Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 67 (2009) 36–49
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

# Acceptance of speech recognition by physicians: A survey of expectations, experiences, and social influence

Alexandre Alapetite[a,b,*], Henning Boje Andersen[a], Morten Hertzum[b]

[a]*Systems Analysis Department, Risø National Laboratory, Technical University of Denmark, DK-4000 Roskilde, Denmark*
[b]*Computer Science, Roskilde University, Universitetsvej 1, P.O. Box 260, DK-4000 Roskilde, Denmark*

Received 30 October 2007; received in revised form 20 June 2008; accepted 15 August 2008
Communicated by C.M. Karat
Available online 22 August 2008

**Abstract**

The present study has surveyed physician views and attitudes before and after the introduction of speech technology as a front end to an electronic medical record. At the hospital where the survey was made, speech technology recently (2006–2007) replaced traditional dictation and subsequent secretarial transcription for all physicians in clinical departments. The aim of the survey was (i) to identify how attitudes and perceptions among physicians affected the acceptance and success of the speech-recognition system and the new work procedures associated with it; and (ii) to assess the degree to which physicians' attitudes and expectations to the use of speech technology changed after actually using it. The survey was based on two questionnaires—one administered when the physicians were about to begin training with the speech-recognition system and another, asking similar questions, when they had had some experience with the system. The survey data were supplemented with performance data from the speech-recognition system. The results show that the surveyed physicians tended to report a more negative view of the system after having used it for some months than before. When judging the system retrospectively, physicians are approximately evenly divided between those who think it was a good idea to introduce speech recognition (33%), those who think it was not (31%) and those who are neutral (36%). In particular, the physicians felt that they spent much more time producing medical records than before, including time correcting the speech recognition, and that the overall quality of records had declined. Nevertheless, workflow improvements and the possibility to access the records immediately after dictation were almost unanimously appreciated. Physicians' affinity with the system seems to be quite dependent on their perception of the associated new work procedures.


*Keywords:* Speech recognition; Technology acceptance; Electronic medical records

## 1. Introduction

Speech recognition has been refined and become more robust in recent years (Lai et al., 2008). The gradual maturation of the technology has been accompanied by adoptions of the technology in the medical domain, where it is used to enter comments into the electronic medical record (EMR), thus replacing the standard way of entering notes by physician dictation and subsequent transcription by medical secretaries or a dedicated service (Zafar et al., 1999; Al-Aynati and Chorneyko, 2003). At the same time as the technology has matured, speech recognition has been developed and implemented for languages spoken by much "smaller" populations, such as Danish (5.4 million speakers).

Vejle and Give Hospital, Denmark, has been one of the first hospitals to introduce speech recognition for all major specialties and departments. Having run a successful project on speech recognition in its radiology department since 2000, this regional hospital (349 beds, and 217 000 outpatients in 2006) began to implement plans for having all physicians in clinical departments use speech recognition to input physician notes and instructions into the

*Corresponding author at: Systems Analysis Department, Risø National Laboratory, Technical University of Denmark, DK-4000 Roskilde, Denmark.
*E-mail address:* alexandre@alapetite.net (A. Alapetite).

1071-5819/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2008.08.004

EMR. The speech-recognition system—software based on Philips Speech Magic, adapted to Danish and deployed by Max Manus A/S—was rolled out in all clinical departments in 2005–2006, and has about 240 physician users as of 2007.

The main purpose of introducing speech recognition across all departments was to ensure a quicker completion of medical record entry and to achieve a higher quality of patient records. The old transcription service was known to sometimes produce backlogs of dictation tapes waiting to be transcribed by medical secretaries, or transcriptions waiting to be checked and approved by physicians. Additionally, an expected consequence was to allow secretaries, who would no longer need to spend time on transcriptions, to take over other duties. It was hoped that the quality of medical records would be enhanced, since physicians would now be going to check and revise their written (speech recognized) record immediately while their intentions were still fresh in memory. While little is known so far about the impacts of speech recognition on the various stages of the writing process and on the quality of outcome (Honeycutt, 2003), the above-mentioned goals fully match criteria such as those reported by Mönnich and Wetter (2000).

The present study had two related objectives: First, to identify physicians' attitudes and expectations about speech recognition that might explain their subsequent level of satisfaction with actual use of the technology. Second, to assess possible changes between prior expectations to and subsequent experience with the technology as a replacement for the traditional mode of producing medical records.

## 2. Related work

Work about the acceptance of speech recognition falls into two main areas: speech recognition and technology acceptance. Studies of speech recognition have predominantly been devoted to recognition of spoken English. However, recognition rates of systems that recognize English are not necessarily transferable to a speech-recognition system for Danish.

### 2.1. Speech recognition

For free-text dictation, speech recognition combines some characteristics of traditional dictation and of word processing (Leijten and Van Waes, 2005): on the one hand, quick and easy use of speech, and on the other, instantaneous graphical feedback and the possibility of jumping back and forth in the text. At the same time, speech recognition has its own advantages and drawbacks.

For transcription of free text, state-of-the-art systems correctly recognize 72–98% of the spoken words according to recent research (Zafar et al., 1999, 2004; Devine et al., 2000; Jungk et al., 2000; Ramaswamy et al., 2000; Kanal et al., 2001; Sears et al., 2001; Al-Aynati and Chorneyko, 2003; Alapetite, 2008), while commercially reported recognition rates are generally above 95%. Several factors contribute to the differences in recognition rates across studies:

- Vocabulary affects speech recognition through its size and domain coverage. Large vocabularies with good domain coverage are attractive, simply because they enable recognition of more words. Conversely, the acoustic distinctiveness of words is larger in small vocabularies, increasing the likelihood of correct recognition. Small vocabularies are, however, mostly relevant for voice navigation. State-of-the-art systems for text transcription have vocabularies comprising tens of thousands of words and optional, add-on vocabularies for specific domains such as the medical domain.
- Speakers influence speech recognition by the clarity and consistency of pronunciation and the degree of fit between their pronunciation and the acoustic model of the system. Speaker-dependent systems achieve higher recognition rates than speaker-independent systems but require one or more training sessions—based on which the system adapts its acoustic model to the speaker—and may be more sensitive to variations of the background noise, microphone, and voice (e.g., due to a cold). Even after training, atypical speakers, including non-natives (Coniam, 1999) as well as children and elderly (Wilpon and Jacobsen, 1996), experience lower recognition rates.
- Noise affects speech recognition in two ways: (a) it distorts the speech signal, making it more difficult to discern the spoken words. (b) In the presence of noise, people alter their voice in an attempt to counter the distortion of the speech signal (the Lombard effect; Lombard, 1911). Ambient noises, such as those heard in hospital wards or emergency rooms, do not significantly affect average speech-recognition rates (Zafar et al., 1999; Alapetite, 2008). However, in spite of numerous noise-cancellation techniques, loud noise, and even moderate levels of noise, may considerably degrade the performance of speech-recognition systems (Gong, 1995; Barker et al., 2005).
- All speech-recognition systems are based on principles of statistical pattern matching (Young, 1996). However, in spite of this commonality, individual systems differ in their parameterization of the speech signal, the acoustic model of each phoneme, and the language model used in predicting the words most likely to follow the preceding words. Thus, different systems make different recognition errors, even when they achieve similar recognition rates. This difference can be used to improve recognition rates by fusing the outputs of multiple systems (Fiscus, 1997; Alapetite, 2008).

Studies of text transcription show that it takes more time for a person to produce a text by voice input followed by correction of the recognition errors than by dictation

followed by proofreading after the text has been typed by a human typist whose time is not included in the comparison (Borowitz, 2001; Al-Aynati and Chorneyko, 2003). Thus, the freeing of typist time for other tasks is achieved at the expense of spending more of the speaker's time. Mohr et al. (2003) studied speech recognition as an aid for typists and found that editing a draft produced by speech recognition took longer than typing the audio-recorded text from scratch. The main time-related advantage of using speech recognition, as opposed to human typists, for text transcription appears to be a considerable reduction of the time from the production of the original dictation until the text is completed (Lai and Vergo, 1997; Ramaswamy et al., 2000; Borowitz, 2001).

Zafar et al. (2004) who reviewed recognition errors made by speech-recognition systems during text transcription found that 9.4% of errors were nonsense errors and 1.6% critical errors. The presence of nonsense and critical errors complicates error correction. Attempts at easing error correction by utilizing the confidence scores generated by speech-recognition systems have yielded mixed results (Suhm et al., 2001; Feng and Sears, 2004). Error correction can be made by voice commands, making text production entirely hands-free, but this is inefficient compared to making the corrections by keyboard and mouse (Suhm et al., 2001). Multimodal methods of text production are also recommended for ergonomic reasons (Juul-Kristensen et al., 2004).

### 2.2. Technology acceptance

Technology acceptance has been studied from many perspectives, including the theory of reasoned action (TRA; Fishbein and Ajzen, 1975), the theory of planned behaviour (TPB; Ajzen, 1985, 1991), diffusion of innovations (DOI; Rogers, 2003), and the technology acceptance model (TAM; Davis, 1989, 1993). These perspectives generally agree that technology acceptance concerns the adoption processes through which individuals decide to acquire and deploy a technology for a specified purpose. They differ, however, in the factors considered to influence the adoption process. Recently, Venkatesh et al. (2003) proposed a model that unified much of the previous work by encompassing an inclusive set of factors:

- *Performance expectancy*: "the degree to which an individual believes that using the system will help him or her attain gains in job performance" (Venkatesh et al., 2003, p. 447). Performance expectancy includes factors such as perceived usefulness (from TAM) and relative advantage (from DOI), which have been the strongest predictors of acceptance in previous studies. In the unified model performance expectancy was, likewise, a determinant of intention to use systems, and more so for men and younger employees.
- *Effort expectancy*: "the degree of ease associated with the use of the system" (Venkatesh et al., 2003, p. 450). Effort expectancy includes ease-of-use factors (from TAM and DOI), which have particularly been found to influence usage behaviour during early use of a system. In the unified model effort expectancy was, likewise, a determinant of intention to use, and more so for women, older employees, and with less experience using the system.
- *Social influence*: "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al., 2003, p. 451). Social influence includes subjective norm (from TRA and TPB) and image (from DOI). In the unified model, social influence was a determinant of intention to use, and more so for women, older employees, with less experience using the system, and when use was mandated.
- *Facilitating conditions*: "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system" (Venkatesh et al., 2003, p. 453). Facilitating conditions include perceived behavioural control (from TPB) and compatibility (from DOI). In the unified model, the effect of facilitating conditions was subsumed by effort expectancy, except for an effect on usage for older employees with experience using the system.

In the study by Venkatesh et al. (2003) the unified TAM explained 70% of the variance in individuals' intention to use systems. Many systems are, however, adopted in organizational contexts, which appear to be somewhat under-recognized in the unified model. Organization-level factors that affect the adoption of technologies include administrative intensity, centralization, external communication, functional differentiation, internal communication, managerial attitude toward change, professionalism, slack resources, specialization, and technical knowledge resources (Damanpour, 1991).

Studies of adoption in organizational contexts often find that it is a two-stage process involving a formal decision to adopt a technology followed by actual deployment of the technology by users (Fichman, 2000; Gallivan, 2001). This creates opportunities for lags between the formal, often organization-level decision, and subsequent local deployment by individuals. One reason for these lags is that the formal decision to adopt a technology and the decisions about actual deployment are typically made by different people, who may disagree. Another reason may be that different considerations are salient to the formal decision and to actual deployment. Specifically, unrealistic expectations during the formal decision to adopt may lead to disappointment among the first employees that actually deploy a technology and these disappointed expectations may, in turn, discourage and delay further deployment (Fichman and Kemerer, 1999). This way, unrealistic expectations produce a subtle combination of performance expectancy and social influence.

## 3. Survey method

A questionnaire was developed and deployed as a survey at Vejle and Give Hospital. The survey was divided into two phases, a prospective phase in which we surveyed physicians' expectations toward speech recognition and a subsequent retrospective phase where physicians' experiences with the technology were surveyed.

### 3.1. Participants

The survey participants were 186 anonymous physicians at Vejle and Give Hospital, about half of whom were introduced to speech recognition in 2005 to replace dictation and subsequent secretarial transcription, and the other half to be introduced to speech recognition as the study progressed during 2006. The departments involved were medicine, neurology, oncology, organ surgery, orthopaedic surgery, and otology.

### 3.2. Survey instrument

The survey instrument was a pair of related and overlapping questionnaires: A prospective one asking respondents about their expectations and attitudes to the use of speech-recognition technology as a front end to the EMR and a retrospective one asking them about their experiences with the technology. The two questionnaires partially overlap, asking respondents the same questions with only changes of tense. This allows us to compare, for each respondent, the answers given before and after the introduction and use of the target technology. The expectations questionnaire contains 23 closed questions (Likert-type or Yes/no) and one open item, and the experiences questionnaire contains 18 closed questions (Likert-type) and seven open items. The two questionnaires shared 10 closed questions, differing only in tense (cf. Appendix A). Prior to its administration, the survey instrument was refined through two rounds of informal pilot testing.

### 3.3. Procedure

The administration of the survey questionnaires followed the schedule for the introduction of the speech-recognition system at the different hospital departments (see Table 1). During 2006, the system was introduced successively into the otology, medicine, and oncology departments. About 1 month prior to their introduction to the system, physicians in each department received emails inviting them to answer the expectations questionnaire. When a department had been using speech recognition for about 4 months, physicians were once again invited by email to participate in the second phase of the survey, this time answering the experiences questionnaire. Physicians in the three departments mentioned completed both the expectations questionnaire and the experiences questionnaire.

Table 1
Schedule for the survey questionnaires

| | Expectations questionnaire | Experiences questionnaire |
|---|---|---|
| Medicine | May 2006 | September 2006 |
| Neurology | | August 2006 |
| Oncology | August 2006 | December 2006 |
| Organ surgery | | August 2006 |
| Orthopaedic surgery | | August 2006 |
| Otology | March 2006 | August 2006 |

Three additional departments completed the experiences questionnaire only. During 2005, speech recognition had been introduced at the orthopaedic surgery, organ surgery, and neurology departments. The physicians in these departments received the experiences questionnaire after they had been using speech recognition for 8–12 months. The additional data consolidate the analysis of the physicians' experiences using voice input.

The physicians were contacted via their professional email address. While participation in the survey was anonymous, each respondent had a unique identifier that enabled us to pair a respondent's expectations and experiences answers. For each of the questionnaires, two email reminders were sent to non-respondents.

### 3.4. Speech contribution rates (SCRs)

In addition to the data collected through the survey, the vendor of the speech-recognition system provided the average SCR for each physician for each month of 2006. The SCR represents the percentage of words that remain unaltered when a physician reviews a document produced by speech recognition and performs any manual corrections and modifications deemed necessary. At Vejle and Give Hospital, approval of the document is the responsibility of the physician who dictated it to the speech-recognition system. Thus, the SCR is similar to, but not identical with, a standard speech-recognition rate. While a speech-recognition rate compares the recognized text with the actual spoken text, the SCR compares the recognized text with the final text entered into the medical records. Thus, the SCR diverges from a speech-recognition rate when a physician not only corrects the recognized text for misrecognitions but also revises it by adding, deleting, or changing formulations compared to the originally spoken text. Physicians may also differ in their willingness to correct inconsequential misrecognitions. Lacking the data required for computing the speech-recognition rate, we find the SCR a useful measure of the system's work-related quality.

### 3.5. Response rate

The survey data will be grouped in two ways during the analysis. First, one set of analyses will investigate the correlations between expectations and experiences. These analyses are based on the data from the 39 physicians who responded to both questionnaires (response rate: 39%). Second, another set of analyses will investigate the respondents' experiences using speech recognition. These analyses are based on the 98 responses, including the 39 responses mentioned above, to the experiences questionnaire (response rate: 53%). A total of 112 questionnaires were received from the 186 physicians to whom invitations were distributed, yielding an overall response rate of 60%. Table 2 gives the response rates for the individual departments.

## 4. Results

### 4.1. Respondents versus non-respondents

To characterise the sample of the population who answered at least one of the two questionnaires when compared to the non-respondents, we compared their respective SCRs and average number of dictations, as reported in Table 3. Respondents had produced, on average, significantly more dictations ($t$-test, $p<.005$, equality of variances not assumed) and achieved significantly higher SCRs ($t$-test, $p<.01$, equality of variances not assumed) than non-respondents. This does not necessarily show that speech recognition worked better for the respondents than for the non-respondents but it indicates that respondents, on average, left more of the speech-recognized text unchanged than non-respondents.

### 4.2. Expectations versus experiences

Ten questions were included in both the expectations and the experiences questionnaire. For each of these questions Fig. 1 shows the percentages of positive (*agree completely* and *agree somewhat*), neutral (*yes-and-no*), and negative (*disagree completely* and *disagree somewhat*) responses as an upper, middle, and lower bar, respectively. The question receiving most positive responses showed that the physicians generally expected and experienced that their department heads approved of the introduction of

Table 2
Response rates for the survey

| Department | Physicians | Expectations questionnaire | | Experiences questionnaire | | Both expectations and experiences questionnaires | |
|---|---|---|---|---|---|---|---|
| | | Respondents | Response rate (%) | Respondents | Response rate (%) | Respondents | Response rate (%) |
| Medicine | 60 | 29 | 48 | 36 | 60 | 23 | 38 |
| Neurology | 24 | – | – | 10 | 42 | – | – |
| Oncology | 23 | 11 | 48 | 8 | 35 | 6 | 26 |
| Organ surgery | 20 | – | – | 11 | 55 | – | – |
| Orthopaedic surgery | 42 | – | – | 20 | 48 | – | – |
| Otology | 17[a] | 13 | 81 | 13 | 76 | 10 | 63 |
| Total | 186 | 53 | 53 | 98 | 53 | 39 | 39 |

[a]During the expectations survey only 16 physicians were employed in the Otology department.

Table 3
Comparison of respondents and non-respondents (in parentheses are the numbers of physician for whom usage data were not available[a])

| Department | Survey respondents | | | Non-respondents | | |
|---|---|---|---|---|---|---|
| | Physicians | Average dictations | Average SCR | Physicians | Average dictations | Average SCR |
| Medicine | 41 (+1) | 2609.6 | 84.6 | 13 (+5) | 1550.1 | 84.7 |
| Neurology | 9 (+1) | 2674.6 | 88.2 | 13 (+1) | 2505.1 | 75.7 |
| Oncology | 8 (+5) | 473.4 | 75.3 | 5 (+5) | 554.0 | 80.6 |
| Organ surgery | 9 (+2) | 3843.3 | 86.3 | 6 (+3) | 3575.5 | 86.8 |
| Orthopaedic surgery | 17 (+3) | 4423.5 | 89.4 | 19 (+3) | 2243.7 | 80.4 |
| Otology | 16 (+0) | 2855.8 | 88.3 | 1 (+0) | 1077.0 | 89.0 |
| Total | 100 (+12) | 2903.3 | 85.7 | 57 (+17) | 2116.6 | 81.2 |

*Note*: SCR—speech contribution rate.
[a]Data about their use of the speech recognition system were available for only 100 respondents (submitting at least one of the two questionnaires) out of 112 (89%) and for 57 non-respondents out of 74 (77%).

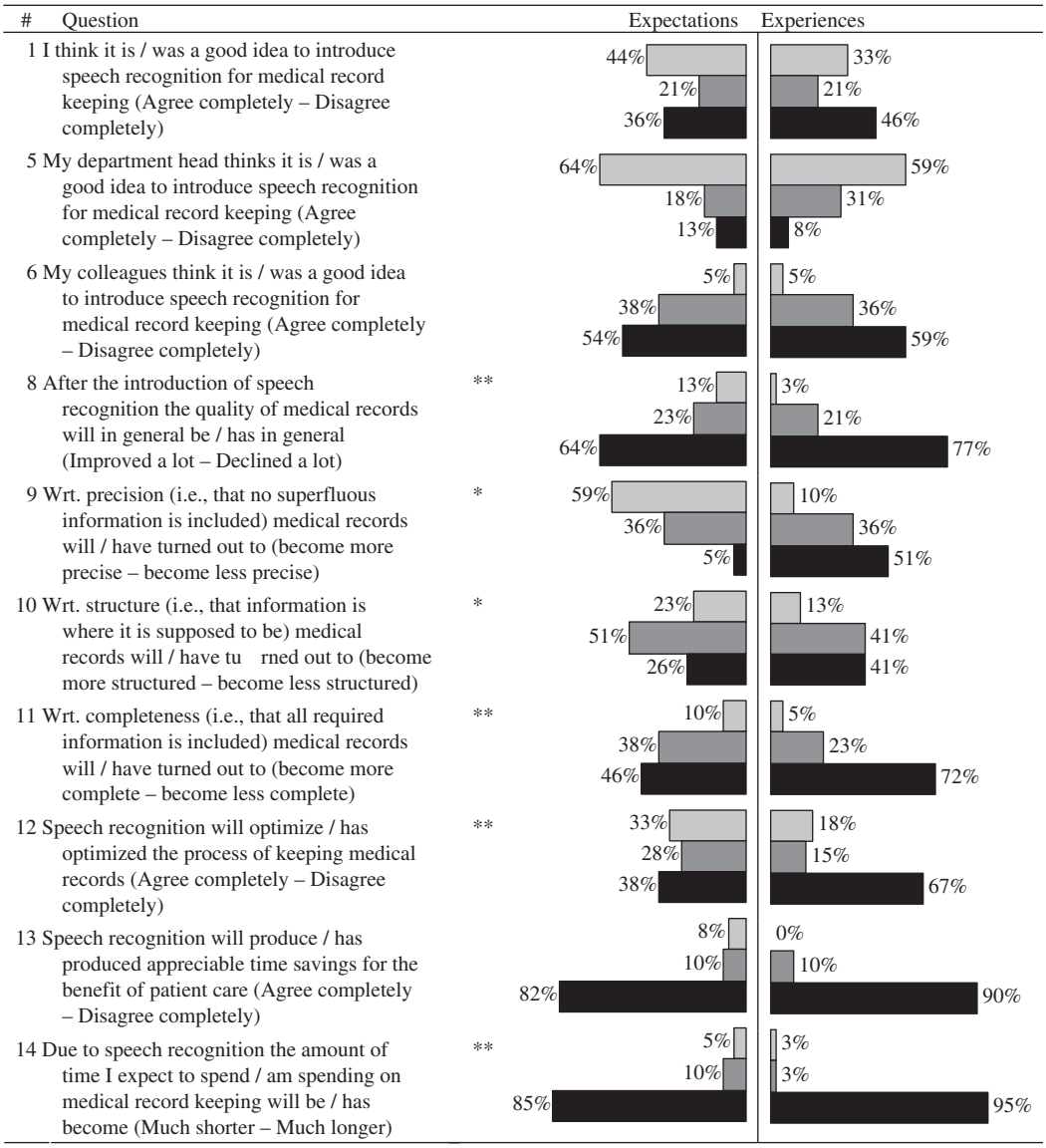


Fig. 1. Comparison of expectations (leftward bars) and experiences (rightward bars), $N = 39$. For each question the upper, light grey bar gives the sum of positive responses (*agree completely* and *agree somewhat*), the middle, dark grey bar gives the neutral responses (*yes-and-no*), and the lower, black bar gives the sum of negative responses (*disagree completely* and *disagree somewhat*). Due to missing values the three bars do not sum to 100% for all questions. Numbers in the leftmost column refer to item numbers in Appendix A. $*p<.05$, $**p<.01$.

speech recognition. However, 4 expectation questions and 6 experience questions received at most 10% positive responses. For 6 of the 10 questions the physicians' experiences differed significantly from their expectations (Wilcoxon test, see Fig. 1); in all cases the change was toward more negative experiences than expectations. That is, physicians' experiences were more negative than their expectations with respect to the general quality, the precision, the structure, and the completeness of medical records and with respect to whether speech recognition had optimized the process of record keeping and whether it had reduced the time physicians spent on record keeping.

The physicians' overall assessment of whether it was a good idea to introduce speech recognition (first question in Fig. 1) shows a significant correlation of .71 between the expectations and experiences questionnaires ($p<.001$, Spearman's rho). That is, the variation in expectations explained ($r^2$) 51% of the variation in experiences. At the same time, the difference in assessment before and after is slightly below the threshold of significance ($p = .051$, Mann–Whitney). For this pair of questions, 4 (10%) physicians had more positive, whereas 11 (28%) had more negative experiences than expectations. Expectations varied across departments: physicians in the oncology department were significantly more negative in their overall assessment before they started using the system compared to physicians in the otology and medicine departments ($p<.005$, Kruskal–Wallis).

### 4.3. Factors influencing overall assessment

In technology-acceptance research, factors that may influence people's acceptance of systems are typically correlated with (self-reported) usage of systems. Because use of the system that we investigated was mandatory, the items included in this study were instead correlated with physicians' overall assessment of whether it was a good idea to introduce speech recognition. Table 4 shows the correlations.

Each of the items concerning performance expectancy was significantly correlated with physicians' overall assessment of speech recognition before starting to use the system (expectations) as well as after having used the system for 4 months or more (experiences). Thus, expectations about improved quality of the contents of medical records and about improved work processes in the production of medical records explained ($r^2$) 22% and 13%, respectively, of the variation in the physicians' overall assessment of speech recognition after 4+ months of use.

Conversely, none of the three items concerning effort expectancy was significantly correlated with physicians' overall assessment after having had experience with the system. Ease of use was, however, significantly correlated with overall assessment before physicians started using the system, suggesting that this item affected physicians' expectations but lost importance as physicians gained experience with the system.

Before they started using the speech-recognition system, physicians' overall assessment correlated significantly with their perception of whether their department head, their colleagues, and the medical secretaries were in favour of the introduction of speech recognition. These three social-influence items explained ($r^2$) 14%, 31%, and 22%, respectively, of the variation in overall assessment before starting to use the system. After having gained personal experience with the system, colleagues was the only one of the three social-influence items that still correlated significantly with overall assessment. Physicians' perception of their colleagues' assessment of the system explained as much of the physicians' overall assessment as their performance expectancy. Conversely, the social influence of department heads and medical secretaries appeared to fade away when the physicians started using the system.

Among the facilitating conditions, the transcription service provided by the medical secretaries was significantly negatively correlated with physicians' overall assessment of speech recognition, explaining 35% of the variation in overall assessment before physicians started using the systems and 19% after they had gained experience using it. It is unsurprising that satisfaction with the previous system for producing medical records tended to co-occur with reluctance toward speech recognition, and vice versa.

We also investigated whether physicians' attitudes toward computers correlated with their overall assessment of whether it was a good idea to introduce speech recognition. Table 5 shows that physicians' computer attitudes were, in general, weakly or very weakly correlated with their overall assessment of speech recognition. Only two correlations were significant. The extent to which physicians like to try out new technology explains ($r^2$) 12% of the variation in their overall assessment of speech recognition before they start to use it but becomes non-significant after they have had experience with the system. Conversely, the extent to which physicians believe that new

Table 4
Correlations (Spearman's rho) of items concerning performance expectancy, effort expectancy, social influence, and facilitating conditions with overall assessment of speech-recognition system, $N = 39$

| # | Items (numbers refer to the questions in Appendix A) | Overall assessment (expectations) | Overall assessment (experiences) |
|---|---|---|---|
| 1 | Overall assessment | 1.00 | .71** |
| | *Performance expectancy* | | |
| 8 | Quality of contents | .50** | .47** |
| 12 | Improved work process | .48** | .36* |
| | *Effort expectancy* | | |
| 3 | Ease of learning | −.01 | −.13 |
| 2 | Ease of use | .46** | .30 |
| 14 | Time spent | .11 | .10 |
| | *Social influence* | | |
| 5 | Department head | .37* | .20 |
| 6 | Colleagues | .56** | .35* |
| 7 | Medical secretaries | .47* | .34 |
| | *Facilitating conditions* | | |
| 4 | Transcription service provided by medical secretaries | −.59** | −.44** |
| 29 | Access to support during introduction+ | .31 | .18 |
| 30 | Quality of support during introduction+ | .42** | .30 |

*Note*: Items are single questions from the expectations questionnaire, except those marked with plusses, which are single questions from the experiences questionnaire. *$p < .05$, **$p < .01$.

Table 5
Correlations (Spearman's rho) of items concerning computer attitudes with overall assessment of speech-recognition system, $N = 39$

| # | Items (numbers refer to the questions in Appendix A) | Overall assessment (expectations) | Overall assessment (experiences) |
|---|---|---|---|
| 17 | I like to try out new technology | .35* | .26 |
| 18 | I am not comfortable when I have to use a new IT system | .08 | .15 |
| 19 | The use of IT during the clinical work will often raise my level of stress | .02 | −.01 |
| 20 | The use of IT will in general lead staff to be more efficient in their clinical work | .16 | .28 |
| 21 | The use of IT will in general make it easier for staff to complete their clinical work | .16 | .19 |
| 22 | When new IT is introduced in our departments/wards, it usually leads to benefits for patients | .33 | .43* |
| 23 | I am often asked for advice about our IT systems by my colleagues | .02 | .05 |

*$p < .05$.

technology usually leads to benefits for patients explains ($r^2$) 18% of the variation in their overall assessment of speech recognition after they have had experience with the system.

### 4.4. Experience of speech recognition

Physicians' experiences with the speech-recognition system were collected when they had used the system for 4 months or more. For each question Fig. 2 shows the percentages of positive (upper bar), neutral (middle bar), and negative (lower bar) responses. The figure also shows the correlation (Spearman's rho) between physicians' responses and their SCR, i.e. the extent to which the user accepts the system-produced text (see Section 3.4). Responses correlated significantly with SCRs for six of the 14 questions, but the correlations were weak.

The overall pattern of responses from the 98 physicians responding to the experiences questionnaire was similar (cf. Figs. 1 and 2) to that of the sub-group of 39 physicians responding to both questionnaires and whose data we have discussed in Sections 4.2 and 4.3. In terms of overall assessment of speech recognition, the 59 physicians who answered only the experiences questionnaire did not differ significantly from the 39 physicians who answered both questionnaires ($p = .08$, Mann–Whitney).

Concerning their overall assessment of whether it was a good idea to introduce speech recognition, respondents were distributed about equally across positive (33%), neutral (36%), and negative responses (31%). Notably, overall assessment was not significantly correlated with SCR. Several other questions indicate that the technical performance of the system was unsatisfactory. Particularly, 69% of physicians disagreed that the number of recognition errors was at an acceptable level, and 76% disagreed that the time and effort they spent correcting recognition errors was at an acceptable level. Unsurprisingly, disagreeing on these questions correlated significantly, though weakly, with low SCRs.

It appears that the introduction of the speech-recognition system has affected medical record keeping negatively in two important ways. First, the time and effort involved in producing medical records is perceived to have increased. Indeed, 94% of physicians found that they now spent more time on medical record keeping, and 83% disagreed that speech recognition had produced timesaving for the benefit of patient care. Second, the quality of the records is perceived to have suffered. Thus, 62% indicate that the general quality of records has declined and 60% that medical records have become less complete. For these two items, there was a significant, though weak, correlation with SCR, indicating that physicians who experienced a decrease in quality and completeness made more changes to the recognized text compared to physicians who experienced an increase in quality and completeness. This suggests that physicians attempted to compensate for the perceived inadequacies of the speech-recognition system. With respect to precision and structure—two other quality attributes—responses were more mixed, but few physicians experienced an improvement (16% and 20%, respectively).

Physicians perceived their department heads as being in favour of the speech-recognition system (48% completely agreed to this item). This suggests strongly that department heads have provided the managerial support necessary to carry through the introduction of the system. Interestingly, physicians perceived their colleagues to be somewhat more negative toward the introduction of the speech-recognition system than their colleagues were in their own overall assessment of the introduction of the system (cf. the first and third questions in Fig. 2). This may suggest that when talking with each other about the system the physicians have highlighted negative aspects. One positive aspect was that 55% of physicians agreed that they knew how the system could learn from their correction of recognition errors. As described below, this led to gradual performance improvements.

### 4.5. Evolution of speech contribution rates

During their first month of using the speech-recognition system, the physicians made an average of 130 dictations. From their second through to their 11th month of using the

| # Question (numbers refer to item numbers in Appendix A) | Experiences | Correlation with SCR |
|---|---|---|
| 1 I think it was a good idea to introduce speech recognition for medical record keeping (Agree completely – Disagree completely) | 33% 36% 31% | 0.11 |
| 5 My department head thinks it was a good idea to introduce speech recognition for medical record keeping (Agree completely – Disagree completely) | 70% 20% 6% | 0.01 |
| 6 My colleagues think it was a good idea to introduce speech recognition for medical record keeping (Agree completely – Disagree completely) | 14% 39% 46% | 0.10 |
| 8 After the introduction of speech recognition the quality of medical records has in general (Improved a lot – Declined a lot) | 15% 21% 62% | 0.25* |
| 9 Wrt. precision (i.e., that no superfluous information is included) medical records have turned out to (become more precise – become less precise) | 16% 37% 43% | 0.23* |
| 10 Wrt. structure (i.e., that information is where it is supposed to be) medical records have turned out to (become more structured – become less structured) | 20% 54% 23% | 0.16 |
| 11 Wrt. completeness (i.e., that all required information is included) medical records have turned out to (become more complete – become less complete) | 11% 29% 60% | 0.28** |
| 12 Speech recognition has optimized the process of keeping medical records (Agree completely – Disagree completely) | 29% 12% 58% | 0.18 |
| 13 Speech recognition has produced appreciable time savings for the benefit of patient care (Agree completely – Disagree completely) | 7% 10% 83% | 0.18 |
| 14 Due to speech recognition the amount of time I am spending on medical record keeping has become (Much shorter – Much longer) | 3% 2% 94% | 0.20 |
| 24 Today the number of recognition errors is at an acceptable level (Agree completely – Disagree completely) | 22% 7% 69% | 0.33** |
| 25 The time and effort I spend correcting recognition errors is at an acceptable level (Agree completely – Disagree completely) | 17% 6% 76% | 0.26* |
| 26 I know how the system can learn from my corrections of recognition errors (Agree completely – Disagree completely) | 55% 21% 18% | 0.08 |
| 27 The system becomes gradually better at recognizing my speech when I mark recognition errors (Agree completely – Disagree completely) | 36% 19% 44% | 0.22* |

Fig. 2. Experience of speech recognition, $N = 98$. For each question the upper, light grey bar gives the sum of positive responses (*agree completely* and *agree somewhat*), the middle, dark grey bar gives the neutral responses (*yes-and-no*), and the lower, black bar gives the sum of negative responses (*disagree completely* and *disagree somewhat*). Due to missing values the three bars do not sum to 100% for all questions. The rightmost column gives the correlation (Spearman's rho) between physicians' responses and their speech contribution rate. $*p<.05$, $**p<.01$.

system the average number of monthly dictations made by a physician was in the range 320–417. This indicates that the system was widely used and that the physicians gained considerable experience. The average duration of a dictation was 17.5 s.

Fig. 3 shows a steady improvement in the SCR for the survey respondents as they gained experience using the system. During their first month of use, they achieved an average SCR of 79%, but after 11 months of usage this had increased to 94%—an average monthly increase of 1.4 percentage points. It should, however, be kept in mind that fewer physicians have used the system for 11 months than for 1 month. This is not an indication that physicians are discontinuing their use of the system but that different departments started using it at different points in time. The highest SCR is achieved by two physicians with 11 months of experience; these two physicians make an average of over 640 dictations per month.

SCRs varied considerably across physicians. Fig. 4 shows that large variation existed even for physicians with the same level of experience with the system. As an example, the bottommost curve shows that after using the system for 1 month three physicians had SCRs below 52%, three above 94%, and the remaining 79 physicians between

60% and 92%. With increasing levels of experience the variation across physicians decreased (standard deviation = 11.9, 6.2, and 3.9 percentage points for 1, 5, and 10 months of experience, respectively). Most of the improvement in average SCR with increasing levels of experience consisted of physicians with low initial SCRs catching up with the other physicians.

### 4.6. Typical comments from the respondents

Of the 98 respondents who answered the experiences questionnaire, 94 expressed comments to at least one of the seven open questions. The results of the above analysis were supported by the free comments, which additionally covered points not addressed by the closed questions. For instance, 33 respondents expressed negative feelings about doing a "secretary's job". As an example, one physician commented:

> "*Why use a high-salary and highly qualified physician, who can type with only two fingers, to do secretarial tasks that could be done better and more cheaply by a secretary who is skilled at touch typing?*"

On the other hand, 14 respondents indicated that the reduced involvement of the secretaries provides independence and removes some errors. Three physicians were concerned that secretaries are no longer there to capture errors or inadequacies, especially with respect to checking reference codes and related documents.

Most respondents found that the new work procedure is optimizing the workflow; as much as 77 respondents offered comments expressing this view (e.g., "*Records are done on the fly*" and "*Records are immediately available for further use*"). However, 83 physicians also indicated that the use of speech recognition takes too much time. Many critical comments concern the integration of the speech-recognition system into the existing EMR system (12 comments) and the user interface, which is seen as too slow and requiring too much mouse interaction (29 comments). Six respondents state that they often avoid using the speech-recognition system and use a keyboard instead. That is, they often type their entries into the EMR system.

Regarding the system's pure speech-recognition capabilities, 61 comments call for improvements. In particular, 20 respondents complain that the system fails to or is much too slow to learn from the corrections they enter, and 18 respondents consider the types of error produced by speech recognition more difficult to spot than previous

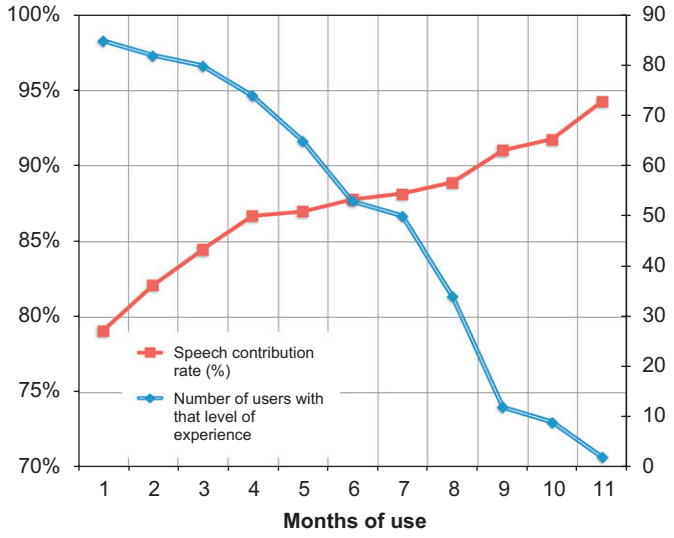


Fig. 3. Speech contribution rate as a function of months of experience using the system, and number of physicians at the different levels of experience.

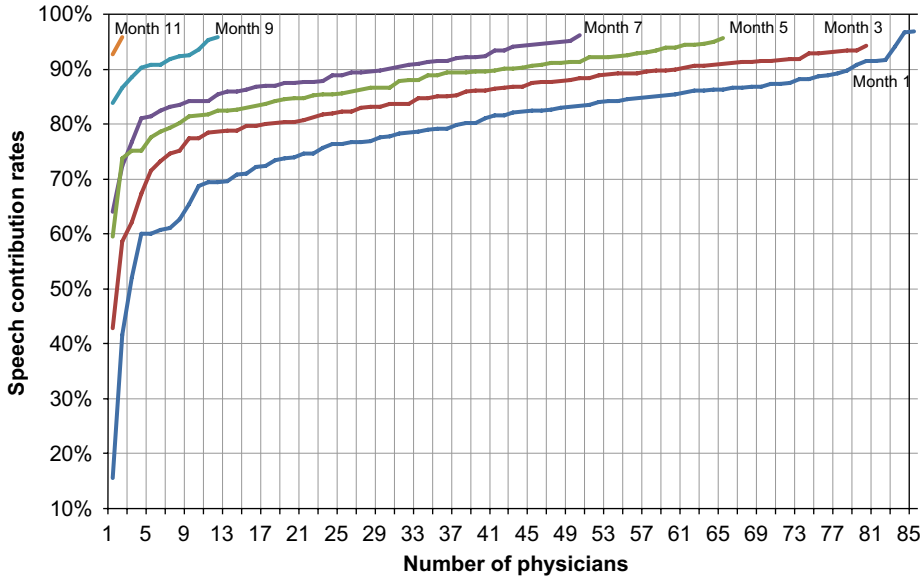


Fig. 4. Accumulated distribution of speech contribution rates at different levels of experience (i.e., months of using the speech-recognition system).

transcription errors and potentially more harmful. Overall, seven respondents express enthusiasm toward the technology, while 26 report that they experience an increase in stress or a decrease in work satisfaction.

## 5. Discussion

### 5.1. Expectations, experiences, and social influence

The two major objectives behind the introduction of speech technology as a front end to the EMR system were to optimize the workflow and thereby achieve a quicker completion of records and to enhance the quality of medical records. The physicians' free-text comments to the open questions suggest that the first objective has been achieved, and information from the hospital corroborates this. Still, physicians' experiences were more negative than their expectations, particularly with respect to the quality of medical records and the time spent producing them. The technical performance of the system was experienced as unsatisfactory, particularly with respect to the number of recognition errors and the time and effort required to correct them. Respondents almost unanimously reported that the time they personally spent producing medical records had increased, and they also agreed that speech recognition had not led to overall time savings for the benefit of patient care. Physicians experienced that the quality of the contents of medical records had declined in general and particularly with respect to record completeness. Finally, respondents are approximately equally divided between those who, in retrospect, think it was a good idea to introduce speech recognition and those who do not.

With respect to predictors of the physicians' acceptance of the system, our results indicate that their overall assessment of speech recognition prior to using it was the strongest among the possible predictors we have tested. This suggests that asking prospective users for their assessment of whether the introduction of a system is a good idea can be used as an early, cheap, and rather reliable indicator of whether they will approve of the system after having used it for some time. This finding discords, however, with Root and Draper (1983), who found little correlation between people's assessments before and after they had experience with a system.

For the predictors identified in previous technology-acceptance studies, we find that performance expectancy and social influence moderately indicated our respondents' overall assessment of speech recognition before they began to use it. After having gained experience with the system, performance expectancy and perception of colleagues' overall assessment of speech recognition still provided some indication of overall assessment. Effort expectancy in terms of perceived ease of use moderately indicated overall assessment before starting to use the system but not after months of use. These results are in agreement with previous technology-acceptance studies with respect to the presence of significant correlations, the general magnitude of correlations, as well as the effect of experience with the system (Davis, 1989, 1993; Adams et al., 1992; Venkatesh et al., 2003). It should be noted that Venkatesh et al. (2003) find that social influence is mainly a predictor of technology acceptance in situations where some people are in a position to mandate that others use a technology, as was the case in our study.

As in previous studies (e.g., Venkatesh et al., 2003) facilitating conditions were perceived rather similarly to effort expectancy, except for the moderate and lasting negative influence of physicians' perception of the transcription service previously provided by medical secretaries. While dissatisfaction with a previous solution may have an only temporary, and supposedly positive, effect on people's assessment of a new technology, our study suggests that a long-lasting and generally well-liked previous solution has a long-term negative effect on people's assessment of a new technology.

With respect to physicians' performance with the speech-recognition system, their SCR correlated only weakly with their assessments of the system. Weak correlations between assessments and performance measures have also been found in previous studies (Frøkjær et al., 2000; Hornbæk and Law, 2007). Physicians' SCR improved over time, particularly for physicians with low initial rates, and after 9 months of use, physicians had an average SCR of 91%. Thus, having used the system for dictating several thousand EMR entries physicians still experienced that they had to revise one in every 11 words of the text produced by the speech-recognition system. This was perceived as unsatisfactory and time consuming, especially because many physicians felt that they were correcting the same errors repeatedly.

The system vendor emphasizes that this is the first generation of their system for recognition of Danish medical speech and they maintain that the second generation, currently under deployment, is faster and has higher recognition accuracy. Initial reports about user experiences with the second generation system are indeed more positive, but time has not allowed a systematic evaluation.

### 5.2. Limitations of the survey

This study has four limitations that should be taken into account in interpreting the results. First, the introduction of the speech-recognition system involved new work procedures as secretary efforts were replaced by physician efforts. This makes it difficult or perhaps impossible to distinguish effects of using the speech-recognition system from effects of the new work procedures, which changed the roles and responsibilities in the production of the medical records. Second, respondents had higher SCRs than non-respondents. This may indicate that respondents have more positive experiences with speech recognition than non-respondents, suggesting that our results be

interpreted as an upper bound of the physicians' acceptance of the speech-recognition system. Third, while a response rate of 60% is comparable with other surveys of technology acceptance (e.g., Adams et al., 1992; Hebert and Benbasat, 1994; Fichman and Kemerer, 1999) it calls for caution in interpreting the results. One indication of a possible bias is the higher SCR of respondents compared to non-respondents. Fourth, physicians who answered both questionnaires received the experiences questionnaire after having used the speech-recognition system for about 4 months, and SCRs were studied over the first 11 months of use. While this entails that the physicians had considerable experience with the system, it remains unknown whether their assessment of the system had stabilized and it appears that their performance was still improving.

## 6. Conclusion

Speech-recognition technology is continuously being refined and is gradually becoming adopted as an alternative to typing text or to dictation and subsequent transcription by secretaries. This study reports the results of a survey of the first hospital to introduce speech recognition in Danish for all clinical specialties and departments. We have found that:

- Physicians' expectations tended to be more positive than their experiences. It is seen as a valuable benefit of the technology that it makes it possible to access records right after their dictation is completed. Yet, the physicians felt that they spent much more time producing medical records with the new system and associated work procedures, that the overall quality of records had declined, and that the performance of the system in terms of recognizing speech was unsatisfactory.
- Performance expectancy, effort expectancy (especially ease of use), social influence, and facilitating conditions were all moderately correlated with physicians' overall assessment of the speech-recognition system before they started using it. While the performance-expectancy items—quality of contents and improved work process—remained significant indicators also after physicians had gained experience with the system, the only other significant items were colleagues (a social influence) and the transcription service previously provided by the medical secretaries (a facilitating condition).
- The percentage of words that remained unaltered when physicians proofread their medical records (the SCR) increased as physicians gained experience with the system. While this indicates a gradual performance improvement, the average SCR after 9 months of use was only 91%. Physicians' SCRs correlated only weakly with their assessment of the system.
- Physicians are approximately equally divided among those who think, in retrospect, that the introduction of speech recognition was a good idea, that it was not, and those who are neutral.

While acknowledging that most physicians in the present study have shown a less than enthusiastic reception of speech-recognition technology, it should not be overlooked that one-third of physicians were positive in their overall assessment of the speech-recognition system after they had gained experience with it. This provides some basis for further efforts to improve speech recognition in Danish and other "relatively small" languages and, specifically, for efforts to introduce it for medical record keeping. It needs to be investigated to which extent longer periods of practice as well as more mature generations of the technology will lead to higher levels of satisfaction among physician users.

## Acknowledgements

The survey instrument was developed by the authors in collaboration with, from Vejle and Give Hospital, leaders of the speech-recognition project, Aase Andreasen and Trine Ankjær. Useful input to the questionnaire was received from the company delivering and implementing the speech-recognition technology, Max Manus A/S.

## Appendix A. Expectations and experiences questionnaires

Questions included in the expectations questionnaire are indicated with plusses in the column B (before). Questions included in the experiences questionnaire are indicated with plusses in the column A (after), and variations in their wording compared to expectations questions are in italics.

All questions have an additional "Don't know" option. Open questions have been left out (see Table A1).

Table A1

| B | A | Question items |
|---|---|---|
| + | + | 1. I think it is [*was*] a good idea to introduce speech recognition for medical record keeping. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 2. I expect it to be easy to use speech recognition once I have become used to it. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 3. I expect to have to spend much effort to become used to working with speech recognition. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 4. The service provided by our secretarial staff is of such high standard that speech recognition will hardly be able to match it. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |

Table A1 (*continued*)

| B | A | Question items |
|---|---|---|
| + | + | 5. My department head thinks it is [*was*] a good idea to introduce speech recognition for medical record keeping. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | + | 6. My colleagues think it [*was*] a good idea to introduce speech recognition for medical record keeping. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 7. Our secretaries think it is a good idea to introduce speech recognition for medical record keeping. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | + | 8. After the introduction of speech recognition the quality of medical records will in general [*has in general turned out to*] … (Improve a lot, Improve somewhat, Remain the same, Decline somewhat, Decline a lot) |
| + | + | 9. With respect to precision (i.e., that no superfluous information is included) medical records will [*have turned out to*] … (Become more precise, Remain at the same level, Become less precise) |
| + | + | 10. With respect to structure (i.e., that information is where it is supposed to be) medical records will [*have turned out to*] … (Become more structured, Remain at the same level, Become less structured) |
| + | + | 11. With respect to completeness (i.e., that all required information is included) medical records will [*have turned out to*] … (Become more complete, Remain at the same level, Become less complete) |
| + | + | 12. I expect that speech recognition will optimize [*Speech recognition has optimized*] the process of keeping the medical record. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | + | 13. I expect speech recognition will produce [*Speech recognition has produced*] appreciable time savings for the benefit of patient care. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | + | 14. I expect that the time I spend on producing medical records in the long run will become [*The time I spend on producing medical records has become*] … (a lot shorter, shorter, the same, longer, a lot longer) |
| + | | 15. Have you talked with colleagues about their experience with speech recognition? (Yes, No) |
| + | | 16. If yes: How was their experience? (Largely positive, Both positive and negative, Largely negative) |
| + | | 17. I like to try out new technology. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 18. I am not comfortable when I have to use a new IT system. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 19. The use of IT during the clinical work will often raise my level of stress. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 20. The use of IT will in general lead staff to be more efficient in their clinical work. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 21. The use of IT will in general make it easier for staff to complete their clinical work. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 22. When new IT is introduced in our departments/wards, it usually leads to benefits for patients. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| + | | 23. I am often asked for advice about our IT systems by my colleagues. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 24. Today the number of recognition errors is at an acceptable level. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 25. *T*he time and effort I spend correcting recognition errors is at an acceptable level. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 26. I know how the system can learn from my corrections of recognition errors. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 27. The system is gradually becoming better at recognizing my speech when I mark recognition errors. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 29. During the introduction of speech recognition the access to support was satisfactory. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | *30. During the introduction of speech recognition the quality of support was satisfactory. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely)* |
| | + | 31. Today the access to support is satisfactory. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |
| | + | 32. Today the quality of support is satisfactory. (Agree completely, Agree somewhat, Yes-and-no, Disagree somewhat, Disagree completely) |

## References


Adams, D.A., Nelson, R.R., Todd, P.A., 1992. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Quarterly 16 (2), 227–247.

Ajzen, I., 1985. From intentions to actions: a theory of planned behavior. In: Kuhl, J., Beckmann, J. (Eds.), Action Control: from Cognition to Behavior. Springer, New York, pp. 11–39.

Ajzen, I., 1991. The theory of planned behaviour. Organizational Behavior and Human Decision Processes 50 (2), 179–211.

Alapetite, A., 2008. Impact of noise and other factors on speech recognition in anaesthesia. International Journal of Medical Informatics 77 (1), 68–77.

Al-Aynati, M.M., Chorneyko, K.A., 2003. Comparison of voice-automated transcription and human transcription in generating pathology reports. Archives of Pathology and Laboratory Medicine 127 (6), 721–725.

Barker, J.P., Cooke, M.P., Ellis, D.P.W., 2005. Decoding speech in the presence of other sources. Speech Communication 45 (1), 5–25.

Borowitz, S.M., 2001. Computer-based speech recognition as an alternative to medical transcription. Journal of the American Medical Informatics Association 8 (1), 101–102.

Coniam, D., 1999. Voice recognition software accuracy with second language speakers of English. System 27 (1), 49–64.

Damanpour, F., 1991. Organizational innovation: a meta-analysis of effects of determinants and moderators. Academy of Management Journal 34 (3), 555–590.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13 (3), 319–340.

Davis, F.D., 1993. User acceptance of information technology: systems characteristics, user perceptions and behavioral impacts. International Journal of Man-Machine Studies 38 (3), 475–487.

Devine, E.G., Gaehde, S.A., Curtis, A.C., 2000. Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. Journal of the American Informatics Association 7 (5), 462–468.

Feng, J., Sears, A., 2004. Using confidence scores to improve hands-free speech based navigation in continuous dictation systems. ACM Transactions on Computer-Human Interaction 11 (4), 329–356.

Fichman, R.G., 2000. The diffusion and assimilation of information technology innovations. In: Zmud, R.W. (Ed.), Framing the Domains of IT Management: Projecting the Future through the Past. Pinnaflex Educational Resources, Cincinnati, OH, pp. 105–127.

Fichman, R.G., Kemerer, C.F., 1999. The illusory diffusion of innovation: an examination of assimilation gaps. Information Systems Research 10 (3), 255–275.

Fiscus, J.G., 1997. A post-processing system to yield reduced word error rates: recognizer output voting error reduction (ROVER). In: Proceedings of the IEEE 1997 Workshop on Automatic Speech Recognition and Understanding. IEEE, Los Alamitos, CA, pp. 347–354.

Fishbein, M., Ajzen, I., 1975. Belief, Attitude, Intention and Behavior: an Introduction to Theory and Research. Addison-Wesley, Reading, MA.

Frøkjær, E., Hertzum, M., Hornbæk, K., 2000. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In: Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems. ACM Press, New York, pp. 345–352.

Gallivan, M.J., 2001. Organizational adoption and assimilation of complex technological innovations: development and application of a new model. ACM SIGMIS Database 32 (3), 51–85.

Gong, Y., 1995. Speech recognition in noisy environments: a survey. Speech Recognition 16 (3), 261–291.

Hebert, M., Benbasat, I., 1994. Adopting information technology in hospitals: the relationship between attitudes/expectations and behaviour. Hospital & Health Services Administration 39 (3), 369–383.

Honeycutt, L., 2003. Researching the use of voice recognition writing software. Computers and Composition 20 (1), 77–95.

Hornbæk, K., Law, E.L.-C., 2007. Meta-analysis of correlations among usability measures. In: Proceedings of the CHI 2007 Conference on Human Factors in Computing Systems. ACM Press, New York, pp. 617–626.

Jungk, A., Thull, B., Fehrle, L., Hoeft, A., Rau, G., 2000. A case study in designing speech interaction with a patient monitor. Journal of Clinical Monitoring and Computing 16 (4), 295–307.

Juul-Kristensen, B., Laursen, B., Pilegaard, M., Jensen, B.R., 2004. Physical workload during use of speech recognition and traditional computer input devices. Ergonomics 47 (2), 119–133.

Kanal, K.M., Hangiandreou, N.J., Sykes, A.M., Eklund, H.E., Araoz, P.A., Leon, J.A., Erickson, B.J., 2001. Initial evaluation of a continuous speech recognition program for radiology. Journal of Digital Imaging 14 (1), 30–37.

Lai, J., Vergo, J., 1997. MedSpeak: report creation with continuous speech recognition. In: Proceedings of the CHI 97 Conference on Human Factors in Computing Systems. ACM Press, New York, pp. 431–438.

Lai, J., Karat, C.-M., Yankelovich, N., 2008. Conversational speech interfaces and technologies. In: Sears, A., Jacko, J.A. (Eds.), The Human–Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, second ed. Erlbaum, New York, pp. 381–391.

Leijten, M., Van Waes, L., 2005. The adaptation process of professional writers with and without dictating experience. Interacting with Computers 17 (6), 736–772.

Lombard, E., 1911. Le signe de l'élévation de la voix. Annales Maladies Oreille, Larynx, Nez, Pharynx 37, 101–119.

Mohr, D.N., Turner, D.W., Pond, G.R., Kamath, J.S., de Vos, C.B., Carpenter, P.C., 2003. Speech recognition as a transcription aid: a randomized comparison with standard transcription. Journal of the American Medical Informatics Association 10 (1), 85–93.

Mönnich, G., Wetter, T., 2000. Requirements for speech recognition to support medical documentation. Methods of Information in Medicine 39 (1), 63–69.

Ramaswamy, M.R., Chaljub, G., Esch, O., Fanning, D.D., van Sonnenberg, E., 2000. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. American Journal of Roentgenology 174 (3), 617–622.

Rogers, E.M., 2003. Diffusion of Innovations, fifth ed. Free Press, New York.

Root, R.W., Draper, S., 1983. Questionnaires as a software evaluation tool. In: Proceedings of the CHI'83 Conference on Human Factors in Computing Systems. ACM Press, New York, pp. 83–87.

Sears, A., Karat, C.-M., Oseitutu, K., Karimullah, A., Feng, J., 2001. Productivity, satisfaction, and interaction strategies of individuals with spinal cord injuries and traditional users interacting with speech recognition software. Universal Access in the Information Society 1 (1), 4–15.

Suhm, B., Myers, B., Waibel, A., 2001. Multimodal error correction for speech user interfaces. ACM Transactions on Computer-Human Interaction 8 (1), 60–98.

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: toward a unified view. MIS Quarterly 27 (3), 425–478.

Wilpon, J.G., Jacobsen, C.N., 1996. A study of speech recognition for children and the elderly. In: ICASSP-96: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, vol. 1. IEEE, Los Alamitos, CA, pp. 349–352.

Young, S., 1996. Review of large-vocabulary continuous-speech recognition. IEEE Signal Processing Magazine 13 (5), 45–57.

Zafar, A., Overhage, J.M., McDonald, C.J., 1999. Continuous speech recognition for clinicians. Journal of the American Informatics Association 6 (3), 195–204.

Zafar, A., Mamlin, B., Perkins, S., Belsito, A.M., Overhage, J.M., McDonald, C.J., 2004. A simple error classification system for understanding sources of error in automatic speech recognition and human transcription. International Journal of Medical Informatics 73 (9–10), 719–730.