Effect of System Configuration on User Performance with Word Prediction: Results for Users with Disabilities

Heidi Horstmann Koester, Koester Performance Research; Richard C. Simpson, TRACLabs

Abstract

Six subjects with physical disabilities transcribed text using two different configurations of a word prediction system. System configuration had a meaningful effect on text entry rate for half of the subjects, with an average of over 60% difference between configurations. For the other half of the subject group, the effect of configuration was less clear, perhaps confounded with the effects of individual variation and increase in skill level over time. Results support the possible benefit of an adaptive system that configures itself to the needs of a given user.

Background

The purpose of this experiment is to examine the effect of system configuration on user performance with word prediction. Word prediction systems can be configured to more closely meet the needs of a given user through the settings of numerous parameters. The system configuration may affect the user's performance with the system, although the effect may not be easily intuited. Various studies have reported on user performance with word prediction (e.g., [1]), but we have found none that has systematically studied the effect of system configuration.

There are two reasons for our interest in how system configuration influences user performance. First, it has direct clinical relevance. Clinicians might be able to configure a system to provide better performance for users if they had a greater understanding of the likely effects of specific settings. Second, it is part of a larger project to develop a word prediction system which will automatically modify, or suggest modifications to, the system configuration in order to maximize the user's text entry rate. A key component in developing these experts is a good understanding of the system parameters that make the most difference to user performance.

Research Question

How does the configuration of a word prediction system affect the text entry performance of an individual user?

Methods

Subjects. Six individuals participated, each with a physical disability with respect to keyboard use. Medical diagnoses included two people with multiple sclerosis, two with cerebral palsy, and two with spinal cord injuries. All were able to use a keyboard with one or more fingers or a mouthstick. All were highly literate, with at least some college education. Only one was an experienced word prediction user; the others were novices with respect to word prediction.

Experimental Design. The study employed an ABA single case design, spaced over two sessions. The general goal was to establish two different word prediction system configurations (A and B) for each subject, then measure and compare performance for each configuration. In the first session, baseline performance with word prediction under Configuration A was measured. In the second session, performance with word prediction under Configuration B was measured, followed by a second measurement of performance under Configuration A. Each phase of the ABA design consisted of a minimum of three trials. Each trial involved transcription of two sentences, with the same sentences used throughout the experiment.

System. A testbed word prediction system was developed for the purposes of this study. This provided a completely flexible method of configuration, a 40,000 word dictionary, and fully integrated data collection facilities. (One person used her own word prediction system, Co:writer, since it was tailored to accommodate her vision problems.)

Definition of System Configurations A and B. The configuration of the word prediction system was defined as the choice of values for the following parameters: when the prediction list is first presented, when the list is removed following first presentation, the number of words in the list, and the minimum number of letters in words in the list. All other system parameters and the word prediction dictionary itself remained constant across all conditions.

The baseline condition, Configuration A, was the same for all six subjects. The prediction list was always displayed, there were 6 words in the list, and there was no minimum length for words included in the list. Configuration B represented an attempt to influence performance by adjusting the configuration parameters based on results of the first A phase. Observations of subject performance during the first A phase were complemented by use of a model equation to predict a configuration that would result in at least a 20% change relative to baseline performance [2]. As an example, one subject's Configuration B used a 3-word prediction list, which appeared on the screen only after the first 2 letters of a word had been entered and contained only words of 5 letters or more.

Data Analysis. The primary measure was text entry rate for each trial, measured in characters per minute. Differences in text entry rate between the A and B phases were assessed graphically for each subject. A visually obvious change in level from the first A phase to the B phase, followed by at least a partial reversal in level with the second A phase, was taken as evidence of a meaningful difference. Additionally, the following quantitative criterion was established for a meaningful difference. For each subject, average text entry rate for each phase was computed using the last two trials in the phase. A percentage difference of greater than 20% for the B phase relative to both the baseline and reversal A phases was taken to be a meaningful difference.

RESULTS

As shown in Table 1, half of the subjects (ML, KM, MK) showed meaningful differences in text entry rate when using different system configurations. For these 3 subjects, text generation rate averaged 61% faster with Configuration B relative to the first phase with Configuration A. The left half of Figure 1 shows the results for one subject in this group.

Subject	Initial Effect (%)	Reversal (%)
ML	77.72	52.36
KM	58.77	27.27
MK	46.02	28.21
MC	28.46	1.26
MR	21.17	5.77
TB	12.17	28.72

Table 1. Change in text entry rate performance between the phases for each subject.

For the remaining three subjects (MC, MR, TB), use of Configuration B had a lower effect on text entry rate relative to baseline (averaging 20% different), and in two cases, a sufficient reversal was not seen on the return to Configuration A. The lack of adequate reversal suggests that some other effect, such as increased skill, rather than system configuration, was primarily responsible for the initial change observed. The right half of Figure 1 shows an example of this for one subject.

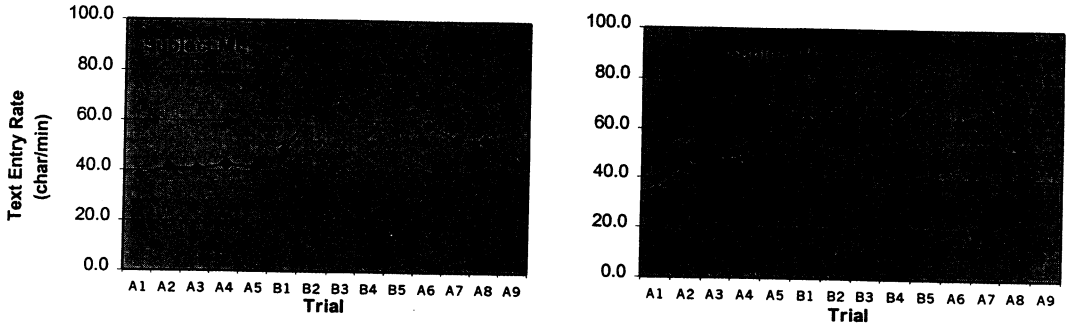

Figure 1. Text entry rate across all trials for two of the six subjects.

DISCUSSION

For some individuals, the configuration of a word prediction system can have a big influence on the resulting text entry rate achieved with the system. This supports the idea of having the system play a role in configuring itself for a given user. One parameter that seemed especially influential was the setting for when the list is displayed. In Configuration A, which is similar to many "default" settings of commercial systems, the list was always on-screen. For most of these subjects, Configuration B delayed the presentation of the prediction list until the user had entered one or more keystrokes. For the faster typists in this subject group, this provided a more advantageous balance between searching and typing, leading to a faster text entry rate. The results suggest that even relatively quick typists (even up to 100 char/min) may benefit from word prediction if the system is configured properly.

The performance of other individuals appears not to be as sensitive to the system configuration. Some people more readily adapt to different configurations. For example, the effect of hiding the list can be mimicked by the user simply ignoring the list during those keystrokes.

The limitations of this study include: relatively few trials per phase, choice of only two configurations for each subject, and natural variation within subjects between the study sessions. A complementary study, also submitted to RESNA 2000, was performed with able-bodied subjects to address at least some of these limitations.

REFERENCES

1. Koester HH, Levine SP. (1996). The effect of a word prediction feature on user performance. Augmentative and Alternative Communication, 12, 155-168.
2. Koester HH, Levine SP. (1997). Keystroke-level models for user performance with word prediction. Augmentative and Alternative Communication, 13, 239-257.

ACKNOWLEDGMENTS

We thank the participants for their generous gifts of time and energy. This research was funded by a Phase I SBIR grant, awarded to TRACLabs by the National Center for Medical Rehabilitation Research of the National Institute of Child Health and Human Development.

ADDRESS

Heidi Horstmann Koester, Ph.D.; Koester Performance Research
2408 Antietam; Ann Arbor MI 48105; hhk@umich.edu