Analysis of the Implementation and Impact of Speech-Recognition Technology in the Healthcare Sector

Save to myBoK

by Ronaldo Parente, PhD, Ned Kock, PhD, and John Sonsini

Abstract

This paper develops a conceptual framework and offers research propositions for understanding the adoption of speech-recognition technology, drawing from Rogers's work on the diffusion of innovation, from interview findings, and from case study analysis. The study's focus was the analysis of the implementation of speech recognition and its impact on performance in the healthcare industry. Our interview findings indicated that, while there is still much room for improvement in the way speech-recognition technology is adopted and implemented, this particular technology has had a significant impact on the ability of healthcare providers to operate more cost effectively and provide a better level of patient care.

Introduction

The healthcare industry has seen dramatic changes in the last decade and has been a major focus of public policy. The industry has undergone, and continues to undergo, radical process changes in the delivery of its products; the effect of these changes on the customer has been dramatic and anything but smooth. Advances in medicine and healthcare have created an environment in which access to quality care and disease prevention is one of the most valued commodities our society has to offer. Therefore, the individual's access to sound healthcare is a major measure of our quality of life.[1]

The advances in technology and innovation are changing business markets, products, and relationships and forcing executives to become familiar with these cutting-edge technologies and innovations in order to jump ahead of their competitors. Cost effectiveness and superior customer service have become key factors for service organizations.[2,3] Thus, the broad aim of this study was to examine the adoption of speech-recognition technology and its impact on performance in the healthcare sector to contribute to the understanding of the implementation and adoption of information technology (IT) in the service industry context.

Since the focus is on theory construction (that is, elaboration of constructs and propositions), it is important to explore a wide range of approaches and perspectives in the context of implementing speech-recognition technology. We collected qualitative data through interviews of healthcare professionals and a case study approach. Based on these data and the literature review, we developed a conceptual framework, depicted in Figure 1, which builds on the theory of diffusion of innovation.[4-7] Our research is qualitative and theory building, specifically in regard to speech recognition in the healthcare sector, and does not intend to test a more general theory.

Figure 1: Conceptual Framework
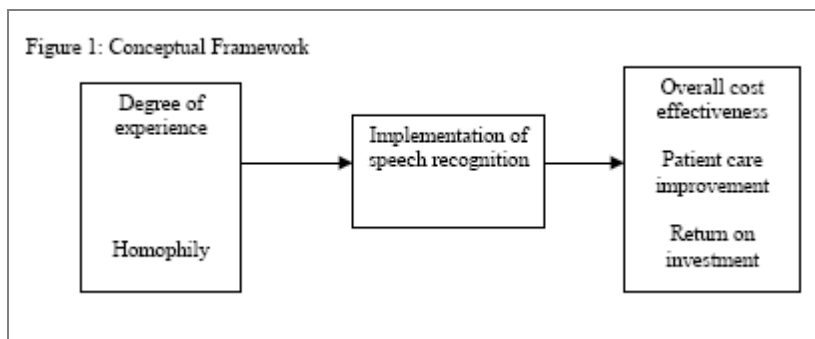
Researchers have argued that there is increasing pressure to reorganize hospitals into more cost-effective environments for information gathering and sharing; in the past hospitals were more like organizations composed of isolated departments and isolated points of delivery.[8] These process changes are being driven by reduced reimbursements for Medicare patients, a shift to managed care and prepaid contracts, and a growing focus on cost effectiveness, quality assurance, and clinical outcomes. Projections from the American Hospital Association indicate that hospital revenues could decrease by as much as 20 to 30 percent at the same time costs increase as a result of the greater volume of required documentation.[9] According to Menduno,[10] competitive pressures in the industry have led to a significant investment in IT systems. In 1997 the healthcare industry spent $15 billion on IT. According to our fieldwork, investment analysts at Volpe Brown Whelan & Company estimate that healthcare wastes as much as $270 million on inefficient IT systems alone.

As a theoretical background our study used Rogers's diffusion theory, and we looked at factors facilitating the adoption of speech-recognition technology and its subsequent impact on performance.[11] Our model was validated through qualitative data collected in the dynamic industry where this emerging technology is introduced. Therefore, the level of analysis was technology in the hospital setting, more specifically, at the point-of-care stage, where doctors and other healthcare providers produce medical reports on the conditions of their patients.

Historical Development of Speech-Recognition Technology

The traditional process by which physicians generate and store patient records can be described as doctors scribbling notes on a patient chart for later transcription, filing, and entry into a database. This database, in turn, contains information about all patient services, including x-rays, blood tests, and reports of other procedures and their outcomes. Many of these forms are templates used to ensure quality of treatment and accuracy of information for reimbursement purposes.[12]

In an environment of decreasing revenues and growing costs, the quality of medical reporting greatly affects how effectively hospitals bill for reimbursement.[13] Hospitals are natural targets for such process improvements as process redesign to increase knowledge flow and knowledge sharing, as well as maintaining or improving the quality of care while reducing costs and liabilities. The challenge facing hospitals is how to improve the process of medical reporting and documentation with innovative technology. For this improvement to work, doctors must give up their customary ways of recording patient care information and adopt new ones.

Even before the radical change in the healthcare industry, healthcare administrators had adopted many innovations as they attempted to improve patient documentation processes. The typewriter may be seen as a first attempt to accomplish this goal, followed by the mainframe computers of the 1950s, which were used primarily for cost and billing purposes. These older systems did not take into account that as much as 75 percent of healthcare costs are determined by the provider of the healthcare. In addition, the systems were incapable of providing data on patient outcomes and instead relied on manual chart reviews.[14]

Speech-recognition software is designed to create text from speech. In other words, instead of using a keyboard to enter text, users talk to the device, and the program types the text. The vision for adapting speech-recognition technology existed long before any real-life practical adaptations were possible. Finally, in the late 1980s and early 1990s speech-recognition technology found its first niche in the marketplace. This niche comprised activities in which users needed to operate computers but did not have a free hand to punch keys or manipulate a mouse.[15, 16] An example from the music industry involves a keyboard player in the recording studio attempting to control other components, particularly rhythm machines, without interrupting his play on the synthesizer. By rigging up an early speech-recognition system to the drum machine input, the

musician was able to change tempos and rhythms by shouting speech commands, which were not picked up by the recording system.[17,18]

By the early 1990s speech-recognition systems found broader use in computerized toys and games, control of instruments and computers, data collection and dictation, and, most significantly, among populations that were denied the use of traditional inputs like keyboards. Persons with disabilities-especially those with hand/arm motor impairments and sight limitations-became a major market for speech-recognition device use.[19]

The healthcare industry first started implementing speech-recognition systems for medical reporting around 1994.[20] However, the first use of such systems was met with skepticism by doctors, who were reluctant to abandon traditional ways of performing their work.[21] The technology needed to be user friendly and accurate in order to carve a niche in the market, but the early systems were perceived as anything but user friendly. These systems still had frustratingly small vocabularies and were not programmed to understand medical terminology. Furthermore, a major barrier to acceptance existed in the industry.[22-24]

Early systems required doctors to learn how to "talk" to the computer rather than the computer learning to "listen" to the doctor. Doctors would be forced to adapt to the technology by changing their way of speaking.[25] Doctors with accents or even head colds found the systems to be frustratingly inaccurate. Compounding the problem were the characteristically slow processing times of the 486 computers of this era and the common inability of these speech-recognition systems to distinguish between background noise and the user's speech. Also, costs of the early systems were excessive. The Voice Med system by Kurzweil was nearly $27,000 per workstation.[26]

Continuous improvement in the technology of speech-recognition systems became imperative for hospitals so that their doctors would come to believe in the value of these systems. Therefore, the vocabularies built into these systems grew tremendously in both size and the degree to which they were tailored to the jargon and terminology of the medical profession.[27] The systems gradually became better at adapting to a particular user's speech, regardless of timbre, speech character, accents, or head colds. Accuracy rates rose dramatically, and doctors were no longer struggling for the "right" words for the system to understand and record.[28] Newer systems even provided each user with an opportunity to "teach" the computer to understand her own speech.[29]

Although the combination of all of these improvements made this technology significantly more user friendly and much better equipped to improve the process of medical reviewing, one final improvement-a continuous speech system-was needed to break down the barriers to user acceptance.

All speech-recognition systems can be broken down into two major categories: discrete speech and continuous speech.[30] Early discrete speech systems recognized individual words one at a time as they were spoken but required the speaker to pause between each word. This presented a huge limitation to a doctor who was under pressure to see many patients and had little time to worry about how long he was pausing between words. Continuous speech systems became available for use by healthcare providers in 1994 and have become common in many areas of hospital and general practitioner operations.[31,32]

The essential difference between discrete and continuous systems is that the continuous systems enable doctors to speak in their normal speech at a normal rate of dictation. This has been the single most important improvement in the technology as far as breaking down barriers to user acceptance in the healthcare industry. Continuous systems have revolutionized the use of speech recognition, and today they can handle up to 160 words per minute, with about 98 percent accuracy.[33] The increased accuracy accrues because these systems are more accustomed to the user's speech patterns, accents, and vocal character. Because the speaker is able to speak naturally, there is a much higher level of consistency in the speech and therefore in the system's ability to recognize the speech.

Research Methods

The purpose of the research was to construct theory. Correspondingly, the research approach was qualitative and the research design was one case study analysis.

Field Interviews

Field interviews explored the extent to which expectations regarding the implementation of speech-recognition technology were met in a hospital setting. Our fieldwork consisted of nearly 40 unstructured interviews over a period of approximately three months in several hospitals adopting and implementing speech-recognition technology. Because the purpose of this study was theory construction, it was important to explore a wide range of approaches and perspectives in the context of adoption and diffusion of speech-recognition technology during qualitative data collection. We made sure the sampling included administrators, staff members, and physicians from several hospital units.

We interviewed 22 physicians, four division administrators, six clerical personnel, and five nurses. For consistency, we also interviewed two outside physicians working in private practice who had previous experience with speech-recognition technology. Our respondents were from relatively diverse cultural and professional backgrounds and worked in several different hospital units. In conjunction with our case study analysis of the orthopedics division of a major pediatric hospital, our conceptual framework tried to link our interview findings to the theory of diffusion of innovation to highlight the paper's contribution to the management information science field and therefore show the potential of speech-recognition technology to change processes in industries besides healthcare.[34]

We tried to follow a relatively general format for the interviews. Interviewees were first provided with a brief description of the research project. Then, most of the terms and constructs used in this study were explained. Questions to each interviewee focused on several issues, as described in [Appendix A]. The personal interviews lasted on average about 30 to 60 minutes and were recorded (unless the interviewee requested otherwise). Interviews were conducted between July and December 2000 in the United States. The information obtained in these interviews, in conjunction with our literature review and case study analysis, provided important insights about the implications of speech recognition in the healthcare sector.

Case Study Analysis

We performed a more in-depth case study analysis of one orthopedics division, having both orthopedists and orthopedic surgeons, of a major pediatric hospital. For the purpose of our case study interview, data were complemented with secondary data in the form of publications about the organization, internal reports, transcripts of internal communications, manuals on procedures and technologies, and a variety of patient records.

Practice

The practice cares for about 20,000 patients per year, with most patients needing one or two visits. Other patients, especially those with spinal cord conditions, would be treated throughout their lifetime. The patients are referred for specialized orthopedic care by pediatricians or general practitioners (primary-care physicians). According to the fieldwork interviews, the referring group of physicians is needed to support the orthopedic practice. Therefore, the goal of the orthopedic staff is to establish and maintain the best working relationship with these referring physicians. The division is a nonprofit operation with revenues of $3.0 million to $3.5 million per year, which are used to offset operating costs. Overall, there were eight physicians, one division administrator, and a group of clerical personnel to support the physicians.

Background of Technology

Speech-recognition software applications comprise several basic components, such as the microphone, sound card, vocabulary, speaker profile, language model, and recognition engine. The microphone and sound card convert analog human speech into a digital waveform. The recognition engine uses speech-recognition codes to statistically match digitized sounds to words. To develop a match between speech and the system's vocabulary, two to three words are analyzed in sequence to determine the most likely word grouping. A large basic vocabulary of 300,000 words and the speaker profile help match the speech of the user and create text by the recognition engine. The speaker profile is a recording of the user's speech matched to a specific text.[35]

Initially the orthopedic division considered four major suppliers of speech-recognition software packages. All four providers offered packages that were designed based on similar technology to translate sounds to text but offered different features. The main providers in the market at the time were IBM, Lernout & Hauspie, Philips Speech Processing, and Dragon Systems. Dragon Naturally Speaking, a software package developed by Dragon Systems, Inc., was offered in a number of versions, including several professional packages for lawyers and physicians.[36] The version adopted at the hospital in the case study was the Dragon Naturally Speaking Mobil Medical Suite.

A medical environment requires a specific complex vocabulary. For Dragon to work in a medical environment, it must contain a separate medical dictionary. This separate special vocabulary of medical words interacts with the recognition engine. Once the speech-recognition software is attuned to the user's speech, documents can be created. In addition to word processing, the Dragon system can understand verbal commands to open applications or edit text.

Dragon uses a microphone and PC for standard dictation. A noise-canceling microphone is recommended to eliminate background noise. Since physicians like to be able to dictate away from their computers, Dragon can be adapted to be used with a handheld recorder. The NORCOM brand of handheld recorder was used in the case study. Many physicians like to dictate while they are treating patients, completing physical examinations, or reviewing laboratory tests. When it is convenient, the physician downloads the recorder into the PC. The Dragon software requires a 350-to 400-HZ processor and 96 MB of RAM.[37]

Dragon contains a 300,000-word dictionary. This vocabulary, in conjunction with a high-speed processor, allows the user to speak at a natural rate and still record speech. The large dictionary by itself, however, is not enough for the required accuracy. The user's speech needs to be compared with the system's vocabulary before it is possible to achieve the necessary accuracy in speech recognition.

The comparison between stored vocabulary and the user's speech is called the speech profile. The speech profile includes 6,000 preferred words in a standard document. For the system to work, the user initially reads the 10- to 15-page document into the system to develop a unique speech profile that the user then edits. During the editing process, the user's words are statistically compared with the words in the system vocabulary, thus telling the system what the pronunciation of the individual actually means.[38]

Case

Templates are available for different types of dictation. For example, if the orthopedic physician is describing the treatment of a leg fracture, she indicates the fracture template to the computer by a speech command or by swiping a bar code. Once the user indicates the type of template being used, the Dragon software uses macros to assist the doctor in filling out the form. The template forms contain all the standard language that is repeated in the type of report needed. Macros provide standardized descriptions of both "normal" and common disease manifestations. These macros can remind the physician to record blood pressure or order a needed test so that key steps in diagnosis and treatment are not missed or left undocumented.[39,40]

In addition to generating reports, Dragon software can be used to create an electronic patient record that-with properly constructed security measures-can either supplement a physician's patient management system or on a larger scale become part of a hospital information system. A clinical documentation system can be implemented for a level of documentation support beyond the creation of individual records.[41] Hospital information systems are designed to post patient information on a hospital-wide system. For individual physician practices Dragon documentation can support a practice management system.[42]

From our interviews and analysis of data, we concluded that some physicians were skeptical about this technology because they were not sure about its usefulness. There was agreement that the technology would require a lot of training before physicians would feel comfortable using it on a large scale. Also, they were not convinced about the actual benefits its implementation would generate. Despite the initial skepticism, the administrator invited representatives from all four manufacturers to make a presentation to the department to answer physicians' questions and explain the benefits of speech-recognition technology. One respondent said that, "Despite the group's skepticism...we have a vested interest in preserving our livelihood,...so we all finally agreed to consider the proposed technology."

As previously stated, after extensive analysis, Dragon Naturally Speaking was the system of choice, along with the medical suite, which provided the medical dictionary. The respondents indicated that several attributes of Dragon seem to add value, including its ability to function as a navigator and be used with equipment from another manufacturer and its compatibility with background recognition. As a navigator it could be used to interface with other programs. This versatility allowed the selection of a cassette recorder that would eliminate background noise, enabling the physician to dictate into it anywhere. According to the interviewees, the Dragon system was less expensive because it did not require the purchase of a major server.

Another respondent described the old process used before the speech-recognition system (see Figure 2): "It needed many different people to complete the report. [The report] went from person to person before finally being copied and sent off to the

main chart, the division chart, the referring physician, and the biller. You can easily see how delays happened."

Figure 2: Old Process
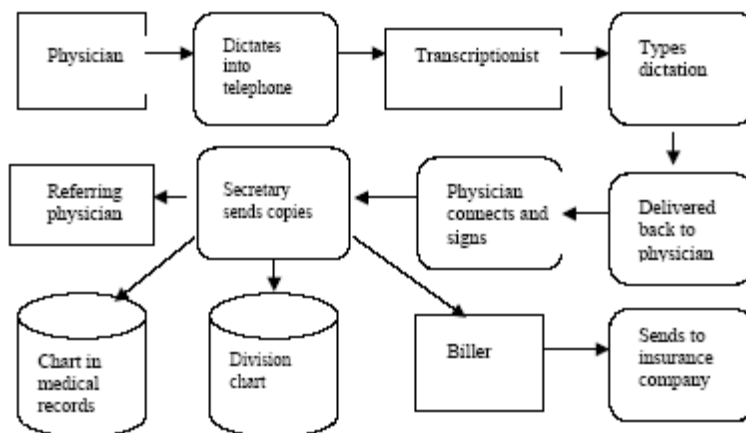
Once the speech-recognition system was implemented, the physicians felt it offered many advantages. First, they were pleased that there was little or no disruption to the way they had grown accustomed to practicing. The use of the cassette recorder and background recognition allowed this. They also saw a marked improvement in report turnaround time, from four days to 24 to 48 hours. An important byproduct was that the referring physicians received reports more quickly.

Moreover, it is easy to notice the reduction in the number of people and steps involved (see Figure 3). Currently secretaries perform the editing. If the practice grows, it may be of value to hire an editor to perform this job for all of the physicians. The division also continues to keep a division chart, which should not be necessary as the comfort level with this system improves; the ability to electronically access the chart through the permanent patient record should completely eliminate the need for double charts.

Figure 3: Current Process
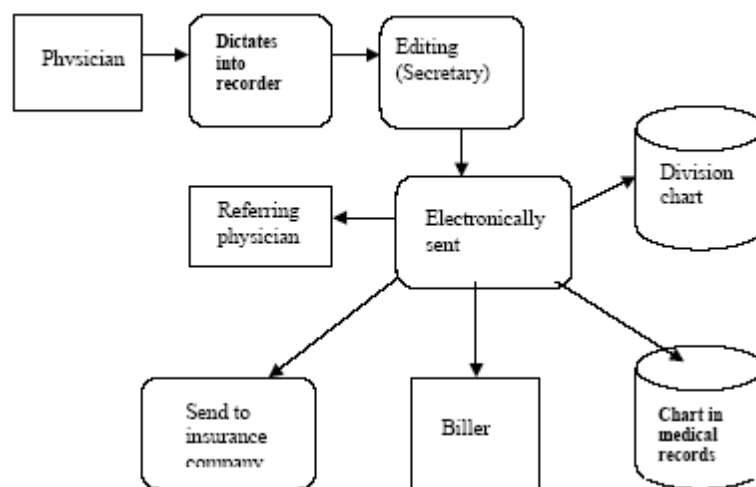
In addition to cost effectiveness considerations, we conclude from the case study that patient care has been enhanced by improvements in accuracy and report turnaround time. More specifically, the adoption of speech-recognition technology improves accuracy of the reports through greater legibility and offers more complete reports through accuracy, point-of-contact reporting, and prompting by technology. Finally, the technology provides quick access to patient data throughout the hospital and to outside providers.

Despite the advantages of the new system, our findings revealed some of the disadvantages of the system during use. Mainly, speech problems such as colds, stuttering, or the use of "ah," "um," and "huh" hinder success. Also, success was best when the

user talked quickly and most naturally. This is difficult for new users because they tend to be self-conscious when using a new system and want to talk slower. The administrators who were interviewed indicated that there was a consensus that these problems should resolve themselves with time. In one administrator's words, "We are confident that these are minor problems and that, as people become familiar with the system, the results will improve."

Conceptual Framework and Research Propositions

The methods and approach in our research are not regarded as theory testing. Our study is founded in a case study analysis and in the findings from interviews and qualitative data. The research is theory building in the sense that we developed a conceptual framework and offered some research propositions. Figure 1 depicts the conceptual model proposed in our study. We used the theory of diffusion of innovation as a guide for developing our research propositions, in combination with the findings from our fieldwork.

In constructing our conceptual framework, we examined some determinant factors revealed by the interview findings that seem to facilitate the adoption and implementation of speech-recognition technology and the impact of the technology on performance outcomes. Our framework also investigated the factors that affect physicians' acceptance and adoption of the technology and how its implementation can provide valuable insights into the challenges of implementing IT in the healthcare sector.

Diffusion of Innovation Theory

The proposed model was derived from our interview findings, case study analysis, and examination of the theory of the diffusion of innovation, which shows how innovations are adopted over a period of time. Diffusion theory is complex, and it is described and applied in a general sense in our study.[43] As defined by Rogers, diffusion is the process of communicating an innovation through certain channels over time among the members of a social system. It is a special type of communication in that the messages are concerned with new ideas. "Communication is a process in which participants create and share information with one another for the purpose of reaching mutual understanding."[44] The four important elements of diffusion are innovation, communication over channels, time, and the members of the social system.[45, 46] According to Rogers, the social system is defined as the set of interrelated units that are engaged in joint problem solving to accomplish a common goal.[47] In our case study, the social system is the orthopedic division of a major pediatric hospital and its physicians, nurses, and administrators. The innovation is the speech-recognition technology. Time is also an important dimension in the diffusion process; it encompasses the process that occurs over time as well as the time taken for an individual to adopt the speech-recognition technology.

The theory of diffusion of innovation allows for understanding the behavior of individual agents, such as physicians, who differ from employees in modern business organizations.[48, 49] The nature of diffusion theory seemed particularly appropriate for the current research, where action and intervention in real-life situations were required through our fieldwork in the orthopedic division of a major pediatric hospital. The development of our research propositions follows our fieldwork and the analysis of several internal documents and secondary data from the division being studied. The information content of the secondary data from the orthopedic division was supported by almost 40 unstructured interviews.

Several researchers have used the theory of innovation diffusion to study implementation problems.[50-52] Others have used organizational theory in the study of the diffusion of software practice.[53] Therefore, we combined what we learned from our qualitative data and from the literature on these theoretical lenses and applied it to the analysis of the implementation and effect of speech-recognition in the healthcare sector.

Technological Experience and Homophily

Our interview findings suggest that two constructs, prior technological experience and homophily, have an important role in facilitating the implementation of speech-recognition technology. Prior technological experience and the degree of user experience and exposure to other technological innovations seem to be major factors for successful implementation. Respondents who had been exposed to other technological innovations seemed to be more open to the adoption of speech-recognition technology than those who had not had such exposure.

The majority of the physicians interviewed indicated that the ease of use (of the Dragon system) played an important role in user acceptance of the new system. One respondent mentioned that, "Its background recognition capability was definitely a

plus, because it meant that physicians would be able to function just as they did before." This background recognition feature was perceived to be easy to use and allowed physicians to simply dictate their findings into a handheld recorder. "Then someone would edit the speech and turn it into text for the report." Often physicians made comments relating the use of speech recognition with their previous experiences with other technologies. One physician stated that, "Those [young physicians] who are more computer savvy and are already familiar with the new software and technologies seemed to express little resistance to the new system." Our interview findings indicated that the perceived ease of use of the technology is associated with previous "technological" experience and that this perceived ease-of-use attribute played an important role in reducing resistance and facilitating the adoption and implementation of the technology by physicians.

Our fieldwork suggested that communication channels, as interpersonal communication with respect to a new idea, are the essence of the diffusion process. Rogers defines "homophily" as the degree to which individuals who interact are alike in terms of beliefs, education, social status, and similar attributes.[54] In agreement with Rogers, our interview findings indicated that more effective communication occurs when two individuals share common meanings and a mutual subculture language and are alike in personal and social characteristics.[55] In other words, they are "homophilous." Based on the arguments above, we suggest the following propositions:

Proposition 1: The level of physicians' exposure to and previous experience in using different information technologies will influence their attitude toward the adoption and implementation of speech-recognition technology.

Proposition 2: The extent to which physicians are homophilous, or share beliefs, education, social status, and other similar attributes, will influence their attitude toward the adoption and implementation of speech-recognition technology.

Speech Recognition and Performance Implications

Our normative framework and propositions suggest that successful implementation of the speech-recognition technology will positively affect performance in the healthcare industry. Our in-depth interviews with physicians and administrators found agreement that emphasis should be put on three basic performance outcomes when considering the evaluation of a new innovation: improving the overall cost effectiveness of the operation, improving patient care, and return on investment.

One interview respondent stated, "The electronic record helps users to meet Health Care Financing Administration [HCFA; now known as the Centers for Medicare & Medicaid Services, CMS] requirements for reimbursement." Another respondent said, "The patient information can be integrated into an electronic patient record with speech-recognition software by using add-ons from value-added resellers." This electronic record can be in the physician's own patient management system or part of a larger hospital information system.[56] In general, respondents agreed that good speech-recognition software should meet CMS standards and provide an interface with feedback to the physician through the examination process. One respondent stated, "Speech recognition would provide the maximum assurance that [CMS] regulations are met. This is a critical issue for maximizing revenues, as the completeness of the report is the key factor in determining reimbursement levels."

The administration of the orthopedic division viewed the adoption and implementation of the speech-recognition technology as an innovative way to lower its operating costs. Collectively, our interviews indicated several major driving forces behind the need to reduce operating costs. First, changes in the reimbursement structure signaled a reduction in revenues. Second, the administration had plans to expand the practice. Third, because of current billing constraints, when a physician joined the practice, there was a six- to eight-month delay in the recovery of billed revenue. There was also some concern over increasing transcription costs, which were around $130,000 per year. In addition, the current system resulted in at least a four-day delay in getting reports returned to the office, which meant that the referring physician might not see the report for up to 10 days.

Physicians and management personnel had scrutinized the transcription costs before the administrator suggested the adoption of speech-recognition technology as a possible solution for lowering the operating costs. The team in charge of the process for finding a solution to reduce costs reviewed the speech-recognition technology and all other systems available at the time. Systems considered included ones developed by Phillips, IBM, Digital, and Dragon. When asked why they chose Dragon, a team member answered that, "All the systems had the same basic function, so we decided to look at the value added by each system and Dragon provided the most value for the cost."

We specifically asked respondents how they thought the speech-recognition technology was affecting operational costs. One respondent said, "We are sure that this technology will have a positive impact in our cost effectiveness, mainly by reducing the cost of transcription and the costs associated with physicians' time spent to complete medical reports." Another respondent

stated, "This technology increases the accessibility of the medical report throughout the hospital system and lowered liability insurance costs by improving thoroughness of reports." In addition, liability insurance providers have come to view speech-recognition systems at the point of patient care as a major contributor to risk mitigation, mainly because these systems are believed to create more comprehensive and accurate reports. Hence, liability premiums are often reduced when speech-recognition technology is introduced to the medical review process.[57]

Drawing from theory and our fieldwork findings, we expect speech-recognition technology to have a positive impact on the healthcare provider's cost effectiveness by reducing the cost of transcription, time spent by physicians, and time spent to complete medical reports; by increasing accessibility of the medical report throughout the hospital system; and by lowering liability insurance costs by improving the thoroughness of the report. Therefore, based on the arguments developed above, we suggest the following proposition:

Proposition 3: The extent to which speech-recognition technology is adopted and implemented will influence the overall operational cost effectiveness of the business unit where it is being implemented.

Patient Care

Another important performance indicator for healthcare providers is the improvement in patient care. Following the same rationale, and again building from our interview findings in the healthcare industry, we expect that the adoption and implementation of speech-recognition technologies should positively affect patient care through some specific process improvements. One respondent mentioned that, "The adoption of speech-recognition technology improves accuracy of the reports through greater legibility and offers more complete reports through accuracy, point-of-contact reporting, and prompting by technology." In other words, all these characteristics and improvements relate to patient care in one way or another. Speech-recognition technology provides quick access to patient data throughout the hospital and to outside providers, which, in turn, should have a positive impact on patient care. Therefore, we propose the following:

Proposition 4: The extent to which speech-recognition technology is adopted and implemented will influence patient care improvement.

Despite the challenges associated with implementing speech-recognition technology, respondents unanimously seemed to believe the technology was worthy of further evaluation, and our case analysis and interview findings support the assessment that use of the technology should enhance return on investment.

Investment and Costs

After assessing the initial investment to implement the speech-recognition system in the orthopedic division, we estimated that the computers that would be used with the new system needed to have sound card and memory upgrades that would cost $300 per computer. Initially only one computer was upgraded, but an additional four would eventually need to be upgraded. In addition, 10 NORCOM cassette recorders were purchased. These devices had the ability to buffer and reduce sound and cost $150 each. Software (that is, the speech file) would cost $900 per user, with the need for the capability of eight speech files.

The division knew that incidental costs would occur throughout the year but did not anticipate they would be significantly high. In addition, the hospital's information system division had agreed to support the system technically. Below, we outline the initial setup costs and note that these costs were relatively low compared to the yearly cost of transcription services without the system (see Table 1) for the orthopedic division. When comparing the costs associated with the implementation of speech-recognition technology to the transcription costs, the majority of respondents indicated that they expected the new speech-recognition system would pay for itself in about a month.

Table 1: Speech-Recognition System Set-up Costs

Item	Initial Cost	Overall Cost
Computer (memory, sound card)	$300	$1,200

Cassette recorders	$300	$1,500
Software (speech file)	$900	$7,200
Total		$9,900

Implementation

Speech-recognition system implementation in the studied orthopedic division started with one physician, who performed his initial standardized dictation in about 30 minutes. Then, an administrator edited the dictation. The physician began using the speech-recognition system as soon as he met with patients, and the administrator continued to edit and create the individual speech file for him. Initial success for the recognition was in the range of 80 percent accuracy. After eight to 10 sessions, it increased to 90 percent and ultimately achieved 95 percent accuracy. Technology standards indicate that a greater than 75 percent success rate will result in payback on investment. Subsequently, each additional physician was added in the same way, and currently six physicians are using the system.

Respondents indicated that the setup costs were paid back in about a month, and implementation of the system resulted in a 50 percent reduction in annualized cost of transcription-related services, amounting to $60,000. Although the goal of the division was a 70 percent reduction, or $84,000, as templates are developed, a further cost savings could be realized if there is a subsequent increase in insurance reimbursement.

In addition, the hospital's bill for a patient service is entirely dependent on the medical report compiled for that patient. Since a speech-recognition system increases the comprehensiveness of the report, the bill itself will be more complete and result in increased revenue.[58] Moreover, liability insurance providers believe speech-recognition systems at the point of patient care to significantly mitigate risk because these systems are thought to create more comprehensive and accurate reports. Therefore, reductions in liability premiums are often realized when speech-recognition technology is introduced to the medical review process.[59] Therefore, we propose the following:

Proposition 5: The extent to which speech-recognition technology is adopted and implemented will influence the return on investment.

Discussion and Implications

Healthcare systems have searched for years for better ways to document and store patient information. Although databases and servers that can store and transmit health information exist, there has been no easily accepted way of getting the data into these systems. Physicians and other healthcare providers must be able to provide care at a bedside or in the office. Unfortunately, they traditionally have wanted to be able to continue to provide care and document results in the same manner through the years.

Speech recognition has been around for some time as a technology, but only recently has it been worth the energy needed for general use. One problem was that the first systems were slow and required speech patterns that were unnatural and too time consuming. These problems hindered acceptance and implementation of the technology. With the advent of continuous speech-recognition systems, however, the tide changed. Newer systems are user friendly, and improvements have made the technology easier for the consumer to use. Therefore, a whole new area of use has opened up for such systems.

This study developed a conceptual framework and provided an analysis of the implementation of speech-recognition technology and its impact on performance. A qualitative approach was used in order to present evidence of the antecedents and benefits of the adoption and implementation of the technology. Our model suggests that two important factors facilitate the adoption and implementation of speech-recognition technology: degree of experience through prior exposure to other technologies and degree to which individuals who interact are alike in terms of beliefs, education, social status, and similar

attributes (that is, homophily). In addition, we examined the potential for the implementation of speech-recognition technology to affect such performance outcomes as reduction in operational cost, improvement in patient care, and return on investment.

From a theoretical perspective, the process of implementation of speech-recognition technology conforms to Rogers's idea that organizations scan continuously for innovations and match a promising innovation with one of their own problems.[60] Our case study findings show that, at first, a small group decided to adopt the innovation by consensus; then, this group put pressure on all other individuals to conform to the decision. Therefore, one important theoretical contribution of this research is the development of a conceptual framework that builds from Rogers's theory of diffusion of innovation and is supported by our fieldwork data.[61] Our findings indicated that this group of change agents assisted the process of implementation of speech-recognition technology and that the implementation process was relatively complex and required a considerable amount of training and support to counter this complexity.

From a managerial perspective, our conceptual framework and qualitative data indicate that to develop individual behavioral acceptance of the use of such an innovative technology, it is important to encourage and cultivate a positive attitude toward using speech-recognition technology as positively affecting performance. Our framework thus should have important implications for managers in the healthcare industry and possibly for managers in other service industries, where the adoption of speech-recognition systems can have a profound effect on overall performance indicators. This can help managers understand speech-recognition dynamics, the process of how to facilitate the technology's adoption, and the mechanisms through which the technology affects performance outcomes.

An important insight from our research is that the need to lower operating costs because of decreasing reimbursement rates from insurance companies has forced healthcare providers to look to nontraditional means of improving efficiency. The implementation of speech-recognition technology in the orthopedics division examined has positively affected cost effectiveness and patient care. Therefore, our research offers some support for the argument that investment in speech-recognition technology to improve the healthcare documentation process is not a waste. In addition, while the initial reluctance of doctors to accept this technology made the transition to computer-based reporting difficult, continuous improvements in the technology finally paved the way for reducing physicians' skepticism regarding the acceptance of new ideas.

Research Limitations and Future Research

The conceptual framework and findings of this research provided important managerial and theoretical insights into the adoption and implementation of speech-recognition technology. Notwithstanding these insights, several limitations should be addressed. First, this study was subject to the usual limitations inherent in a qualitative research design. Second, we must be cautious in drawing conclusions outside the scope of this research. These results are from industry-specific research, with data collected from one specific case study. Therefore, we must be careful in trying to generalize its findings. Applying the framework to different technologies and professional user groups over a period of time would allow for the generalization of the conceptual model.

Therefore, future research is warranted to address these limitations and expand the theoretical validity of our framework. One opportunity for future investigation that deserves further attention is the potential for conducting comparative multiple-case analyses. This could be a series of case studies exploring a variety of contexts adopting speech-recognition technology in different industries. This technology currently is evolving, but the question of whether implementation of speech-recognition technology can be relevant in other industries remains. New, more advanced, and easy-to-use systems are likely to be developed in the near future. In turn, these technological advancements should have applications not only in the healthcare service industry, but also in many other information-intensive industries where shared knowledge can be used as leverage to improve firm performance and competitive advantage. The congruence between our interview findings and qualitative data and Rogers's diffusion theory suggests that this generalization may be possible.

Researchers could also try to develop a specific measurement scale of the constructs in our framework as well as try to incorporate other potential relevant constructs to extend our conceptual study model. The next step would be to perform a large-sample survey to collect empirical data and validate and test the proposed framework.

Conclusion

In the future, speech-recognition technology could also greatly benefit the provider who must create documentation at the inpatient bedside. Therefore, continued use of speech-recognition technology outside this area will need to occur in order to validate its worth. Perusal of the current literature on speech recognition alone provides an appreciation of how it is being applied in healthcare. Could this be a technology that 10 years from now will be the primary way healthcare professionals document patient care? Although we cannot now answer this question with complete confidence, it would not be a far-fetched notion. Just think of how the future development of computer systems will even further influence the use of this technology. Every hospital and healthcare organization should be making progress in the use of this technology in order to prepare for the future.

Ronaldo Parente, PhD, is assistant professor of management in the Franklin Perdue School of Business, Salisbury University in Salisbury, MD. Ned Kock, PhD, is associate professor of management information systems in the College of Business Administration at Texas A&M International University in Laredo, TX. John Sonsini received his MBA from Temple University in Philadelphia, PA.

Acknowledgements

Notes

1. Kramer, Teresa. "ROI for an Electronic Medical Record." Behavioral Health Management 19, no. 1 (1999): 28-31.
2. Graham, Ian. TQM in Service Industries: A Practitioner's Manual. New York: Technical Communications, 1992.
3. Kettinger, William J. "Global Measures of Information Service Quality: A Cross-National Study." Decision Sciences 26, no. 5 (1995): 569-588.
4. Rogers, Everett M. Diffusion of Innovations. New York: The Free Press, 1962.
5. Rogers, Everett M. Diffusion of Innovations, ed. 3. New York: The Free Press, 1983.
6. Rogers, Everett M. Diffusion of Innovations, ed. 4. New York: The Free Press, 1995.
7. Rogers, Everett M., and F. F. Shoemaker. Communication of Innovations: A Cross-Cultural Approach, ed. 2. New York: The Free Press, 1971.
8. Asper, Guilhermo, Carlos Santana, and Ned Kock. "Strategies for Building Networks: Toward a Brazilian Health-Care Value Chain." In: Muffatto, Moreno and Kulwant S. Pawar (Editors). Proceedings of the Seventh International Symposium on Logistics. Padova, Italy, 2001, pp. 87-93.
9. Pierce, Garvin. "Can We Talk? Voice Recognition Technology Is Speaking to Evaluation and Management Documentation." Healthcare Informatics 16, no. 2 (1999): 157-158.
10. Menduno, Michael. "Why Internet Technology Is the Next Medical Breakthrough." Hospital & Health Networks 72, no. 21 (1998): 28-35.
11. Rogers, Everett M. Diffusion of Innovations. 1995.
12. Pierce, Garvin. "Can We Talk?"
13. Asper, Guilhermo et al. "Strategies for Building Networks."
14. Darnell, Tim. "Devices as Dynamic as the People Who Use Them." Health Management Technology 18, no. 2 (1997): 78-83.
15. Bennett, Robert. "Operating Computers by Voice." Accent on Living 39, Spring (1995): 78.
16. O'Leary, Stephanie. "Voice Recognition Systems: Part 2." Paraplegia News 49, no. 9 (1997): 65-66.
17. Bennett, Robert. "Operating Computers by Voice."
18. Gold, Ben and Nelson Morgan. Speech and Audio Signal Processing: Processing and Perception of Speech and Music. New York: John Wiley and Sons, 1999.
19. Bennett, Robert. "Operating Computers by Voice."
20. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition." American Medical News 40, no. 10 (1997): 19-21.
21. Hoffman, Thomas. "Techno-Phobic MDs Refuse to Say 'Ah!' (Doctors' Reluctance to Use Computer Technology for Making Patient Records)." Computerworld 31, no. 8 (1997): 75-76.
22. Bennett, Robert. "Operating Computers by Voice."

23. Williams, Laurie. "Technologies in Other Industries Can Save Dollars in Healthcare." Health Management Technology 16, no. 10 (1995): 16-18.
24. D'Allegro, Joseph. "Voice Recognition Systems Gain with Doctors." National Underwriter 102, no. 45, Nov. 9 (1998): 7-16.
25. DeLambo, Chris. "Pick the Right Input Device for Your POC Systems." Health Management Technology 17, no. 6 (1996): 24-27.
26. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition."
27. Padilla, Enrique. (1997). "Voice Recognition Technology Overview." Available at: http://www.pbol.com/voice_recognition/voice_article.html.
28. Pierce, Garvin. "Can We Talk?"
29. Gold, Ben and Nelson Morgan. Speech and Audio Signal Processing.
30. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition."
31. Bennett, Robert. "Operating Computers by Voice."
32. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition."
33. Ibid.
34. Rogers, Everett M. Diffusion of Innovations. 1995.
35. Pierce, Garvin. "Can We Talk?"
36. Ibid.
37. DeLambo, Chris. "Pick the Right Input Device for Your POC Systems."
38. Gold, Ben and Nelson Morgan. Speech and Audio Signal Processing.
39. D'Allegro, Joseph. "Voice Recognition Systems Gain with Doctors."
40. Borzo, Greg. "What You Say Is What You Get." American Medical News 36, no. 10 (1995): 29.
41. Pierce, Garvin. "Can We Talk?"
42. Ibid.
43. Rogers, Everett M. Diffusion of Innovations. 1995.
44. Ibid.
45. Ibid.
46. West, Linn et al. "Beef Producers Online: Diffusion Theory Applied." Information Technology & People 12, no. 1 (1999): 71.
47. Rogers, Everett M. Diffusion of Innovations. 1995.
48. Ibid.
49. West, Linn et al. "Beef Producers Online."
50. Brancheau, James Clayton and James C. Wetherbe. "The Adoption of Spreadsheet Software: Testing Innovation Diffusion Theory in the Context of End-User Computing." Information Systems Research 1 no. 2 (1992): 115-143.
51. Johnson, Bonnie McDaniel, and Ronald E. Rice. Managing Organizational Innovation. New York: Columbia University Press, 1987.
52. Moore, Gary. "End-User Computing and Office Automation: A Diffusion of Innovations Perspective." INFOR 25, no. 3 (1987): 214-235.
53. Zmud, Robert W. "Diffusion of Modern Software Practices: Influence of Centralization and Formalization." Management Science 28, no. 12 (1982): 1421-1431.
54. Rogers, Everett M. Diffusion of Innovations. 1995.
55. Ibid.
56. Cole-Gomolski, Barb. "Hartford Hospital Seeks Competitive Edge with $1.7m Net." Computerworld 33, no. 7, Feb. 15 (1999): 43.
57. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition."
58. Hoffman, Thomas. "Techno-Phobic MDs Refuse to Say 'Ah!'"
59. Fishman, Eric. "Speech-Recognition Software Improves, Gains Recognition."
60. Rogers, Everett M. Diffusion of Innovations. 1995.
61. Ibid.

Appendix A

Type of questions asked of interviewees

- Perceptions regarding competition, future market trends, and customer demands
- Cultural backgrounds and specific training
- Understanding about the meaning of terms such as homophily, speech-recognition technology, and diffusion of innovation (we provided them with definitions of these terms if their understanding of the terms differed from our definitions)
- Specific operational arrangements that a hospital adopting speech recognition uses
- Factors that foster or discourage the adoption of speech recognition
- Their impressions of and perspectives on the acceptance of the technology
- Whether they felt that speech recognition had been a major improvement
- What they considered to be major advantages and disadvantages of the technology
- What they perceived to be the major determinants facilitating acceptance of the technology
- Their perception of the value of speech-recognition technology
- Their prior experience with and exposure to other information technologies
- How labor and human resources issues affect the speech-recognition decision
- The positive and negative implications of speech recognition
- The characteristics of the buyer-supplier contracts and relationships with the technology providers
- The level of responsibility transferred to physicians
- Specific knowledge management tools used in the speech-recognition arrangement
- Specific types of organizational arrangements dealing with the implementation of speech-recognition technology
- Issues related to the capital investments for speech recognition

---