Radiological Reporting That Combine Continuous Speech Recognition with Error Correction by Transcriptionists

Tamaki ICHIKAWA[1], Takashi KITANOSONO[2], Jun KOIZUMI[1], Yoichi OGUSHI[3], Osamu TANAKA[4], Jun ENDO[1], Takeshi HASHIMOTO[1], Shuichi KAWADA[1], Midori SAITO[1], Makiko KOBAYASHI[1], Yutaka IMAI[1]

Department of [1]Radiology and [3]Medical Informatics, School of Medicine, Tokai University, 143 Shimokasuya, Isehara, Kanagawa, 259-1193*, Japan;*
[2] *Department of Vascular and Interventional Radiology, University of Rochester Medical Center, Rochester, New York,* 14642 *USA; and*
[4] *Department of Radiology, Jichi Medical University, Omiya Medical Center, Saitama, Japan*

(Received August 8, 2007; Accepted October 29, 2007)

Objective: We evaluated the usefulness of radiological reporting that combines continuous speech recognition (CSR) and error correction by transcriptionists.
Materials and Methods: Four transcriptionists (two with more than 10 years' and two with less than 3 months' transcription experience) listened to the same 100 dictation files and created radiological reports using conventional transcription and a method that combined CSR with manual error correction by the transcriptionists. We compared the 2 groups using the 2 methods for accuracy and report creation time and evaluated the transcriptionists' inter-personal dependence on accuracy rate and report creation time. We used a CSR system that did not require the training of the system to recognize the user's voice.
Results: We observed no significant difference in accuracy between the 2 groups and 2 methods that we tested, though transcriptionists with greater experience transcribed faster than those with less experience using conventional transcription. Using the combined method, error correction speed was not significantly different between two groups of transcriptionists with different levels of experience.
Conclusion: Combining CSR and manual error correction by transcriptionists enabled convenient and accurate radiological reporting.

Key words: radiological reporting, transcriptionist, continuous speech recognition

INTRODUCTION

First described in 1981 [1], speech recognition systems are becoming increasingly popular in radiology departments across North America and Europe [2-7]. Recent systems do not require discrete speech, and continuous speech recognition (CSR) systems are available for radiological dictation [8-10] that allow radiologists to speak more naturally and faster than with earlier systems [9]. Although picture archival and communication systems (PACS) are popularized in many radiology departments [11, 12] and radiology images are delivered to clinicians immediately after examinations, the lag between image availability and PACS reports is longer than with conventional dictation, transcription and report creation. As busy radiology departments install PACS to improve the efficiency and quality of their operations, this gap is the most important problem to overcome [11]. To overcome this issue in the era of PACS, newer CSR systems offer a palatable alternative to discrete speech recognition and increase recognition speed. However, accuracy remains an issue, the editing process distracts most radiologists from image interpretation; most would prefer a user interface that more closely simulates conventional dictation [1, 8].

We evaluate the usefulness of a method in which reports created using CSR corrected by transcriptionists, a method that we suspect will allow transcriptionists of greater and lesser experience to transcribe and create radiological reports speedily and accurately.

MATERIAL AND METHODS

Two transcriptionists each were assigned to 2 groups according to their years of transcription experience: the two in Group A had more than 10 years' experience, and the two in Group B had less than 3 months' experience. All four listened to the same 100 dictation files (voice linked text files) and created radiological reports using both conventional human transcription and a combined method in which they corrected errors in reports created using CSR. The transcriptionists were allowed unlimited opportunities to listen to the dictation and playing and recording times of dictated speech were the same. A hundred dictation files (7349 words) were divided into 10 series, each of which comprised 5 reports of plain radiographs, 3 reports of computed tomographic (CT) examination and 2 reports of magnetic resonance (MR) imaging examination: each file ranged from 615 to 863 words. During one month, the transcriptionists completed 200 reports; each day, each

Tamaki Ichikawa, Department of Radiology,School of Medicine, Tokai University, 143 Shimokasuya, Isehara, Kanagawa, 259-1193, Japan
Telephone number: 0463-93-1121, Fax: 0463-93-6827 E-mail address: tamaki-i@mars.sannet.ne.jp

Table 1 Error words and report creating time on each series of each transcriptinist on both methods

TR1

No of series	error words 1	error words 2	RC time1	RC time2
1	6	2	522	317
2	0	1	577	396
3	0	1	692	415
4	1	2	640	406
5	0	0	595	432
6	1	0	594	417
7	1	0	488	336
8	0	0	548	337
9	1	1	622	438
10	2	2	471	438

TR2

No of series	error words 1	error words 2	RC time1	RC time2
1	2	3	550	331
2	4	3	580	454
3	1	1	678	468
4	3	3	734	429
5	2	0	668	442
6	3	1	590	449
7	2	1	534	338
8	0	0	580	367
9	2	2	682	436
10	2	2	501	381

TR3

No of series	error words 1	error words 2	RC time1	RC time2
1	4	2	1243	366
2	1	1	1257	401
3	1	1	1458	452
4	3	2	1654	430
5	1	1	1369	445
6	2	0	1205	450
7	1	1	1106	340
8	0	0	1289	337
9	2	1	1254	450
10	2	2	1078	420

TR4

No of series	error words 1	error words 2	RC time1	RC time2
1	2	2	1390	345
2	3	3	1450	450
3	1	1	1782	480
4	3	3	1632	434
5	2	2	1775	458
6	3	3	1483	487
7	2	2	1232	345
8	0	0	1345	370
9	2	1	1296	440
10	2	2	1145	429

TR1, 2: transcriptinist of Group A
TR3, 4: transcriptinist of Group B
Group A: transcriptionists with more than 10 years' transcription experience
Group B: transcriptionists with less than 3 months' transcription experience

Error words 1: error words using conventional transcription
Error words 2: error words using combined method
RC time 1: Report creation time using conventional transcription.
RC time 2: Report creation time using combined method

transcriptionist created one series of reports using one method. None of the transcriptionists was experienced using CSR. All speech files were dictated in Japanese by the author using CSR system, AmiVoice RadScribe version 5.14 (Advanced Media, Inc., Tokyo, Japan) that did not require the training of the system to recognize the user's voice (so-called enrollment). The dedicated dictionary and language model were created from 3-megabyte radiological report texts. We used this software on computers with Windows operating systems. The speech recognition speed of the CSR system was set at highest level (0.1 seconds). Three hundred and eleven of 7349 words were misrecognized, which corresponded to a 95.8% rate of average speech recognition accuracy for free-style dictation using CRS. The following equation was used to establish the rate of accuracy of word recognition based on the word-counting system of AmiVoice: accuracy rate of recognition = (total words − misrecognized words − dictation error words − insertion error words) / total words × 100%. Total speech recognition time was 2116 seconds. We compared the accuracy rate and report creation time using conventional transcription with those using the combined method each 20 series for both groups of transcriptionists. Statistical difference was tested by Mann-Whitney U-test. Additionally, using post hoc testing, we compared inter-personal dependence of transcriptinists on accuracy rate and report creation time in both methods respectively. Statistical significance was defined as $P < 0.05$.

RESULT

Error words and report creation time on each series of each transcriptinist on both methods are shown on Table 1. The accuracy rate and total report creation time for conventional transcription and combined methods are shown for both groups in Table 2a. The accuracy rate in both methods is more than 99.7% and not significantly different between the 2 groups of transcriptionists (conventional transcription: P=0.383, combined method: P=0.445). Using conventional transcription, transcriptionists of Group A were faster than members of Group B ($P < 0.0001$), but there was no significant difference in error correction speed between transcriptionists with different levels of experience comprising the 2 groups using the combined method (P=0.192). All transcriptionists created reports faster using the combined method ($P < 0.05$). The transcriptionists demonstrated no inter-personal dependence on accuracy rate in either method or on report creation time using the combined method.

DISCUSSION

CSR is steadily replacing conventional transcription in radiology departments as PACS and CSR systems are installed to improve the efficiency and quality of operations by reducing departmental operational costs and report turnaround time. [11, 12].

In large departments, virtually all reported installations indicate substantial cost savings [9]. Time saving is another important advantage of the newer CSR for radiological reporting. Since its implementation was

Table 2a. Accuracy rate and report creation time for both methods

No of TR	Group	AR 1 (%)	AR 2 (%)	Total Time 1 (seconds)	Total Time 2 (seconds)
1	A	99.8	99.8	5749	3932
2	A	99.7	99.8	6097	4095
3	B	99.7	99.8	12913	4091
4	B	99.7	99.7	14530	4238

No of TR: Number of transcriptinists
Group A: transcriptionists with more than 10 years' transcription experience
Group B: transcriptionists with less than 3 months' transcription experience
AR 1: accuracy rate on conventional transcription
AR 2: accuracy rate on method combined method
Total Time 1: total report creation time using conventional transcription
Total Time 2: total report creation time using combined method

Table 2b Comparison of percentage of time saved in report creation between the combined and conventional transcription methods.

Number	Group	Time Saving (%)
1	A	32.7
2	A	33.1
3	B	60.1
4	B	71.3

Reports were created significantly faster using the combined method than using conventional transcription.

improved, CSR has been demonstrated to be faster than conventional dictation for CSR total reporting times and report turnaround time [9, 13].

However, with CSR, recognition accuracy still remains problematic, at about 95% for free-style dictation in radiological reporting and error correction time is longer than dictation time [14, 15], and CSR is troublesome for radiologists used to conventional dictation. Radiologists prefer to spend their time performing examinations and interpreting images than correcting errors on a computer. CSR will not prove successful if it seems to force them to look at reports and images made using CSR systems. Already, some implementations allow the radiologists to use CSR as they would conventional transcription, dictating reports and passing them to a transcriptionist for editing [15, 16]. However, Mohr's study demonstrated that speech recognition did not improve the productivity of secretaries or transcriptionists [16], whereas Vorbech' s study indicated that typist' performance contributed to overall time savings with CSR and that on average, reports edited with CSR were generated 19% faster than those created by conventional text editing in German [15]. In our study, in Japanese, the average time saved using the combined method was about 50% of that using conventional transcription. Time saving for an unskilled transcriptionist was 70.8% using the combined method compared with conventional transcription (Table 2b). Even for the most experienced transcriptionist, time saving was 31.6%. However, the time saving may vary and depend on topics, speaker and languages.

Although in Japan at present, radiological reports are created by transcriptionists from tape dictation, the method is more popular in North America and Europe. However, there are fewer than 100 members of the Japan Medical Transcriptionist Association. In this study, transcriptionists performing conventional dictation achieved more than 99.7% accuracy, the same as previously reported [9]. Experienced transcriptionists are more accurate, but slower than CSR in creating reports and unskilled transcriptionists take longer to create accurate reports. To our knowledge, such comparison of speed and accuracy of report preparation has not been compared for among transcrionists in English. In Japanese, all transcriptionists can create reports speedily and accurately using a combined method (Table. 2).

Radiologists, including the author, who are used to conventional dictation are in the habit of dictating long sentences while looking at images and not observing reports on computer screens. To avoid system recognition errors of continuous speech caused by word coarticulation, wherein the beginning and ending phonemes of a word are affected by the surrounding words [13], radiologists who use CSR must dictate long sentences clearly and fluently and remember to speak clearly at the beginning and ending of sentences. Care and time are necessary to correct errors in long sentences to create accurate reports. In addition, radiologists used to conventional dictation find it troublesome to verify the accuracy of what is being recorded by moving their eyes between radiological reports on the computer screen and images they are describing.

Transcription accuracy remains below 100% using either dictation method because transcriptionists mishear something or those dictating include spontaneous speech "disfluencies" such as stammering, slurring, hesitation, filled pauses, repairs, fragments, and interruptions, that are part of human nature[9]. The CSR system will not compensate for the poor articulation of the radiologists. Those human problems must be solved by humans. However, the combined method helps resolve most of these issues.

Our study is limited because only one radiologist (the author) dictated the speech files and 5 years' experience with the CSR system after, so accuracy was high. In real life, many radiologists will use the system with better or worse recognition rates than each others. An extensive evaluation of a wide range of radiologists would be useful to further determine the usefulness of the combined method.

Another limitation is the contents of the dictated files. Reports of only 3 types of examination, plain radiographs, CT, and MR imaging, were dictated. Different modalities and anatomical locations examined will require different sets of technical terms and may affect.

Finally, this study was not performed in a working environment, where such factors as background noise, interpersonal communications, phone calls and pages could distract the process both dictating and transcribing. A study in a more authentic working environment is warranted.

A disadvantage of the combined method is that it is a not complete real-time reporting system. Radiologists must send their dictation speech files to transcriptionists for editing, but transcriptinisrs are not available around the clock to edit reports. Thus, 24-hour report turnaround is not possible.

In conclusion, using a transcription method that combined CSR with error correction by transcriptionists, error correction speed was not significantly different between groups consisting of transcriptionists with different levels of experience, but generation of reports was faster by all transcriptionists. Therefore, this combined system is felt to be useful, convenient, and accurate for radiological reporting.

ACKNOWLEDGEMENT

The authors thank transcriptionists Bandou Y, Murai R and Iizuka Y of Department of Radiology, Tokai University, Watari S and Ohtaka N of the Department of Radiology, Mizonokuchi hospital, Teikyo University, Kishi Y and Takahara T for their assistance.

REFERENCES

1) Leeming BW, Porter D, Jackson JD, Bleich HL, Simon M. Computerized radiologic reporting with voice data-entry. Radiology 1981; 138: 585-588
2) Robbins AH, Horowitz DM, Srinivasan MK, *et al*. Speech-controlled generation of radiology reports. Radiology 1987; 164: 569-573
3) Robbins AH, Vincent ME, Shaffer K, Maietta R, Srinivasan MK. Radiology reports: assessment of a 5,000-word speech recognizer. Radiology 1988; 167: 853-855
4) Herman SJ. Accuracy of a voice-to-text personal dictation system in the generation of radiology reports. Am J Roentgenol 1995; 165: 177-180
5) Zemmel NJ, Park SM, Maurer EJ, Leslie LF, Edlich RF, Evaluation of Voice Type Dictation for Windows for the radiologist. Med Prog Technol 1997; 21: 177-180
6) Dietz U, Rupprecht HJ, Espinola-Klein C, Meyer J. Automatic report documentation in cardiology using a speech recognition system. Z Kardiol 1996; 85: 684-688 (article in German)
7) Kolles H, Feiden W, Computer-assisted speech recognition inn diagnostic pathology. Development of Dragon Dictate 30-k system for documentation. Pathology 1995; 16: 439-442 (article in German)
8) Rosenthal DI, Chew FS, Dupuy DE, Kattapuram SV, Palmer WE, Yap RM, Levie LA, Computer-based speech recognition as a replacement for medical transcription. Am J Roentgenol 1998; 170: 23-25
9) Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E. Continuous speech recognition in MR imaging reporting: advantage, disadvantages, and impact. Am J Roentgenol 2000; 174: 617-622
10) Schwarz LH, Kijewski P, Hertogen H, Poossin PS, Castellino R. Voice recognition in radiology reporting. Am J Roentgenol 1997; 169: 27-29
11) White KS. Speech recognition implementation in radiology. Pediatr Radiol 2005; 35: 841-846
12) Bramson RT, Bramson RA. Overcomng obstacles to work-changing technology such as PACS and voice recognition. Am J Roentgenol 2005; 184: 1727-1730
13) Seltzer SE, Kelly P, Adams DF, *et al*. Expenditing the turnaround of radiology reports in a teaching hospital setting. Am J Roentgenol 1997; 168: 889-893
14) Ichikawa T, Koizumi J, Takahara T, *et al*. Continuous speech recognition system for radiological reporting: comparison with experience of dictation. Nippon Acta Radiologica. 2005; 65: 384-386 (article in Japanese)
15) Vorbech F, Ba-Ssalamah A, Kettenbach J, Huebsch P. Report generation using digital speech recognition in radiology. Eur. Radiol. 2000; 10: 1976-1982
16) Mohr DN, Turner DW, Pond GR, Kamath JS, De Vos CB, Carpenter PC. Speech recognition as a transcription aid: a randomized comparison with standard transcription. J Am Med Inform Assoc. 2003; 10: 85-93