Assistive Technology Lab
3-59 Corbett Hall – University of Alberta Edmonton, AB T6G 2G4
Ph: (780) 492-5422 Fx: (780) 492-1696 Email: atlab@ualberta.ca

Alberta Health Services

UNIVERSITY OF ALBERTA

Published citation:
Higginbotham, D. J., Bisantz, A.M., Sunm, M., Adams, K.D and Yik, F. (2009). The Effect of Context Priming and Task Type on Augmentative Communication Performance, Augmentative and Alternative Communication.

The Effect of Context Priming and Task Type on Augmentative Communication Performance

D. Jeffrey Higginbotham, Ann M. Bisantz, Michelle Sunm, and Kimberley Adams

Department of Communicative Disorders and Sciences
Department of Industrial Engineering
University at Buffalo, The State University of NY
Faculty of Rehabilitation Medicine, University of Alberta
Glenrose Rehabilitation Hospital, Edmonton, Alberta

Abstract

Augmentative and Alternative Communication (AAC) devices are special purpose electronic devices that generate speech output and are used by individuals who have specialized communication needs instead of voice communication. Word prediction, including context specific prediction has been proposed to help overcome barriers using these devices (e.g. slow communication rates and limited access to situation related vocabulary) but has not been tested in terms of effects during actual task performance. In this study, we compared AAC device use, task performance, and user perceptions across three tasks, in conditions where the AAC device used either was, or was not, primed with task specific vocabularies. While context priming had a significant effect on AAC device use as measured by keystroke savings, these advantages did not translate into higher level measures of rate, task performance, or user perceptions. Importantly, this indicates to AAC device designers and users that fine-grained measures of device use may not be predictive of high level performance.

Introduction

Over the past thirty years, advances in Augmentative and Alternative Communication (AAC) have provided individuals with complex communication needs (e.g., persons with cerebral palsy, head injuries, autism, mental retardation, ALS, stroke, etc.) with special purpose electronic devices for communication (Beukelman & Mirenda, 2005). These devices generate speech output and are used by individuals instead of voice communication. These devices employ customized input interfaces, sophisticated vocabulary databases and search algorithms, synthetic speech output, and access to mainstream computer programs. Such innovations provide the technological infrastructure required for successful societal integration.

Despite these advancements, significant technologically-related performance barriers exist that significantly impede educational, vocational and social achievement. Documented problems include:

•Slow communication rates (<10 wpm) that limit education and workplace capabilities.
•Frequent problems accessing and controlling the AAC system during conversation tasks.
•Lack of linguistic precision and limited access to situation related vocabulary.
•Difficulties learning and using current AAC interfaces.

Word prediction and other rate enhancement approaches have been proposed to address the first three issues listed above. Recent developments in augmentative communication have focused on using the web and other sources to prime word prediction systems with topic specific vocabulary to make communication faster and more efficient. The use of pattern recognition and other AI-related algorithms hold the promise for optimizing prediction vocabularies based on past user performance. However, it is not clear how these context-specific strategies will actually benefit the interactive communication performance of augmented speakers. For example, as shown by Koester and Levin (1996) dynamic display devices, ie. changing word prediction lists, impose considerable workload-related constraints on user performance.

The purpose of this study was to examine communication performance and human factors related to the use of context-enhanced word prediction systems.

Method

Participants

Twenty-four pairs of non-disabled adults agreed to participate in this study. Participants were between the ages of 17 and 60 ($M = 23.15$, $SD = 7.531$). Participants were fluent English speakers with no prior AAC device experience, no diagnosis or treatment for hearing loss, and no diagnosis or treatment for language disorder. One participant in each pair was randomly selected to be the AAC user, the other participant served as the communication partner.

Equipment

A Fujitsu tablet computer equipped with a touch screen interface was used in this study. Augmentative communication software (the Portable Tablet IMPACT, produced by Enkidu Research, Inc; Moulton, 2002) which allowed users to key in letters and produce voice output was used; the software also performed word prediction from an extensive dictionary based on letters that the user typed. The device was programmed with a 1-second key-press delay to better simulate conditions of typical augmentative device use (e.g., many individuals who use AAC devices have physical disabilities and cannot key or type fluently).

Tasks

Participants completed three tasks during the experiment: a narrative task, a map task, and a tangram puzzle. During the narrative task, the AAC user silently read a story taken from the Resource Allocation Paradigm of Pittsburgh (Doyle and McNeil, 1998), and then retold the story to the communication partner. The AAC user could refer to pictures illustrating the story but not the text during re-telling. A 10-question test of comprehension was then given to the communication partner. During the map task (taken from the HCRC Map Task Corpus; HCRC, 1992), both partners were given a simplified map showing landmarks; the AAC user's map had a route drawn on it. The AAC user instructed the partner how to draw the route (with respect to the landmarks) on the map. To increase task complexity, the maps were slightly different (e.g., some landmarks were changed or moved). For the tangram task, the pairs collaborating on solving a tangram (7 piece shape) puzzle. They each were given three of the pieces, and were instructed not to move their own pieces, but to take turns instructing their partner as to how to move the pieces in order to match the goal puzzle shape.

Context Priming

Prior to the experiment, 10 pairs of participants performed each of the three tasks, twice, using the device. The AAC device captured the vocabulary generated during the task into a logfile. These logfiles were then used to "prime" the device, separately for each task, for some experimental conditions. Priming the device with the task specific vocabulary increased the chances that context specific words would be predicted by the device, during the task.

Independent Variables

Two independent variables were tested in a mixed experimental design. The device was either primed with the task vocabulary for the task ("context-on") or not ("context-off"); this was a between-subjects variable. Task (narrative, map, tangram) was a with-in subjects variable. Task order was balanced across pairs.

Procedure

To complete the task, the AAC user was first given 15 minutes of training with the device. Pairs then completed the three tasks. Participants completed the NASA-TLX workload assessment after each task. After all tasks were completed, participants completed a user satisfaction questionnaire with questions drawn from the Questionnaire for User Interaction Satisfaction (Version 7) (QUIS) (Norman et al., 1998).

Dependent Measures

For each task, task specific performance measures were assessed. These included: score on the comprehension test for the narrative task; number of pairs completing, percent of path length completed, and error in route (computed as the area between the intended and drawn routes according to published procedure) for the map task; puzzle completions for the tangram task. Task time and NASA-TLX scores were also captured for each task. User satisfaction measures were collected at the end of all three tasks.

Task process measures were also computed to understand the manner in which participants used the AAC device. Logfiles of user were automatically captured and analyzed (Lesher et al, 2003) to determine the total of words output (e.g., spoken) by the device; the words used per minute (with pauses of 10 s or more removed); and keystroke savings (the ratio of keystrokes input to the device divided by the keystrokes that would have been required to produce the words, with no word prediction capabilities available).

Results

Results were analyzed using one-way (for task specific measures) or two-way mixed ANOVAs. Data from two pairs' narrative task were excluded from some analyses because they had been given incorrect instructions regarding the maximum length of time for the task.

Participants in the context-on group had a significantly higher percentage of keystroke savings than participants in the context-off group ($F(1, 20) = 4.310$, $p = 0.051$). This difference, however, did not translate into significant differences between context condition for words per minute and total number of words. There were no significant effects of context condition on any of the task specific performance measures, nor for workload or usability measures.

There were significant main effects of task on words per minute and total number of words generated. Post-hoc analyses indicated that the narrative and map tasks had significantly higher word per minute scores compared with the tangram task, and that the number of words used for the map task was significantly higher than the other two tasks. There were also significant main effects of task on task completion time and workload (narrative < map < tangram for both).

Discussion & Conclusions

Studies that examine the impact of AAC device characteristics, such as context-based priming, on task performance are critical because ultimately, the intent of AAC devices is to aid users in functional (home, school, community) tasks and settings. Reducing keystrokes, or improving communicative rate are important indicators of the effectiveness with which word prediction algorithms or interface manipulations have been implemented. However, direct links between these measures (used commonly in AAC research) and task performance measures have not typically been investigated.

In our study, we did find that context-based vocabulary priming improved a measure of keystroke savings. However this effect did not translate to higher level measures of device use reflective of a conversation (words per minute or total number of words generated) nor to task performance measures. Thus, demonstrating improvements in keystroke use may not be sufficient evidence to conclude that such improvements will follow in other measures of communication speed and word use in conversations, nor to measures of task performance, workload, or user satisfaction. We also demonstrated that measures of AAC use are sensitive to the task the AAC user is performing. The relative utility of AAC devices in different task environments is a fertile ground for future study.

A limitation of this study is that it used constrained laboratory tasks performed with able bodied participants, who were not familiar with AAC in general and had limited device training. We did manipulate the device (through a 1 s keystroke delay) to simulate restrictions in use experienced by some typical AAC device users. Further work using actual AAC device users, in real task contexts, is necessary to validate improvements in AAC device design.

Finally, a simulation study (not described here) based on the context vocabulary gathered during the training phase indicated that under simulated optimal choices, context priming had an even greater effect on keystroke savings. The fact that this effect was reduced during the actual study indicates that theoretical improvements in efficiencies may be attenuated during actual use (i.e., users may not always select predicted words as a whole, but may type each letter instead), perhaps due to factors such as the interface itself, training, or shared attention across the device, the task, and other communication partners. These factors should be studied explicitly in further research.

References

Beukelman, D.R. & Mirenda, P. (2005). Augmentative and Alternative Communication, Third Edition, Baltimore: Brookes Publishing..

Hart, S.G. Straveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and Theoretical research. Human Mental Workload. P. A. H. a. N. Meshkati: 139-183.

Human Communication Research Centre, H. C. R. C. (1992). The HCRC Map Task Corpus. Edinburgh, Scotland.

Koester, H. H. and S. P. Levine (1996). "Effect of a Word Prediction Feature on User Performance." Augmentative and Alternative Communication 12(3): 155-168.

McNeil, M. R. (1998). Resource Allocation Paradigms of Pittsburgh (RAPP), University of Pittsburgh.

Moulton, B. J., Lesher, G. (2002). Tablet Portable IMPACT. Spencerport, NY, Enkidu Research, Inc.

Lesher, G.W., Moulton, B.J., Rinkus, G., & Higginbotham, D.J. (2003). Software tools for emulation and analysis of augmented communication. CSUN 2003, California State University, Northridge.

Norman, K., Schneiderman, B, Harper, B. & Slaughter, L. (1998). Questionnaire for User Interaction Satisfaction (version 7). College Park, Maryland, University of Maryland.