# DigitCHAT: Enabling AAC Input at Conversational Speed

Karl Wiegand
Northeastern University
360 Huntington Ave
Boston, MA 02115, USA
wiegand@ccs.neu.edu

Rupal Patel
Northeastern University
360 Huntington Ave
Boston, MA 02115, USA
r.patel@neu.edu

## ABSTRACT

Augmentative and alternative communication (AAC) systems are used by many different types of people. While almost all AAC users have speech impairments that preclude the use of verbal communication, they may also have varying levels of vision or motor impairments, perhaps due to age or the particular nature of their disorder. Speed, expressiveness, and ease of communication are key factors in choosing an appropriate system; however, there are social considerations that are often overlooked. AAC systems are increasingly being used on mobile devices with smaller screens, in part because ambulatory AAC users may feel uncomfortable carrying around large or unusual machines. DigitCHAT is a prototype AAC system designed for fast and expressive communication by literate AAC users with minimal upper limb motor impairments. DigitCHAT's interface was designed to be used discretely on a mobile phone and supports continuous motion input using a small set of visually separated buttons.

## Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Graphical User Interfaces; K.4.2 [**Social Issues**]: Assistive Technologies for Persons with Disabilities

## General Terms

Design, Human Factors

## Keywords

AAC, Continuous Motion, Mobile

## 1. MOTIVATION

People who use augmentative and alternative communication (AAC) systems often have speech impairments severe enough to prevent the use of verbal communication [1]. Depending upon the nature of their disorders, many AAC users may have accompanying motor impairments, such as tremors or reduced mobility of their hands and arms [2]. They may also have vision impairments that make it difficult to see small font sizes. AAC systems that operate on small mobile devices often use on-screen keyboards that were not designed for people with bigger hands or people with upper limb motor impairments. These keyboards usually occupy less than half of the available screen real estate and have small buttons that are positioned adjacent to each other. Many elderly users, much less users with diagnosed motor or vision impairments, have difficulty with these keyboards because of the button and font sizes [4]. Additionally, these systems tend to focus on the creation and use of stored utterances instead of real-time composition, unnecessarily reducing the flexibility of conversation. The current work aims to address the need for an AAC system that can be used at conversational speeds on a small-screen mobile device by ambulatory users with mild upper limb motor impairments. Although intended for AAC users, this prototype system also has potential for non-AAC users who may be temporarily unable to use their voices.


*ASSETS'14,* October 20–22, 2014, Rochester, NY, USA.
ACM 978-1-4503-2720-6/14/10.
http://dx.doi.org/10.1145/2661334.2661422.

## 2. APPROACH

DigitCHAT enables rapid, face-to-face communication via small touch-screen devices, such as mobile phones. The system uses large buttons to assist users who may have difficulty making precise movements. Buttons are visually separated to maximize visibility when a user naturally obscures part of the screen by touching it. The interface is organized as a telephone number pad, resulting in higher lexical ambiguity compared to commercial systems like Swype or SwiftKey; however, this design provides familiarity, especially for older users, and reduces the time required to learn the layout.

To further increase communication speed and assist users with upper limb motor impairments, DigitCHAT supports two types of input: mixed and continuous. In mixed mode, users can provide a combination of discrete taps or non-contiguous path segments to specify the desired word. At any given time, the most likely word is displayed at the top of the screen, and can be selected by tapping it or ending a path on it. In continuous mode, users draw a single line through all desired buttons. As the user's finger or stylus moves over the screen, the most likely word is displayed at the top of the screen. When the user disconnects from the screen's surface, the most likely unigram is spoken aloud immediately. Users can cancel the current path, without speaking the displayed word, by ending on the Stop or Cancel sign. With many current AAC systems, listeners must

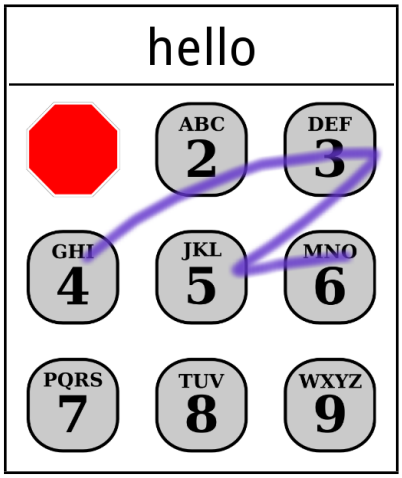


**Figure 1: Example input path for the word "hello."**

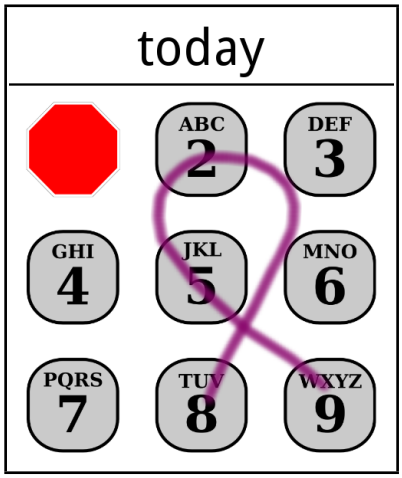


**Figure 2: Example input path for the word "today."**

wait while the user composes a complete utterance [3, 6]. This waiting period places increased pressure on the AAC user to generate utterances as quickly as possible, creating uncomfortable silences and often encouraging the use of telegraphic utterances. By automatically speaking each word as it is completed, DigitCHAT can significantly reduce these gaps and facilitate conversational turn-taking.

DigitCHAT uses a predetermined vocabulary and dictionary-based implementation with unigram statistics based on the Crowdsourced AAC-Like Corpus [5]. Every word in the dictionary is converted into a physical path traversing a standard telephone number pad. The width of this path is incrementally varied up to the size of a standard button and ordered collisions are recorded as possible sequences that a user might take to specify a given word. These paths are then reverse-indexed, so that DigitCHAT can look up the user's provided path and retrieve the set of words, with unigram probabilities, that the path could indicate.

Words that share the same sequence of buttons are sometimes called "textonyms." For example, the words "bat" and "cat" are textonyms because they are both specified with the discrete numeric sequence 2-8 or the continuous motion sequence 2-5-8. DigitCHAT implements two approaches to resolving textonyms. In the first approach, the most likely textonym is displayed and users can scribble over the last button in their numeric sequence to rotate through possible textonyms. Users can disconnect from the screen to speak the displayed word aloud or end their scribbling on the Stop sign to cancel the utterance. In the second approach, DigitCHAT implements basic learning and remembers the user's preferred textonyms for any given path.

## 3. FEEDBACK AND FUTURE WORK

We have made DigitCHAT freely available for Android devices on the Google Play Store in order to gauge interest and elicit feedback. Thus far, DigitCHAT has undergone two design and development iterations based on suggestions from users in the target population. In addition to informal feedback from ad-hoc testers, we have received narrative emails from three users and are preparing for a formal study with participants at a clinical facility that serves individuals with chronic neuromotor disorders. A common request, which has since been implemented, was to allow cancellation in order to prevent unexpected or undesirable words from being spoken. We have created several user-configurable options, such as the movement threshold for textonym rotation, but it may be possible to implement a learning algorithm to discover the ideal values for these settings automatically. While the current version of DigitCHAT relies primarily on unigram statistics, we intend to look at potential improvements from using skip-grams or implementing phrasal prediction. We are also experimenting with different methods to efficiently add and remove words from the dictionary to allow for full vocabulary customization.

## 4. REFERENCES

[1] J. Higginbotham, H. Shane, S. Russell, and K. Caves. Access to AAC: Present, past, and future. *Augmentative and Alternative Communication*, 23(3):243–257, Jan. 2007.

[2] J. Light, D. Beukelman, and J. Reichle. *Communicative competence for individuals who use AAC: From research to effective practice*. Paul H. Brookes Publishing Co., 2003.

[3] J. Light and K. Drager. AAC technologies for young children with complex communication needs: State of the science and future research directions. *Augmentative and alternative communication*, 23(3):204–216, 2007.

[4] H. Nicolau and J. Jorge. Elderly text-entry performance on touchscreens. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '12, pages 127–134, New York, NY, USA, 2012. ACM.

[5] K. Vertanen and P. O. Kristensson. The imagination of crowds: Conversational AAC language modeling using crowdsourcing and large data sources. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 700–711. ACL, 2011.

[6] B. Wisenburn and J. J. Higginbotham. An AAC application using speaking partner speech recognition to automatically produce contextually relevant utterances: objective results. *Augmentative and alternative communication*, 24(2):100–109, 2008.