FINAL PROGRESS REPORT - 1R43HD068026-01A1

Company: Koester Performance Research
Principal Investigator: Heidi Koester, PhD
April 30, 2013

ABSTRACT

This report summarizes the results achieved in the Phase I SBIR project, "Model-Based Software for Configuring Single-Switch Scanning Systems." We developed an algorithm for modifying a single-switch scanning interface to increase a user's text entry rate (TER). We evaluated that algorithm in a study of nine single-switch scanners. Text entry rates improved by an average of 120% (p=.003). All nine subjects increased their TER by at least 40% and five of the nine increased their TER by over 100%.

ORIGINAL SPECIFIC AIMS

As stated in the original proposal:

> *Our Specific Aim in Phase I is to develop a software tool that allows clinicians to identify the most appropriate configuration for a single-switch scanning system...*
>
> *... we propose to develop a software tool that will assess each user's abilities and provide a near optimal set of parameters for any specific scanning interface for that specific user. Our software will help clinicians determine a highly effective configuration for the scanning system in order to maximize TER for that individual...*
>
> *Test of Feasibility: We must show that our tool will arrive at scanning system settings that will increase TER for subjects by 100% or more. We must also show that the simulation model within our software predicts TER to within a 10% error or less, across a variety of scanning configurations.*

PROGRESS TOWARD SPECIFIC AIMS

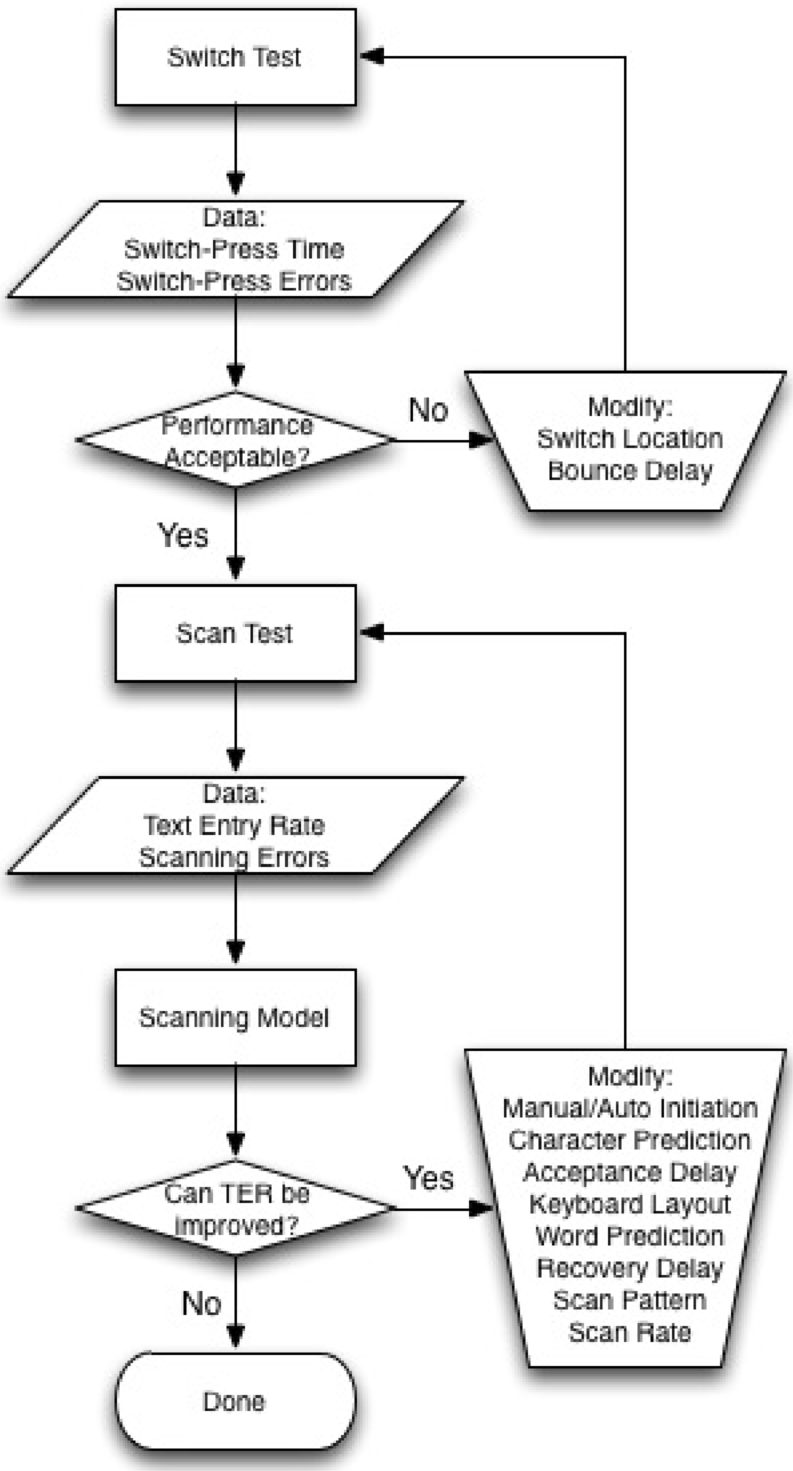
Algorithm Development

We developed and evaluated an algorithm for enhancing a client's single-switch scanning interface (see Figure 1).

Step 1 - Switch Test. The switch test is currently implemented within a special build of our Compass assessment software. The switch test asks the user to complete a series of single-, double- and triple-clicks. The average switch press time and the frequency of switch press errors (specifically errors from "bouncing" on the switch) are calculated from the data. The switch test can be repeated with different switches or switch activation sites until performance is acceptable to the clinician and the client.

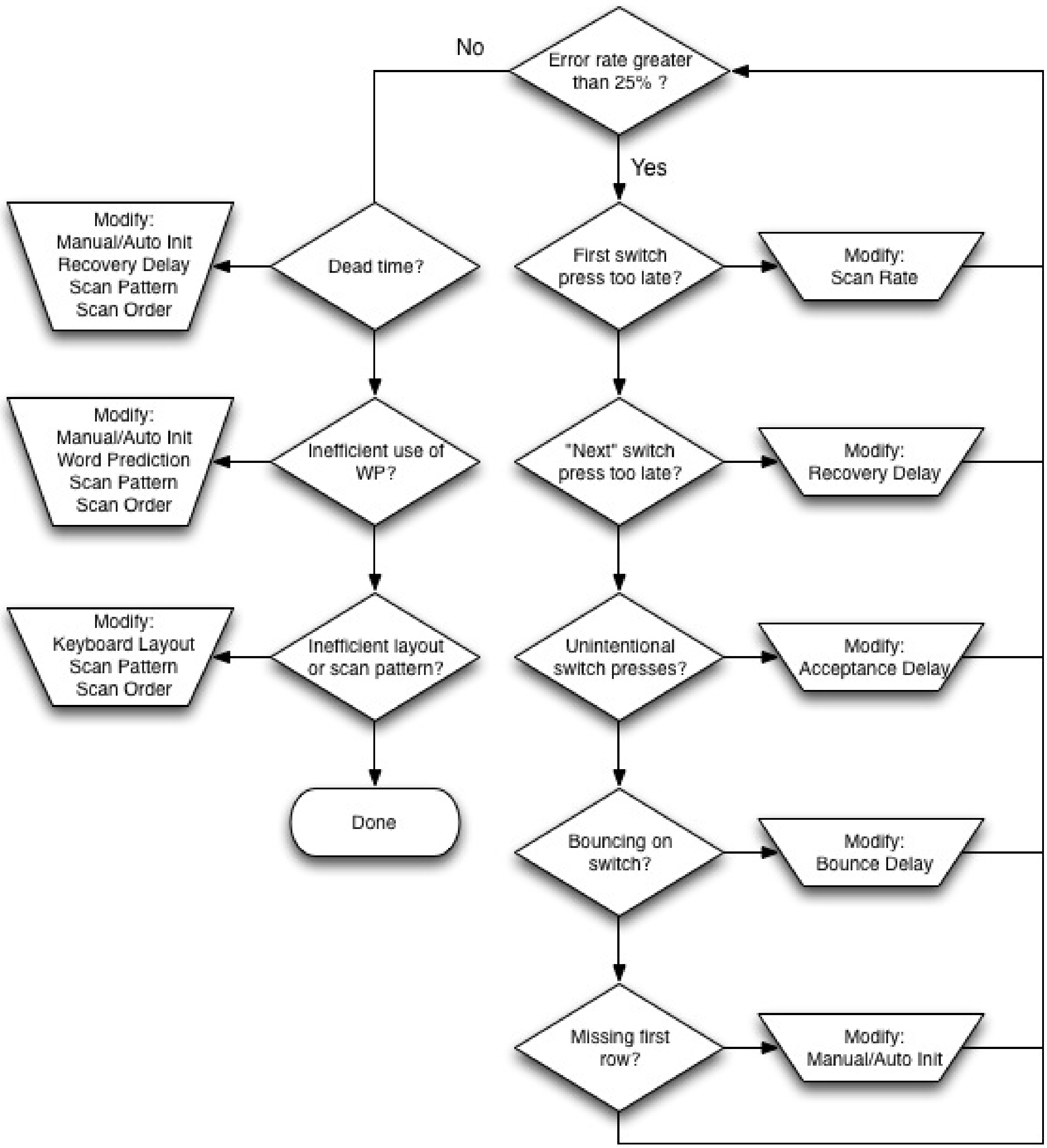
Step 2 - Scan Test. The scan test is currently performed on the user's AAC device. Clients are asked to enter two sentences on their device. Raw "keystroke" data are collected through the language activity monitoring (LAM) features built into some AAC devices and a video recording of the user's screen. This data are analyzed manually to identify scanning errors and to calculate TER.

Figure 1. Algorithm for modifying the configuration of a single-switch scanning interface.



Step 3 - Scanning Model. Data from the switch and scan tests are used to identify specific modifications to the client's scanning interface (see Figure 2). If the user's scanning error rate is greater than 25%, modifications are made to reduce errors. Once the error rate has been reduced, additional modifications are made to increase the client's overall efficiency. The scanning model is used to predict the effect of modifications before they are made.

Figure 2. Decision algorithm for making modifications



Evaluation

We compared the performance of single switch scanners with their existing configuration to their performance with the configuration identified with our algorithm. A longitudinal ABA study design was used.

Baseline. Each subject's baseline TER was measured using their current scanning system and configuration. Subjects transcribed two sentences using their AAC device in its normal configuration. Subjects also transcribed two sentences using a letters-only keyboard with the other configuration parameters (e.g., scan rate, recovery delay) identical to their original configuration.

Modification. The algorithm was used to identify an optimal configuration for each subject. The predicted optimal settings were entered into each subject's scanning system, for them to use for the remainder of the study.

Intervention. Each subject completed four intervention sessions, with at least one week between each session. In each session, subjects transcribed two sentences using their AAC device with the configuration identified by the algorithm and two sentences using a letters-only keyboard (with the other configuration parameters identical to the configuration identified by the algorithm).

Reversal. After four sessions were completed, each subject's scanning system was restored to its original settings. Subjects transcribed two sentences using their AAC device in its original configuration and two sentences using a letters-only keyboard.

Nine subjects completed the protocol. As shown in Figure 3, all subjects increased their TER by at least 40% and five of the nine subjects increased their TER by over 100%. As shown in Figure 4, all nine subjects returned to baseline performance during the reversal phase. Average performance in the baseline and reversal phases was compared to performance during the intervention phase using a paired t-test. The difference between original and modified configurations was significant ($p = .003$).

Figure 3. Increase in Text Entry Rate for each subject

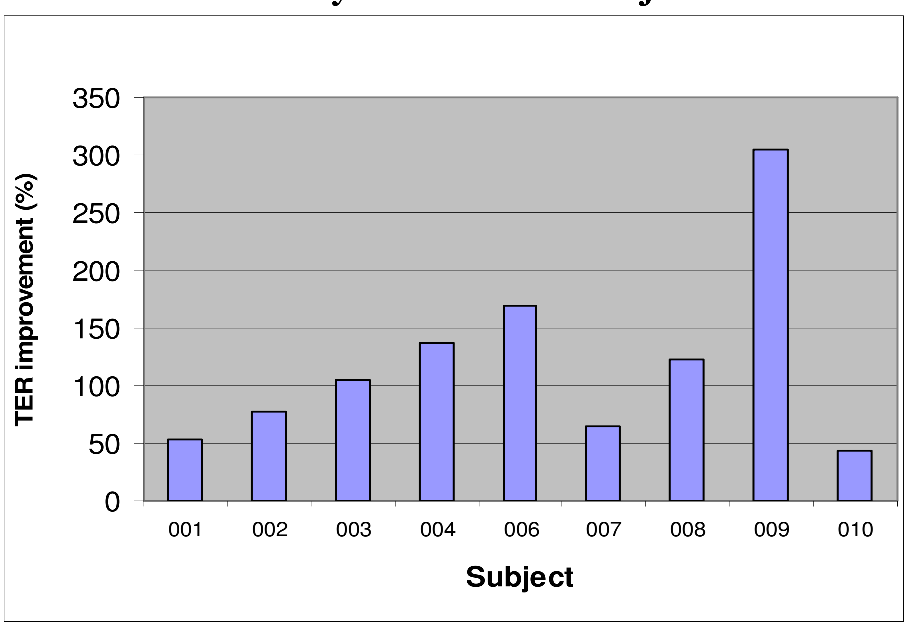
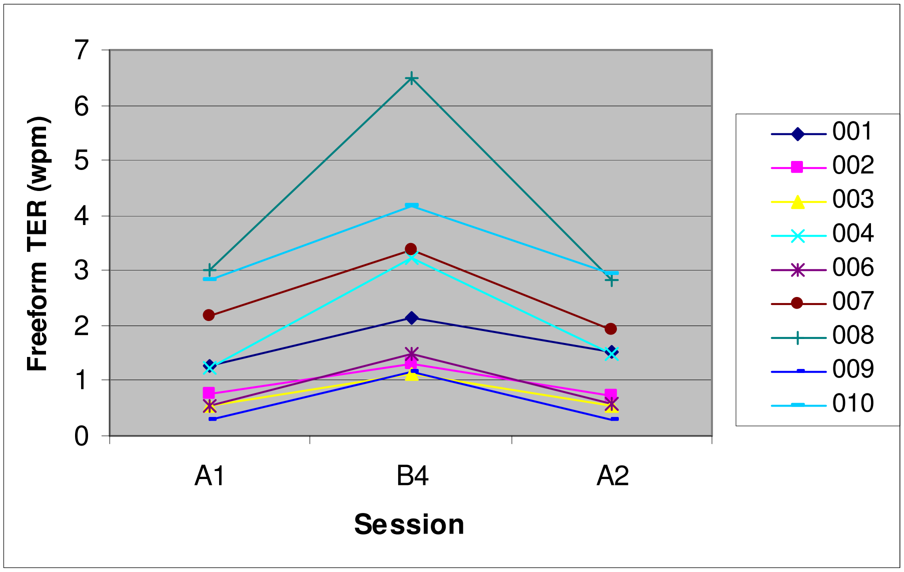


Figure 4. Comparison between baseline (A1), intervention (B4) and reversal (A2) conditions for each subject. Intervention data was collected during the fourth session.



Model Accuracy

A key aspect of demonstrating feasibility of this approach is establishing the accuracy of the model's predictions. We have thus far focused on modeling performance with the letters-only (LO) keyboard, because the model has not been extended to cover all of the text entry features used by subjects in the free-form (FF) condition. The model predicts a subject's TER based on the user's behavior (e.g., frequency of errors occurring) and the system's configuration (e.g., letter layout, scan rate, recovery delay).

As shown in Table 1, six different data sets were used in the model to predict a subject's TER when using the LO keyboard during the baseline session (A1), the fourth intervention session (B4), and the reversal session (A2):

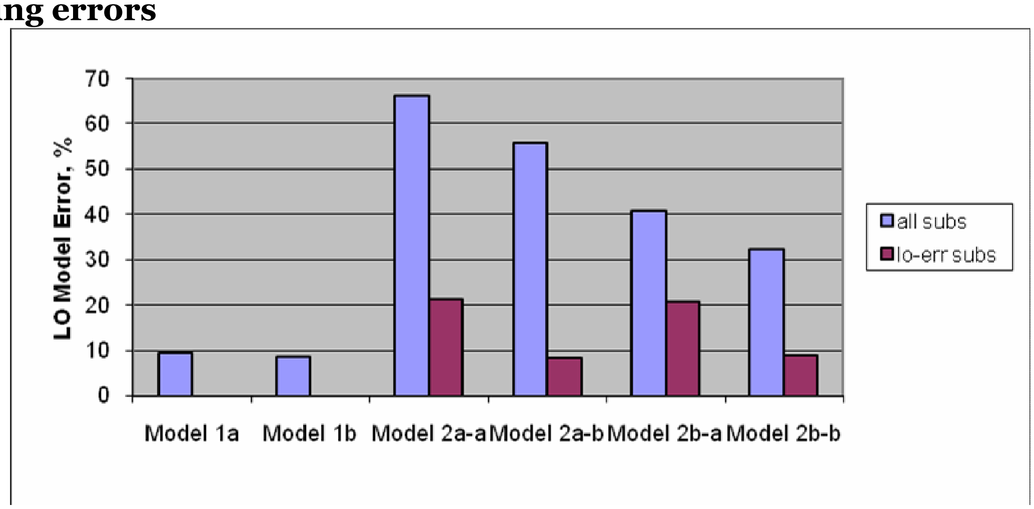
- Model 1a: Actual error rates for A1, B4, and A2. Switch press time taken from Switch test measurements
- Model 1b: Actual error rates. Switch press time obtained by multiplying the system's Scan Rate by 0.65 (in effect, reversing the 0.65 rule).
- Model 2a-a and 2a-b: Error rates from the Scan test conducted prior to Baseline, using a scanning system configured with the same timing parameters as the subject's Baseline system.
- Model 2b-a and 2b-b: Error rates from the Scan test conducted prior to Baseline were used to predict performance at A1 and A2. Error rates from B1 (the first session in which the reconfigured system was used) were used to predict performance at B4.

Table 1. Data sets used to make model predictions for Sessions A1, B4, A2.

Model Flavor	Scanning Error Rates	Switch Press Time	Settings
Model 1a	A1, A2, B4	Switch test	Each session
Model 1b	A2, A2, B4	0.65 * scan rate	Each session
Model 2a-a	Scan test	Switch test	Each session
Model 2a-b	Scan test	0.65 * scan rate	Each session
Model 2b-a	Scan Test for A1 and A2 B1 error rates for B4	Switch test	Each session
Model 2b-b	Scan Test for A1 and A2 B1 error rates for B4	0.65 * scan rate	Each session

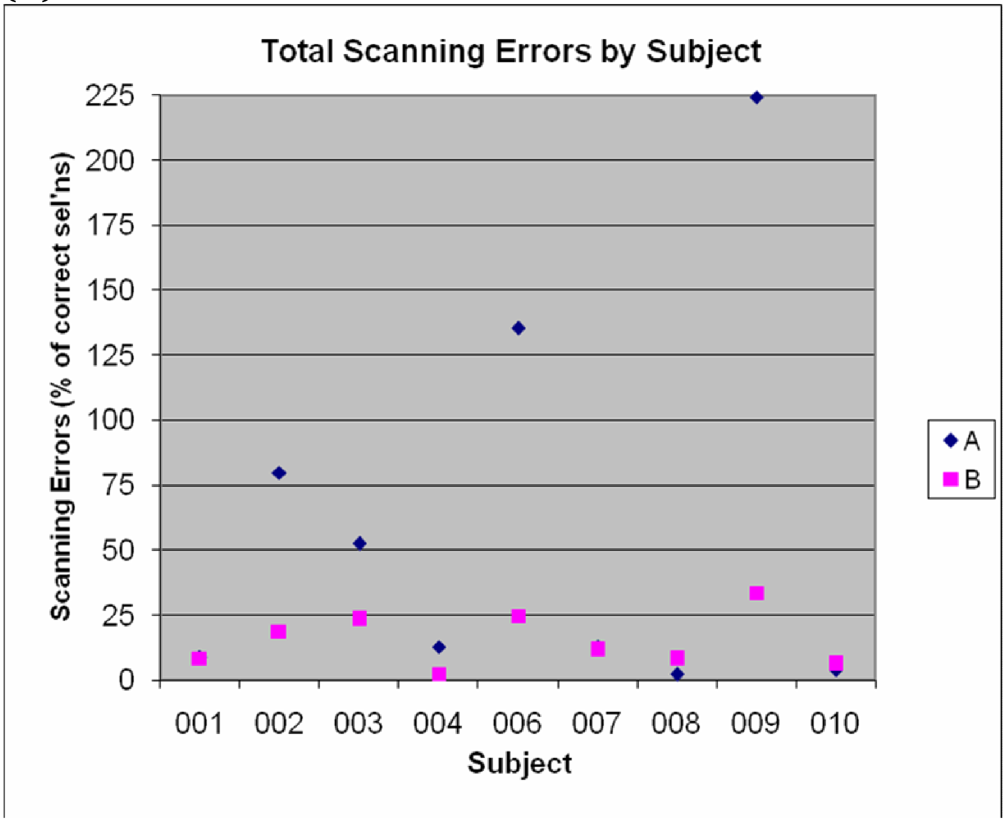
As shown in Figure 5, both Models 1a and 1b had high overall accuracy, with error averaging less than 10% across all subjects. For the Model 2a flavors, which used fixed error rates from a separate Scan test, model error averaged more than 50%. However, looking just at the 4 subjects whose scanning errors were below 25%, Model 2a was much more accurate, with Model 2a-b the best at 9.81% error.

Figure 5. Model accuracy for all subjects and for subjects with less than 25% scanning errors



As expected, the accuracy of the model's predictions is determined by the accuracy of the input data it receives. The model can be given precise values for the system's configuration, but the user may change his or her behavior in ways the model cannot anticipate. This is especially true when the configuration of the scanning system is changed to reduce errors. As shown in Figure , four subjects experienced dramatic drops in error rate, the exact magnitude of which were difficult to predict a priori.

Figure 6. Error rates for each subject averaged across all (A) baseline and (B) intervention sessions.



<u>Test of Feasibility</u>

From our Phase I proposal, our feasibility goals were: 1) to increase text entry rate for subjects by 100% or more; and 2) model TER to within a 10% error or less. We met the first goal, with an average TER improvement of 120%. This strongly suggests that our method for enhancing text entry rate yields successful results across a variety of individuals and AAC systems. We met the second goal for Models 1a and 1b. This has two implications. First, for individuals who already use their AAC system with very few errors (about half of our study sample), the model can accurately predict the impact of various changes to the system configuration. Second, for individuals whose current system is difficult for them to use, the initial intervention is to decrease those scanning errors; then our model can accurately simulate how a person's TER will change in response to changes in scanning errors. Refinements to this model in Phase II will allow clinicians to apply our algorithm to their single-switch scanning users more efficiently and effectively.

List of Publications and Presentations

Presentations

1. Koester, H. Be an AT Quant! AT Practitioner Rounds, Wexner Medical Center, Ohio State University, 11/14/2012.

2. Koester, H. Better Computer Access Solutions through the Use of Evidence. Invited lecture in OT 7411 graduate course, Ohio State University, 2/6/2013.

3. Koester, H., DiGiovine, C., Wenninger, B. Be an AT Quant! RESNA 2013 Conference, to be presented 6/22/2013.

Publications

We have not completed any manuscripts to date. We are preparing two manuscripts for submission in the 2nd quarter of this year.