User-Centered Development of a Gesture-Based American Sign Language Game

Seungyon Lee, Valerie Henderson, Helene Brashear, Thad Starner, Steven Hamilton
College of Computing, Georgia Institute of Technology, Atlanta, GA 30332 USA
{sylee, vlh, brashear, thad, shamilto}@cc.gatech.edu

Harley Hamilton
Atlanta Area School for the Deaf, Clarkson, GA 30021 USA
HHamilto@doe.k12.ga.us

ABSTRACT

We present our on-going development of *CopyCat*, a gesture-based computer game for deaf children. Using gesture recognition techniques, *CopyCat* allows deaf children to communicate with the computer using American Sign Language (ASL) and encourages them to practice signing in an enjoyable way. Our goal is to help young children to transition from disjoint single-sign communication to phrases that represent complete thoughts. Focusing on a user-centered design method, we use pilot studies with deaf children to improve the usability and enjoyment of the game as well as to develop our computer gesture recognition.

We use a Wizard of Oz (WOz) technique which allows us to separate interface testing from gesture recognition development, thus allowing us to focus on how deaf children interact with the gesture-based user interface. This paper describes the configurations and results of each pilot study and shows the design process based on the children's feedback.

KEYWORDS

Deaf, children, ASL, language acquisition, computer game, user-centered development, gesture recognition, human-computer interaction (HCI), wizard of Oz method

BACKGROUND

Ninety percent of deaf children are born to hearing parents. Most of these parents do no not know or are not fluent in sign language ("Annual Survey of Deaf," 2001). Often a child's first exposure to signing is at school. Since early childhood is a critical period for language acquisition, early exposure to ASL is key for deaf children's linguistic development (Mayberry 1991, Newport 1990). Children construct their own understanding through interaction with their environment (Meadows

A paper presented at the
Instructional Technology and Education of the Deaf Symposium
National Technical Institute for the Deaf
Rochester, NY 14623 USA
June 27-30, 2005
http://www.rit.edu/~techsym

2002), and the ability to communicate determines the pace of linguistic development. Hearing children learning a spoken language combine words in their expressive communication by two years of age (Wells 1985). Deaf children of deaf parents also combine signs to communicate by one and a half years of age. Deaf children of hearing parents acquire language in the same sequence as the first two groups but at a much slower pace. The slower linguistic development of this third group has been ascribed to incomplete language models and a lack of daily interaction using a language (Hamilton and Lillo-Martin 1986, Spencer and Lederberg 1997). Hamilton and Holzman (1989) also linked delayed language acquisition with delayed short-term memory development.

Figure 1. Screenshot of ASL Game Interface
A) Tutor video B) Live video feed C) Attention button D) Animated character and environment E) Action buttons

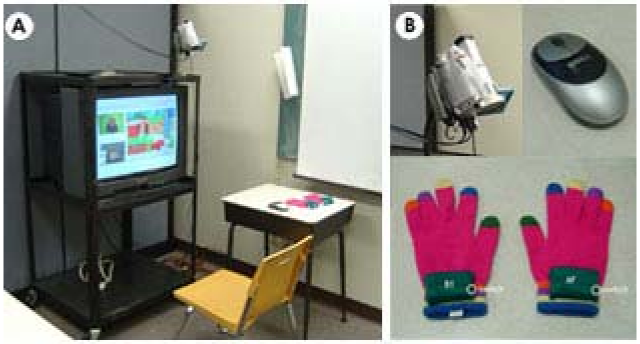

Figure 2. Input Devices for User
A) Hardware and furniture layout B) Input devices (video camera, mouse, and colored-gloves with wrist-mounted accelerometers)

GAME DEVELOPMENT

CopyCat is a research prototype combining an interactive computer game with sign language recognition technology. This study aims to assist young deaf children's language acquisition by interactive tutoring and real-time evaluation. Unlike many educational games relying on English grammar skills or spoken audio files, *CopyCat* provides English-free interface. The gesture recognition system supports ASL-based communication between the user and the computer game. The child is asked to wear colored gloves with wrist-mounted accelerometers and sit in front of the computer equipped with a video camera for computer vision recognition system. While playing the game, the child communicates with the animated character, Iris the cat, through ASL. This game is both mentally and physically engaging and allows the child to practice ASL in an enjoyable way.

The game interface includes Action Buttons (Fig. 1E) which are assigned to phrases that the child signs to Iris, the pre-recorded tutor video (Fig. 1A) which demonstrates given phrases, a live-video feed (Fig. 1B) which reflects user's gesture in real-time, an Attention Button (Fig. 1C) which segments the data for the gesture recognition system, and the character animation (Fig. 1D) which communicates with the user through ASL. In addition, a series of attention drawing visual indicators and help button guide user to understand the situation for each stage of interaction.

When a child clicks a pictorial Action Button (Fig. 1E), a pre-recorded video tutor (Fig. 1A) signing assigned ASL phrase is loaded and played automatically. After the video plays, the Attention Button (Fig. 1C) flashes red, cueing the child to press it to wake Iris up for next step of interaction. Once Iris wakes up, the icon on the Attention Button is changed and a square red box around the live-video feed (Fig. 1B) flashes. This indicates that Iris pays attention to user's signing. The user signs to Iris and watches himself signing through the mirror-like live-video window. After the child finishes signing, he again clicks the Attention Button (Fig. 1C) to tell Iris to execute his command. If the sign is correct and clear, Iris will understand it and follow the command. Otherwise she will look puzzled and encourage the user to try again. Children can replay the tutor video as many times as they want. Figure2 shows the required devices for users.

The purpose of our initial studies is to

1) create a compelling game that works within the scope of our gesture recognition technology.
2) collect signing examples for training the sign recognition engine.

To reach these goals, we performed four different pilot studies with deaf children at Atlanta Area School for the Deaf (AASD). In several on-site simulations, we evaluated participants' performance qualitatively for game interface development and quantitatively for sign language recognizer development.

GESTURE RECOGNITION ENGINE

The gesture recognition engine for this game is part of an on-going sign language recognition project at Georgia Tech. Our previous work demonstrates that the recognition accuracy of ASL is greatly increased by combining accelerometer data with computer vision (Brashear et al. 2003). We have engineered wireless, Bluetooth accelerometers that are encased in boxes about the size of a matchbox. These devices are worn on the wrist in pockets on the gloves, as described previously (Fig. 2).

One common difficulty in speech and gesture recognition systems is known when to attend to the user's activities. We address this issue through a "push to sign" mechanism, similar to the "push to talk" mechanism used in many speech recognition systems. Our push to sign mechanism is integrated into the game design. When the user clicks to get Iris's attention and clicks to finish signing, these actions have two effects. First, they activate game cues (such as Iris's behavior) and second, they segment the data stream. In this manner we greatly reduce the complexity of both the recognition problem and our data collection.

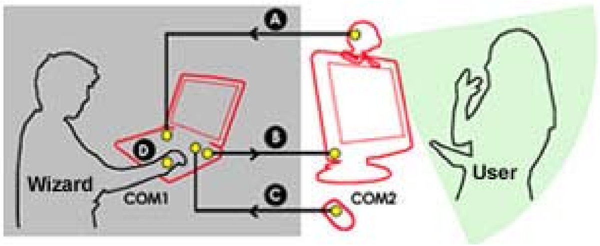

Figure 3. Test Setting for 1st and 2nd Pilot Studies
COM1) System interface for wizard COM2) Game interface for user A) Live camera feed B) Interface output split between wizard and user C) Child's mouse and D) Wizard's response

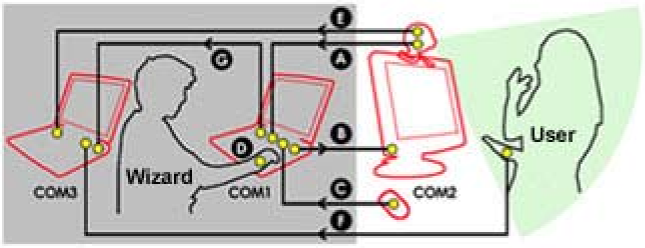

Figure 4. Test Setting for 3rd and 4th Pilot Studies
COM1) System interface for wizard COM2) Game interface for user COM3) Gesture recognition engine A+E) Live camera feed B) Interface output split between wizard and user C) Child's mouse D) Wizard's response F) Data from accelerometers and G) Network between the game interface and the gesture recognition engine

Glossed ASL	English Translation
A. You like mouse?	Do you like mice?
B. You happy?	Do you feel happy?
C. You hungry now?	Are you hungry now?
D. You go play balloon.	Go play with the balloon.
E. You make flowers grow. Go-on.	Go make the flowers grow.
F. You go catch butterfly	Go catch the butterfly.
G. Who your best friend?	Who is your best friend?
H. Look there Iris, mouse over-there.	Look, Iris! A mouse, over there!

Table 1. Phrase Set for 1st and 2nd Pilot Studies

Color	Locations	Predator
Blue	Behind wagon	Snake
Green	Under chair	Spider
White	In flowers	Alligator
Orange	In bedroom	
Black	On wall	

Template1 : <Color> kitten <Location>
Template2 : Go chase <Predator>
(Examples : Blue kitten behind wagon ; Go chase snake)

Table 2. Phrase Set for 3rd and 4th Pilot Studies

PILOT STUDY

The wizard of Oz (WOz) technique is an evaluation method which uses a human "Wizard" to simulate system functionality. One of the benefits of this technique is that researchers obtain feedback during the design process from people for whom the system is intended. The designer can evaluate the system's usability through user participation even with a prototype with limited functionality.

During the simulation, a participant performs tasks using the prototype. The Wizard is situated out of sight of the subject, receives the subject's input, and controls the system manually, emulating the missing functionality of the system (Dix et al. 2004). The subject is not aware of the Wizard's presence and believes that he is using a fully functioning system. In our scenario, while the child plays the game, a human Wizard simulates the computer recognizer and evaluates the correctness of the player's sign with the help of the interpreter.

Focusing on a user-centered design approach, we have iterated the game interface and development strategy through several studies with deaf children at AASD using the WOz method. From November 2004 to March 2005, seven deaf children ages 9-11 at AASD participated in the tests. While slightly older than our targeted age of 6-8, the older children were able to provide more detailed feedback. In addition, this strategy left our primary subject pool uncontaminated for the longitudinal study of the final version of the project (full functioning game interface and sign language recognizer).

With the assistance of educational technology specialists and linguists at AASD, we developed a list of eight age appropriate phrases for our game (Table. 1). The first and second pilot studies were performed with the initial hardware setting shown in Figure 3. When the user asked Iris a yes or no question, she nodded and responded with a thought bubble. For the rest of the questions, Iris executed particular actions based on the user's command. The gesture recognition system was added for the third and fourth pilot studies (Fig. 4) to test the functionality of the system. The game was improved, and a new phrase set was added as shown in Table 2.

First Pilot Study (November, 2004)

Configuration

Three children participated in the test with the setting shown in Figure 3 and eight phrases shown in Table 1. We asked these children to play the game as long as they wanted. Our goal was to evaluate how well the children could control the game, including the push-to-sign mechanism. After they finished the game, we interviewed them about their experiences.

Results

We found that the children had no trouble with the push-to-sign mechanism. At the beginning of the test, our facilitator demonstrated the mechanism, and the children easily understood how to manipulate the interface. Although they sometimes forgot to press the Attention Button before or after signing, the children realized their mistake immediately and corrected it quickly. One problem discovered during this test was the variability of the interpreters. The interpreters graded each child differently based on their knowledge of the child's linguistic skills. We also discovered that the furniture arrangement obscured the children's hands as they signed.

Second Pilot Study (December, 2004)

Configuration

Two children participated in the test with the setting shown in Figure 3, and eight phrases shown in Table 1. We asked the children to wear colored gloves with wrist-mounted plastic blocks in order to evaluate the comfort of the wrist-mounted accelerometers. We moved the desk to the right side of the user so that it would not block any signing. For the second pilot study, a single interpreter was asked to evaluate the correctness of the sign using a consistent level of strictness.

Results

We found that the colored gloves with wrist-mounted plastic block didn't bother the children.

Interactive showed that action oriented responses (such as Go catch the butterfly, Go make the flowers grow) were preferred by the children over passive responses (such as Do you feel happy? Are you hungry now?).

Third Pilot Study (January, 2005)

Configuration

Based on the results of the prior pilot study and the children's preference for a more competitive, action oriented game, we revamped the prototype with more levels and a goal oriented, narrative flow. In the new prototype, "Kitten Escape!", Iris's multicolored kittens have jumped out of the basket and hidden in the background where various animals such as snakes, spiders, and alligators lived. Unlike the previous phrase set with unrelated vocabularies, the combination of relevant vocabularies explaining the location of multicolored kittens (Table 2.) direct Iris to find and return each kitten to the basket. During the third pilot study, we tested the new hardware and software configurations. Accelerometers were mounted on the wrist-pocket of the colored gloves. The gesture recognition engine was connected to the game for sensor testing and preliminary data collection. (Fig. 4)

Results

It appeared that the new prototype, which was revised to a goal and action-oriented narrative flow, motivated the children more than the previous prototype. The modified WOz method allowed us to efficiently gather data, track collection progress, and rapidly assess data quality.

Fourth Pilot Study (February ~ March, 2005)

Configuration

The fourth pilot study was used to collect a complete set of signing samples from multiple users. The prototype "Kitten Escape!" was expanded to three levels. Five children were asked to wear colored gloves with wrist-mounted accelerometers and play all three levels of the game in revised setting (Fig. 4). The test was repeated five times for each child. For this fourth pilot study, we clarified that this was a competition and showed the score of all players during the test.

Results

We found that the competition prominently affects children's level of engagement. Some children used the live video feed window as a mirror to rehearse their signing before they actually signed to Iris. They practiced signing, watching and correcting themselves, to make sure their signing to Iris was as clear as possible. Samples gathered from this trial were used to determine what kind of variance there might be in user data (both in children's signing and their interaction with the game) and as a training set for our ASL recognizer.

RESULTS AND DISCUSSION

We have adjusted the game contents and test settings for data collecting through the iterative design process. The WOz method allowed us to obtain user's rapid feedback from on-site simulation. The results of pilot studies showed that the goal-oriented contents encouraged deaf children to practice ASL as motivated learners. At the end of the fourth trial we have 541 samples of correctly signed phrases, and 54 samples of signed phrases with mistakes. Our future work will include modeling and analysis of these phrases, and incorporating these models into the recognition engine. We plan to work further with educational specialist to create more educational content for the game and evaluate its impact on learning. As the recognition engine improves we will continue to work toward a fully participative game with automatic ASL recognition.

ACKNOWLEDGEMENT

We would particularly like to thank the children, staff, and administration at Atlanta Area School for the Deaf for their enthusiasm and support of this project. This work is funded in part by NSF Career Grant #0093291, an NSF Graduate Fellowship, and the NIDRR Wireless RERC under the U.S. Department of Education grant #H133E010804. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Department of Education.

REFERENCES

Gallaudet Univ. (2001, January). *Regional and National Summary Report of Data from the 1999-2000 Annual Survey of Deaf and Hard of Hearing Children and Youth*, Technical Report, Washington, D.C.

Mayberry R.I. & Eichen E.B. (1991). The Long-Lasting Advantage of Learning Sign Language in Childhood : Another Look at the Critical Period for Language Acquisition, *Journal of Memory and Language,* 30. pp. 486-498

Newport E.L. (1990). Maturational Constraints on Language Learning, *Cognitive Science* 14. pp. 11-28.

Meadows S. (2002). *The child as thinker : The Development and Acquisition of Cognition in*

Childhood, Routledge, London UK

Hamilton H. & Holzman T. (1989). Linguistic Encoding in Short-Term Memory as a Function of Stimulus Type, *Memory and Cognition*. 17 (5). pp. 542-550

Hamilton H. & Lillo-Martin D. (1986). Imitative Production of Verbs of Movement and Location: A Comparative Study. *Sign Language Studies*. 50. pp. 29-57

Spencer P. & Lederberg A. (1997). Different Modes, Different Model: Communication and Language of Young Deaf Children and Their Mothers. In M. Romski (Eds.), *Communication and Language: Discoveries from Atypical Development.* Harvard University Press. pp. 203-230

Wells G. (1985). *Language Development in Preschool Years*, Harvard Univ. Press, Cambridge, MA

Dix A. et al. (2004*) Human-Computer Interaction* (3rd ed.). Prentice Hall

Brashear H., Starner T., Lukowicz P. & Junker H. (2003, Oct.). *Using Multiple Sensors for Mobile Sign Langusge Recognition.*, In Proceedings of the 7th IEEE International Symposium on Wearable Computers, pp. 45-52.