Usage, performance, and satisfaction outcomes for experienced users of automatic speech recognition

Heidi Horstmann Koester, PhD

Rehabilitation Engineering Research Center on Ergonomics, University of Michigan, Ann Arbor, MI

Abstract—This paper presents a variety of outcomes data from 24 experienced users of automatic speech recognition (ASR) as a means of computer access. To assess usage and satisfaction, we conducted an in-person survey interview. For those participants who had a choice of computer input methods, 48% reported using ASR for 25% or less of their computer tasks, while 37% used ASR for more than half of their computer tasks. Users' overall satisfaction with ASR was somewhat above neutral (averaging 63 out of 100), and the most important role for ASR was as a means of reducing upper-limb pain and fatigue. To measure user performance, we asked users to perform a series of word processing and operating system tasks with their ASR systems. For 18 of these users, performance without speech was also measured. The time for nontext tasks was significantly slower with speech ($p < 0.05$). The average rate for entering text was no different with or without speech. Text entry rate with speech varied widely, from 3 to 32 words per minute, as did recognition accuracy, from 72% to 94%. Users who had the best performance tended to be those who employed the best correction strategies while using ASR.

Key words: assistive technology, communication aids for disabled, computer access, outcomes, physical disability, speech recognition, user performance.

INTRODUCTION

Automatic speech recognition (ASR) systems for computer access allow users to enter text and commands into the computer using their voice. These systems have the potential to greatly improve the productivity and comfort of performing computer-based tasks for a wide variety of users. ASR may be particularly attractive to people with physical disabilities, when nonspeech methods of computer input, such as the keyboard or mouse, may be too slow or too painful to fully meet their needs. This study explores how well ASR systems are meeting the needs of users who have physical disabilities. It not only focuses on user performance measures such as speed and accuracy, but also assesses qualitative aspects of using ASR such as usage and satisfaction. The following literature review summarizes what is known about ASR usage, performance, and satisfaction for users with physical disabilities. A more detailed literature review was published previously [1].

Usage

Usage refers to both the frequency and purpose of ASR system use. Frequency is a measure of how often the ASR system is used, and can be specified more precisely when one considers how often ASR is selected for use by individuals who have a choice of input methods. The purpose of use addresses the sorts of tasks for which speech recognition is used, again particularly by individuals who have a choice of input methods.

Two studies have reported frequency of ASR system usage for users with disabilities [2,3]. In a telephone survey of 28 users of discrete speech recognition systems (which is an earlier generation of ASR technology no longer in common use), 21 percent of the participants reported using their ASR system one to three times a week, while 54 percent used it four or more times a week [2]. The remaining 25 percent had stopped using ASR. The survey did not consider task-specific usage of ASR but did report that 61 percent of participants had non-ASR input methods available. DeRosier interviewed 10 ASR users [3]. Of all participants, 30 percent reported using their system every day, 30 percent several times a week, and 40 percent once a week or less. Of these participants, 80 percent had non-speech input methods available to them, but the study did not report the relative usage of ASR as compared to these other methods. A third study discusses the concepts of use and achievements through use, measuring both constructs for 40 current and prior ASR users with disabilities [4]. However, the actual usage data are not presented in reports on the study.

These studies suggest that a majority of ASR users who have physical disabilities use their ASR system regularly. However, an additional, significant minority appears to use ASR seldom or not at all. A goal of this study is to understand ASR usage patterns in more depth.

Performance

Performance primarily refers to the speed and accuracy with which tasks can be performed with speech recognition. Text entry rate in words per minute (wpm) is the primary speed measure discussed in this paper, because text entry is a necessary element of general-purpose computer access. Accuracy for speech recognition is typically reported as the recognition accuracy of the system itself, indicating how well the system recognizes the user's voice. Recognition accuracy is reported as the number of words correctly recognized, as a percentage of the number of words spoken [5]. **Table 1** summarizes published data

Table 1.
Performance with continuous speech recognition systems for participants without disabilities. Text entry speeds include time required to correct recognition errors. Transcription rate for these participants using standard keyboard averaged 32.5 words per minute (wpm).

User Experience	Recognition Accuracy (%)	Text Entry Rate (wpm)
Initial Use	85–93	14
Extended Use	94	25–30

on text transcription performance for people without disabilities using commercially available continuous speech recognition systems [5–7]. Reported text entry rates, after approximately 20 hours of experience, have averaged 25 to 30 wpm, with recognition accuracy of 94 percent [6,7]. For these users without disabilities, speeds on text entry tasks using speech input were generally slower than speeds on the same tasks using keyboard and mouse. However, to make a fair comparison to more commonly used methods such as the keyboard can be difficult, since many of the participants had already developed a high degree of skill with these methods.

The available data show that text entry rates with ASR, even at fairly high levels of recognition accuracy, are much slower than voice dictation speeds, which are generally at least 150 wpm [8]. The primary reason for this slower speed is the need to correct recognition errors [9]. Experience with ASR appears to enhance performance, but beyond that, we have little data about factors that influence performance for better or worse.

A more significant issue is that the data in **Table 1** represent only a handful of subjects who do not have disabilities. Little performance data exist for users with physical disabilities and unimpaired speech. For measurements of text entry rate and recognition accuracy, two studies involved users of discrete speech recognition systems, the earlier generation of ASR technology. In Kotler and Tam's study, six users of discrete speech recognition had text entry rates ranging from 9 to 15 wpm, with recognition accuracy from 62 to 84 percent [10]. The second study reports on a single well-trained user who achieved 20 wpm [11]. A third study involving ASR users with physical disabilities measured speech production behaviors as an indicator of the physical workload required for ASR use but did not focus on basic speech and accuracy measurements [8]. Given the paucity of performance data, a goal of this study is to help understand the range of productivity that an ASR user can expect, particularly

with newer continuous speech recognition systems, and to begin to address the issues of how to enhance ASR performance.

Satisfaction

Measures of satisfaction attempt to capture how well an individual's ASR system is meeting his or her particular needs. Usage and performance may correlate with satisfaction to some extent, but theoretically, high satisfaction is possible even in situations of low usage and performance, and vice versa. Satisfaction includes satisfaction with the ASR system itself as well as with the services associated with the system, including assessment, training, maintenance, and follow-up [3,12].

The surveys just described in the review of usage also assessed satisfaction [2,3]. In Schwartz and Johnson's survey, 75 percent of the 28 participants rated their ASR system as "good" or better [2]. The remaining participants were no longer using their system. Regarding satisfaction with training, 67 percent thought the 2 days of training they received were "helpful" or better, but 75 percent of participants would have preferred additional training. DeRosier, using the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST) instrument [12], reports that on average, the 10 participants were "more or less satisfied" with their ASR system and "quite satisfied" with the services provided with their system, such as training [3]. When participants were asked what they liked most about ASR, they volunteered the following: (1) ASR provides an opportunity to access the computer, (2) ASR increases efficiency, and (3) ASR provides an alternative to manual input. The aspects that participants disliked most were inconsistent recognition accuracy and technical problems. Goette also discusses satisfaction in her interview study but does not report specific results on the level of satisfaction [4].

These studies suggest that most users are reasonably satisfied with their ASR systems but that there is room for improvement, particularly with respect to recognition accuracy, technical problems, and training. Given that most of these data are from a single clinical site [2], a goal of this study is to gain further insight into the satisfaction of ASR users who have physical disabilities.

Research Goals

This literature review shows that we need to know more about how well ASR is actually meeting the needs of users with physical disabilities, what sorts of barriers exist to user success, and what are the best ways to overcome those barriers. This study seeks to provide a baseline understanding by addressing the following specific aims:

1. Determine who is using speech recognition, how often, and for what kinds of tasks.
2. Determine what nonspeech input methods (if any) ASR users employ, how often they employ them, and for what purpose.
3. Determine the performance and satisfaction of current users of speech recognition systems.
4. Assess the performance costs related to speech recognition errors.
5. Compare usage, performance, and satisfaction for ASR and nonspeech input methods.

METHODS

Overview

For this study, data were typically collected in a single two-part session. (For three users, two sessions were scheduled.) To help us assess user performance, participants used each of two input conditions to perform a prescribed series of computer tasks: the "Speech-Plus" condition involved the use of speech input, and the "No-Speech" condition prohibited the use of speech input. For usage, satisfaction, and other subjective information, participants answered a 53-item survey during an in-person discussion before task performance, as well as a 7-item questionnaire following task performance.

Participants

All 24 participants had physical disabilities that affected their ability to use the standard keyboard and mouse, and all had at least 6 months of ASR experience. Performance data were obtained from 23 participants (one participant chose to discontinue participation in the study following the survey). Of those, 18 could perform the tasks with a nonspeech alternative. Seventeen of these typed directly on the standard keyboard, and one used an on-screen keyboard. The remaining five participants entered text only with speech.

Procedures

Sessions occurred in the participant's home or office, on his or her computer. A researcher verbally asked each

of the 53 items in the survey and recorded the responses on a hard copy. The survey covered the following topics:

1. General information, including demographics, education, and employment status.

2. Background in computer use and information about their computers and computer usage, including applications used and relative time spent on each.

3. Type of ASR system, reasons for using ASR, training received, and likes and dislikes about ASR.

4. Other input methods used, reasons for using them, training received, and likes and dislikes.

5. Relative usage of ASR as compared to other input methods for various computer tasks.

After the survey was completed, the task performance portion of the session began. Six word processing and operating system tasks were defined for each input condition. Two were text entry tasks: transcription of a paragraph from hard copy and a short composition on a supplied topic. The four remaining nontext tasks required opening, saving, and moving files; simple text formatting; and browsing and creating folders. The tasks were identical for each input condition, except that the transcription text and the composition topic were comparable but not the same. The order of input conditions was counterbalanced across participants.

Instructions for each task were presented in hard copy, one task per page. Participants began each task when they were ready and proceeded at their own pace. In the Speech-Plus condition, participants entered text with speech, but they were also allowed to choose nonspeech methods, such as direct control of the mouse to execute commands and make corrections. They were instructed to perform the tasks in the way that they "usually" do.

After completing the tasks, participants completed a seven-item questionnaire regarding their opinions about ASR. Each item was a statement, and participants rated their agreement with the statement on a scale of 1 to 7, with 1 = strongly disagree and 7 = strongly agree. For example, item 1 read, "It is easy to use speech recognition."

Data Collection and Analysis

Survey Data

For survey data, responses to quantitative items were entered into a spreadsheet and mean responses were calculated across participants. For a comparison of responses to different survey items, paired statistics were used: paired *t*-tests for items coded as ordinal variables, and chi-square tests for items coded as categorical variables. Open-ended comments during the survey were also recorded to provide further insight into participant responses.

Task Log Files

For task performance data, the participant's computer screen and speech were recorded on videotape to allow for detailed analysis of user actions. With the use of the videotapes, a time-stamped log of user actions and system responses was produced for each participant. Each user action was coded for the following: (1) whether it involved speech, keyboard, or mouse input; (2) for the Speech-Plus condition, whether the action was part of a correction episode to fix a recognition error; and (3) if so, which method of error correction strategy was used. We counted speech recognition errors manually by comparing the user's spoken input to the system's output for each utterance and entered into the log. If a user committed one or more errors during an action, those were counted and entered into the log for that action as well.

Dependent Variables for Performance

The task log files formed the basis for calculation of the dependent variables for performance:

1. Completion time for each of the six tasks: The times for each user action and system response for a task were summed.

2. Text entry rate: The number of correct characters produced, divided by the number of minutes required to produce them. This measure in units of characters per minute was then divided by 5.5 characters per word to yield text entry rate in words per minute. Note that text entry rate includes time required to correct speech recognition errors and user errors.

3. Recognition accuracy: This applies only to the Speech-Plus input condition. For text entry tasks, this was calculated as the total number of recognition errors for text, divided by the number of words spoken. For spoken commands, this was calculated as the total number of recognition errors for commands, divided by the number of commands spoken. The overall recognition accuracy was the weighted average of these two component measures.

4. Task errors: These are errors made by the user, which distinguishes them from recognition errors made by the ASR system. Task errors cover a whole range of

user errors, such as hitting the wrong key during typing, speaking an incorrect command, or ignoring one of the task instructions. The total number of task errors was summed from the log file. Additionally, the number of net errors remaining, when all tasks were completed, was also counted manually and recorded. Net errors reflect how well the user's final product matched that produced by error-free completion of each task. At task completion, any error of word omission, inclusion, or spelling counted as a single net error, as did any uncorrected formatting or file manipulation error.

5. Correction of recognition errors: This applies only to the Speech-Plus condition. Macros were written to count the total number of times each correction strategy was used, across all six tasks. Additionally, we calculated the amount of time required for each correction episode by the summing the times for each action in that episode.

We used paired *t*-tests to compare performance with and without speech. We computed 95 percent confidence intervals (CIs) for the means of ASR performance variables to create bounded estimates for population means. In the following sections, 95 percent CIs are shown in square brackets throughout when means are presented.

Satisfaction Data

Responses to the postsession questionnaire were entered and averaged. We computed an overall ASR "satisfaction score" for each participant by summing the responses to the six items that related to satisfaction with ASR (after first reversing the sense of responses to negative items), then scaling the sums to a range of 0 to 100. We computed a 95 percent CI for the mean of the satisfaction score to create bounded estimates for the satisfaction score.

SURVEY RESULTS

Participant Characteristics

Participants were recruited through three assistive technology service providers. Individuals were asked to participate if they had been using an ASR system for at least 6 months to accommodate some form of physical disability. The group included 15 men and 9 women, ranging in age from 15 to 59 years, and averaging 36.8 years old. Clinical diagnoses included 11 with cervical spinal cord injuries, 4 with multiple sclerosis or amyotrophic lateral sclerosis, 3 with upper-limb repetitive stress injuries, 2 with cerebral palsy, 2 with muscular dystrophy, 1 with stroke, and 1 with arthrogryposis. Fifteen participants had a bachelor's degree or higher; nine did not. No participant reported difficulty with reading and writing. Fifteen participants were working or going to school at least part-time; nine were doing neither.

Computer Usage

Overall, participants used their computers extensively, averaging about 10 hours a week. Thirteen required use of a computer for their work or school; eleven did not. Perceived need for the computer was high: average agreement to the statement "Successful use of a computer is important to my life goals" was 6.1 [5.7, 6.6][*], on a scale of 1 to 7.

As shown in **Table 2**, when participants used their computer, surfing the World Wide Web was the most popular activity, followed by word processing and email. All participants used all three applications at least some of the time. All other applications, such as finance, games, or graphics, were used much less often.

---

[*]95 percent confidence interval (CI).

Table 2.
Self-reported usage of various computer applications. Each cell shows number of participants (out of 24) who reported using application for given portion of their total computer time.

Usage Time (%)	Browser	Word Processing	Email	Finance	Games	Other
Do Not Use	0	0	0	14	18	11
0–25	9	18	15	8	5	9
26–50	8	3	9	2	1	3
51–75	5	2	0	0	0	1
76–100	2	1	0	0	0	0

Speech Recognition System: Needs, Training, Experience

All participants used ASR to some extent to access their computer. Two used DragonDictate (Dragon Systems, Inc., no longer manufactured), one used VoiceXpress (Lernout & Hauspie, Inc., no longer manufactured), and the remaining twenty-one used a version of Dragon NaturallySpeaking (ScanSoft, Inc., Peabody, MA). Five used tabletop microphones and nineteen used headsets. Six had used some form of ASR for less than 1 year (but more than 6 months), ten for 1 to 3 years, and eight for more than 3 years.

All but two participants received their ASR systems after consultation with a professional experienced in ASR and computer access applications, and their hardware and software configurations were those either provided by the professional at delivery or subsequently upgraded. However, not everyone received extensive training. Ten reported receiving 2 or fewer hours of training; eight had between 2 and 10 hours, and six received more than 10 hours of training from someone experienced in ASR. Despite the variation in amount of training, participants were generally satisfied with their ASR training. The "sufficiency" of ASR training in meeting their needs received an average rating of 5.5 [4.6, 6.3][*] on a scale of 1 to 7.

To determine why participants decided to use ASR, we asked them to rate the importance of eight different reasons, on a scale of 1 to 7, with 7 being highest in importance. The top reason for deciding to use ASR was that it would cause less fatigue or pain than other methods, followed closely by a need for computer input that was easy, fast, and accurate, as well as a perception that the participants' alternatives to ASR were limited (**Table 3**).

Other Input Methods: Needs, Training, Experience

In addition to their ASR systems, most, but not all, participants had other input methods they also used. Of the 24 participants, 19 used some form of nonspeech keyboard access. Fourteen used a standard keyboard, with no adaptations. Access in these cases was most typically with single-finger typing, although two of the participants who had repetitive stress injuries could touch-type with all 10 fingers as pain and fatigue allowed. Four participants used a standard keyboard with typing splints

Table 3.

Importance of eight reasons for using automatic speech recognition (ASR), rated on scale from 1 to 7.

Reason for Using ASR	Mean Importance	95% Confidence Interval
Less Fatigue or Pain	6.2[*]	5.8, 6.6
Ease of Use	5.6[*]	4.9, 6.4
Limited Alternatives	5.5[*]	4.8, 6.3
Speed	5.4[*]	4.7, 6.1
Accuracy	5.3[*]	4.6, 6.0
Recommended for Me	4.9[*]	4.0, 4.8
Personal Preference	4.7	3.8, 5.5
Cool Technology	2.9	2.1, 3.7

[*]Reasons with importance significantly greater than neutral rating of 4.0 ($p < 0.05$).

such as a palmar band, and one used an on-screen keyboard with a head-controlled mouse. Of the 24 participants, 21 had nonspeech access to mouse functions: 7 could use the standard mouse with no adaptations, 13 used a trackball, and 1 used a head-controlled mouse.

All participants had used their other input methods for at least a year. Of the 19 keyboard users, 9 reported receiving specific training on keyboard use and rated its sufficiency 5.1 [3.5, 6.7][†] on a scale of 1 to 7. For mouse access, 10 of the 21 pointing-device users reported receiving specific training, with an average sufficiency rating of 5.2 [3.8, 6.6][†].

Usage of ASR and Nonspeech Input Methods

As noted in the previous section, 21 out of 24 participants had nonspeech access to keyboard or mouse functions in addition to speech recognition. To help us understand how these 21 users with a choice of input methods made their choice, participants assessed the relative amount of time spent with each input method. **Table 4** shows participants' reports of the amount of time they spent using each of their input methods in three different task conditions: across all of their computer tasks, for text input tasks only, and for command input tasks only.

Across all computer tasks, no single input method dominated for participants who had a choice. Of particular interest is that about half reported using their ASR

[*]95 percent CI.

[†]95 percent CI.

Table 4.
Relative usage of automatic speech recognition (ASR), keyboard (KBD), and pointer (PTR) for three categories of computer tasks. Each cell shows number of participants (out of 21) who reported using input device for specified portion of time. Results are reported for those 21 participants who had choice of input methods.

Usage Time (%)	All Computer Tasks			Text Input Tasks			Command Input Tasks		
	ASR	KBD	PTR	ASR	KBD	PTR	ASR	KBD	PTR
Do Not Use	0	2	0	0	2	18	8	8	0
0–25	10	8	7	5	7	3	6	10	2
26–50	3	7	7	5	4	0	4	2	5
51–75	5	3	5	3	3	0	2	1	5
76–100	3	1	2	8	5	0	1	0	9

system for 25 percent or less of their computer tasks. A chi-square test showed no significant difference in the usage pattern for the three input methods ($p = 0.67$).

For text input tasks, reported usage changed significantly from the all-tasks scenario. Usage of ASR as well as keyboards increased, as would be expected, given that they are primarily designed for text input. Of the participants who had a choice, 52 percent used ASR for more than half their text input and 38 percent used keyboards for more than half their text input. A chi-square test comparing ASR and keyboard usage showed no significant difference ($p = 0.53$). Very few users reported using their pointing device to enter text, as would be expected in this population where only one person used an on-screen keyboard.

For command input tasks, pointing device usage dominated (chi-square $p < 0.001$). Of participants with a choice, 38 percent did not use ASR for any commands. Most participants used ASR and the keyboard to enter commands only occasionally, with a few users relying on them heavily. No significant difference emerged between ASR and keyboard usage for commands.

To help understand the relative usage data, we also asked participants about their reasons for choosing nonspeech input methods. They rated the importance of six reasons on a scale of 1 to 7, with 7 being most important. As shown in **Table 5**, the top two reasons were when other input methods were judged to be easier and/or faster. As an example, most participants believed that pointer use was quicker and easier than using spoken commands. Given that the most frequently used application was Web browsing and that Web browsing is dominated by point-and-click operations, the relatively low usage of ASR overall may not be that surprising. A second example cited by several participants is the belief that manually typing "short" text (a sentence or two)

Table 5.
Importance of reasons for choosing nonspeech input methods instead of automatic speech recognition.

Reasons for Choosing Nonspeech Input Method	Mean	95% Confidence Interval
They Are Easier	6.3*	5.7, 6.8
They Are Faster	6.1*	5.5, 6.8
Less Setup Involved	4.6	3.5, 5.7
Frustration With Speech	4.4	3.4, 5.3
To Rest My Voice	2.0	1.2, 2.8
Just For Variety	1.9	1.2, 2.6

*Indicates reasons with importance significantly greater than neutral rating of 4.0 ($p < 0.05$).

would be quicker and easier; these participants reserved ASR for longer text.

The second "tier" of reasons in **Table 5** includes other input methods that often require less-involved setup than ASR. Donning a headset or loading the ASR software into memory (which some participants kept unloaded to preserve system resources) can present a barrier. Participants also went to other input methods when ASR was not working well and caused frustration, either for a particular command or in general. Less important reasons were the need to rest one's voice or to switch methods just for variety.

An important reason that emerged from participant comments was technical incompatibility between ASR and some applications. When ASR works poorly or not at all with a particular application, participants had no choice but to use alternative methods. Of 24 participants, 12 volunteered this reason. Participants had particular problems getting ASR to properly work with their email programs, which is significant because email is one of the three most frequently used applications.

Subjective Opinions About ASR and Nonspeech Input Methods

Participants rated their satisfaction with 11 usability indicators for both ASR and nonspeech alternatives. The results, shown in **Tables 6** and **7**, support the reasons stated in the "Speech Recognition System: Needs, Training, Experience" section, regarding why participants used ASR. ASR succeeds at reducing physical effort, the top reason for using ASR and its most-liked benefit. Eighty-three percent liked the effort required to use ASR, and only seventeen percent found it fatiguing, vocally or otherwise. ASR also provides acceptable speed to most participants: 75 percent liked the speed they achieved. Participants' largest complaint with ASR related to recognition accuracy. Only 54 percent liked the recognition accuracy they achieved, and fixing mistakes ranked as the top dislike at 75 percent. The second-most-frequent dislike was loss of privacy during use of ASR at 37.5 percent of participants.

The pattern of usability ratings shows some notable differences between speech recognition and nonspeech input methods, particularly with regard to fatigue and accuracy. While lower effort is the most-appreciated benefit of ASR, only 32 percent of participants liked the effort involved in their nonspeech input methods. Corroborating this result, 64 percent disliked the fatigue that results from using nonspeech methods. However, more people liked the accuracy of their nonspeech methods (73%) compared to ASR (54%).

To further examine subjective opinion of ASR compared to nonspeech input methods, we asked participants to rate their agreement to several statements regarding learnability, ease of use, reliability, speed, and fun. Agreement was rated on a scale of 1 to 7, from "strongly disagree" to "strongly agree." **Table 8** shows the results for each category of input method.

Although neither was judged especially difficult to learn, speech recognition rated significantly more difficult to learn than other input methods. ASR was also judged to have less consistent accuracy compared to other nonspeech input methods. The source of inconsistent recognition accuracy with ASR was unclear to most participants. The inconsistency was attributed to "elves" in the system ("some days it's fine; others it's not"), the application being used, having a cold, thinking about what to say, and forgetting to dictate clearly. Inconsistency with other input methods was easier to pinpoint; most who reported inconsistency said that it came with physical fatigue.

Table 6.
Percentage of participants who liked particular aspects of their automatic speech recognition (ASR) system and nonspeech input methods.

Likes	% Responding Yes	
	ASR	Nonspeech
Effort	83.3	31.8
Speed	75.0	68.2
Ease	75.0	72.7
Fun	66.7	50.0
Accuracy	54.2	72.7
Cool	37.5	23.8

Table 7.
Percentage of participants who disliked particular aspects of their automatic speech recognition (ASR) system and nonspeech input methods.

Dislikes	% Responding Yes	
	ASR	Nonspeech
Fixing Mistakes	75.0	—
Privacy	37.5	—
Setup	20.8	22.7
Disturbs Others	20.8	—
Fatigue	16.7	63.6
Too Much Thinking	16.7	0.0

Summary of Survey Results

Web browsing was the most frequent computer task for these participants, followed by email and word processing. The top reason for using ASR, and its most appreciated benefit, was to reduce fatigue and pain associated with manual input methods. Speed was secondary, but still important, and 75 percent of users liked the speed they were able to achieve with ASR. Users' main concern with ASR was the inconsistent recognition accuracy and the need to fix recognition mistakes. Secondary concerns relate to technical glitches, loss of privacy, and disturbance to others.

Almost all the participants had a nonspeech input method for text entry (19 of 24 participants) and pointing (21 of 24 participants). These individuals chose to use ASR primarily for text input tasks, in which keyboard and ASR input was used with roughly the same frequency, rather than command input tasks, in which manual pointing was the dominant method.

Table 8.
Participants' level of agreement to statements about their automatic speech recognition (ASR) systems and nonspeech input methods. Ratings are on scale of 1 to 7, from "strongly disagree" to "strongly agree," with 4.0 representing neutral rating.

Statement	ASR		Nonspeech Input Methods	
	Mean	**95% CI**	**Mean**	**95% CI**
It was difficult to learn to use [this system].	3.3	2.5, 4.2	1.8[*]	1.1, 2.4
I expected using [this system] to be easier than it actually is.	3.8	3.0, 4.6	2.1[*]	1.5, 2.7
My accuracy with [this system] seems to get worse at times.	4.6	3.7, 5.5	2.7[*]	1.9, 3.5
I expected using [this system] to be faster than it actually is.	3.6	2.8, 4.5	2.4[*]	1.7, 3.1
I enjoy using [this system] to access my computer.	5.3	4.6, 6.0	4.2	3.6, 4.9

[*]Significant difference between ASR and nonspeech input methods ($p < 0.05$).   CI = confidence interval

TASK PERFORMANCE RESULTS

Of the 24 survey participants, 23 completed the 6 tasks used to measure performance with ASR. The results follow.

Comparison of Performance With and Without ASR
Performance data both with and without ASR are available for the 18 participants who had nonspeech access to a keyboard and mouse. Speed and accuracy of task performance are presented in the following.

Speed With and Without ASR
The overall time required to perform the tasks averaged 28 percent slower in the Speech-Plus condition as compared to No-Speech, but this difference was not statistically significant. For tasks that did not involve text entry, involving formatting, saving, and copying, a large and consistent difference was seen in performance with and without ASR. Participants averaged 61 percent slower on these nontext tasks in the Speech-Plus condition ([25.3, 96.6][*], $p < 0.05$).

When we considered text entry tasks alone, no significant difference existed between text entry rates with and without ASR across all participants. Text entry rate with ASR averaged 16.9 wpm [13.5, 20.3][*]. In the No-Speech condition, text entry rate averaged 15.0 wpm [9.5, 20.4][*]. For 11 of the 18 participants, use of ASR provided higher text entry rates. These participants enjoyed an average enhancement of 108.9 percent, or 8.9 wpm. The remaining seven had slower text entry rates with speech compared to without speech, averaging 24.9 percent, or 5.8 wpm, slower.

Participants who enjoyed enhanced rate with ASR compared to their nonspeech input method tended to be those whose text entry rate without ASR was relatively low. A rough cutoff point seems to be a nonspeech typing speed of 15 wpm. Of the 11 participants who typed slower than 15 wpm, 10 enjoyed substantial gains in text entry rate when using speech input. Similarly, of the seven participants who typed faster than 15 wpm, only one achieved any gain in text entry rate when using speech input. **Figure 1** shows the relationship between the enhancement in text entry rate with ASR and the nonspeech text entry rate.

Task Accuracy With and Without ASR
With respect to errors committed during performance of the task, participants made significantly fewer errors during the Speech-Plus condition, averaging 34.4 percent, or 5.7 fewer errors than in No-Speech ($p < 0.05$). However, when all tasks were completed, those performed during the Speech-Plus condition had significantly more errors remaining ($p < 0.05$). Of the 18 participants, 12 left two or more errors in the Speech-Plus condition; only 6 of 18 met this standard in the No-Speech condition. Participants fixed significantly more errors in the No-Speech condition, an average of 84 percent, compared to 58 percent of errors successfully fixed in the Speech-Plus condition ($p < 0.05$). In other words, participants made more errors using their No-Speech methods, particularly typographical errors on the keyboard, but they fixed almost all of them. (In both input conditions, participants typically identified and attempted to fix all errors, but in the Speech-Plus condition, these attempts were not always successful.) The next section focuses on performance in the Speech-Plus condition in greater detail.
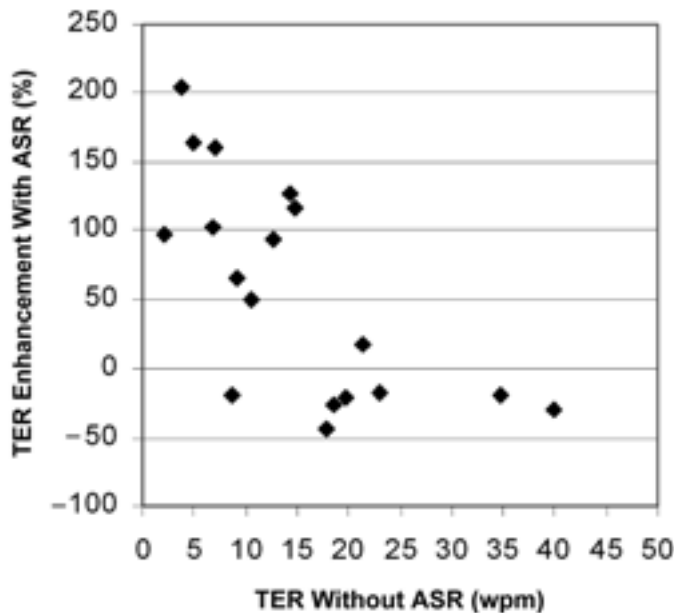
[*]95 percent CI.

Figure 1.
Improvement in text entry rate (TER) provided by automatic speech recognition (ASR), relative to nonspeech text input with Scratch-That technique, as function of TER without ASR. Graph shows that slower nonspeech typists tend to enjoy significant improvement in TER when using speech (wpm = words per minute).

Performance With ASR

Text Entry Rate and Recognition Accuracy

Speech recognition performance varied widely between participants, even though all were long-term users who had unimpaired speech. Text entry rate with speech ranged from 3.5 to 32.2 wpm, with an average of 16.9 wpm [13.5, 20.3]*. Recognition accuracy for text ranged from 72.1 to 93.8 percent, averaging 85.0 percent [82.2, 87.9]*. Recognition accuracy for commands ranged from 63 to 100 percent, averaging 87.2 percent [82.5, 91.9]*.

Figure 2 shows the relationship between text entry rate and recognition accuracy for text. Higher accuracy tends to be associated with higher text entry rate, but the correlation, at 0.62, is not as strong as might be expected. **Figure 2** also illustrates the wide range of observed performance. The data points can be grouped into three clusters: Cluster A, with six participants, had the best overall performance, with text entry rates at 25 wpm or above.
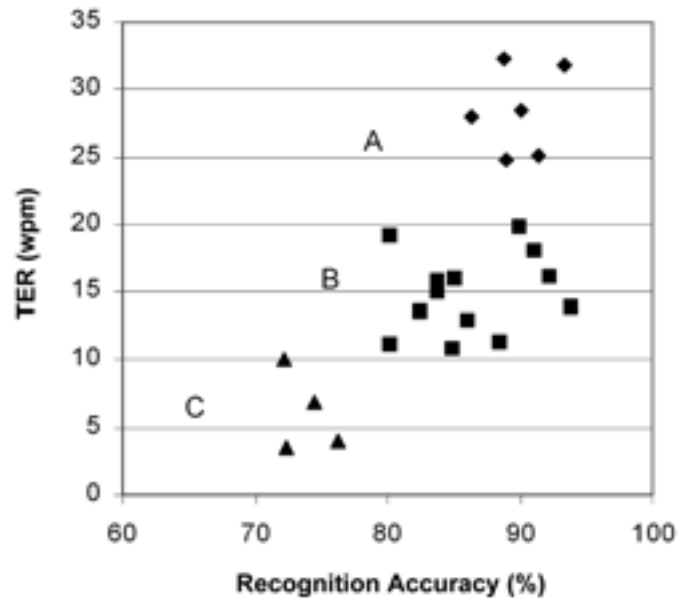
---

*95 percent CI.

Figure 2.
Text entry rate (TER) with automatic speech recognition (ASR) as function of recognition accuracy. Wide range of participant performance is shown, as well as positive correlation between recognition accuracy and TER (wpm = words per minute). Cluster A, 6 participants, had TERs at 25 wpm or above. Cluster B, 13 participants, had TERs between 10 and 20 wpm; Cluster C, 4 participants, had TERs at 10 wpm or below.

Cluster B, with 13 participants, achieved text entry rates between 10 and 20 wpm, while the 4 participants in Cluster C entered text at 10 wpm or below.

Handling Recognition Errors

Participants spent substantial time fixing ASR recognition errors. On average, 56 percent of the text entry time in the Speech-Plus condition was directly involved in correcting recognition errors. Correcting each recognition error required an average of 23 seconds and 1.8 attempts. Participants used four possible methods to correct recognition errors:

1. Scratch-That. Use a "Scratch That" or "Undo" command to erase the immediately previous utterance. Then redictate the utterance.

2. Select-Redictate. Select the erroneous word or phrase by a voice command (e.g., "Select 'license that'"). Then redictate the correct word or phrase in its place (e.g., "eyes and sat").

3. Correction-Dialog. Select the erroneous word or phrase by a voice command and open the error correction

dialog (e.g., "Correct 'license that'"). Then choose from one of the following correction methods:

a. Selecting the correction from a pick list of alternate recognitions.

b. Spelling the word by voice.

c. Spelling the word by another input method such as the keyboard.

For either of the spelling correction methods, the pick-list of alternates will change as the user types in letters for the word. If the desired word or phrase appears in the numbered pick list, it can be selected at any time with the appropriate voice command (e.g., "Choose 3").

4. Direct-Edit. Fix the error directly by using another input method such as the keyboard.

The Correction-Dialog strategy was used most often, in an average of 37 percent of correction episodes across all participants, followed by Select-Redictate (25.3%), Scratch-That (20.9%), and Direct-edit (17.0%). That Correction-Dialog was most often used is encouraging, because it is generally the most appropriate strategy to use. However, the Scratch-That strategy was also frequently used to correct recognition errors, although it is intended primarily to correct dictation errors (e.g., when the user misspeaks or coughs). The problem with using Scratch-That to correct recognition errors is that it limits the ASR system's capability to learn from its recognition mistakes. And indeed, overuse of the Scratch-That strategy is associated with degraded text entry performance, as shown in **Figure 3**. Those who used Scratch-That for more than 25 percent of their correction episodes had lower text entry rates.

Summary of Task Performance Results

In comparing participants' performance in the Speech-Plus and No-Speech conditions, the following key results were observed:

1. No significant difference resulted in overall task completion time.

2. Tasks that did not involve text entry were significantly slower with speech than without.

3. No significant difference resulted in text entry rate. Some participants enjoyed substantial rate enhancement with speech; some (particularly those who typed faster than 15 wpm) did not.

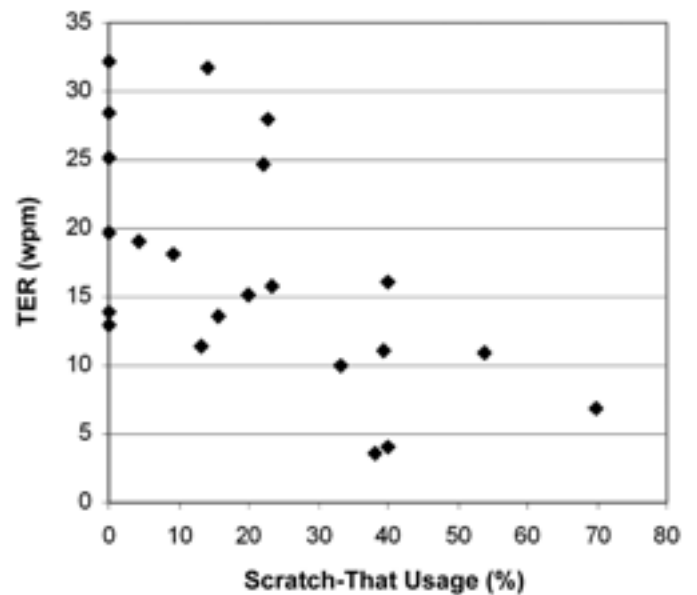4. Task accuracy on completion was significantly lower with speech than without.

Figure 3.
Text entry rate (TER) with automatic speech recognition (ASR) as function of Scratch-That usage. Participants who use Scratch-That strategy had slower TERs (wpm = words per minute).

Looking at ASR performance alone, we observed the following key results:

1. Average text entry rate was 16.9 wpm [13.5, 20.3][*], range 3.5 to 32.2 wpm.

2. Average recognition accuracy was 85.0 percent [82.2, 87.9][*], range 72 to 94 percent.

3. The majority of text entry time with ASR was spent correcting recognition errors.

4. Users employ the Correction-Dialog method most frequently, which is appropriate, but the Scratch-That technique is overused.

POSTSESSION QUESTIONNAIRE RESULTS

Table 9 summarizes participants' level of agreement with the seven statements in the post-task questionnaire. Results showed high agreement that ASR provides the fastest text input method and is easy to use. Participants also had significantly greater than neutral agreement that ASR provides more accurate text entry than any other

---

[*]95 percent CI.

Table 9.
Ratings of agreement to six statements in the posttask questionnaire on automatic speech recognition (ASR) use on scale of 1 to 7, from "strongly disagree" to "strongly agree."

Statement	Level of Agreement	
	Mean	95% CI
I can enter text more quickly with speech recognition than with any other method.	6.0*	5.2, 6.7
It is easy to use speech recognition.	5.4*	4.8, 6.0
I can enter text more accurately with speech recognition than with any other method.	5.1*	4.4, 5.8
Using speech recognition can be a frustrating experience.	4.8	3.9, 5.6
The system correctly recognizes almost everything I say.	4.5*	4.0, 5.1
It is difficult to correct errors made by the speech recognition system.	3.5	2.7, 4.2
I was tired by the end of the session.	2.7	1.7, 3.7

*Agreement significantly greater than neutral rating of 4.0 ($p < 0.05$).
CI = confidence interval

method and that their ASR systems correctly recognize almost everything they say.

When scores were combined for the first six statements into an overall satisfaction score, the average was 63.1 [55.4, 70.8]*, on a scale of 0 to 100. A score of 100 represents a strong agreement with all positive statements and a strong disagreement with all negative statements, while a score of 50 represents a neutral response to all statements. Thus, this average is significantly more positive than neutral, but not overwhelmingly positive.

As with the task performance data, a wide range of satisfaction scores resulted—from a low of 27.8 to a high of 88.9. To uncover a sense of what accounts for this variation, we found that text entry rate with ASR has no relationship with the satisfaction score ($R^2 = 0.002$). However, those participants with higher recognition accuracy tend to have higher satisfaction scores ($R^2 = 0.23$, $p = 0.027$).

DISCUSSION

The results reported in this study cover a broad range of objective and subjective aspects of speech recognition use. The following discussion attempts to synthesize the results across the domains of usage, performance, and satisfaction and to interpret their clinical significance. Overall, the results suggest that use of ASR provides some consistent benefits but is also associated with some

significant limitations. These are discussed in more detail in the following paragraphs.

Primary Benefits of ASR

For the participants in this study, use of ASR has a meaningful role in attaining comfortable and efficient computer access. ASR appears to be particularly successful as a means of reducing the pain and fatigue that can be associated with manual input methods. Participants typically used ASR as a complement to, rather than a total substitution for, nonspeech input methods. These 24 users employed ASR primarily when entering text, and it provided sufficient text entry rate for that purpose, although it was not always faster than users' manual input methods. The results suggest that users who type faster than 15 wpm are less likely to enjoy a speed enhancement using ASR, although they may still enjoy the fatigue- and pain-reduction benefits.

Participants in this study enjoyed using ASR, for the most part, and their satisfaction with it was statistically independent of the text entry rate that it provides. The moderate satisfaction observed here is higher than has been reported for novice users without physical disabilities, who were highly dissatisfied [9], even though their absolute performance was quite similar to that achieved by these users. This finding tends to support the suggestion that users who have physical disabilities may be more tolerant of ASR's shortcomings [9]. Previous findings of moderately high satisfaction among ASR users with physical disabilities corroborate this as well, at least among those users who stick with ASR [2,3].

---

*95 percent CI.

Limitations Associated With ASR

The benefits provided by ASR often come with some limitations as well. These limitations are not necessarily inherent in the technology itself but may also relate to limits in users' understanding of how to effectively apply ASR to meet their needs.

Limitations Related to Recognition Accuracy

The major theme in ASR limitations is recognition accuracy. Participants cited inconsistent recognition accuracy as their primary dislike, consistent with DeRosier's study [3]. As measured in the performance of text entry tasks, the average recognition accuracy was poorer than might be expected, averaging only 85 percent compared to the 92 percent reported previously for four experienced users without disabilities [6]. Indeed, only 6 of 23 users exceeded 90 percent recognition accuracy.

Recognition errors can have serious consequences for performance time, user satisfaction, and net accuracy of the "finished product." For these users, fixing recognition errors consumed 56 percent of the text entry task time. This is a primary reason why participants' text entry rate was good, but not great, averaging 16.9 wpm. Only the six participants in Cluster A entered text as quickly as previously reported for long-term users without disabilities [6,7]. However, this level of 25 wpm and above reported in the Karat et al. studies may overestimate the "typical" performance that is achieved with experience, because it is based primarily on data from four professional researchers in speech recognition [6,7].

With respect to user satisfaction, those with lower recognition accuracy tended to be less satisfied with ASR. We attribute this to users' dislike for the disruption of fixing recognition errors, rather than their effect on text entry rate, since text entry rate by itself had no statistical relationship to satisfaction in this study.

Finally, recognition errors with ASR resulted in a finished product that was less accurate than participants produced in the No-Speech condition. One reason for this is that errors produced with speech recognition may be harder for users to detect than errors produced with manual methods such as typing [7]. Users cannot readily "feel" when an error has been committed and so must rely heavily on proofreading to check accuracy. An additional source for the less accurate final output is occasional difficulty in correcting errors that were detected. Participants occasionally noticed an error, attempted to correct it by voice, then either did not ensure that it was

fixed correctly or simply gave up if the attempt was not successful.

Limitations Related to Task-Technology Fit

Task-technology fit refers to the degree of compatibility between the requirements of the task and the capabilities of a technology to help complete that task [4]. In this case, the "task" is not just one task; rather, "task" refers to the whole range of activities involved in the use of a personal computer. For these participants, Web browsing, using email, and word processing were the primary applications, with their associated subtasks of executing commands and entering text. While the ASR systems used by these participants are designed to support command execution as well as text entry, only 14 percent of participants reported using ASR for the majority of their command inputs. The dominant use of ASR by far was for text entry, primarily within a word processor. Several (at least five) participants volunteered that they use ASR only for longer text.

The limited use of ASR for commands means that although Web browsing is the most-used application, very few participants used ASR regularly for Web browsing, citing either technical difficulties or preference for manual methods. Is this a self-limitation on the application of ASR or is it a function of a true problem with the technology? This study did not specifically ask participants why they did or did not use ASR for Web browsing, but participant comments provide some hints. For those who do not use ASR with Web browsing because they never tried it, their assumption was that their manual method for command input provided better performance than ASR. Future studies would be necessary for one to determine exactly when that assumption is correct, but anecdotally, many of these participants had good control over their pointing device and had no real reason to seek an alternative. Four participants reported technical incompatibilities between their browser and ASR system, and six stated that their manual methods were faster and easier for them. Also possible is that the preference for manual command input may relate to cognitive issues about remembering the syntax and content of the spoken commands.

Striking Variability of Results

Many of the major dependent variables measured in this study showed a wide variation between participants. ASR usage for all computer tasks ranged from less than

25 percent to more than 75 percent of tasks. Text entry rate with ASR ranged from 3.5 to 32.2 wpm, and recognition accuracy ranged from 72 percent to 94 percent. Satisfaction scores ranged from 28 to 89, on a scale of 0 to 100.

What accounts for this wide range? The influence of factors such as school/employment status, amount of ASR training, nonspeech typing ability, and ASR correction techniques will be assessed in the next phase of this work. A preliminary insight related to usage is that those who use ASR for most of their computer tasks tend to be those with a compelling physical reason to do so, either because manual typing is so slow (around 5 wpm) or painful. This appears to be a more important factor for usage than the performance that ASR provides. A similar principle may hold for satisfaction scores. The results just mentioned show that satisfaction does not depend on absolute speed achieved with ASR, although recognition accuracy does have a positive relationship with satisfaction. In addition to recognition accuracy, a greater physical need for an alternative to manual input may increase satisfaction with ASR. These hypotheses need further exploration.

For performance measures, the correction strategies employed by users appear to have a strong influence on their resulting performance [7,9]. Fixing ASR errors with the system's correction dialogue allows the system to improve its voice model, while use of the general-purpose Scratch-That command actually can degrade the model. Our results support this, because the higher-performing Cluster A participants used Scratch-That only about 10 percent of the time, while Cluster C participants used it almost half the time. Enforcing appropriate correction strategies through clinical interventions is likely to yield enhanced performance with ASR. The clinicians in this study all report coaching users on the correction dialogue, and their efforts appear to have had some effect. Use of the correction dialogue is the modal strategy for these users, which is a sharp improvement to Halverson et al.'s new untutored users, who used the correction dialogue only 8 percent of the time [9]. However, significant room for improvement still exists.

In addition to use of appropriate correction strategies, early impressions are that those who enjoyed the best performance with ASR (Cluster A) tend to have the fastest manual typing speed, and the fastest hardware, the highest level of formal education, and the highest vocational or educational need for productivity. They are not necessarily those with the most formal training in the use of ASR.

Limitations of the Study

The study presented here has a number of strengths: employment of current and experienced users of ASR who have physical disabilities, sufficient number of participants to draw some statistical conclusions, collection of subjective as well as objective data, careful design of experimental tasks, and detailed analysis of user actions and their associated times. However, its design has some limitations as well. Relative to the qualitative data, much of our data about usage rely on self-reported information about how often ASR was used as compared to other input methods. This comparison is not something that users tend to consciously think about, so estimating this in a precise way may not be an easy and reliable task. However, users appeared to be able to reconstruct their usage patterns and confidently estimate the nearest quartile of relative use. For the satisfaction data, the score used was a combination of survey item responses. Use of an established satisfaction instrument may have been better, such as the QUEST, either instead of or in addition to the survey items [12]. These "homemade" scores reflect several constructs represented in the QUEST, particularly those related to the system (e.g., simplicity of use, comfort, and effectiveness), but does not incorporate satisfaction with ASR-related services. The main reason for constructing a custom survey just for this study was to address a wide range of issues that is unique to ASR and its typical context of use.

Regarding the user pool, the inclusion criteria were relatively broad: anyone with unimpaired speech who is currently using ASR to accommodate a physical disability and has used their ASR system for at least 6 months. These criteria have the advantage of yielding a fairly wide cross-section of users, which is appropriate for an initial baseline study, but it may reduce our ability to draw conclusions for specific sets of user conditions, whether for a particular disability diagnosis, specific ASR system, specific ASR application, etc. An advantage to our broad criteria is that we were able to gather 24 relatively local participants, although additional participants would have been certainly welcome.

One should also note that, consistent with the research goal of determining a baseline for typical ASR outcomes, the conditions of ASR use in this study were taken "as is" for each participant and were not modified in any way prior to data collection. While the hardware and software configurations used by 22 of the 24 participants were provided by experienced assistive technology

professionals, they were not necessarily the best possible, and we may have been able to improve some individuals' performance by upgrading or reconfiguring their hardware or software. This means that these results may best be generalized to the performance of individuals who receive ASR-related services (including assessment, system configuration, training, and follow-up) that are appropriate, but not necessarily optimal.

The specific tasks participants were asked to do were artificial, in order to provide a consistent task set that would allow performance to be pooled across participants. We attempted to define representative tasks that would be familiar and required for any personal computer user, but any artificial task set has some differences relative to a user's real-life tasks. In particular, the performance results here do not reflect the composition of a meaningful or complex text. A text composition task was included in the protocol, but the topics were so trivial (e.g., what is your favorite food?) that the task probably did not represent true composition. Therefore, the text entry rates for text reported here probably best reflect transcription or simple composition. Data from other studies suggest that true composition would be likely to be slower both with and without speech [7,9].

Future Work

Immediate future work to be performed on this data set is to identify the influential factors for performance and satisfaction. If we can better understand why certain users had better ASR outcomes than others, we may be able to improve future ASR interventions to provide better average outcomes to everyone.

An important area of future work is to add to this initial data set of 24 users. Additional data would provide added confidence in these baseline measures and also enhance our ability to identify influential factors in ASR outcomes.

An unanswered question in this work relates to the user's choice of input methods. Study participants who could choose between speech and nonspeech input certainly exercised that choice. But we do not know much about how people choose between input methods, and whether they do so optimally. Studies in other areas of human-computer interaction suggest that people do not always make optimal decisions about how to do particular tasks [13]. If we knew more about what the optimal choice was in this case and how users' decisions compare to optimal, we might be better able to help users make more efficient choices.

Improved ASR interventions will require a clinical study evaluating the success of those interventions. If we believe that certain clinical practices will yield better results (e.g., more intense coaching on appropriate correction strategies), we need to test this hypothesis.

CONCLUSIONS

While many questions regarding effective use of speech recognition systems remain, results from this study can be used to inform clinical ASR interventions in several ways. Some specific insights are strongly supported by the data:

1. Users and practitioners should have realistic performance expectations. The data suggest that an "average" ASR user will enter text at 17 wpm, with recognition accuracy of 85 percent. A high-performing user may achieve approximately 30 wpm. This rate may be considerably below the performance level that a new user may expect.

2. Practitioners should coach appropriate correction strategies. When fixing recognition errors, using the appropriate strategy increases performance and, ultimately, user satisfaction. Teach users to employ the correction dialogue for almost every recognition error and to avoid using Scratch-That unless they misspeak.

3. Users and practitioners should remember that small increases in recognition accuracy can lead to big gains in text entry rate. Each percentage point increase may improve text entry rate by almost 1 wpm.

4. Practitioners should recognize that ASR users who have nonspeech input methods will frequently use them instead of, or in conjunction with, the use of speech. Acknowledge this up front, and help users determine how to best combine their input methods to meet their needs.

5. Practitioners should measure users' outcomes regularly. Simple measurements of usage, satisfaction, speed, and accuracy can be very valuable in determining whether an intervention is meeting expectations. The QuickMAP procedure is a primarily paper-and-pencil procedure that can be used to measure speed and accuracy with ASR within clinical settings [14].

Clearly, we have a long way to go before we have a thorough understanding of how well ASR meets the needs of people with physical disabilities, but this study takes an initial step toward forming a complete, evidence-based set of best practice guidelines that will lead to improved application of ASR.

REFERENCES

1. Koester HH. User performance with speech recognition: a literature review. Assist. Technol. 2001;13(2):116–30.
2. Schwartz P, Johnson J. The effectiveness of speech recognition technology. Proceedings of RESNA '99 Conference; 1999; Long Beach (CA). Washington (DC): RESNA Press; 1999. p. 77–79.
3. DeRosier R. Speech recognition software as an assistive device: a study of user satisfaction and psychosocial impact [M.S. Thesis]. Philadelphia (TX): Temple University; 2002.
4. Goette T. Factors leading to the successful use of voice recognition technology. Proceedings of the 3rd International ACM Conference on Assistive Technologies; 1998; Marina del Rey (CA). New York: ACM Press; 1998. p. 189–96.
5. Devine EG, Gaehde SA, Curtis AC. Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. J Am Med Inform Assoc. 2000;7:462–68.
6. Karat J, Horn DB, Halverson CA, Karat C. Overcoming unusability: Developing efficient strategies in speech recognition systems. Poster at Computer-Human Interaction 2000, ACM Conference on Human Factors in Computer Systems. The Hague, Netherlands; 2000 Apr 1–4.
7. Karat C, Halverson CA, Horn DB, Karat J. Patterns of entry and correction in large vocabulary continuous speech recognition systems. Proceedings of the CHI '99 Conference; 1999; Boston (MA). New York: ACM Press; 1999. p. 568–74.
8. Rieger JM. The effect of automatic speech recognition systems on speaking workload and task efficiency. Disabil Rehabil. 2003;25:224–35.
9. Halverson CA, Horn DB, Karat C, Karat J. The beauty of errors: Patterns of error correction in desktop speech systems. Proceedings of Human-Computer Interaction, INTERACT '99. Burke (VA): IOS Press; 1999. p. 133–40.
10. Kotler A, Tam C. Effectiveness of using discrete utterance speech recognition software. Augmentative Altern Commun. 2002;18(3):137–46.
11. Dalton JR, Peterson CQ. The use of voice recognition as a control interface for word processing. Occup Ther Health Care. 1997;11:75–81.
12. Demers L, Lambrou RW, Ska B. Item analysis of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). Assist Technol. 2000;12(2):96–105.
13. Olson J, Nilsen E. Analysis of the cognition involved in spreadsheet software interaction. Int J Hum Comput Interact. 1988;5:221–65.
14. Koester HH. A method for measuring client performance with speech recognition. Proceedings of RESNA 2002 Conference; 2002; Minneapolis (MN). Washington (DC): RESNA Press; 2002. p. 115–17.