Journal of Construction Education
Spring 2001, Vol. 6, No. 1, pp. 34-42
Copyright 2001 by the Associated Schools of Construction
1522-8150/01/$3.00/General Manuscript

An Initial Look at Voice Recognition

Steve Williams
Department of Building Science
Auburn University

The author looked into the potential of voice recognition technology as a possible alternative to the traditional keyboard. His experiences with this new technology, along with the determination of baseline typing speed for construction students, are presented. In addition, the potential for the use of this technology in the construction industry, and the strong need for further study is discussed.

Key Words: Voice Recognition, Speech Recognition, Information Technology, Construction

Introduction

Periodically, departmental faculty members meet with representatives from construction companies to review issues that are of interest to the industry. Historically, these meetings have been interesting and informative, and have helped faculty keep in touch with problems facing the industry.

In August of 1996, a small group of faculty members had a series of meetings with contractors in the region. A recurring theme in these discussions involved problems contractors had regarding communications from the field to the home office. Almost all the contractors mentioned problems involving daily reports and other forms of periodic communications.

Another problem frequently mentioned involved the continually increasing paperwork required of project managers and construction executives. The contractors discussed the importance of written communication skills in construction students, and emphasized the surprisingly large percentage of their workday that required written correspondence.

After the meetings with the contractors, the faculty members met to summarize the meetings, and to decide what course of action (if any) needed to be taken by the department. The general consensus of the faculty members was that voice recognition software could potentially alleviate the two problems mentioned by the contractors.

Literature Research

A literature search indicated promising developments in speech recognition. The United States Government has invested heavily in this technology (Weinstein, 1995). Several industry analysts along with the heads of large corporations addressed voice recognition in editorials and reviews.

Greater affordability, along with enthusiastic industry and governmental commitment, were common threads in the discussions and articles.

There are occasional references in the literature to some of the problems that exist with current voice technology (Mehm, 1996), and one author that predicts that progress will be "painfully slow" (Levinson, 1995), but the overwhelming majority of articles exhibit an aura of optimism and commitment A sampling is listed below. (Machrone, 1996 and Mayor, 1996)

The dream of being able to talk to your computer and have it respond intelligently is an old one. But for years, the technology was relegated to the high end of the marketplace, with PC applications initially nonexistent and then too expensive and bulky to capture much of the business market. Now an improved breed of out-of-the-box, PC based software is attempting once again to seed corporate offices with the voice commanded PC. (Mayor, 1996)

Industry analyses are equally enthusiastic regarding the future of voice systems. Some of the major impediments of the past, such as RAM availability and processor speed, we rapidly falling away, making what some refer to as the voice typewriter a near-term reality (Machrone, 1996).

> “In at least the 10-year time scale, that there will be widespread voice recognition. You will talk to your computer. And once we can do that, it won't be tong before your computer will talk back to you with the information you want.” Gordon Moore, -- author of Moore's Law and co-bounder of Fairchild Semiconductor and Intel. (Schmit, 1996)
>
> Q: So in five to 10 years, what kind of company will Microsoft be?
>
> A: "...In 10 years, a lot of that will be speech recognition, speech synthesis and vision. We our second-biggest business and probably still will be. "Bill Gates, founder and CEO of the Microsoft Corporation (Maney, 1996)

Selection of Voice Recognition Software

As a result of a literature search and subsequent discussion among faculty, a decision was made by the department head to fund the purchase of voice recognition software. A faculty member would be charged with looming the software and reporting whether further investigation and investment of resources was warranted.

The software chosen for the project was DragoaDictate for Windows 2.5 - Classic Edition. The cost of the software (for use with a broad range of applications) was approximately 500 dollars, but the same company is now offering application specific software for as little as 100 dollars. (Vardlamudi, 1996)

The software is based on discrete dictation, which requires the user in insert a slight pause between dictated words. Speech is analyzed using both an acoustic mode (where a word is compared with speech samples from thousands of other people) and a contextual mode where the software attempts to place the next word in proper context. Continuous speech recognition

products (which recognize normal speech patterns, and require no pauses) are becoming available, but are generally regarded to lack maturity, and are usually applicable to a very specific profession, such as radiology. (Schwartz, 1996) Several recent comparisons of voice recognition software me available in the literature. (Mehra, 1996)

Informal Interviews

As software was being evaluated and purchased, the author conducted several informal interviews with people in various occupations about the importance/relevance of direct electronic dictation. Each person was asked what he or she was most interested in teaming about electronic dictation. In the course of the discussions, several people voiced some form of one or both of the following questions:

1. Is electronic dictation more efficient than typing?
2. How many hours of warning are required to achieve competency?

One of the individuals interviewed opinion that "people will use electronic dictation only if it is faster and easier than the form of data input that they currently use." Obviously, voice input is of high interest in a person suffering from cereal tunnel syndrome or some other disability involving the hands or arms, while a person capable of typing 75 words per minute may be content with the traditional keyboard. It was noteworthy that most of the people interviewed were not satisfied with their keyboarding skills, and seemed very interested in other forms of input.

Initial Evaluation Parameters

Although a statistically significant study was not considered, certain anecdotal information was gathered to serve as a basis for decisions regarding further study/investment into construction applications of voice recognition technology.

The determination of the average typing speed of construction students is obviously an important variable in any discussion regarding the merits of voice recognition systems in construction. Additionally, a reporting of the experiences of an individual in teaming and using the technology would be of interest, and would serve as background for future decisions. Similarly, a compilation of industry reviews and commentary would also bolster future decisions.

Typing

Thirty construction students were asked in participate in a five minute typing test. The text for the rest came from a commonly used college textbook on reinforced concrete. Students used Microsoft Word for the word processing software. The students were instructed to correct errors as they trade them (although they did not always do so), and to try to hand in a perfect paper. The

students were in their third or fourth year of college. The results of the typing test are compiled in Table 1 and Table 2.

Table 1

Typing Data

Student Number	Gender	Year	Instructor	Time (in)	# of Words	# of Errors	WPM
1	M	3	Love	5	116	2	22.8
2	M	4	Love	5	90	0	18
3	M	3	Love	5	122	0	24.4
4	F	3	Love	5	92	0	18.4
5	F	4	Love	5	82	0	16.4
6	M	4	Love	5	103	0	20.6
7	M	4	Love	5	109	0	21.8
8	M	4	Love	5	102	0	20.4
9	M	3	Love	5	121	0	24.2
10	M	4	Love	5	95	0	17
11	M	4	Love	5	82	0	16.4
12	M	4	Love	5	77	1	15.2
13	M	4	Williams	5	79	1	15.6
14	M	4	Williams	5	100	0	20
15	M	4	Williams	5	100	0	20
16	M	4	Williams	5	84	0	16.8
17	M	3	Williams	5	78	0	15.6
18	M	4	Williams	5	105	1	20.8
19	M	4	Williams	5	117	0	23.4
20	F	6	Corely	5	105	0	21
21	M	3	Hein	5	84	0	16.8
22	M	3	Hein	5	135	4	26.2
23	M	4	Hein	5	154	3	30.2
24	M	4	Hein	5	98	1	19.4
25	M	4	Hein	5	83	0	t6.6
26	M	4	Hein	5	128	1	25.4
27	M	4	Hein	5	103	1	20.4
28	M	3	Hein	5	178	0	3 5

Table 2

Typing Test Summary

Average Words Per Minute	20.27
Mean Words Per Minute	20.00
Minimum Words Per Minute	13.20
Maximum Words Per Minute	35.60
Standard Deviation	4.83

Voice Dictation

After determining average typing speeds, the author set up a dictation test using the same text the students used. At this time, the author had approximately twenty hours of practice with electronic dictation. The author had never "practiced" on this particular text in advance. The author

intended to correct all errors as they occurred, but later analysis of the test yielded 2 to 4 errors in each trial. The dictation test was conducted in a quiet classroom. Since one of advertised feature of voice recognition software is its capacity to "learn," the author repeated the test five times to see if a teaming trend was evident. The results are reported in Table 3, and a comparison of voice and keyboard speeds is shown in Figure 1.

Table 3

Voice Data

Trial #	Time (in)	# of Words	# of Errors	Dictated Words per Minute
1	5	76	2	15.2
2	5	134	2	26.8
3	5	152	4	30.4
4	5	165	3	33
5	5	182	2	36A

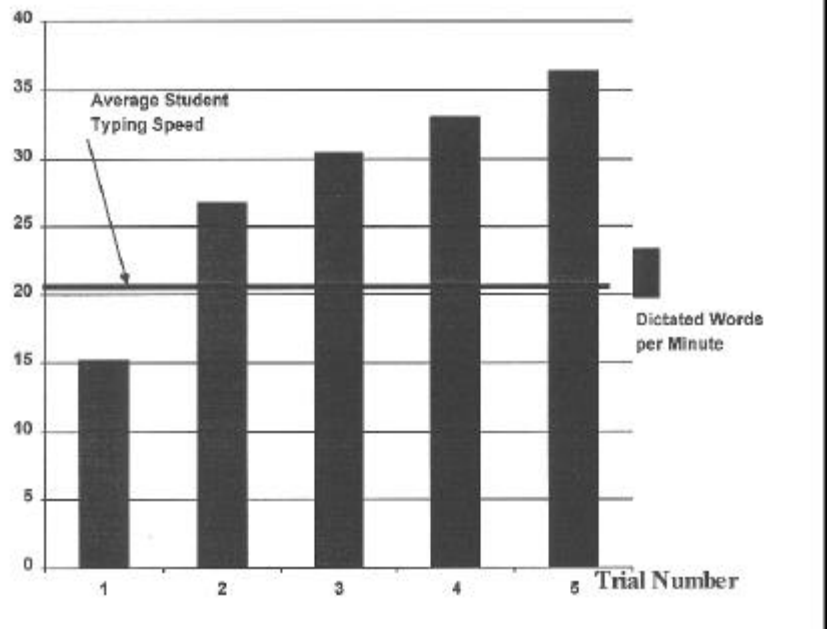

Figure 1. Comparison of voice and keyboard speed.

Discussion of Typing and Voice Tests

The average typing speed for the construction students tested was about 20 words per minute. The first electronic dictation yielded a speed of 15 words per minute, or about 75 percent of the average typing speed of the students.

The voice recognition software demonstrated ability to "team." By the third repetition the initial voice dictation speed was doubled. This increase is significant because it indicates that the speed of all electronic dictation will increase as a function of usage. Dictation speeds continued to increase as a function of the number of trials, with a period of diminishing returns noticeable in the fourth and fifth repetitions.

Conclusion

The author's initial experience with voice recognition suggests strong potential for voice recognition and a need for further study regarding construction applications. The literature indicates that the quality of the software is rapidly evolving, and that the price of the software is rapidly decreasing. The literature bather indicates that leaders of major information technology corporations see voice recognition as a prominent pan of their business within the next ten years. Using commercially available software with minimal training, the author was able to approach the average typing speed of a group of thirty construction students the first time the text was dictated. Further repetitions yielded significant increases in speed, indicating that dictation speed would increase as the software "learned" the user's voice patterns.

Other than occasional editing, the entire body of this document was electronically dictated using DragonDictme for Windows, Version 2.5, Classic Edition.

References

Berst, Jesse (December 15,1996) "*Why I've Gone Back m the Future*", Windows Sources Magazine, Ziff-Davis Publishing Company, Boulder, CO.

Levinson, Stephen (October 1995) "*Speech Recognition Technology: A Critique*", Proceedings, National Academy of Sciences, Volume 92, pp.9953-9955

Machrone, Bill (September 30, 1996*)" Bill and Jesse's Time Line of the Future*", PC Week Online, Ziff-Davis Publishing Company, Boulder, CO.

Maney, Kevin, (October 14,1996) "*What Will Microsoft be Like in 10 years*?", Interview with Bill Gates in USA Today, Gannett Publishing Co., Inc., Arlington, VA.

Mayor, Tracy (July 1, 1996) "*Look Ma, No Hands*", PC Week Magazine, Ziff-Davis Publishing Company, Boulder, CO.

Mehm, Dime (September 30, 1996, Vol. 18, Issue 40) "*A Comparison of DragonDictate Classic Edition for Windows 2.5, IBM Voice Type 3.0 for Windows 95, and Kurzweil Voice for Windows, Release 2.0*", InfoWorld Magazine, InfoWorld Publishing Co., Boulder, CO.

Mills, Elinor (November 22, 1996) "*Voice Pilot Hawks Web-based Voice Translator*", InfoWorld Electric Magazine, InfoWorld Publishing Co., Boulder, CO.

Schmit, Julie (November 8, 1996) "*Intel Co-founder: Chip Evolution Continues*", Interview with Gordon Moore in USA Today, Gannett Publishing Co., Inc., Arlington, VA.

Schwartz, Ephraim (September 16,1996, Vol. 18, Issue 38) "*IBM Unveils its Speech Technology*", InfoWorld Magazine, InfoWorld Publishing Co., Boulder, CO.

Vaillansudi, Pardhu (November 11, 1996, Vol. 19, Issue 46) "*Speech Recognition Software gets Application Specific*", InfoWorld Magazine, InfoWorld Publishing Co., Boulder, CO.

Wilson, Carol (December 15, 1996) "*Computers Learn to Listen*", Interactive Week Magazine, Interactive Enterprises, LLC.

Weinstein, Clifford J. (October 1995) "*Military and Government Applications of Human-Machine Communication by Voice*", Proceedings, National Academy of Sciences, Volume 92, pp.1001 1-10016.

Appendix I

Journal Describing Early Lemming Process

Note: Entries in this journal contain spelling, usage, and grammar errors. In the spirit of retaining original thoughts and opinions, no subsequent editing was done.

Entry I - Tuesday, October 1, 1996
This journal is being dictated rising Dragon Dicame. I installed the software this morning, I dictated most of a letter of reference for a colleague the same afternoon. I also went through the tutorial after software was installed.
My initial reaction is mixed. It is fun to dictate letters. The software does seem to be a little slow to me. I am running it on a Pentium 90 computer, with 16 megabytes of RAM.
I need to go back and review some of the correction features. When I make a mistake that occurs about three words back, I have trouble fixing it.
I have spent about three hours with the software.

Entry 2 - Wednesday, October 2
1 got up early this morning and answered my e-mail using Dragon Dictate. It seems the software is already beginning to recognize my voice. I am really bothered that the word display does not keep pace with my speaking voice. I wonder if a faster computer would reduce this problem.
I returned home in the afternoon and answered e-mail with dictated responses. I felt much better about my own skill level and about the voice recognition software ability to understand my words. I spent three hours dictating on this second day. This stuff may work!

Entry 3 - Thursday, October 3
1 got up early again this morning and answered my e-mail via voice. I did compose a long message the be technology committee and when I changed to command mode and told the program to hit send, the program shut down and my long message was lost. That was frustrating.
I also read an article in InfoWorld that rated dictation systems, Of the three systems rated, Dragon Dictate was third. I need to remember to get an electronic copy of that article and pat it in my files. I told a friend today that I thought I was getting as fast at voice dictation as I using the keyboard.

Entry 4 -Tuesday, October 8
I have been remiss in my Journal writing for the last few days. I dictated a letter this morning and looked at the statistic's section for the first time. I dictated 17 words per minute at an accuracy rate of 91 percent. I need to read from the newspaper soon so that I will get a feet for how fast I can go. I am suit very optimistic about the technology, I need to get better with the error correction. I do feet like I can dictate as fast as I can type. One of the things that I have noticed is how comfortable (physically) it is to dictate as opposed to typing. I sit with my hands behind my head and my feet propped up on the desk when I dictate. I have accumulated about 11 hours training so far.

Entry 5-Thursday, October 10
I continue to team. I called typical support for the fast time yesterday. They were friendly and competent. I am having trouble with the file save command, especially when used with my e-mail system. I am going to call them again this morning. It will be interesting to see what Ashley (a graduate student also working with voice) thinks about product.
It definitely takes some time and some patience to learn. You really have to be committed to learning to use the dictation system. T am a little discouraged today, but I am going to keep plugging,

Entry 6--Friday, October 11
I took a typing test yesterday. I took an article out of USA Today and read to the computer. I really had a hard time. I typed only 10 words per minute and my accuracy was only 76 percent. There were difficult words in the article but I was disappointed that speed and accuracy were so low.

I was able to talk to Keith Byerly at Dragon Systems. We had a great discussion about the future of voice recognition systems. Keith told me that he dictated 90 words per minute at Comdex last year. We got along very well and I plan to talk to him again. I have about 15 hours invested so far.

Entry 7--Tuesday, October 15
My wife and I had a race yesterday. She was able to type at a rate 60 words per minute. I was able to dictate at a rate of fifty-one words per minute. I am getting more comfortable every day with the software. I find that it is helpful if I enunciate very carefully. I will be giving a short presentation this Friday for the industry advisory counsel. I want to make sure that I understand voice macros before I give the presentation.

Entry 8--Monday, October 21
I gave my presentation to the Industry Advisory Council last Friday. I felt like it went very well. Ashley and I practiced a good bit. We used some voice macros to enhance what we showed them. Voice macros are really slick. I can see a professional doing some very impressive things with them. All in all, I felt like they were impressed. I need to remind myself to dictate more frequently, I have invested about 20 hours so far.