Incorporating Peephole Interactions into Children's Second Language Learning Activities on Mobile Devices

Brenna McNally[1], Mona Leigh Guha[1], Leyla Norooz[1], Emily Rhodes[2], Leah Findlater[1]

[1] Human-Computer Interaction Lab
College of Information Studies
University of Maryland
College Park, MD 20742, USA
bmcnally@umd.edu, mona@cs.umd.edu,
leylan@umd.edu, leahkf@umd.edu

[2] Usability Lab
College of Arts and Sciences
University of Baltimore
Baltimore, MD 21201
emily@emilyrhodes.com

ABSTRACT
Physical movement has the potential to enhance learning activities. To investigate how movement can be incorporated into children's mobile language learning, we designed and evaluated two versions of a German vocabulary game called *Scenic Words*. The first version used movement-based dynamic peephole navigation, which requires physical movement of the arms, while the second version used touch-based static peephole navigation, which only requires standard touchscreen interactions; static peepholes are the *status quo* interaction technique for navigation, commonly found, for example, in map applications and games. To compare the two types of navigation and to assess children's reactions to dynamic peepholes, we conducted an in-home study with 16 children (ages 8–9). The children participated in pairs but individually played each version of the game on a mobile device. While results showed that the more familiar static peepholes were the preferred interaction style overall, participants became accustomed to the movement-based dynamic peepholes during the study. Participants noted that the dynamic peephole interaction became easier over time, and that it had some advantages such as for dragging-and-dropping elements in the game.

Categories and Subject Descriptors
H.5.2 [Information interfaces and presentation]: User interfaces – *user-centered design*.

General Terms
Design, Human Factors.

Keywords
Children, mobile, physical movement, peepholes, learning.

1. INTRODUCTION
Engaging in physical movement has been shown to enhance children's learning, for example, in gesturing while solving math problems [22] or learning new vocabulary [34]. Physical movement has also been incorporated into educational software with motion sensing technologies such as the Nintendo WiiMote [21] and the Microsoft Kinect [24].

IDC'14, June 17–20, 2014, Aarhus, Denmark.

http://dx.doi.org/10.1145/2593968.2593982

Figure 1. Participants using the Scenic Words application to learn German vocabulary with movement-based dynamic peephole navigation (left) and touch-based static peephole navigation (right).

We investigate how ubiquitous, off-the-shelf mobile devices can be used to support physical movement in children's second language learning activities. Current mobile devices can sense a variety of movement-based interactions through the touchscreen itself or with accelerometers, gyroscopes and even cameras. These devices are also increasingly accessible to today's youth: an international survey by the Groupe Speciale Mobile Association found that 65% of children ages 8 to 18 use a mobile phone, 27% of whom are smartphone users [19]. In the US, eight of ten children ages 5–8 have used smartphone or tablet devices [26]. This ready access to sensor-laden mobile devices gives us the opportunity to incorporate physical movement into learning activities, promoting the goal of learning anytime, anywhere.

Children's use of mobile device interaction techniques has been studied primarily in terms of stationary use. Focusing on the touchscreen, for example, studies have examined touch accuracy for tapping different objects [7, 5], issues children have with movement-based gestures (e.g., swipe and drag-and-drop) [8], and how children draw using gestural input with and without visual feedback [3]. These works noted a marked difference in performance between adults and children [7, 5], highlighting the importance of understanding and designing for children's interactions specifically.

In this paper, we investigate *peephole* interactions [18, 25, 35, 39] as a means of incorporating movement into a children's mobile learning application (Figure 1). Peepholes enable a mobile device to act as a lens into a virtual world. With *static peepholes* children explore a virtual scene by swiping the touchscreen to bring different areas into view; static peepholes are the status quo navigation technique found in many map applications and games. *Dynamic peepholes* are controlled more actively, by physically moving the device around as if it were a small window, while the virtual world remains still. While studies on peephole interaction with adults have explored performance and preference [35, 18, 25], to our knowledge no research has examined the design of such interaction with children.

We employed peephole interactions as a starting point to investigate the design of movement-based second language learning software for children. Our primary research questions were: (1) How do children explore and interact with static and dynamic peepholes? (2) Do dynamic peepholes have the potential to elicit movement in children's mobile interactions, toward our goal of augmenting learning activities?

We conducted a study with sixteen children aged 8–9 to compare use of a dynamic peephole interface (with arm movement) to a static peephole interface (touchscreen gestures only). To do so, we first designed and built a German vocabulary learning game, called *Scenic Words*, for a mobile phone platform. In Scenic Words, users discover words hidden in clouds in the sky and categorize these words by dragging them into labeled jars at the bottom of the scene. We created two versions of the game, one that uses static peephole navigation to move about the scene, and one that uses dynamic peephole navigation. Children then participated in pairs in an in-home study session that was designed to closely approximate a context where natural play would occur. For both types of peepholes, we looked at how participants positioned the device during use, how they explored the virtual space, their preferences, and how they performed on vocabulary recall and matching tests.

Results showed that the *status quo* static peepholes were the preferred interaction style overall. However, results also highlight the potential of using dynamic peepholes to elicit movement in children's mobile interaction. Participants became accustomed to the movement-based dynamic peephole navigation over the course of the study, noting that it became easier to use, and that it had some advantages over the static peepholes (e.g., for dragging-and-dropping elements in the game). The primary contributions of this work are insights into how children use dynamic and static peephole interaction techniques, how children think about each technique and why they prefer one technique over the other, and a comparison of vocabulary recall and matching in mobile learning activities that incorporate peephole interactions.

2. RELATED WORK

Many areas of literature inform the current work. We reviewed work on Mobile Assisted Language Learning, children's mobile interaction, movement and children's learning activities, and research (with adults) on peephole interactions.

2.1 Mobile Assisted Language Learning

Being fluent in multiple languages is becoming a critical skill for children in the US and elsewhere; the US National Standards for Foreign Language Education envisions a future where all students will be able to speak English and one other language to ensure they can communicate with a pluralistic American society and abroad [2]. Accordingly, Mobile Assisted Language Learning (MALL) is a growing field. MALL is a subset of M-learning—learning facilitated by mobile technologies that are potentially available anytime, anywhere [10, 29]. A review of MALL research from 2007-2012 found that these technologies have been used to explore many concepts in second language learning—such as language learning outside the classroom and game-based learning— supporting the idea that MALL activities can enrich learners' second language acquisition [38].

Mobility of learning generates new modes of educational delivery: personalized, learner-centered, situated, collaborative, ubiquitous, and lifelong learning [37]. Kumar's work on mobile games with rural children in India used speech recognition to help them read content with understanding [30], and looked at how children voluntarily incorporate the use of mobile games to learn oral, written, and vocabulary concepts [31]. Mobile devices have also been used to support collaborative activities among peers with reading, improving collaboration and promoting motivation [32]. Educational delivery in MALL has the potential to aid learners in understanding a variety of language concepts.

2.2 Children's Mobile Interactions

The marked differences in performance found between how adults and children using touchscreen technologies suggest a need to study and design touch interactions specifically for children [5, 7].

Research on children's mobile device interaction has outlined several challenges and developed suggestions for overcoming them. For instance, in studying how to interpret children's gesture input on mobile devices, Anthony *et al*. [5] found that attempts to tap .25" square targets failed 30% of the time with children age 7-10. Touch accuracy could be improved by using larger targets or by increasing the active area around the target to allow slightly out-of-bounds targets to count as a hit [4, 5]. Work by Brown et al. [7] also noted that children had difficulty tapping small targets. Additionally, children in that study preferred to use individual strokes to create gestures, such as a square, rather than continuous strokes. We build off of these findings in the design of Scenic Words, for example, ensuring the touchscreen targets (clouds, jars) are larger than the recommended minimum size.

2.3 Movement in Children's Learning Technologies

Using technology to incorporate movement into children's learning activities can be accomplished in many ways. Mobile technologies that encourage users to move and explore their environments have been used to facilitate the investigation of scientific questions in real-world contexts [1, 28] and to bridge informal learning contexts such as parks and museums with formal classroom learning contexts [9, 27]. Movement has also been incorporated into shared storytelling activities to facilitate creating stories in varied contexts [14]. While these mobile learning technologies allowed movement, other research required the movement of the mobile devices in order for participants to complete objectives such as data collection [36].

Other learning technologies more directly tie physical movement to learning, which is also our focus. Antle [6], for example, explored how to teach children musical concepts such as harmony, melody, and rhythm through whole-body interfaces that create music. While this study was successful at teaching children musical concepts, certain interactions were not discoverable for some children and showed that there can be an initial learning period when including new movement-based interactions. Motion sensing technologies, such as the Nintendo WiiMote or the Microsoft Kinect, have also been explored for their ability to enhance learning. For instance, the Mathematical Imagery Trainer [21] leverages a WiiMote to teach proportional equivalence to children. In *The Potential of Kinect in Education* [24], Hsu discusses benefits such as interesting interaction types and

Figure 2. The virtual world in Scenic Words, where users collect clouds displaying vocabulary words into jars displaying associated categories.

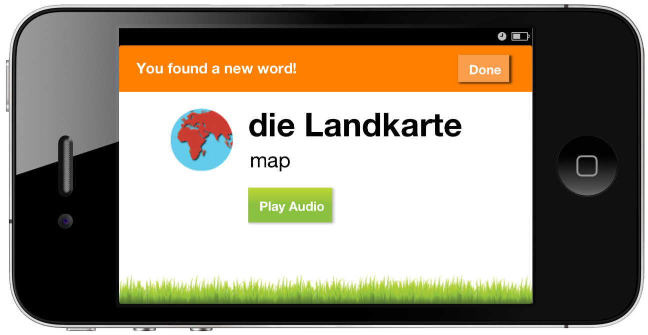

Figure 3. The Word Card in Scenic Words, which appears when a cloud or bin is tapped.

promoting learning via multi-sensory input as well as constraints such as large space requirements and calibration time. Xdigit [33], for example, is a gesture-based children's game for the Kinect whose goal is to enhance arithmetic learning. Dindler [11] explored using a Hydroscope, technically a large dynamic peephole device that users pushed along the floor to view a digital ocean, at a marine center to convey concepts about river beds and the properties of previously constructed fish. In contrast to these systems, mobile technologies (our focus) do not have large space requirements and users expect systems to work with minimal or no calibration.

2.4 Peephole Interaction Comparisons with Adults

Numerous studies with adults have investigated the differences between static peepholes (where the device remains still) and dynamic peepholes (where the user physically moves the device). Mehra et al. [35] found that dynamic peephole interactions were more natural than static peepholes as they allowed users to rely on spatial memory. Users were also able to complete tasks faster and more accurately with dynamic peepholes. Hürst and Bilyalov [25] compared the use of static and dynamic peephole navigation to explore 360-degree panoramic images, finding that dynamic peepholes were preferred and participants performed tasks better when using this condition. However, when participants were seated and unable to utilize the full rotation of the device, most opted for static peephole interactions.

Other results have been less straightforward and suggest that context may impact use of peephole interactions. Wenig et al. [39] conducted a field study of pedestrian navigation along a specific route, where participants used a static peephole, a dynamic peephole, or a static map image. While no significant difference was found between the two peephole conditions, they both outperformed the regular photographs. Grubert's work [17, 18], in contrast, found that peephole interactions alter between contexts. Static peepholes and magic lens peepholes, a version of peepholes that uses augmented reality to provide an overlay of information to what the user sees through a device, were evaluated by playing a find-and-select game at a public transportation stop. Findings in an initial study indicated that participants preferred magic lens interactions, while findings in a repeat study indicated that magic lens interactions were used and preferred less. The social context, such as one public transportation stop being a transit area and the other being a waiting area, could have influenced the participants' preference of interfaces [18].

While the literature on peephole interactions demonstrates affordances that may be context dependent, these interactions have yet to be explored with children. Additionally, as children are still advancing through stages of development that are different from those of an adult, their interactions and preferences for the different peephole conditions may differ greatly.

3. SYSTEM DESIGN

We designed a mobile application named *Scenic Words*. The Scenic Words mobile application is a German vocabulary learning game designed for iPhone and iPod touch and intended for use by children ages 8 to 9. It is written in the Objective C programming language. We iteratively designed the game with children and built it to support the two different versions of navigation: dynamic and static peepholes.

3.1 Design Process

The Scenic Words application was developed using an iterative design process. Two Cooperative Inquiry [12, 13, 15, 20] design sessions were held at different points in the development process: one in the formative stage, and one in the evaluation stage of our iterative design process. The same children, ages 7 to 11, participated in both design sessions.

During the formative design session, the adult and child design partners worked in small groups to envision how they would use a peephole environment to learn a new language. For each group, three walls in a small room were covered in paper so that the virtual environment could be drawn, and mobile device outlines and art supplies were used to illustrate what would be displayed on the mobile devices as the virtual worlds were explored. Groups presented their ideas on the virtual world, how the mobile device would be used, and how learners would use this space to learn a foreign language. Design ideas from this session suggested that an application using peephole interactions to promote language learning activities should 1) combine realistic and imaginary elements in the virtual space, 2) teach a second language in small parts, 3) use a variety of physical movements, and 4) include game-like elements.

Following this initial design session we built the first iteration of the Scenic Words application. We incorporated findings from our literature review, such as making sure our touchscreen targets were large enough and increasing the active area, as well as the outcomes from the early design session, such as combining realistic and imaginary elements in the virtual space by having the children catch "clouds" in "jars".

We then presented the prototype application to adult and child design partners in a second design session, which focused on evaluation of a prototype within the iterative design process. Small groups were given mobile devices with a prototype of the Scenic Words application running, and were asked to give their

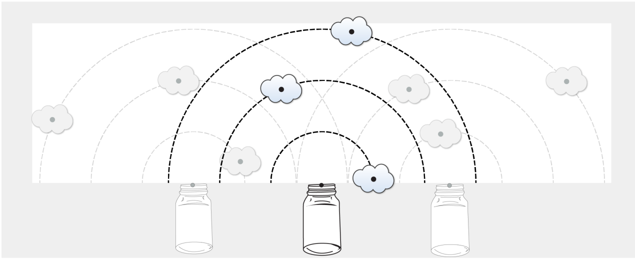

Figure 4. To ensure that each user had to move clouds to jars from a variety of distances, clouds were randomly placed along three evenly spaced arcs that were centered at the corresponding category jar.

feedback on both the dynamic and static conditions of gameplay. No direction was given regarding how to interact with the dynamic condition, and only one group was able to independently discern how to interact with the device to see the virtual space after exploring the application for about a minute. After an adult designer suggested, “now try using the app by standing and using your arms” there was a collective “oooh!” from the children, and the activity commenced in earnest. In addition to insights into how children are likely to explore the application and the direction they would need, the feedback from the evaluative session resulted in numerous simplifications to the application. Most notably, these simplifications included the removal of vocabulary word definitions because they were considered distracting and unnecessary in such a short activity.

3.2 Gameplay

The goal of Scenic Words is for the child to learn German vocabulary words. Clouds holding words are distributed across a sky and jars labeled with word categories (e.g., “school”, “colors”) appear on the ground; see Figure 2. Users earn points for correctly categorizing words by dragging clouds to the appropriate jars. For instance, the German word for “apple” would go in the jar labeled with the German word for “food”, and a user’s score would increase by one point.

Clouds and jars initially display only a question mark and are unmovable (Figure 2). To activate a cloud and jar, the user taps it to bring up a Word Card (Figure 3). A Word Card shows the new vocabulary word 1) in German, 2) in English, 3) with a picture of the item, and 4) with an audio clip of a native speaker saying the word. Once the user closes the Word Card, the now-activated cloud or jar displays the German word and, for clouds, the picture as well. From this point, the cloud can be moved and the jar can have items placed in it. The user makes a drag-and-drop interaction to move an activated cloud to a jar. The game ends when all nine clouds have been placed in the correct jars.

When the game loads, three jars (categories) and nine clouds (words) appear, such that there are three words for each category.

3.3 Navigation and Placement

To locate the word clouds and the category jars, the user must explore the virtual world with peephole interactions. In the static peephole version of the game, users swipe the touchscreen to explore the virtual world, scrolling the scene as they locate and categorize the clouds. The dynamic peephole version requires users to explore the same virtual space by physically moving their mobile device; touchscreen gestures do not scroll the scene. This dynamic peephole interaction was implemented through the use of the 3-axis accelerometer and the gyroscope that are built in to the iPhone. Beyond these navigation differences, both versions of the game were the same.

Table 1. Pairs of participants were randomly assigned to an order. They were assigned either the static or dynamic condition first, paired with either word set 1 or word set 2.

	Static First	Dynamic Second
Order 1	Word Set (WS) 1	WS 2
Order 2	WS 2	WS 1
	Dynamic First	**Static Second**
Order 3	WS 1	WS 2
Order 4	WS 2	WS 1

To ensure that users would need to explore different areas of the scene and move clouds from a variety of locations to the jars, we controlled item placement as follows. Jars, corresponding to categories, were randomly assigned to the left, middle, or right position. The three clouds containing words associated with that jar were then randomly placed at three 1.2” (200 pixel) intervals from the top-center point of the jar, falling along three 180-degree arcs (Figure 4). The scene itself was 5.2 times wider than the screen and 3.1 times taller, where the device screen measured 2.94” x 1.96”, respectively. The system was designed to shift the scene at approximately the same rate in both peephole conditions, that is, panning 4” in the static condition corresponds approximately to moving the device 4” in dynamic condition.

4. STUDY METHOD

To test Scenic Words in a context where we believe that children would use it in practice, we conducted in-home studies with eight pairs of children (sixteen participants in total). While we employed a mixed-methods approach, our focus was on use of peephole interactions *during* the language learning activity. Thus, this study was designed primarily to capture qualitative observations and self-report data, rather than precise performance measures of speed and accuracy in using peephole navigation.

German vocabulary was used for this study because courses in German are not available to students in the local public school system until high school, and we wanted to use a language with which participants would not have previous knowledge.

4.1 Pilot Tests of Method

Before we began the study, three potential methods for the in-home sessions were pilot tested with five children ages 8–9, three females and two males. We conducted three sessions such that the number of participants and presentation of the dynamic and static conditions varied as follows:

- Single participant
- Paired participants: Individual initial exploration of each condition, with each child having a device
- Paired participants: Joint initial exploration of each condition, with each child having a device

Having a single participant use Scenic Words and complete the study— including surveys, interviews, and the free recall tests— was an uncomplicated procedure. However, the feedback elicited from that participant during the survey and interview was terse and did not provide much insight into his interactions with the application. Testing the application with paired participants, where each child had the opportunity to explore each condition alone and then played again together, elicited more feedback as a dialogue developed between the participants regarding their use of the application. However, the relocation of the participants that this method required was logistically difficult and disruptive to the activities. Testing the application with paired participants where they played each condition together from start to finish elicited the most conversation between participants and with researchers, caused less disruption, and led to more unique interactions—such as using the device upside down. Accordingly, we chose this lattermost method for the study.

4.2 Participants

Sixteen children were recruited in pairs of two through advertisements to parents on listservs and through word of mouth. None had participated in the design sessions or pilot study. All children were 8 or 9 years old, with thirteen male and three female. All children had at least some previous experience using mobile touchscreen devices. Ten reported using these devices daily, while eight personally owned a mobile touchscreen device.

No children who had previous experience with German participated in the study, as German vocabulary learning was the focus of Scenic Words and we wanted to limit prior knowledge as much as possible. Nine of the participants had previously participated in second language learning activities, and reported varying degrees of fluency in eight languages including Spanish, French, Farsi, Arabic, Chinese, and American Sign Language.

4.3 Study Design

The study was designed as a 2×2 within-subjects factorial design with two factors:

- **Word Set:** Two sets of 9 vocabulary words
- **Peephole Type:** *Dynamic*, movement-based, or *static*, touch based, peephole interactions

To allow participants to learn new words when playing with each type of navigation, we created two word sets of nine words each (three words for each of three categories). The two word sets were then paired with the navigation conditions such that participants used word set 1 with dynamic peephole navigation and word set 2 with static peephole navigation or vice versa (Table 1). The orders of presentation for both factors were fully counterbalanced and pairs or participants were randomly assigned to an order.

The categories of German vocabulary words used in this study were chosen by consulting local curriculum for topic areas, such as "food". Vocabulary words that fell under the selected categories, such as "cheese" for "food", were then selected from a child's German language learning book [16] and were revised based on feedback from the second design session with children.

Each word set contained nine words: three vocabulary words in each of the three categories. Following the study, no statistically significant effects on the participants' recall or matching scores were found when comparing word set 1 to word set 2.

4.4 Procedure

Pairs of participants completed a single session that lasted from 45 minutes to an hour. Sessions began with a brief introduction to the project, a description of the Scenic Words game, and a survey on the participant's background. Parental permission and the children's assent were obtained, including for recording of audio, video, and photographs. Additionally, two researchers attended each session and recorded observational notes regarding the use of the system.

The static and dynamic peephole conditions were then presented in counterbalanced order. For each peephole condition, participants played the game twice, followed by a free recall test and a matching text. For the gameplay, we were interested in understanding the initial reactions of participants to the peephole conditions. As such, we provided no initial instruction on how to navigate the scene. If a participant became frustrated during the first attempt at a condition and remained stuck for more than 90 seconds, that attempt was abandoned and additional instruction was given. For the second repetition of the task in the dynamic condition, the instruction, "This time, I would like you to hold the device in front of you and use your arms to move around" was given to all participants, as pilot testing had shown that initial explorations of this condition varied and that this instruction was helpful in explaining how to explore the virtual space.

Table 2. Code set used to assess the participant's use of static and dynamic peephole conditions.

Navigation Styles	Expressed Likes and Dislikes
Attempt to interact with the background and cloud simultaneously	How long it took to complete the condition (like/ dislike)
Attempt to scroll background	Time to complete game being displayed (like/ dislike)
Tilt	Ability to drag (like/ dislike)
Body or arm rotation	Liked how easily words could be moved (like/ dislike)
Drag cloud and background separately	Disliked how tiring it was to complete the condition
Bump cloud against edges	Disliked Recalibrating
Unique interactions	Disliked Bouncing
Loss of navigation	Liked learning
Attempt to zoom	Liked Word Card elements
Device Locations	**Social Interactions**
Held up, free arms	Peeking
Held up, propped arms	Partner instruction
Held low in lap	Comparison of scores
Propped on a surface	
Flat on the floor	

Once the game had been completed twice for a given peephole condition, the two participants were separated into different rooms and each completed the free recall and matching tests. During the free recall test participants were asked to write as many words as they could recall of the nine new words, either as pairs of words (German and English) or individual words (just German or just English). Participants were informed that spelling did not count, and that if they were not comfortable writing the researcher would write for them. The free recall test ended when the participant indicated that they did not recall any more words. For the matching test, participants were asked to draw a line between any of the nine German-English pairs of words that they remembered.

After both peephole conditions and associated tests were finished, each participant individually completed a follow-up survey, which asked about perceptions of static and dynamic peepholes and the Scenic Words application in general. Finally, the researchers led a joint semi-structured interview with both children regarding their attitudes and preferences toward the application.

5. ANALYSIS

Multiple forms of data were collected. As this was a mixed-methods study, the data were analyzed qualitatively and quantitatively. All user input on the device (*e.g.*, touch events, word placement, time in word card) and device movement were logged automatically by the Scenic Words software.

To analyze the video data, interview responses, and observational notes, we developed a code set following the iterative process described by Hruschka et al. [23]. A single researcher first reviewed two participants' data and developed an initial code set, refining it after discussion with two other researchers. The code set was then further refined after analyzing two additional participants' data. Finally, two researchers independently coded two more randomly selected participants' data, noting whether an interaction was either present or not present. Agreement was 100% across all codes with the exception of body or arm rotation, where one disagreement occurred and Cohen's kappa was 0.5. The final code set is shown in Table 2.

Navigation styles, likes or dislikes expressed by the participants, and social interactions were coded using researcher notes, relevant survey responses, and relevant interview responses. Changes in device position were coded based on the videos of the sessions. Videos were reviewed in full; however, due to technical issues for one pair of participants (trial 5), only 7 of the 8 sessions were videotaped and able to be reviewed. In addition to recording when and how the participant changed the position of the device in each condition, we noted whether the researcher had to intervene to correct the participant's position.

The results of the free recall test were assessed on a 0-2 point scale, looking at word pairs: 0 points for not recalling any part of the word pair, 1 point for partial recall of the word pair (e.g. the English word and part of the German word, just the English word), and 2 points for recalling the full word pair. Spelling did not count toward the assessment (e.g. groon in place of grün). Category words were omitted from all calculations as they were seen across both conditions, and were therefore more frequently encountered by participants in comparison to the word sets. Pairs of words on the matching test were marked correct, incorrect, or blank. We compared scores on both tests between the static and dynamic conditions using Wilcoxon's signed rank test.

6. FINDINGS

The focus of this study was to investigate peephole interaction in the context of language learning activities, rather than on measuring precise performance (speed, accuracy) with the two peephole techniques. Thus, while this section begins with a brief overview of performance, our primary focus is on physical device position, observation and self-report data of issues such as confusion and frustration, and interaction with the language elements of the game (e.g., word cards). Throughout this section we identify individual participants by combining the trial or pair number (Tx) and participant number within the pair (Px): T4P2, for example, refers to trial number 4, participant number 2.

6.1 Overall Time and Errors

Overall, it took a similar amount of time to complete the game in each peephole condition. Note that participant T2P1 is excluded from the measures of total time spent in the peephole conditions, time in word cards, and number of times audio was played. This participant played audio clips 329 times in the first instance of the first condition (static peepholes), which was at outlier at more than three standard deviations away from the mean number of audio plays for all participants in the first instance of the static condition (M = 29, SD = 81.1).

Participants spent an average of 3.8 minutes (SD = 1.7) to complete each static condition and an average of 3.6 minutes (SD = 1.4) to complete each dynamic condition. Participants spent more of their time reviewing word cards in dynamic condition, an average of 3.5 seconds (SD =4.25) accounting for 26.2% of their time, than they did in static condition, which had an average of 2.7 seconds (SD = 3.5) accounting for 22.8% of their time. Thirteen incorrect word placements occurred in static condition, while seven occurred in the dynamic condition.

Table 3. Comparisons of device position during gameplay.

Differences in Device Positions	Static (%)	Dynamic (%)
At least one position change	21.4	64.3
Dominant and final position match	96.4	92.9
Initial and final positions match, 1st use	71.8	35.7
Initial and final positions match, 2nd use	85.7	78.6

The time it took participants to complete each gameplay suggests the second use was easier to complete for both peephole conditions. The first use of static peepholes took participants an average of 4.6 minutes (SD = 1.8) to complete, while it only took 2.9 minutes (SD = 0.9) to complete the second time. Similarly, the first use of dynamic peepholes took participants an average of 4.3 minutes (SD = 1.3) for the first use and 3.2 minutes (SD = 1.2) for the second use.

The number of condition failures, where the researcher intervened after ~90s to stop the condition because the participant exhibited frustration or chose to stop the condition, was higher in the dynamic than the static condition: 8 versus 2 failures respectively.

6.2 Device Position

We examined how participants held the device during gameplay in each peephole condition to understand how participant behavior changes over time, perhaps with participants finding a preferred or optimized position over the course of the session. Videos were reviewed to code changes in device position, including initial, final, and dominant positions, with the dominant position being the position used the majority of the time.

We generally saw experimentation during the beginning of a session followed by settling in to one position for the duration of gameplay. Participant T7P2 was an exception with position shifts throughout gameplay in both conditions, but there was one position in each peephole condition that the participant kept returning to.

Gameplay with static peepholes was largely stable in regard to the position of the device (Table 2), with participants favoring positions where the device was propped on a surface or laid down flat on the floor. Six instances (21.4%) of at least one position change of the device during gameplay were seen. The dominant position of the device during gameplay matched the final position 96.4% of the time. In the first use of the static condition, initial and final positions were the same 71.1% of the time. During the second use, after participants had become acclimated to the peephole interaction required to complete the game, initial device position matched the final position 85.7% of the time. Three of the four differences in initial and final position of the device during the second use were changes from holding the device to placing it on a surface or the floor. The fourth was a change from laying the device on the floor to having it propped on a surface.

Dynamic peephole interactions required more exploration than static peephole interactions. Participants were much more likely to have shifts in the device position when using the game in the dynamic peephole condition (Table 3), and participants ultimately favored positions where the device was held up. There were eighteen instances (64.3%) of at least one position change during dynamic gameplay. The dominant position of the device during gameplay matched the final position 92.9% of the time. Researchers intervened twice to correct position, both during the gameplay using dynamic peepholes: once during the second use to remind a participant to "use their arms", and once during the first use when a participant kept having to move a device around a table, suggesting the participant push back her chair. That noted, in the first use of the condition, initial and final positions were only the same 35.7% of the time. The second use, after participants had become acclimated to the peephole interaction required to complete the game and received additional instruction, found that initial device position matched the final position 78.6% of the time. Position changes during the second use of the dynamic condition, after the participants had received additional instruction, fluctuated almost entirely between the two dominant positions: having the device held up with arms free, and having the device held up with propped arms.

6.3 Recall and Matching

For the recall tests, each question was assessed using a scale from 0 to 2 points, for a maximum test score of 18. For the matching tests, each question was marked as correct, incorrect, or blank, for a maximum test score of 9. For the free recall test, participants averaged similar scores in both peephole conditions: 5.4 points (*SD* = 2.3) with static peepholes and of 5.3 (*SD* = 2.7) with dynamic peepholes. For the matching test, participants averaged 4.1 (*SD* = 2.2) correct and .56 (*SD* = .89) incorrect answers with the static peephole condition. With the dynamic peephole condition, the average was 3.7 (*SD* = 1.9) correct and .56 (*SD* = 1.2) incorrect on the matching test. No statistically significant differences were found with Wilcoxon signed rank tests for any of these scores.

While no statistically significant effect on recall or matching was found when comparing static to dynamic peephole interaction techniques, it is encouraging to see how well participants were able to remember vocabulary from a new language after using the application for only a short time.

6.4 Preference Between Peephole Conditions

While survey responses to the question asking what version of the application participants preferred showed support for using static peepholes, a deeper look in to comments and use of the application suggest a less straightforward view of preferences. When surveyed, eleven participants preferred using the static condition, four participants preferred the dynamic condition, and participant T5P1 said he liked "both, because of the advantages and disadvantages of each condition".

Dragging interactions, which require a user to touch an interface element and maintain contact with the screen while moving the element, appear to be less complicated when using dynamic peepholes. With dynamic peepholes, participants could drag a cloud while viewing the virtual space by moving the device, whereas viewing the virtual space using static peepholes requires touch interaction. Although we did not explicitly ask participants about dragging interactions, six participants remarked on dragging the clouds during gameplay, the survey, or the interview. Five said that they disliked how dragging worked in the static condition and single participant disliked how dragging worked in the dynamic condition. T8P1, for example, terminated his use of static condition saying, "I don't want to do this. I don't want to drag the clouds to the jars when it's far."

When looking at the overall ability to place words, and not just the drag interaction, participants were divided in their opinions. Comments about disliking how difficult it was to move words were mentioned by five participants regarding static peephole navigation by three participants regarding dynamic peephole navigation. Participants also expressed that they liked how easily words could be moved, which was expressed by four participants regarding the static peephole navigation and by seven regarding dynamic peephole navigation.

These conflicts may be due to unfamiliarity with dynamic peepholes. Dynamic Peepholes took the most getting used to, as evidenced by six participants noting that this condition was "easier the second time" or that they "got the hang of it". Participant T1P1, for example, ended up preferring the dynamic condition and he was quoted saying, "This is so much easier" even though, initially, this participant had some difficulty, saying "I'm stuck, can't really go up and down." T5P2 found the static peephole game easier to play because he "could move around easily and [he] knew exactly how to".

6.5 Interaction Styles During Peephole Use

Participants attempted to view the virtual world in different ways depending on whether they were using dynamic or static peepholes. When exploring gameplay with dynamic peephole interactions, participants attempted to use tilt or to scroll the background (as you would in the static condition), as well as variety of unique explorations. In the static condition, participants attempted multi-touch gestures to complete the game.

Overall, placing clouds required almost twice as many touchscreen interactions (e.g., taps, drags, swipes) in the static condition compared to the dynamic condition, with 735 and 385 total touchscreen interactions, respectively. One participant noted how tiring it was to complete the static condition. During the static condition, half of the participants attempted to interact with the background and the cloud simultaneously, using multi-touch gestures to drag the cloud while also scrolling the background. Most commonly, participants (8) would drag the cloud to the edge of the screen and use the bump of the cloud against the edge to scroll the background. Fewer participants (5) would drag the cloud partway, let go, and then drag the background separately.

During gameplay with dynamic peepholes, 87.5% of participants attempted to use a tilt interaction 71.4% attempted to scroll the background. Researchers most commonly noted these interactions during the first use of the condition, before the instruction to "use your arms" was given. There were numerous unique interactions noted by researchers as participants initially explored how to play the game using dynamic peephole interactions, such as turning the phone diagonally, upside down, an attempt at using 3 to 4 finger multi-touch gestures, and pushing the device backward and forward. Three participants disliked how tiring it was to complete the dynamic condition, such as T3P1 who said, "My arms hurt from holding [the device] out." However, other participants such as T8P2 noted the benefits of the moving condition, and preferred the physical interactions: "[Moving was] easier to touch [the cloud] and put it where you want it- like you're taking a picture."

6.6 Partner Interaction

We deliberately recruited children in pairs so that more dialogue would be elicited during gameplay and interviews. However, despite Scenic Words being a game that is used individually, partner interactions had the ability to influence the use of the application as well as convey items of interest. Additionally, children often play in social groups, even if they are using games intended for individual use. We coded for these interactions to understand how use may have been modified due to having a partner present.

Most sessions (6 of the 8) exhibited some form of partner instruction. There were verbal requests where one partner asked what should be done, such as T1P1 who asked his partner for direction on how to get clouds into the jars. There was also more direct instruction, such as participant T8P2 who played his partner's audio for him, or participant T4P2 who physically grabbed his partner's wrist during the dynamic condition to move it up and down and said, "look up". Peeking was noted in two groups, where one participant would look over the shoulder of their partner in order to observe some element of gameplay, such as how the audio was played. Game scores, based on points earned by correctly categorizing vocabulary, were displayed during gameplay and comparisons of scores between partners were noted in two groups, once as a casual remark and once as an expression of surprise that the partner was so far ahead.

Upon completion of gameplay the number of seconds it took the participant to complete the game was displayed to participants, causing some participants to feel as though they were in a race with their partner. Two participants in different sessions did not like time being displayed because it felt like they were racing each other as opposed to learning, including T6P1 who said, "The timer was like a race, so we won't memorize the words. It's not good for learning. It should be optional." However, five participants liked this feature, several noting their faster times between their first try at a condition to the second try, or the difference in their times between using the static and dynamic conditions.

7. DISCUSSION

In this work we investigated the design of a second language learning application for children called Scenic Words. Specifically, we investigated how children explore and interact with static and dynamic peepholes on a mobile device in an authentic context, and whether these interactions show potential for learning activities.

7.1 Incorporation of Peephole Interactions into Mobile Learning Activities for Children

We believe that incorporating movement into learning activities on mobile devices has the potential to benefit young learners, and that the results of this study support this notion. While mostly stating that they prefer static peepholes, several participants noted that the use of dynamic peepholes became easier over time, even with the short timeframe they used the application. This is to be expected, as static peephole interactions are more commonly found on mobile applications; therefore, users are likely more comfortable with static peepholes. Some participants also noted the benefits of using dynamic peepholes, such as only having to control the word with touch and being able to view the virtual space through moving. Using drag and drop interactions appeared to be easier in the dynamic condition. Given the interaction styles we noted, it could be useful to incorporate tilt in the dynamic condition and multi-touch gestures in the static condition.

More constrained measures, regarding user speed and accuracy with both peephole environments, would have the potential to provide additional information on how participants performed using the different interaction styles. However, we designed our study task to focus on language learning rather than interaction performance, and to measure performance with the interaction techniques we would need a more constrained task than those used in this study's vocabulary learning game.

Incorporating dynamic, movement-based peephole interaction into mobile applications poses challenges. Participants most often required additional instruction, the brief direction they were given during this study to "use their arms", to understand and make use of the dynamic peephole interaction technique to complete their tasks. This was reflected in how participants positioned the device during their use in each peephole condition, with device position being more stable when using static peepholes than dynamic peepholes. The tendency to resort to a single manner of positioning the devices also indicates that with additional practice with either peephole technique the mobile device may be positioned in a consistent manner during the entirety of use. Given this tendency it may also be beneficial to create virtual worlds that can be explored from a single position, such as the 180-degree scope used in Scenic Words.

7.2 Impact on Vocabulary Gain

One goal of this work was to examine if peephole interaction techniques, particularly movement-based dynamic peephole interactions, had an impact on participants' recall or matching scores. Although we did not find statistically significant differences between static and dynamic conditions on these scores, we are encouraged by how well participants were able to remember vocabulary from a new language after using the application for only a short time. We are also intrigued by trends that started to emerge, such as an increased number of errors in the static condition and a greater percent of total time being spent in word cards in the dynamic condition. However, to measure potential learning gains through using peephole interactions on mobile devices our scope would need to be expanded beyond a single session with immediate post-tests.

8. LIMITATIONS AND FUTURE WORK

We recognize that the study results are limited. We believe that these limitations are appropriate for an initial study, and that in future work many of these can be ameliorated. This was an initial study with sixteen children working in pairs of two, where we took a mixed-methods approach to analyzing our results. Future work using a larger sample size, a more equal distribution of genders, or more semantically meaningful gestures could lead to the ability to learn additional information, such as potential correlation between recall and matching tests and peephole type as well as more generalizable results.

Additionally, this study looked specifically at how children interact with static and dynamic peepholes when in their home. Given that prior work with peepholes have noted that context of use impacts how participants use the different peephole interactions, how children use peephole interactions may vary from our findings when the context is outside of the home. We hypothesize that how children use the Scenic Words application will vary from context to context; the use patterns we found in-home may be quite different from what we would see in schools or outside. Future work should explore how context changes how children the peephole interactions by investigating peepholes in different contexts.

Furthermore, the study method may have been a factor in our results. Participants completed each condition twice using the same word sets, and the familiarity with the vocabulary could, at least in part, explain why participants completed the second trial more quickly. Also, while we used immediate post-tests, full retention of vocabulary would need to be tested for at a later date, as testing immediately after the activity might lead to deceptively high recall rates. We also did not use a pre-test and instead asked participants if they knew any other languages and accepted from those responses that the participants did not know German. Future work should include a pre-test component as well as a delayed post-test component.

Finally, we recognize that limitations in the technologies that were used for this study may have had an effect on our results, as issues with recalibration during the dynamic condition and the "bouncing" effect of scrolling that wasn't completely smooth were noted by three of our participants.

Our observations suggest there may be other types of movement that could be explored in future work. The many ways children initially explored the virtual world using dynamic peephole navigation suggests a willingness to try a variety of interactions, such as multi-touch gestures, pushing the device backward and forward, turning the device upside- down, and turning the device diagonally. Again, since the dynamic peephole was a new interaction method, in contrast to the more familiar static peephole method, the relative success of the dynamic peepholes makes us optimistic in regard to future novel interaction methods.

9. CONCLUSION

In this work, we explored the possibilities of using dynamic and static peepholes on mobile devise to support children's second language learning. We found that participants completed the activity in similar amounts of time, preferring static peephole interactions but being willing to explore and becoming more accustomed to dynamic peephole interactions over time.

10. ACKNOWLEDGMENTS

We would like to thank the adult and child design partners at Kidsteam for their contributions to this work during the design sessions. We would like to thank Elizabeth Foss and Alex Kuhn for their help and recommendations on this work. Finally, we would also like to acknowledge the families, particularly the children, who participated in this study.

11. REFERENCES

[1] Ahn, J., Gubbels, M., Kim, J. and Wu, J. SINQ: Scientific INQuiry Learning using Social Media. In *Proc. CHI* (2012), 2081-2086.

[2] American Council on the Teaching of Foreign Languages. *National Standards for Foreign Language Education*. A Collaborative Project of the ACTFL, AATF, AATG, AATI, AATSP, ACL, ACTR, CLASS and NCJLT-ATJ. Web. 2014.

[3] Anthony, L., Brown, Q., Nias, J. and Tate, B. Examining the Need for Visual Feedback during Gesture Interaction on Mobile Touchscreen Devices for Kids. In *Proc. IDC* (2013), 157-164.

[4] Anthony, L., Brown, Q., Nias, J., Tate, B. and Mohan, S. Interaction and Recognition Challenges in Interpreting Children's Touch and Gesture Input on Mobile Devices. In *Proc. ITS* (2012), 225-234.

[5] Anthony, L., Brown, Q., Tate, B., Nias, J., Brewer, R. and Irwin, G. Designing Smarter Touch-Based Interfaces for Educational Contexts. *Journal of Personal and Ubiquitous Computing* (2013), 1-13.

[6] Antle, A., Corness, G. and Droumeva, M. What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. *The Interdisciplinary Journal of Human-Computer Interaction*, 21, 1-2 (2009), 66-75.

[7] Brown, Q. and Anthony, L. Toward Comparing the Touchscreen Interaction Patterns of Kids and Adults. In *Proc. Proceedings of the SIGCHI Workshop on Educational Software, Interfaces and Technology* (2012).

[8] Brown, Q., Bonsignore, E., Hatley, L., Druin, A., Walsh, G., Foss, E., Brewer, R., Hammer, J. and Golub, E. Clear Panels: A Technique to Design Mobile Application Interactivity. In *Proc. DIS* (2010), 360-363.

[9] Cahill, C., Kuhn, A., Schmoll, S., Pompe, A. and Quintana, C. Zydeco: using mobile and web technologies to support seamless inquiry between museum and school contexts. In *Proc. IDC* (2010), 174-177.

[10] Chinnery, G. EMERGING TECHNOLOGIES: Going to the MALL: Mobile Assisted Language Learning. *Language Learning and Technology*, 10, 1 (2006), 9-16.

[11] Dindler, C., Krogh, P. G., Beck, S., Stenfelt, L., Nielsen, K. R. and Grønbæk, K. Peephole experiences: field experiments with mixed reality hydroscopes in a marine center. *Proceedings of the 2007 Conference on Designing for User eXperiences* (2007), p.20.

[12] Druin, A. Cooperative inquiry: developing new technologies for children with children. In *Proc. CHI* (1999) 592-599.

[13] Druin, A. The role of children in the design of new technology. *Behaviour and Information Technology*, 21, 1 (2002), 1-25.

[14] Fails, J., Druin, A. and Guha, M. L. Interactive Storytelling: Interacting with People, Environment, and Technology. *International Journal of Arts and Technology*, 7, 1 (June 9-12 2014), 112-124.

[15] Fails, J., Guha, M. and Druin, A. Methods and Techniques for Involving Children in the Design of New Technology for Children. *Human–Computer Interaction*, 6, 2 (2012), 85-166.

[16] Goodman, M. *Let's Learn German Picture Dictonary*. McGraw-Hill Companies, Inc, Hong Kong, 1991.

[17] Grubert, J., Morrison, A., Munz, H. and Reitmayr, G. Playing it Real: Magic Lens and Static Peephole Interfaces for Games in a Public Space. In *Proc. MobileHCI* (2012), 231-240.

[18] Grubert, J. and Schmalstieg, D. Playing it Real Again: A Repeated Evaluation of Magic Lens and Static Peephole Interfaces in Public Space. In *Proc. MobileHCI* (2013), 99-102.

[19] GSM Association and NTT DOCOMO *Children's use of mobile phones: An international comparison 2012*. Japan (2013).

[20] Guha, M. L., Druin, A. and Fails, J. A. Cooperative inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction*, 1, 1 (2013), 14-23.

[21] Howison, M., Trninic, D., Reinholz, D. and Abrahamson, D. The Mathematical Imagery Trainer: From Embodied Interaction to Conceptual Learning. In *Proc. CHI* (2011), 1989-1998.

[22] Howison, M., Trninic, D., Reinholz, D. and Abrahamson, D. The Mathematical Imagery Trainer: from embodied

interaction to conceptual learning. In *Proc. CHI* (2011), 1989-1998.

[23] Hruschka, D. J., Schwartz, D., John, D. C. S., Picone-Decaro, E., Jenkins, R. A. and Carey, J. W. Reliability in coding open-ended data: Lessons learned from HIV behavioral research. *Field Methods*, 16, 3 (2004), 307-331.

[24] Hsu, H. The Potential of Kinect in Education. *International Journal of Information and Education Technology*, 1, 5 (December 2011).

[25] Hürst, W. and Bilyalov, T. Dynamic versus Static Peephole Navigation of VR Panoramas on Handheld Devices. In *Proc. MUM* (2010), p.25.

[26] Kang, C. *Survey: For young children, mobile devices such as tablets, smartphones now a mainstay*. The Washington Post, Washington DC, 2013.

[27] Kuhn, A., Cahill, C., Quintana, C. and Schmoll, S. Using tags to encourage reflection and annotation on data during nomadic inquiry. In *Proc.CHI* (2011), 667-670.

[28] Kuhn, A., Cahill, C., Quintana, C. and Soloway, E. Scaffolding science inquiry in museums with Zydeco. In *Proc. CHI* (2010), 3373-3378.

[29] Kukulska-Hulme, A. and Shield, L. An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20, 03 (2008), 271-289.

[30] Kumar, A., Reddy, P., Tewari, A., Agrawal, R. and Kam, M. Improving Literacy in Developing Countries Using Speech Recognition-Supported Games on Mobile Devices. In *Proc. CHI* (2012), 1149-1158.

[31] Kumar, A., Tewari, A., Shroff, G., Chittamuru, D., Kam, M. and Canny, J. An Exploratory Study of Unsupervised Mobile Learning in Rural India. In *Proc. CHI* (2010), 743-752.

[32] Lan, Y.-J., Sung, Y.-T. and CHang, K.-E. A Mobile-Device-Supported Peer-Assisted Learning System for Collaborative Early EFL Reading. *Language Learning and Technology*, 11, 3 (2007), 130-151.

[33] Lee, E., Liu, X. and Zhang, X. "Xdigit: An Arithmetic Kinect Game to Enhance Math Learning Experiences." *Retrieved February* 14 (2012): 2013.

[34] Macedonia, M. and Kriegstein, K. Gestures Enhance Foreign Language Learning. *Biolinguistics*, 6, 3-4 (2012), 393-416.

[35] Mehra, S., Werkhoven, P. and Worring, M. Navigating on Handheld Displays: Dynamic versus Static Peephole Navigation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13, 4 (2006), 448-457.

[36] Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O'Malley, C., Stanton, D., Thompson, M. and Weal, M. Ambient wood: designing new forms of digital augmentation for learning outdoors. In *Proc. Proceedings of the 2004 conference on Interaction design and children: building a community* (2004), 3-10.

[37] Sharples, M., Taylor, J. and Vavoula, G. Towards a theory of mobile learning. *Proceedings of mLearn 2005*, 1, 1 (2005), 1-9.

[38] Viberg, O. and Grönlund, A. Mobile Assisted Language Learning: A Literature Review. In *Proc. Mobile and COntextual Learning (mLearn)* (2012), 9-12.

[39] Wenig, D., Nulpa, T., Malaka, R. and Lawo, M. An Evaluation of Peephole Interaction with Panoramic Photographs for Pedestrian Navigation. In *Proc. GI Zeitgeist* (2012), 23-32.