Visualizations: Speech, Language & Autistic Spectrum Disorder

Joshua Hailpern
University of Illinois at Urbana Champaign
201 N Goodwin Ave.
Urbana, IL 61802 USA
Jhailpe2@cs.uiuc.edu

Laura DeThorne
University of Illinois at Urbana Champaign
901 South Sixth St
Champaign, IL 61820 USA
lauras@uiuc.edu

Karrie G. Karahalios
University of Illinois at Urbana Champaign
201 N Goodwin Ave.
Urbana, IL 61802 USA
kkarahal@cs.uiuc.edu

Mary-Kelsey Coletto
University of Illinois at Urbana Champaign
901 South Sixth St
Champaign, IL 61820 USA
mcoletto@uiuc.edu

Jim Halle
University of Illinois at Urbana Champaign
1310 S. 6th St.
Urbana, IL 61820 USA
halle@uiuc.edu

CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 978-1-60558-012-8/08/04.

Abstract
Without speech, we can have great difficulty communicating wants, emotions, needs, and interacting with society at large. During typical child development, an infant acquires language skills without explicit teaching. However, some children, including those with Autistic Spectrum Disorder (ASD) have explicit difficulty developing these skills in the context of everyday interactions. HCI is situated to help by developing technology and techniques to teach speech and language skills to children with ASD through the use of visual and auditory feedback. This paper examines preliminary results from a study, as well as describes new directions of research.

ACM Classification Keywords
H 5.2 [Information Interfaces and Presentation]: Screen design, Voice I/O. K4.2 [Social Issues]: Assistive technologies for persons with disabilities

General Terms
Design, Experimentation, Human Fators

Keywords
Accessibility, Visualization, Autism, Children, Speech, Vocalization

Introduction

Autistic Spectrum Disorder (ASD), a developmental disorder affecting 1 in 150 births[7], can affect the social interaction of individuals, as well as their communication skills. Whereas neurologically typical children will develop speech and language skills without any exceptional effort by parents, doctors or practitioners, some children (e.g., those with ASD) do not acquire these skills as quickly, if at all. This can be crippling for many children, in that language is "a unique characteristic of human behavior… [that] contributes in a major way to human thought and reasoning"[16]. Though we are learning more and more about ASD every day, a cure or way to prevent ASD is unlikely in the near future. While awaiting a cure, research must be conducted to help affected individuals live lives that are as productive and happy as possible. This means helping with social interaction, as well as teaching skills to aid in communication.

Autistic Spectrum Disorder & Treatments

With the publication of Kanner's case studies in 1943 [11], ASD became a well-documented disorder. It continues to affect children and adults over 60 years later. The descriptions of the eleven children in that seminal work still capture many of the interpersonal and communicative difficulties those with ASD face today. Specifically, ASD is typified by delays in communication, empathy, social functioning, and expression. The Autism Society of America describes ASD as "insistence on sameness… Preference to being alone… spinning objects [and] obsessive attachments to objects"[3]. Children with ASD are not a homogeneous group. As a "spectrum" disorder, individuals who are considered *low functioning* have greater difficulty with social and communicative skills than those who are considered *higher functioning.*

In the 1960s, Ivar Lovaas adopted a technique called "applied behavior analysis," in which target behavior is rewarded with food or toys. These rewards are then "faded" or removed gradually over time, so that the behavior is maintained by naturally occurring events. This technique has been shown to be effective, especially in teaching speech [16]. While successful, this system requires many expensive and time-consuming sessions with trained professionals, intense attention and prolonged contact with practitioner/parent, and potentially anxiety-causing human-to-human contact as the main mode of education (the detached and alone feeling of many children with ASD [5, 11] causes a degree of difficulty for practitioners and subjects).

HCI & Autistic Spectrum Disorder

Work in the HCI community has provided tools to greatly improve the ability to diagnose [9, 13, 14], as well as to teach, human-to-human interaction to high functioning children with ASD [12, 22]. In addition to the skill-improvement and diagnosis based approaches, HCI researchers have also explored the encouragement of general "play" via technology. This work has demonstrated that technology and computers appear to reduce the apprehension caused by human-to-human interaction [15, 17, 19]. However, there appear to be few HCI studies focusing on either teaching speech/vocalization skills, or the targeting of low functioning children with ASD.

Feedback to Encourage Vocalization

Based on the existing approaches, we hypothesize that computers and HCI techniques are capable of not only acting as a means to support communication (e.g., AACDs, emails, IM) but also as a method to help augment practitioners by *teaching* and *reinforcing both* vocalization and speech skills in children who are on the autism spectrum. Due to the current limitations of speech recognition software [18, 21], solutions must be designed to aid and supplement practitioners and researchers rather than replace such experts. There exists evidence in both HCI and Speech and Hearing Science literature to support the notion that computerized visualizations can influence the communication of individuals [1, 4, 6, 20], and specifically the behavior children with ASD [12, 15, 17, 19, 22]. By taking these findings a step further, we hope to explore methods and technology that can help facilitate the speech and vocalization education process for children with communication skill deficits. Specifically we intend to use contingent visual and auditory feedback to (a) motivate and reward vocalization and (b) provide information about the acoustic properties of vocalizations.

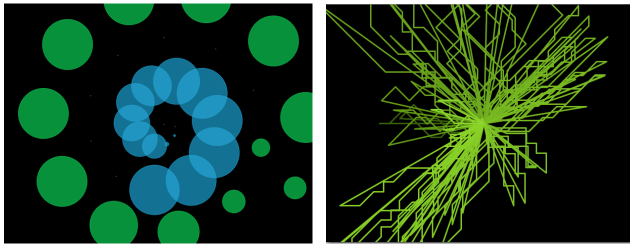

Figure 1. Samples of Visualizations Used in the 2007 Study

Fall 2007 Pilot Study

The first phase in establishing this new research direction was determining whether digital vocal visualizations (Figure 1) and/or auditory feedback could be constructed to influence the vocalizations of children on the spectrum. Further, it was necessary to determine what permutations of auditory and visual feedback would have the greatest effect (be it on an individual or group level). These questions were explored in an initial, four-month pilot study that took place from August to November 2007. Five "low functioning" children (3-8 years) with autism were enrolled in this pilot study. During six, 40-50 minute sessions, the children were presented with a variety of visual and auditory feedback for any sounds they made. Data on each child's response to computer visual and auditory feedback were collected. In addition, a video of each session was made (with parental permission).

Data from the pilot study are currently being analyzed. Analyses are focusing on visualization preferences, as well as the influences on the visualizations on sound production, attention, and affect. Variables will be subjected to within-subject comparisons across trials and conditions. With over 24 hours of video to analyze, this process is labor-intensive and time consuming.

As we distill the quantitative data, qualitative observations suggest positive results. All the children appeared to have an idiosyncratic combination (or 2) of auditory and/or visual feedback to which they responded optimally. Further, there appeared to be unique preferences for types of audio and visual stimuli that were associated with increases in child vocalizations and engagement with the tasks. Children generally appeared excited, if not elated, to participate and to interact with the software. Parents echoed these observations. To further capture the comments and reactions of parents, we offered an optional questionnaire to parents about their child's experience. One parent's perspective characterized that of the group:

> My child's reaction is one of excitement and looking forward to see what was next to come. Applause on your study. You may be onto something here.

One child's performance exceeded that of the other children in the study. We created a "Wizard-of-Oz" system in which our software prompted him with a word or phrase (e.g., "Juice"). Once the child imitated the target word, as gauged by the researcher, he was rewarded with a visual and auditory display for which he had appeared to show preference previously. During two extra sessions, this child learned the "game" and quickly began mimicking the sounds he heard, including ones his mother had not heard him say before although she had modeled them in the past). Although this N=1 study is not conclusive, it does indicates a potential contribution that warrants further investigation. We plan to focus on future studies toward (a) providing more specific visual feedback and (b) generalizing speech production into more naturalistic contexts.

New Directions: Graphic Matching

In regard to the specificity of feedback, most of the contingent visualizations utilized in the Fall 2007 study were abstract. They provided no goal or target sound. Rather, these visualizations were simply a set of reinforcing set of audio and visual stimuli. Yet in the past, software solutions have attempted to provide a goal-oriented system to aid in speech creation [10]. To build upon the positive reaction of the "Wizard-of-Oz" system in the 2007 study, we are exploring other forms of goal-oriented visualizations.

Recording studio software products, such as Pro Tools [8] and Logic [2], have made use of the waveform form of visualization. Further, the waveform is also used in tools to assess speech [23]. These representations allow the user to visualize audio data, and come to new understandings about the sound itself.

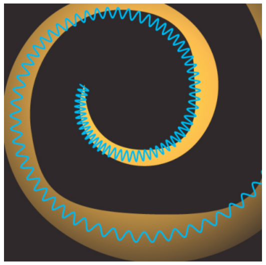

Figure 2 A Mock-Up of Graphic Matching; aChild must make his voice follow a path (time) with a given volume (width).

Unlike the waveform which has been used for many years, our work steps away from the scientific looking visuals of waveforms and moves towards abstract graphics that are playful and less intimidating. The premise behind the graphic-matching is to create an outline on the computer screen of a target shape (be it of pitch, volume, speech-rate, etc., over time). As the child makes a sound, he or she will attempt to "fill in" the outline based on his/her own sound (Figure 2). If the outline is sufficiently filled-in, as determined by the software, a clinician, or some combination of the two, the software will reward the child with enjoyable visual and/or auditory feedback. Using target speech patterns, and input speech patters, we can calculate a single or multidimensional differences in a variety of forms, and visualize them in order to modify the child's speech.

New Directions: Toy-Like Interfaces

A second, unexplored, direction leverages the research done by Abowd and Kientz [13] and the qualitative observations from the 2007 pilot study. Specifically, we are investigating toy-like interfaces that could be brought home by the child. Such a device could allow children to experiment with vocalizations across contexts (i.e., outside of speech therapy) and within the constant need of a therapist or investigator. The primary medium of interaction with the toy would be through vocalization, and it responds through different forms of visualization akin to the goal matching metaphor. It could respond auditably, flashing eyes (if anthropomorphic), a small display, or interface with a TV (Figure 3). Because the toys would be digital, they could record sound and interaction data, providing these data to clinicians and researchers either via a dock at home, or by bringing it back to an office. By having more, objective and empirical data, clinicians can target their lessons more easily to skills the children acquiring or are having trouble acquiring.

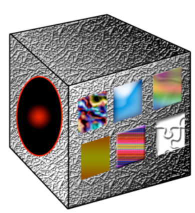

Figure 3 A Mock-Up of a toy-like device

Conclusion

With early qualitative results appearing to be promising, utilizing computer auditory and visual feedback appears to be an exciting new area of research for the HCI community. These software techniques not only have the potential to aid practitioners, but also to help improve the lives and communicative abilities of those with ASD.

Citations

1. Adler-Block, M., Bernhardt, B.M., Gick, B. and PBacsfalvi, P. The Use of Ultrasound in Remediation of North American English /r/ in 2 Adolescents. *American Journal of Speech-Language Pathology*, *16* (May 2007). 128-139.
2. Apple.Inc. Logic Pro, 2007.
3. Autism Society of America, A. Autism Society of America, http://www.autism-society.org, 2007.
4. Barry, R.M. EPG from square one: An overview of electropalatography as an aid to therapy. *Clinical Linguistics & Phonetics*, *3* (1). 81-91.
5. Baskett, C.B. The effect of live interactive video on the communicative behavior in children with autism, University of North Carolina at Chapel Hill, Chapel Hill, 1996.
6. Bergstrom, T. and Karahalios, K., Seeing More: Visualizing Audio Cues. in *INTERACT*, (Rio de Janerio, Brasil, 2007).
7. Center for Disease Control and Prevention, C. Autism Information Center, DD, NCBDDD, CDC, Atlanta, April 25, 2007.
8. DigiDesign. Pro Tools, 1991-2007.
9. Hayes, G.R., Kientz, J.A., Truong, K.N., White, D.R., Abowd, G.D. and Pering, T., Designing Capture Applications to Support the Education of Children with Autism in

International Conference on Ubiquitous Computing, (Nottingham, England, 2004).

10. IBM. Speech Viewer III, 1997.
11. Kanner, L. Autistic Disturbances of Affective Contact. in Kanner, L. ed. *Nervous Child 2*, V.H. Winston, 1943, 217-250.
12. Kerr, S.J., Neale, H.R. and Cobb, S.V.G. Virtual environments for social skills training: the importance of scaffolding in practice *Proceedings of the fifth international ACM conference on Assistive technologies*, ACM Press, Edinburgh, Scotland, 2002.
13. Kientz, J.A., Arriaga, R.I., Chetty, M., Hayes, G.R., Richardson, J., Patel, S.N. and Abowd, G.D. Grow and know: understanding record-keeping needs for tracking the development of young children *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, San Jose, California, USA, 2007.
14. Kientz, J.A., Hayes, G.R., Abowd, G.D. and Grinter, R.E. From the war room to the living room: decision support for home-based therapy teams *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, ACM Press, Banff, Alberta, Canada, 2006.
15. Lehman, J.F. Toward the use of speech and natural language technology in intervention for a language-disordered population *Proceedings of the third international ACM conference on Assistive technologies*, ACM Press, Marina del Rey, California, United States, 1998.
16. Lovaas, I.I. *The Autistic Child*. John Wiley & Sons, Inc, New York, 1977.
17. Michaud, F. and Théberge-Turmel, C. Mobile robotic toys and autism. in Dautenhahn, K. ed. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*, Springer, 2002, 125-132.
18. Nakagawa, S. A Survey on Automatic Speech Recognition. *IEICE TRANSACTIONS on Information and Systems*, *E85-D* (3). 465-486.
19. Parés, N., Carreras, A., Durany, J., Ferrer, J., Freixa, P., Gómez, D., Kruglanski, O., Parés, R., Ribas, J.I., Soler, M. and Sanjurjo, A. Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment *Proceeding of the 2005 conference on Interaction design and children*, ACM Press, Boulder, Colorado, 2005.
20. Shuster, L.I. and Ruscello, D.M. Evoking [r] Using Visual Feedback. *American Journal of Speech-Language Pathology*, *1* (May 1992). 29-34.
21. Strik, H. and Cucchiarini, C. Modeling pronunciation variation for ASR: A survey of the literature. *Speech Communication*, *29* (2-4). 225-246.
22. Tartaro, A. Storytelling with a virtual peer as an intervention for children with autism. *SIGACCESS Access. Comput.* (84). 42-44.
23. Whittaker, S. and Amento, B., Semantic Speech Editing. in *SIGCHI conference on Human factors in computing systems*, (Vienna, Austria, 2004), ACM.