

## Voice Recognition Software Versus a Traditional Transcription Service for Physician Charting in the ED

ROBERT G. ZICK, MD AND JON OLSEN, MD

This study was conducted to compare real-time voice recognition software to a traditional transcription service. Two emergency department (ED) physicians dictated 47 charts using a voice dictation software program and a traditional transcription service. Accuracy, word per minute dictation time and turnaround time were calculated from the data. The transcription service used in our study was more accurate than the voice recognition program with an accuracy of 99.7 percent versus 98.5 percent for the voice recognition program. The average number of corrections per chart was 2.5 for the voice recognition program and 1.2 for the traditional transcription service. Turnaround time was much better using the computer voice recognition program with an average turnaround time of 3.65 minutes versus a turnaround time of 39.6 minutes for the traditionally transcribed charts. The charts dictated using the voice recognition program were considerably less costly than the manually transcribed charts. In summary, computer voice recognition is nearly as accurate as traditional transcription, it has a much shorter turnaround time and is less expensive than traditional transcription. We recommend its use as a tool for physician charting in the ED. (*Am J Emerg Med* 2001; 19:295-298. Copyright © 2001 by W.B. Saunders Company)

Patient charting is a very time-consuming and tedious task. Physicians in general, and emergency physicians in particular are often overwhelmed by this onerous task. Handwritten charts are difficult to read and often incomplete. In cases involving litigation, illegible handwritten charts are often difficult to defend. Using transcription services to transcribe charts is costly and the turnaround time can be long. With the advent of real-time continuous voice recognition software the possibility exists to have accurate, quick, legible, complete charts at a very low per-chart cost. This study was designed to look at one of the newest voice

recognition programs for medical dictation (Dragon NaturallySpeaking Medical Suite version 4.0) and compare it with a traditional transcription service with regard to accuracy, speed of dictation and cost.

### METHODS

A total of 47 charts were dictated by 2 ED physicians at a suburban level 1 trauma center with an annual census of 45,000. One of the physicians was an "advanced" user, having several years of experience with and having dictated several hundred charts with the software. The second physician was a "basic" user with approximately 2 weeks experience and having dictated approximately 20 charts with the software.

Dragon NaturallySpeaking Medical suite version 4 was installed onto a 450 MHz Pentium II computer with a 10 GB hard drive and 128 MB of RAM. The software required each user to have a training session so the program could learn the voice pattern of each user. The overall training time was approximately 30 minutes, with 10 minutes of actual dictation time speaking to the computer, and 20 minutes for the computer to process the new speech files.

The advanced user dictated 27 charts and the basic user dictated 20 charts. Each chart was dictated twice, first using the voice recognition software, and then a second time using the traditional voice transcription service, for a total of 94 charts. The document that was generated initially using the voice recognition program was used as the basis for the traditionally transcribed document, thus both charts included the same text.

For the computer-generated voice recognition charts, templates were used which contained headings for the patient's chief complaint, history of present illness, past medical history, past surgical history, medications, allergies, family history, social history, and the physician's "normal" physical examination. The stopwatch was then started and the chart was dictated. Any portion of the physical examination that was different than that described on the template was dictated over the default value. Any portion of the physical examination that was performed by the physician that was not on the template was added to the chart. Any portion of the physical examination that was listed on the template but was not performed by the physician was removed from the chart. The time to dictate and correct the chart using the voice recognition software was recorded

---

From the Lutheran General Hospital Department of Emergency Medicine, Park Ridge, IL and the University of Chicago Hospitals Section of Emergency Medicine, Chicago, IL.

Manuscript received October 6, 2000, returned November 26, 2000, revision received December 6, 2000, accepted December 28, 2000.

Supported in part by software donated by Dragon Systems, Inc., Newton, MA. Hardware for the study was donated by Lutheran General Hospital's Department of Emergency Medicine.

Address reprint requests to Robert G. Zick, MD, MBA, 355 Ridge Ave, Evanston, IL, 60202.

*Key Words:* Emergency department transcription, voice dictation, charting.

Copyright © 2001 by W.B. Saunders Company  
0735-6757/01/1904-0011\$35.00/0  
doi:10.1053/ajem.2001.24487

(dictation time and correction time were not recorded separately because people tend to correct mistakes that the computer makes right after they notice them), as were the number of errors (an error being any mistake that required the user to redictate or manually type in a correction on the chart). The number of words that the user dictated were counted and the percentage accuracy turnaround time and word-per-minute dictation rate were calculated.

The traditional dictation service transcriptions were evaluated in the following manner: The voice recognition chart was used as a basis for the traditional transcription chart. The transcription service telephone number was called and the entire chart that was created using the voice recognition program was read to the transcription service. The total time to dictate the chart was recorded using a stopwatch, and the time that the chart was sent to the transcription service was also recorded. After the transcription service finished transcribing the chart, it was sent back to a computer in the ED via modem. The computer generated a time stamp when the file was received. The transcribed chart was then corrected and the time to correct the chart was also recorded. The number of errors on the transcription service's transcribed document was also recorded. An error on the transcription service chart was defined as any discrepancy between the computer-generated chart and the transcription service-generated chart. The total turnaround time for the traditional transcription service charts was then calculated by adding together the dictation time, the time between the completion of the dictation and the arrival of the transcribed chart, and the time to correct the transcription service dictation. The number of words in the transcription service document were then counted and the percentage accuracy and word per minute dictation time were calculated. The word per minute dictation time was defined as the total number of words dictated divided by the number of minutes turnaround time.

Ninety-five percent confidence intervals were calculated for the means and the difference in means between the computer transcribed and traditionally transcribed charts. The normal distribution was used to calculate the confidence intervals when analyzing the data as a whole, and the 2-tailed t test was used to calculate the confidence intervals when comparing the advanced users' performance to the basic users'.

## RESULTS

Our data comparing the voice recognition program to the traditional transcription service with regard to accuracy, average number of errors, average turnaround time and word per minute dictation time is listed in Table 1.

Comparison of the results between the 2 users is shown in Table 2.

## DISCUSSION

There are several ways to create ED records: handwritten charts, handwritten templates, traditional dictation services, and now computer-generated voice recognition systems. Handwritten charts are time-consuming, fatiguing, and often difficult to read. Marill, et al found that handwritten template charts (ie, the T system) is associated with higher gross billing and physician satisfaction, but no significant decrease in emergency physician total evaluation time.<sup>1</sup> Traditionally transcribed charts are expensive and the turnaround time is long.

There has been little written in the emergency medicine literature regarding the use of voice recognition software. In 1989, Spacone analyzed a software program for voice-controlled text creation that was used on a microcomputer in the emergency medical department of a large hospital.<sup>2</sup> It was found that automated speech recognition (ASR) permitted doctors to produce their reports faster and more accurately than handwriting or dictation to tape. In addition, ASR reporting was found to be less expensive than dictation to tape. Other advantages with voice recognition included the ability to create computer-stored patient files that could be used for statistical purposes. Spacone found a few minor drawbacks, including the need for a procedure for training the system to recognize the user's voice.

In 1996, Zemmer, et al assessed the accuracy of IBM VoiceType Dictation emergency medicine software package and found that the vocabulary domain, microphone position, and background noise markedly reduced dictation accuracy. They found that the VoiceType Dictation system was impractical in an ED setting.<sup>3</sup>

There has been more research on voice recognition in the radiology literature than in the emergency medicine literature. Antiles et al from the Radiology Consulting Group at Massachusetts General Hospital in Boston found that computerized voice recognition systems (VR) reduced costs and

**TABLE 1.** Comparison of Voice Recognition and Traditional Transcription Service Dictations

	Voice Recognition (95% CI)	Transcription (95% CI)	Difference (95% CI)
Accuracy (%)	98.5 (98.2-98.9)	99.7 (99.6-99.8)	1.2 (0.8-1.5)
Average no. errors/chart	2.5 (2-3)	1.2 (0.9-1.5)	1.3 (.67-1.88)
Average dictation and correction time (min)	3.65 (3.35-3.95)	3.77 (3.43-4.10)	0.12 (-0.34-0.58)
Average turnaround time for receipt of a completed document (min)	3.65 (3.35-3.95)	39.6 (28.23-50.97)	35.95 (24.59-47.31)
Throughput words/min	54.5 (49.6-59.4)	14.1 (11.1-17.2)	40.4 (34.4-46.39)

NOTE. Voice recognition was slightly less accurate than traditional transcription (98.5 v 99.7%) with an average of 2.5 errors/chart compared with 1.2 errors/chart with traditional transcription. Average turnaround time to obtain a copy of the medical record was much longer for transcription (39.6 minutes) than for voice recognition (3.65 minutes) because of the time needed for the transcribed chart to be typed off site and sent back to the emergency department via modem. The voice recognition program had a much higher throughput (54.5 v 14.1 words per minute) than the traditional transcription service.

**TABLE 2.** Comparison of Advanced Versus Basic User Dictations

	Advanced User (95% CI)	Basic User (95% CI)	Difference (95% CI)
Accuracy—voice recognition (%)	99.1 (98.8-99.4)	97.7 (97.1-98.2)	1.49 (.92-2.05)
Accuracy—transcription (%)	99.7 (99.6-99.8)	99.6 (99.4-99.8)	0.1 (-0.06-0.33)
Voice recognition throughput words/min	64.2 (60-68.5)	40.6 (34.1-47.1)	23.6 (15.95-31.27)
Transcription throughput words/min	19.5 (15.9-23)	8.4 (5.7-11.1)	11.1 (5.75-16.41)
Turnaround time-voice recognition (min)	3.45 (3.08-3.82)	4.0 (3.47-4.53)	.6 (-1-1.2)
Turnaround time-transcription (min)	26.18 (21.3-31.1)	63.45 (35.8-91.1)	37.3 (14.44-60.1)
Transcription dictation rate words/min*	91.9 (89.1-94.7)	111.2 (100.9-121.5)	19.3 (8.41-30.19)

NOTE. The advanced user had a higher percentage accuracy using the voice recognition program than the basic user (99.1 v 97.7%). The difference in percentage accuracy using the transcription service between the advanced and basic users was not significant (99.7 v 99.6%). Voice recognition throughput for the advanced user was 23.6 words per minute higher (64.2 v 40.6 words per minute) than the basic user. The advanced user also had greater transcription throughput using the transcription service (19.5 v 8.4 words per minute) than the basic user. The advanced user had a lower voice recognition turnaround time than the basic user (3.45 v 4.0 minutes) which was not significant. The advanced users' turnaround time using the transcription service was much faster than the basic users (26.18 v 63.45 minutes).

\* The transcription dictation rate is the rate of dictation using the transcription service without taking into account transmission or transcription time. It is the rate of dictation calculated by determining the number of words dictated to the transcription service, divided by the number of minutes the physician was on the telephone to the dictation service. The transcription dictation rate was faster for the basic user, who had more experience dictating using the transcription service.

enhanced service. Massachusetts General Hospital in Boston implemented VR over a 2-year period and achieved an annual savings of \$530,000 and a 50% decrease in report throughput. Those accomplishments required solid planning and implementation strategies, training, and sustainment programs.<sup>4</sup>

It can be seen from our data that traditional voice transcription is slightly more accurate than computer-generated transcription. There were an additional 1.3 errors per chart using the voice recognition program compared with using the traditional transcription service. The traditional transcription service was slightly more accurate than the voice recognition system, but the turnaround time using the traditional transcription service was much longer.

For the advanced user, there was a slight percentage accuracy increase in traditional transcription over voice recognition. This is likely attributable to the advanced user's familiarity with the voice recognition software. The basic user, on the other hand, had a higher difference in percentage accuracy of traditionally transcribed over computer-generated charts. This is likely because of the basic user's increased familiarity with the traditional transcription service and decreased familiarity with the computer program.

The accuracy of the basic user using the voice recognition program was less than that of the advanced user by 1.49%. This is probably attributable to the basic user's lack of familiarity with the software. Both the advanced and basic user had similar accuracy rates (99.7 v 99.6%) using the transcription service.

The advanced user's dictation and correction time of 3.45 minutes using the voice recognition software was less than the basic user's which was 4 minutes. This difference, however, was not significant.

The cost of using real-time voice recognition software is significantly less than using a transcription service for ED documentation. The per-chart cost of using the transcription service at our institution is \$7.50 per chart. An ED with an annual census of 45,000 patients per year could conceivably incur transcription costs of \$337,500 per year assuming all

of the ED charts were transcribed. Obviously, the cost would be less based on the percentage of charts that are dictated versus hand-written or transcribed using other means. The cost of a computer capable of running the voice recognition software and an inexpensive printer is approximately \$2,000, and the cost of Dragon NaturallySpeaking Medical Suite Version 4 is approximately \$1,000, for a total start-up cost of \$3,000. This amounts to a potential savings to the ED of \$334,500 during the first year alone. The per chart cost during the first year would be \$3,000/45,000 charts, or 7 cents per chart, which is much less expensive than using a transcription service.

In addition to transcription costs, the cost of the physician's time is also very significant in the ED. The average time for the physician to dictate and correct a chart (not taking into account transcription turnaround time) was 3.65 minutes using the voice recognition program and 3.77 minutes using the transcription service. This amounts to a time savings of 0.12 minutes per chart by using the voice recognition program. The difference between these 2 times was not significant, however. If you take into account this 0.12 minute per chart time savings into the cost analysis, however, at a physician hourly rate of \$110/hour, that amounts to a 22 cent per chart cost savings to the ED by using voice recognition. For an ED with an annual census of 45,000, that amounts to a potential savings of \$9,900 in physician time in addition to the transcription cost savings described earlier.

One drawback of using templates for physician charting is that it is easy to leave on the template a default value for the physical examination that was either not done or was different than the patient's actual examination. In our study, if we neglected to perform a part of the physical examination, that portion of the template was removed from the document (using the mouse and the delete key) before the printing of the final chart. Portions of the examination that were performed but were different than the default value were dictated over the default value. The use of macros could alleviate the problem of leaving on the chart portions of the physical examination that were not actually done. The

Dragon NaturallySpeaking software has macro capability, so you could conceivably use a template with a blank physical examination section and create a voice command macro to insert an entire normal physical examination if the examination was indeed normal. You could also create small macros to insert individually each normal body part that was examined, and free-text dictate the portions of the physical examination that were abnormal. Macros could also be used to delete portions of the physical examination that were not performed on the patient.

One of the limitations of our study was that we recorded data for only 2 users. Time constraints caused by one of the authors' graduation from residency and subsequent loss of clinical privileges at the host institution contributed to this limitation. We do, however, feel that the Dragon NaturallySpeaking software is a viable alternative to traditional transcription for both novice and experienced users, and we would expect similar results and conclusions if more users had been entered on the study.

The software is very accurate right out of the box, but it also learns from its mistakes and its accuracy improves over time. The users' accuracy and speed should also improve as the user becomes more familiar with the product. This was shown by the results of the advanced user, who was more familiar with the product and who consistently had more accurate transcriptions than the basic user.

## CONCLUSION

Computer voice recognition transcription using real-time voice recognition software is an economical and accurate way to generate ED records. The software is nearly as accurate as traditional transcription, it has a much shorter turnaround time and it is less expensive. We recommend its use as a tool for physician charting in the ED.

Special thanks to Douglas Propp, MD for obtaining the hardware for this study, and also special thanks to Myrna Greenfield and Dragon Systems Inc. for donating the software for use in this study.

## REFERENCES

1. Marill K, Gauharou E, Nelson B, et al: Prospective randomized trial of template-assisted versus undirected written recording of physician records in the emergency department. *Ann Emerg Med* 1999; 33: 500-9
2. Spacone AB: Microcomputer voice-recognition program in a hospital emergency department. *J Soc Health Syst* 1989; 1:111-8
3. Zimmel NJ, Park SM, Schweitzer J, et al: Status of voicetype dictation for windows for the emergency physician. *J Emerg Med* 1996; 14: 511-5
4. Antiles S, Couris J, Schweitzer A, et al: Project planning, training, measurement and sustainment: The successful implementation of voice recognition. *Radiol Manag* 2000; 22:18-31