

SYSTEM DESCRIPTION

Hardware

Our applications run on an 85.6 cm by 64.2 cm DiamondTouch table [6] with a top-projected 1280 x 1024 pixel display. The DiamondTouch is a touch-sensitive input device, which is combined with a ceiling-mounted projector so that a display is co-located with the input. Four users can simultaneously interact with the device. Users sit on pads that are electrically coupled to the table, so the table is able to associate each touch with a particular user. Some of our applications deliver private audio messages to members of the group based on items they are interacting with on the display. Users receive this audio through one-eared headsets, using a system similar to that described in [12].

Software

Our software is written in Java, using the DiamondSpin tabletop interface toolkit [22], which facilitates building interfaces that accept simultaneous input from multiple users and with components that are oriented toward arbitrary angles.

We have created three language-learning applications with flexible structures that can be adapted to fit varying content. The three applications are the ClassificationTable, the MatchingTable, and the PoetryTable.

The ClassificationTable (Figures 2 and 3) presents users with a set of virtual "clues." A clue can vary in length from a single word to several sentences, based on the current lesson. The four corners of the table are labeled according to four different categories (e.g., countries, characters from a novel, authors, vocabulary themes, number of syllables, etc.), and the task for the group is to classify each of the clues into one of these four categories. Users can touch clues with their fingers and drag them around the table. Clues re-orient themselves to face the nearest table edge in order to facilitate legibility. Users can drag a clue into one of the corner regions, and receive feedback from the system regarding the correctness of their classification (the form of this feedback is described in the "design variations" section). Users work together with their teammates to decide on the correct classifications.

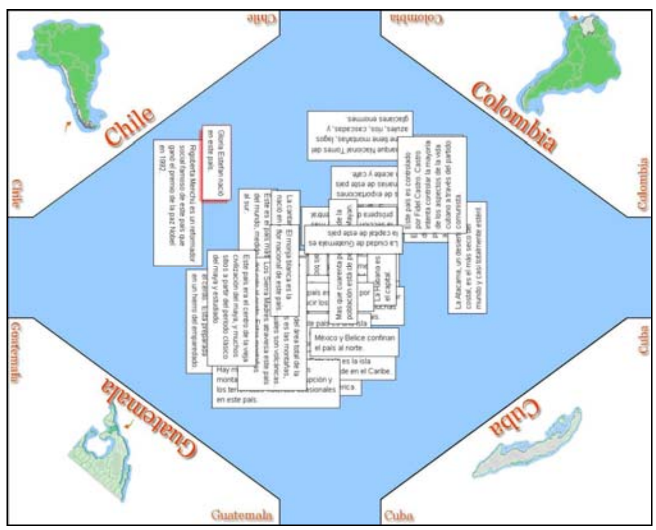

Figure 2. This screenshot shows the ClassificationTable at the beginning of a task for students learning Spanish. Clues are piled in the center of the table. Each clue is a Spanish-language fact pertaining to one of the four countries depicted in the table's corner areas. Students drag clues around the table with their fingertips and drop them onto the appropriate corner, then receive feedback about the correctness of the classification.

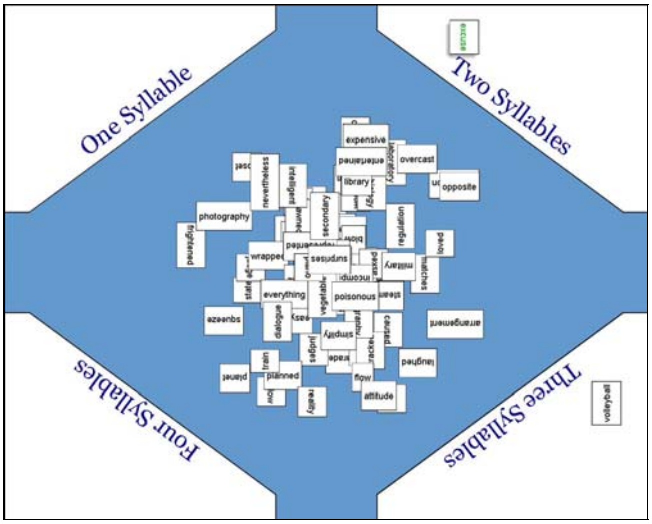

Figure 3. The ClassificationTable allows students to sort individual vocabulary words or longer sentences into one of four corners of the table, based on various properties. In this example, students learning English work together to classify the English words according to the number of syllables they contain. Double-tapping a word allows a student to hear it pronounced through a private headset. The words turn green when they have been properly classified, and red when they are placed incorrectly.

The MatchingTable (Figure 4) allows students to match words and phrases with images. Students can move words and photos around the table by touching them with a finger and dragging them. Words can be associated with a photo by dragging them on top of the photo and letting go. The words are then attached to that photo, and if the photo is subsequently moved, the words will stick to it. Words can be removed from a photo by touching the words and dragging them away. The words turn green when they have been correctly matched and red when they have been incorrectly matched.

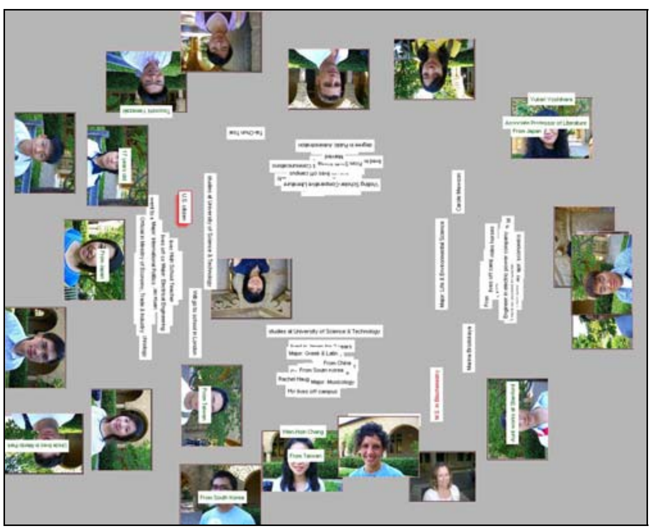

Figure 4. The MatchingTable application allows students to match words and phrases with images. In this example, students learning English wrote descriptions of themselves during class, and then a few days later tried to match their classmates' photos to the phrases they had used to describe themselves.

The PoetryTable (Figure 5) allows students to create free-form sentences and phrases by moving word tiles around the table with their fingers. Words can be conjugated by the addition of prefixes and suffixes, which are available as choices from a menu invoked by double-tapping a word tile. The activity is made more challenging by presenting both correct and incorrect options for students to choose from in the conjugation menus. This application is a descendant of the entertainment-oriented tabletop poetry application mentioned in [22].

Design Rationale

The design of our software has been influenced by previous findings regarding design guidelines for tabletop applications. Research on the SoundTracker system [12] found that private feedback helped encourage shy group members to interact more with the tabletop. Since one of our goals is to encourage all students, including underperformers, to be engaged with the learning activity, we have experimented with the use of private feedback in our applications' design. We have also taken into account Scott et al.'s findings on territorial tendencies [17] – we placed the ClassificationTable's "category areas" in the corners of the table, rather than along the sides directly in front of each seat, in order to reduce each individual user's sense of "ownership" over each area, so that all group members would feel comfortable placing clues in any region. Finally, we used the DiamondSpin tabletop toolkit [22] to design an interface that gracefully handles issues of orientation on the tabletop: to facilitate readability, items on the table turn to face the closest table edge.

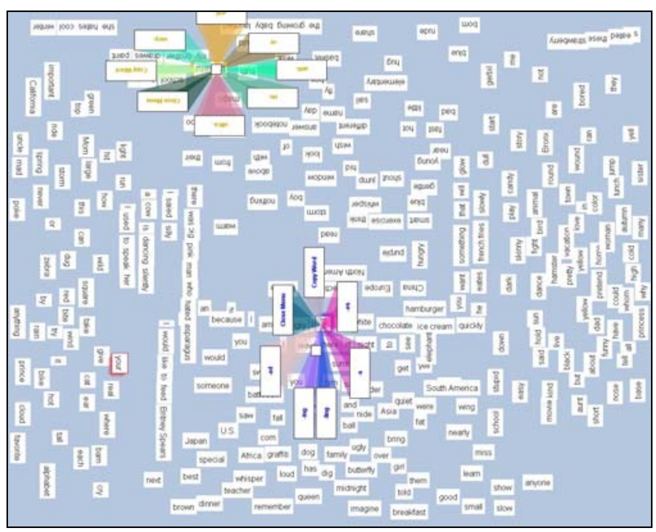

Figure 5. The PoetryTable application allows students to create free-form sentences using current vocabulary words. Double-tapping a word tile invokes a menu that allows the student to alter the word by adding prefixes and suffixes.

Our system differs from other language-learning technologies, as the majority of those technologies are tools for rote memorization, conjugation, and translation (e.g., "Study Spanish," an online tutorial that emphasizes vocabulary and verb conjugation drills [www.studyspanish.com]). These applications emphasize memorization of language mechanics but lack focus on the conversational skills that are central to developing fluency. Based on interviews we conducted with two professors of beginner- and intermediate-level foreign language courses at our university, *fluency*, the willingness to engage in conversation, is the main educational goal for beginner and intermediate foreign language students at our university,

Our three tabletop applications provide foreign language learners with an opportunity to engage with peers in conversation while completing activities related to vocabulary, pronunciation, literature, or cultural knowledge. Our software gives learners immediate feedback on the accuracy of answers, thus heading off any misconceptions that may arise with delayed feedback. Our applications give students an opportunity to explain, justify, and debate answers with other group members. The content of each activity serves as a conversational prompt, thus encouraging learners to develop fluency through engaging in conversation with one another. Our emphasis on fluency and language acquisition aligns with the pedagogical approach for foreign language within our university's foreign language program.

DESIGN VARIATIONS

Two factors of particular interest for educational applications is how they impact students' levels of participation in the activity and how they facilitate awareness of one's own and others' contributions. Participation can measure either direct interactions with the software itself or the amount of foreign-language conversation produced by each group member, since both of these actions align with desired learning outcomes. Increasing the amount of and equitability of participation among group members is a concern for educators, since one drawback of group work is the tendency for the strongest students to complete work while underperforming group members hardly participate. This problem is known as the "free rider" problem [9]. Increased awareness of one's own and others' contributions to an activity can impact participation, and result in more accurate self assessment. Accurate self-assessment has valuable implications for second-language learning, as described in [15].

We are interested in how the pedagogical value of educational tabletop groupware can be improved by subtly altering the user interface to impact participation and awareness. To explore this issue, we created several variants of our tabletop activities. We then conducted preliminary evaluations with students at our university in order to determine whether our design variations had a perceptible impact on participation equity and self-assessment accuracy in order to identify which design variants showed promise for more extensive exploration and evaluation. These variants fall into four main groups: feedback modality, feedback privacy, spatial configuration, and interaction visualizations.

Feedback Modality

One variation we explored was the modality through which we provided feedback regarding the correctness of clue classifications in the ClassificationTable application. We explored two alternatives – visual feedback (the clues' text turned either green or red to indicate correct or incorrect placement), and audio feedback (either an upbeat or discordant tone was played to indicate correct or incorrect placement). We hypothesized that audio feedback would increase the amount of conversation among group members (important for rehearsing a foreign language) by "breaking the ice" and inserting noise into the environment, thereby making it less awkward for students to generate their own "noise" by talking. Audio feedback should also increase awareness since it "pushes" information to users.

Feedback Privacy

Another interface variation we explored was whether feedback regarding the correctness of clue placement in the ClassificationTable activity was conveyed publicly (to the entire group) or privately (only to the group member who moved a particular clue). To explore private feedback in the context of the shared environment, we used individually-targeted audio feedback via one-eared headsets (a setup similar to that described in [12]). We hypothesized that private feedback would increase participation equity by reducing the potential for embarrassment over incorrect answers, and thereby encouraging shy and underperforming students to contribute more to the activity. We also hypothesized that private feedback would increase the accuracy of students' self-assessments of performance by drawing more attention to their individual contributions.

Spatial Configuration

We altered the initial configuration of clues and photos in our three tabletop activities in order to explore the impact of initial layout on participation. In the "four piles" design, all objects (clues, word tiles, photos, etc.) were initially placed into four random, equally-sized virtual "piles" near the four users' seats around the borders of the table. In the "central pile" design, all objects were initially placed into a single virtual pile in the center of the table. We hypothesized that the four piles design would increase participation equity as compared to the centralized design by making under-contributors feel more responsibility for the items that started out nearest them, and by making over-participators hesitant to reach out and take responsibility for objects that originated near others. This hypothesis is in accordance with studies of tabletop group work [17] that have found that the central area of a table is considered a group-owned, public space, while the areas directly in front of each user are considered personal or private zones.

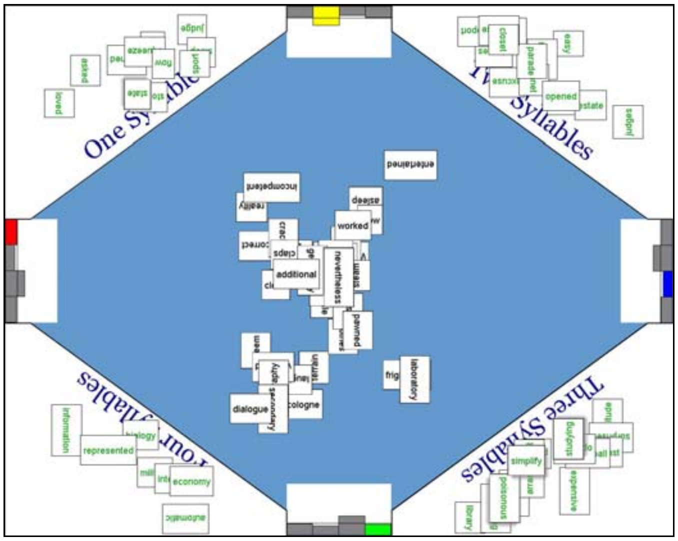

Figure 6. Interaction Visualizations in front of each user reflect his contribution to the activity (the number of attempted answers) relative to the rest of the team.

Interaction Visualizations

DiMicco et al. [7] explored the impact of a real-time visualization on group participation in a planning task (where participation referred to the amount each person contributed to a conversation). For the planning task, they found that over-contributors spoke less in the presence of the visualization, but that under-contributors did not increase their participation because they did not believe the display was accurate. Inspired by this study, we have integrated real-time histograms into some versions of our tabletop activities (Figure 6). The histograms appear on the table in the region directly in front of each user, and reflect the number of answers contributed by each group member based on tracking touch interactions with the table. We hypothesize that the histograms in our application will increase participation equity and awareness, and will have greater impact than in the setting in [7], because they should have more credibility to users in the context of a computer-mediated activity, since they track the number of answers each group member has contributed. We have also altered the design of the visualizations described in [7] to make them more appropriate for a collaborative educational activity by having a customized visualization for each group member (instead of a single visualization for viewing by the entire group); we highlighted the current user's bar in color and grayed out the bars representing the other three users, so that students would feel they were comparing themselves more to a group average than competing directly against other group members.

EVALUATION

To evaluate the utility of tabletop groupware for foreign-language learning activities, and to explore the impact of our design variations on participation and awareness, we observed the use of our system in two distinct contexts. The first context was a controlled laboratory setting where subjects who were learning a foreign language (but were not all members of the same class, or at the same level of proficiency) used the table for a single, one-hour session. The second use context was as part of a language course at our university. This latter set of students used the table during two sessions over the course of two weeks, completing activities that tied in to the current lessons in their class.

Context 1: Controlled Lab Setting

In the lab setting, we used the ClassificationTable software with four variations of an activity where clues contained Spanish-language descriptions of geographical and cultural facts about Spanish-speaking countries, which had to be grouped according to the country they described. This study explored our first two design variants – feedback modality and feedback privacy.

We recruited thirty-two paid subjects from within our university, who performed the experiment in groups of four (for a total of eight groups). Subjects' ages ranged from 18 - 28; 17 were men and 15 were women. Subjects spoke English as their primary language and had formally studied Spanish by taking at least one year of classes, and in some cases had completed several years of study.

The study employed a within-groups design, with each group using the table in each of four feedback conditions (no feedback, visual feedback, public audio feedback, and private audio feedback). In each condition, groups sorted different sets of clues (about 32 clues per set) about distinct collections of Spanish-speaking countries. To avoid ordering and learning effects, the presentation order of clue sets and feedback types were balanced among groups using a Latin Square design.

Users were instructed to converse in Spanish during the activity to simulate use in a real language-learning environment. After each of the four activities, participants individually completed a questionnaire asking them to rate several aspects of their interaction in the current condition. After the final condition, each participant completed an additional questionnaire that asked them to make comparisons among the four feedback design alternatives they experienced.

Each session was directly observed by two of the authors (who took notes) as well as videotaped for post-hoc analysis. All user interactions with the DiamondTouch table (i.e., who moved clues and where they were placed) were logged by our software.

Context 2: Authentic Classroom Use

One drawback of the laboratory session was participants' lack of investment in the activity – in particular, many concerns that would be present in a real classroom (such as concerns about peers' opinions of performance) were not present in the lab setting. Additionally, we wanted to observe repeated use of the tabletop and be able to discuss with teachers the impact of the activities on student performance in the classroom. Accordingly, we integrated our technology into the curriculum of an "English as a Second Language" course at our university. We met with the course's instructors, who provided us with curriculum-related content to use in the ClassificationTable, MatchingTable, and PoetryTable activities. The students in this course were from a variety of backgrounds, speaking several different languages natively (Japanese, Chinese, Korean, or Italian), and were all learning to speak, read, and write English. The course had twenty students, most of whom completed a total of five tabletop activities over a two-week period. During these sessions, some of the activities employed the spatial layout variants and interaction visualizations in order to allow us to observe the impact of these designs on participation and awareness. Due to scheduling difficulties, illness, and other absences, not all students completed all activities – four groups of four students each completed the three activities contrasting the "four piles" and center layouts, and three groups of four students each completed the two activities exploring the impact of interaction visualizations. Because of the small size of the class, we cannot report statistically significant analyses from this context, but instead report trends that we observed which would be interesting to study further with larger groups in a more formal context.

FINDINGS

Visual vs. Audio Feedback

Overall, audio feedback promoted increased awareness of contributions to and performance in the classification activity. Audio feedback increased the accuracy of individuals' self-assessments of their own contribution to the activity – on the questionnaires given after each task, each subject was asked to estimate what percent of items he had sorted correctly on his first try. By comparing these estimates to logs of actual activity, we can see that the estimates were significantly more accurate with private and public audio feedback as compared to no feedback at all, while the visual feedback did not provide any improvement over the baseline (no feedback) case. Mean deviation from true values for estimates of the percent of clues a user sorted correctly on her first try: No feedback = 24.46%, Private audio = 13.42%, Public audio = 14.44%, Visual = 16.27%. Repeated Measures ANOVA: ($F(1,3)=2.91, p<.05$). Paired-sample T-tests: no feedback/private audio ($t(24)=2.49$, $p<.03$), no feedback/public audio ($t(24)=2.15$, $p<.05$).

Measuring the amount of time each group spent conversing by automated audio analysis of the videotapes revealed that groups spent significantly more time conversing with each other when they received public audio feedback (74.32%) as compared to with visual feedback (60.83%). Mean percent of time spent talking in each session: Public audio feedback = 74.32%, Visual feedback = 60.83%. Paired-sample T-test: ($t(7)=2.80$, $p<.03$).

Visual feedback resulted in skewed participation in terms of total sorting actions as compared to private audio, which resulted in more equitable participation within groups. Participation equity can be measured by comparing the standard deviation of the number of answers contributed by each member within a group; lower standard deviations reflect a more equitable distribution of interactions. Mean standard deviation among groups' total sorting actions: private audio = 4.13, visual feedback = 6.16. Repeated Measures ANOVA: ($F(1,3)=3.19$, $p<.05$). Paired-sample T-tests: private audio/visual feedback ($t(7)=2.37$, $p<.05$).

Public vs. Private Feedback

When using the private audio played over individual users' earpieces, users communicated to the group the feedback that they were receiving privately, through both vocalizations and gestures. Positive feedback was typically clearly emphasized to the group via thumbs-up gestures or exclamations of "Sí," "Bueno," and "Yeah!" Negative feedback was typically acknowledged more subtly, by a slight head-nodding "no" or the user moving the incorrect clue back into the central region without commenting. This lack of drawing attention to the negative feedback supports our design motivation for providing such feedback privately – to reduce potential embarrassment over incorrect answers by not pointing them out to the entire group.

Likert-scale responses to the statement "I felt self-conscious when other people at the table knew whether or not my answer was correct" indicated that users felt less self-conscious about their performance with the private audio feedback as compared to either visual feedback or even a lack of any feedback at all. The responses also indicated that private audio made people feel less self-conscious about their answers than the public audio feedback, but with marginal statistical significance. Mean 7-point Likert scale scores (7 = strongly agree, 1 = strongly disagree): No feedback = 2.86, Private audio = 2.25, Public audio = 2.78, Visual feedback = 2.71. Repeated Measures ANOVA: ($F(1,3)=3.63$ $p<.02$). Paired-sample T-tests: private audio/no feedback ($t(27)=1.97$, $p\leq.05$), private audio/public audio ($t(21)=1.81$, $p=.08$), private audio/visual feedback ($t(30)=2.89$, $p<.01$).

The increased comfort with private audio was also reflected in users' comments on the questionnaires, such as "Private audio gave me instant feedback without everyone knowing I got it wrong," and "I prefer private audio, so I can know what I got correct or incorrect with a sense of confidentiality," and "I would guess more [with private feedback] since others couldn't observe me." In contrast, comments like "Public Audio: guessing [was] embarrassing" indicate reduced comfort for underperformers in the public audio feedback condition.

Typically, subjects first selected clues from the central pile and read them (sometimes silently and sometimes out loud). If a user was confident that he knew the answer, he immediately classified the clue into one of the corner areas.

Otherwise, he read the clue out loud and polled other group members for advice regarding its placement. Work tended to be more parallel near the beginning of each session, when users had many clues to explore, and reverted to a more serial strategy as the "easier" clues were sorted and groups were forced to discuss and debate the placement of more difficult items. The public audio feedback engendered a more serial strategy than other feedback styles, because when more than one clue was sorted simultaneously it became difficult for users to disambiguate which feedback sounds were associated with which clues. As a result, users sometimes needed to re-sort the same clue in order to replay the sound. Because of this difficulty, groups consciously attempted to work more serially with the public audio design.

Piles vs. Centered

Trends in the data collected during classroom use of the ClassificationTable, MatchingTable, and PoetryTable applications support our hypotheses that laying out information in four piles, one near each group member, rather than in the center of the table, seemed to encourage more equitable participation. We measured participation along two dimensions – the number of touch interactions on the DiamondTouch surface and the percent of foreign-language conversation contributed by each group member. Both of these measures were gathered automatically, based on touch data recorded by the DiamondTouch and voice data recorded by the microphone headsets worn by each participant. For each group, we calculated the standard deviation of the percent of touch events contributed by each group member, and of the percent of talking time contributed by each group member. Lower standard deviations reflect more equitable contributions among group members. For each of the three activities completed, groups had lower standard deviations for both the percent of touch interactions contributed and the percent of conversation contributed under the "four piles" condition. (ClassificationTable: mean stdev touches (center) = .096, mean stdev touches (piles) = .060, mean stdev talking (center) = .169, mean stdev talking (piles) = .161; MatchingTable: mean stdev touches (center) = .137, mean stdev touches (piles) = .122, mean stdev talking (center) = .195, mean stdev talking (piles) = .141; PoetryTable: mean stdev touches (center) = .138, mean stdev touches (piles) = .076).

Students' comments on questionnaires distributed after the activity reflected a potential drawback of the "four piles" layout, however, suggesting that it detracted from the collaborative feel of the activity. Some students indicated in their comments that the piles were "more dependent of yourself [sic]," while the center layout made them think more of "teamwork." Another student wrote about the centered layout, "We can discuss together and work together. I think it's more interesting than four piles." These statements indicate that the center layout may have been more successful in achieving the goal of a cooperative educational activity, by allowing the students to reach a shared understanding through conversation and collaboration.

Interaction Visualizations

No consistent impact of the presence or absence of the interaction visualizations on the equitability of touch or speaking interactions was found. However, due to the limited number of students in the class, it is premature to draw conclusions from this data.

Student reactions to the presence of the visualizations were mixed. Some students enjoyed the competitive feel that the visualizations lent to the activity, commenting "make me be competitive – encourage [sic]," while others found this intimidating, as indicated by comments like "I become too self-conscious. Concerned too much about the graph [sic]."

General Observations

Overall, the ClassificationTable, MatchingTable, and PoeteryTable were easy for subjects to learn and use. Groups spent an average of only 1.06 minutes learning to use the applications in the tutorials (they were allowed to remain in the tutorials as long as they felt necessary), despite the fact that 41 of the 48 (85.4%) had never used a DiamondTouch table before. Subjects were also very engaged in the educational activity, speaking in the foreign language throughout. Many commented to each other during the task that they found the activity fun and entertaining. No major usability problems were observed, other than the difficulty of disambiguating the target of the public audio feedback for near-simultaneous sorting actions in context 1.

DISCUSSION

The results of our evaluations supported several of our initial hypotheses regarding the impact of our four design variants on participation equity and self-assessment accuracy. Private feedback reduced embarrassment over contributing incorrect answers to the group activity, and resulted in modest increases in participation equity. Audio feedback increased conversation levels, and promoted more accurate self-assessment as compared with visual feedback. Laying out clues in piles near each of the four users, rather than in the center of the table, seemed to increase participation equity, although it had the unanticipated drawback of reducing the collaborative feel of the application. It would be beneficial to explore these designs through more extensive laboratory and classroom use to confirm these effects with greater statistical confidence.

Overall, we found that the table was an engaging platform for foreign-language education activities, promoting face-to-face discussion and providing students with feedback regarding their progress without necessitating the presence of an instructor. Participants found the table easy to use, and both students and teachers were excited about the new technology.

Methodological Limitations

A few of our initial design questions remain unanswered, due to methodological limitations of our evaluation strategy. We plan to pursue these questions in future work. There were drawbacks of both the laboratory-style evaluation approach of context 1 and of the naturalistic-setting evaluation of context 2. Groups in the laboratory context, since they did not know each other beforehand and were not completing the activities as part of an actual class, were not motivated by the same concerns as students in an authentic context – in particular, they may have been less concerned with how their peers perceived them, and so were not affected by embarrassment over producing incorrect answers, which we suspect would play more of a role in an authentic learning environment. The nature of our second observational context, as part of the curriculum of an actual foreign-language course at our university, made it difficult to gather statistically significant data regarding the impact of our design variations, since there were only a small number of students in the class, and the students' inexperience with English (they spoke several different native languages) made the gathering of subjective data from questionnaires unreliable. Regulations regarding educational fairness and privacy at our institution made assessing the impact of the use of our systems on students' grades in the course infeasible.

In context 2, we attempted to assess the impact of both the piles/center variant and the interaction visualizations on students' awareness of the amount they had contributed to the activity. On questionnaires following the tabletop activities, we asked students to rank their level of contribution relative to other group members; we planned to compare these assessments to the records of contributions kept by our software. However, for both of these design variants, all students who responded to the question ranked themselves as "average," and many didn't respond at all. We suspect that this phenomenon is due to the difficulty many of our participants had in understanding English and therefore in completing our questionnaire. As a result, we were unable to assess our designs' impact on self-assessment accuracy in this particular use context – we plan to address this issue in a re-designed study, as discussed in our Future Work section.

Feedback from Foreign Language Instructors

The three instructors who co-taught the English-as-a-Second-Language course that utilized our technology were generally positive about the potential for tabletop technology in their classroom after their students used our applications. They described the benefits of the technology for their students as, "...a chance to work in groups: share ideas, collaborate, communicate, problem solve, look for clues." They also noted, "the students said they learned from it." Also, they mentioned that digital technology enabling co-located group work helps to "extend the learning period without the pressure of having the instructor," thus creating a comfortable learning environment for students.

Regarding fee-rider issues, the instructors informed us that they were "very much" concerned about under-participation in group activities. However, they felt that real-time interaction visualizations might not be appropriate during a group activity, noting that "...at least for our program, the information may be a little too sensitive to share." However, they suggested that, rather than potentially adding a competitive feel to a within-group activity by showing interaction visualizations, adding a competitive feel to between-group activities, by showing groups how well they performed compared to other groups, might increase engagement and participation without the potential of stigmatizing individual members of the class. The instructors expressed interest in using the DiamondTouch table and our ClassificationTable, MatchingTable, and PoetryTable software again in future iterations of their course.

Future Work

Because of the limitations due to the small class sizes at our university on our ability to draw statistically significant conclusions regarding the impact of our interface design variations on participation and self-assessment, we plan to conduct an additional formal experiment to explore the impact of the "four piles" vs. "center" designs and of the presence of interaction visualizations. We are also interested in how the concept of interaction visualizations can be adjusted to avoid stigmatizing the performance of specific students; designs that include private displays for each group member (such as PDAs) is one possibility. We also plan to continue to support authentic use of our tabletop applications at our university, to gather additional feedback from students and instructors. We plan to continue to pursue the theme of exploring how subtle aspects of interface design can impact issues of pedagogical importance.

CONCLUSION

We have presented a discussion of design issues for tabletop interfaces that support cooperative language learning. The affordances of interactive tables for supporting face-to-face group work, providing immediate feedback to students regarding their progress, and recording interaction histories for instructors to review, make tabletop technology an exciting new platform for educational software. We developed three tabletop CSCL applications targeted at foreign-language education: the ClassificationTable, MatchingTable, and PoetryTable.

We explored how the properties of software for interactive tables can be tuned to achieve specific pedagogical goals, such as increasing participation equity among group members and improving the accuracy of students' self-assessments. The four design variants (feedback modality, feedback privacy, spatial configuration, and interaction visualizations) were evaluated in two contexts – a laboratory evaluation, and an authentic classroom setting. The results of these evaluations cast light on the impact of these design choices, such as the participation-equalization effect of private feedback and the self-assessment accuracy increase of audio feedback. Some results (such as the potential equalization effect of a distributed rather than centralized placement of materials) are preliminary, and highlight issues to be explored in future work.

Although interactive table technology isn't yet available to most educators, the match of the technology's affordances with the educational goals of small-group work suggest that understanding issues regarding the design of cooperative software for interactive tables is a valuable investment in what may become an important educational technology platform a few years down the road.

ACKNOWLEDGMENTS

Acknowledgements have been removed for blind review.

REFERENCES

1. Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development*, 59, 87–96.
2. Bos, M. C. (1937). Experimental study of productive collaboration. *Acta Psychologica*, 3, 315–426.
3. Cappelletti, A., Gelmini, G., Pianesi, F., Rossi, F., and Zancanaro, M. Enforcing Cooperative Storytelling: First Studies. *Proceedings of ICALT 2004 (4th IEEE International Conference on Advanced Learning Technologies)*.
4. Coleman, E. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of Learning Sciences*, 7, 387–427.
5. Corbett, A. T., Anderson, J. R. Locus of feedback control in computer-based tutoring: impact on learning rate, achievement and attitudes. *Proceedings of CHI 2001*, 245-252.
6. Dietz, P., and Leigh, D. DiamondTouch: A Multi-User Touch Technology. *Proceedings of UIST 2001*, 219-226.
7. DiMicco, J.M., Pandolfo, A., and Bender, W. Influencing Group Participation with a Shared Display. *Proceedings of CSCW 2004*, 614-623.
8. Johnson, P. and Arenas, J. (1995). Students' Perceptions of Foreign Language Communicative Teaching and Class Size. *Education Resources Information Center* (ERIC) #ED396560 (http://www.eric.ed.gov).
9. Joyce, W.B. On the Free-Rider Problem in Cooperative Learning. *Journal of Educational Business*. 74(5), 271-274, May-June 1999.
10. King, A. (1990). Facilitating elaborative learning in the classroom through reciprocal questioning. *American Educational Research Journal*, 27, 664–687.
11. Matsushita, M., Iida, M., Ohguro, T., Shirai, Y., Kakehi, Y., and Naemura, T. Lumisight Table: A Face-to-Face Collaboration Support System that Optimizes Direction of Projected Information to Each Stakeholder. *Proceedings of CSCW 2004*, 274-283.
12. Morris, M.R., Morris, D., and Winograd, T. Individual Audio Channels with Single Display Groupware: Effects on Communication and Task Strategy. *Proceedings of CSCW 2004*, 242-251.
13. Rogers, Y. and Lindley, S. Collaborating Around Large Interactive Displays: Which Way is Best to Meet? *Interacting with Computers*, 2004.
14. Rogers, Y., Hazlewood, W., Blevis, E., and Lim, Y. Finger Talk: Collaborative Decision-Making Using Talk and Fingertip Interaction Around a Tabletop Display. *Proceedings of CHI 2004 Extended Abstracts*, 1271-1274.
15. Ross, S. Self-Assessment in Second Language Testing: A Meta-Analysis of Experiential Factors. *Language Testing*. 15(1), 1-20, March 1998.
16. Ryall, K., Forlines, C., Shen, C., and Morris, M.R. Exploring the Effects of Group Size and Table Size on Interactions with Tabletop Shared-Display Groupware. *Proceedings of CSCW 2004*, 284-293.
17. Scott, S.D., Carpendale, M.S.T., and Inkpen, K. Territoriality in Collaborative Tabletop Workspaces. *Proceedings of CSCW 2004*, 294-303.
18. Scott, S.D., Grant, K., and Mandryk, R. System Guidelines for Co-located Collaborative Work on a Tabletop Display. *Proceedings of ECSCW 2003*, 159-178.
19. Sluis, R., Weevers, I., van Schijndel, C., Kolos-Mazuryk, L., Fitrianie, S., and Martens, J. Read-It: Five-to-Seven-Year-Old Children Learn to Read in a Tabletop Environment. *Proceedings of Interaction Design and Children 2004*, 73-80.
20. Shen, C., Everitt, K., and Ryall, K. UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces. *Proceedings of UbiComp 2003*, 281-288.
21. Shen, C., Lesh, N., Forlines, C., and Vernier, F. Sharing and Building Digital Group Histories. *Proceedings of CSCW 2002*, 324-333.
22. Shen, C., Vernier, F., Forlines, C., and Ringel, M. DiamondSpin: An Extensible Toolkit for Around-the-Table Interaction. *Proceedings of CHI 2004*, 167-174.
23. Stewart, J., Bederson, B., and Druin, A. Single Display Groupware: A Model for Co-present Collaboration. *Proceedings of CHI 1999*, 286-293.
24. Streitz, N., Geiβler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. i-LAND: An Interactive Landscape for Creativity and Innovation. *Proceedings of CHI 1999*, 120-127.
25. Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13, 21–39.
26. Wu, M. and Balakrishnan, R. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. *Proceedings of UIST 2003*, 193-202.