

# Design Goals for a System for Enhancing AAC with Personalized Video

Katie O’Leary<sup>1</sup>, Charles Delahunt<sup>2</sup>, Patricia Dowden<sup>3</sup>, Ivan Darmansya<sup>4</sup>, Jiaqi Heng<sup>4</sup>,  
Eve A. Riskin<sup>2</sup>, Richard E. Ladner<sup>4</sup>, Jacob O. Wobbrock<sup>1</sup>

<sup>1</sup>Information School  
DUB Group  
{kathlo, wobbrock}@uw.edu

<sup>2</sup>Electrical Engineering  
delahunt@uw.edu  
riskin@ee.washington.edu

<sup>3</sup>Speech & Hearing Sciences  
dowden@uw.edu

<sup>4</sup>Computer Science  
& Engineering  
{blinkerz, jiaqih, ladner}  
@cs.washington.edu

## ABSTRACT

Enabling end-users of Augmentative and Alternative Communication (AAC) systems to add personalized video content at runtime holds promise for improving communication, but the requirements for such systems are as yet unclear. To explore this issue, we present *Vid2Speech*, a prototype AAC system for children with complex communication needs (CCN) that uses personalized video to enhance representations of action words. We describe three design goals that guided the integration of personalized video to enhance AAC in our early-stage prototype: 1) Providing social-temporal navigation; 2) Enhancing comprehension; and 3) Enabling customization in real time. Our system concept represents one approach to realizing these goals, however, we contribute the goals and the system as a starting point for future innovations in personalized video-based AAC.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces-*Graphical user interfaces (GUI)*. K.4.2 [Computers and Society]: Social Issues-*Assistive technologies for persons with disabilities*.

## General Terms

Design, Human Factors.

## Keywords

Complex Communication Needs, Video, AAC

## 1. INTRODUCTION

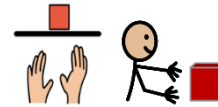
Since the release of the Apple iPad in 2010, speech-generating Augmentative and Alternative Communication (AAC) applications for children with Complex Communication Needs (CCN) have been a rapidly growing consumer-driven trend [1]. Popular examples of AAC applications include Proloquo2Go, Tap to Talk and Voice4u. Like *Vid2Speech*, these applications target beginning communicators with CCN who are preliterate and preverbal, but they feature only static symbols (Figure 1).

Like written words, static AAC symbols are representations of concepts for communication. But for children with CCN, written words are less recognizable, and therefore more difficult to learn and use, than iconic representations like color photographs or drawings [9]. For example, a photograph of a ball is easier to recognize than the written word “ball.” Action words like “want,” however, are not well-represented in static iconic form because they are dynamic and ephemeral, rather than concrete and material [7].

Copyright is held by the author/owner(s).

*ASSETS’12*, October 22-24, 2012, Boulder, Colorado, USA.  
ACM 978-1-4503-1321-6/12/10.

Designing AAC systems that overcome the limitations of static symbols for representing action words is crucial because there is evidence that action words are the most substantial part of a young child’s vocabulary [3]. Furthermore, with modern smartphones and tablets, we have the ability to capture video in everyday settings relevant to specific users, enabling real-time personalization of AAC content. However, owing to the lack of video-based AAC, the requirements for AAC systems using personalized video content are as yet unclear.



**Figure 1: Typical static symbols for the action word “want” found in current AAC applications.** Left image courtesy of The Picture Communication Symbols ©1981–2011 by Mayer-Johnson LLC. All Rights Reserved Worldwide. Used with permission. Right image courtesy of SYMBOLSTIX, LLC. Used with permission.



**Figure 2. Current *Vid2Speech* interface.**

*Vid2Speech*, our AAC prototype, exploits the potential of personalized video for creating recognizable representations of action words. *Vid2Speech* is the first prototype tablet-based speech-generating system for children with CCN that enhances AAC with personalized video of action words. By illuminating the design goals for *Vid2Speech*, we provide a framework for successful integration of personalized video in AAC applications.

## 2. RELATED WORK

Previous work has explored the value of motion effects in AAC [5]. Findings suggest that cartoon animations of action words are easier to recognize than static icons [10,11]; black-and-white

video is easier to recognize than animated line drawings [8]; and animation accelerates learning of static icons [4]. Finally, work by Light & McNaughton [6] suggests that personalization and “just-in-time” capture of language concepts should be key priorities for future AAC systems. Our system is designed to build on these findings by combining motion effects and real-time personalization to increase transparency of action words.

### 3. VID2SPEECH DESIGN GOALS

The Vid2Speech interface displays photos of the child’s scheduled activities across the top of the screen. The user navigates through the schedule to access communication concepts categorized by activity. For example, the park activity contains personalized videos and photos of actions like swing, see saw, and slide (Figure 2). The photos are pressed to play the videos; pressing the photo for swing will play a video of the user swinging.

To develop the early stage prototype of Vid2Speech, we created the following three design goals for video-based AAC:

- Providing social-temporal navigation
- Enhancing comprehension
- Enabling customization in real time

These design goals are described in detail below.

#### 3.1 Providing Social-Temporal Navigation

Young children with CCN may have difficulty understanding and navigating traditional grid displays [2,6]. Research suggests that when asked, children represent language concepts holistically, in context and based on familiar experiences, rather than atomized and hierarchically categorized in discrete parts (i.e. a wagging dog tail categorized under animals) [6]. To facilitate access to videos in Vid2Speech, we have categorized videos according to time of day, place and social situation. For example, the video of “swinging” is categorized under the social activity “park” which is accessed at a particular time of day. This social-temporal navigation provides an accessible and predictable scaffold for adding and using personalized content.

#### 3.2 Enhancing Comprehension

While static representations of action words are less effective than animated or moving representations, there are compelling reasons to retain rather than abandon static symbols. Firstly, displaying multiple moving videos simultaneously on the screen has the potential to maximize the cognitive load for the user and become distracting [5]. Retaining static icons for action words, while correlating them with personalized videos, leverages the advantages of video for illuminating the meanings of static symbols to aid comprehension. The videos help the user learn the correlated static symbols – perhaps even written words – so that the videos can be faded as the user develops into a fluent communicator.

#### 3.3 Enabling Customization in Real Time

Caregivers of children with CCN, such as parents, speech therapists and teachers, do not typically have access to or time for specialized video editing or recording equipment for AAC. Vid2Speech is designed to help caregivers capitalize on serendipitous moments for language learning and growth by capturing personalized videos in real time using a tablet’s built-in cameras. While we expect the initial customization of the software to take some time, the software can be customized incrementally as situations and communication opportunities arise.

## 4. FUTURE WORK

Our next goal is to investigate the ways in which personalized video compares to standardized animated graphics for enhancing AAC for children with CCN. By developing a relationship with caregivers and children, we hope to engage these important stakeholders in a participatory design process before evaluating a final version of the Vid2Speech system.

## 5. ACKNOWLEDGMENTS

This project has been developed with funding from Intel.

## 6. REFERENCES

- [1] Announcing the AAC-RERC White Paper on Mobile Devices and Communication Apps (2011). *Augmentative and Alternative Communication*, 27, 2, 131-132.
- [2] Beukelman, D. R., & Mirenda, P. (1998). *Augmentative and alternative communication: Management of severe communication disorders in children and adults*. Baltimore: P.H. Brookes Pub.
- [3] Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley
- [4] Fujisawa, K., Inoue, T., Yamana, Y., & Hayashi, H. 2011. The Effect of Animation on Learning Action Icons by Individuals with Intellectual Disabilities. *Augmentative and Alternative Communication*, 27, 1, 53-60.
- [5] Jagaroo, V., & Wilkinson, K. 2008. Further Considerations of visual cognitive neuroscience in aided AAC: The potential role of motion perception systems in maximizing design display. *Augmentative and Alternative Communication*, 24, 1, 29-42.
- [6] Light J. & McNaughton D. 2012. Supporting the communication, language, and literacy development of children with complex communication needs: State of the science and future research priorities. *Assistive Technology*, 24, 1, 34-44.
- [7] Lloyd, L. L., Fuller, D. R., & Arvidson, H. H. (1997). *Augmentative and alternative communication : A handbook of principles and practices*. Boston: Allyn and Bacon.
- [8] Mineo, B., Peischl, D., & Pennington, C. 2008. Moving targets: The effect of animation on identification of action word representations. *Augmentative and Alternative Communication*, 24, 2, 162-173.
- [9] Mirenda, P., & Locke, P. A. 1989. A comparison of icon transparency in nonspeaking persons with intellectual disabilities. *Journal of Speech and Hearing Disorders*, 54, 2, 131-40.
- [10] Schlosser, R. W., Shane, H., Sorce, J., Koul, R., & Bloomfield, E. 2011. Identifying Performing and Under Performing Graphic Icons for Verbs and Prepositions in Animated and Static Formats: A Research Note. *Augmentative and Alternative Communication*, 27, 3, 205-214.
- [11] Schlosser R.W., Bloomfield E., Debrowski L., Deluca T., Miller S., Schneider D., Neff A., Shane H., Sorce J., & Koul R. 2012. Animation of graphic symbols representing verbs and prepositions: Effects on transparency, name agreement, and identification. *Journal of Speech, Language, and Hearing Research*, 55, 2, 342-358.