Speech Technology in Dutch Health Care: A Qualitative Study

Ellen Luchies, Marco Spruit and Marjan Askari

Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

Keywords: Speech Technology, Speech Recognition, Natural Language Processing, Health Care, Dutch Health Care.

Abstract: This study investigates the opportunities of speech technology in Dutch hospitals, and to what extent speech technology can be used for documentation. Furthermore, we clarify why speech technology is used only marginally by Dutch hospital staff. We performed interviews where speech technology users, managers in hospitals and software suppliers were contacted as participants. We then transcribed our interviews and synthesized the pros and cons of speech technology as well as major barriers for the adoption. Our results show various influencing factors that could be clarifications for the fact that only 1% of the medical staff uses speech technology in the Netherlands. The major reasons we found are: speech technology usage at only radiology and pathology departments, *smarttexts* and *smartphrases* of the Electronic Health Record (EHR) compete with speech technology, caregivers have to adjust their way of working which evokes resistance, lack of central authorization at Dutch hospitals and finally, financial barriers. Our results show that speech technology works for radiology and pathology as a tool for documentation, but is found less useful for other departments. For the remaining departments, different applications show potential, such as structured reporting.

1 INTRODUCTION

Caregivers, specifically nurses and physicians, experience the highest work load compared to many other professionals (NOS, 2017). According to Schumacher (2017), it is to be expected that physicians's workload will increase even more in the upcoming years, among other things due to the increasing amount of people who need care in western countries. This increase is caused by aging of the population (Schumacher, 2017). The groups aged 65+ and 80+ are increasing rapidly, as shown in Figure 1. As people age, they need more care as a result of physiological and psychological weaknesses, as is shown in Figure 2. This phenomenon is called multimorbidity.

A more efficient way of working is therefore necessary to be able to cope with the increasing workload in health care. Speech technology can offer a solution for this problem (Ajami, 2016). This technology has seen major improvements in the last decade (Parente, Kock and Sonsini, 2004; Ajami, 2016).
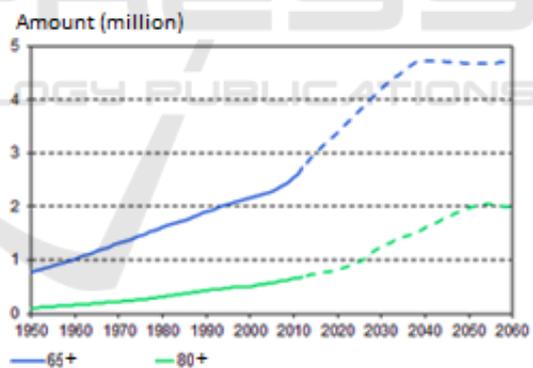

Figure 1: The aging Dutch population (Schumacher, 2017).

According to Ajami (2016), speech technology can contribute to a more efficient way of working. Physicians are able to document faster, and make reports available faster (Ajami, 2016). Nowadays, many systems reach an accuracy up to 98% (Parente, Kock and Sonsini, 2004; Johnson *et al.*, 2014; Ajami, 2016). Nevertheless, speech technology is used by only 1% of the Dutch hospital staff (Nuance, 2015). The aim in this study was therefore to investigate the barriers and potentials of speech technology in Dutch health care.
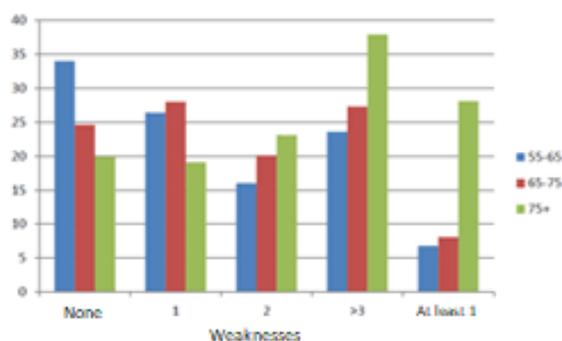
Figure 2: Amount of weaknesses per agegroup (Schumacher, 2017).

2 BACKGROUND

2.1 How Does Speech Technology Work?

Digital dictation can be seen as the predecessor of speech technology. It worked as follows: the doctor dictated the report, sent the dictation to the secretary who transcribed the dictation. This document was then sent back to the doctor for validation, and the report was made available (S1, see Table 2 below). However, since speech technology is available to doctors, they are able to dictate using a computer. The users speaks, and the system converts the speech into words on the screen. After this phase, the doctor has to correct the document to filter out errors. Previously, this was done by the secretary (S1).

Speech technology uses different tools to support a qualitative conversion from speech to text.

- **Acoustic model** (Koivikko, Kauppinen, and Ahovuo, 2008): this model defines how sounds are pronounced. (Renckens, 2009).
- **Speech corpus**: this corpus defines the different ways so every sound and phoneme can be recognized (Renckens, 2009) despite the fact that each individual pronounces a sound different (Ajami, 2016).
- **Lexicon** (Koivikko, Kauppinen, and Ahovuo, 2008) (Ajami, 2016): this is the dictionary of the computer (Renckens, 2009). It contains the words that are recognizable by the system. If a word does not occur in the dictionary, the system is not aware of the existence of that

word, and is therefore not able to recognize it (S2). In addition, a phonetical transcription is available for each word in the dictionary (Renckens, 2009). The quality of speech technology strongly depends on the dictionary that is used. When a dictionary contains many words, the system is more likely to confuse words with each other, which leads to more mistakes and a lower accuracy (Ajami, 2016).

- **Language model** (Koivikko, Kauppinen, and Ahovuo, 2008; Ajami, 2016): the language model is a statistical model. It calculates the likelihood that words are related and occur in a certain sequence, based on previously spoken reports in the database (Renckens, 2009). An advantage of this is the possibility to construct a word or sentence based on statistics when the system is not able to do this based on speech. A disadvantage of this is the fact that uncommon words will not be chosen because common words are more likely to be used according to the databases (Vervoort, 2009).

2.1.1 The Process from Speech to Text

The process to construct a word from speech is shown in figure 3. This figure is based on studies of Vervoort (2017), Renckens (2009) and Geitgey (2016). It starts by recording speech with a digital voice recorder. These sound waves are segmented by the computer (Renckens, 2009; Vervoort, 2017).

In the next step, the segments are converted into numbers by *sampling*. This is a technique that measures the height of the sound wave on equally scattered points in the wave (Vervoort, 2017). After sampling, the signal must be filtered to reduce background noise (Vervoort, 2017). By measuring the amount of energy in the sound waves, a spectrogram is created (Vervoort, 2017). This is seen as a fingerprint of the dictate (Geitgey, 2016).

Subsequently, the spectrogram is used as the input for a neural network. The output represents the likelihood per phoneme. To compute this likelihood, the neural network uses the acoustic model, speech corpus, lexicon and language model (Renckens, 2009).

For speech technology, a Recurrent Neural Network (RNN) is used that saves previous calculations to influence future calculations (Geitgey, 2016). This way, speech technology is a learning system which improves itself (Geitgey, 2016).
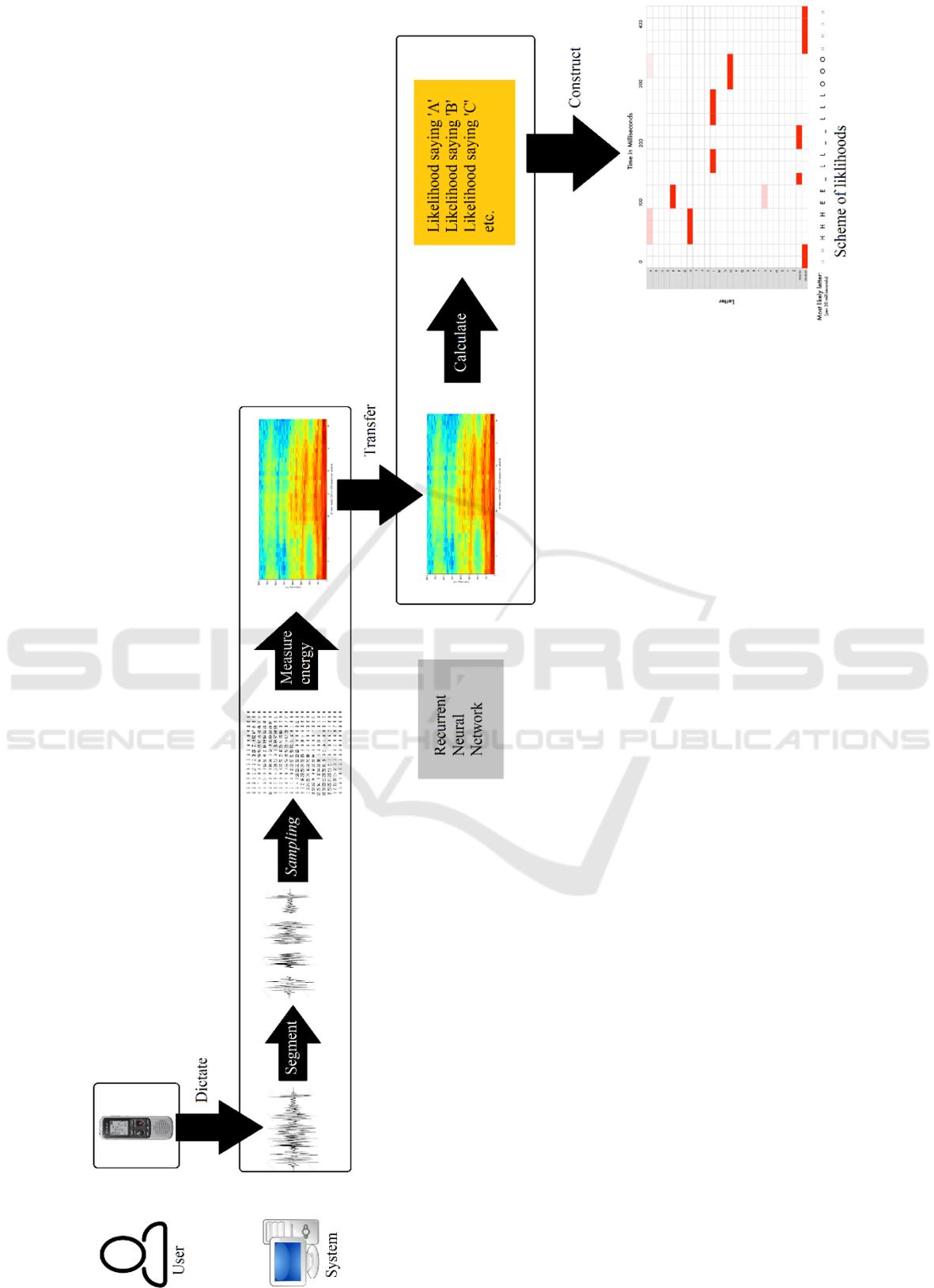
Figure 3: Process model from speech to text.

After the calculations of the likelihood per phoneme, a scheme representing these likelihoods is constructed (Geitgey, 2016). This is the last step as shown in Figure 3. In this example, the word 'hello' is constructed. The predictions of this scheme are sequenced with double characters and gaps in the word. When these are filtered, three words are still possible: 'hello', 'hullo' and 'aullo'. Since 'hello' is more likely because it occurs more often in the database than the other two options, 'hello' is chosen (Geitgey, 2016). In case another word was meant, the user has to correct it manually. This improvement will then be saved for future predictions (Renckens, 2009).

2.2 Pros and Cons of Speech Technology

In this section we state the pros and cons according to literature enlightened by interviewees. The main advantage of speech technology is time reduction (Ajami, 2016; Koivikko, Kauppinen, and Ahovuo, 2008). According to a study of Nuance, people can type 40 words per minute at best, whereas people can speak 120 words per minute (Nuance, 2008). Furthermore, Nuance (2015) states that doctors are documenting 13.3 hours a week on average. For nurses, this is 8.7 hours per week (Nuance, 2015). This concerns an estimated 30% of the working week, therefore speech technology could be very profitable.

Different studies show that radiology and pathology benefit most from speech technology (Ajami, 2016; Johnson *et al.*, 2014). This is clarified by the fact that radiology and pathology can cut down on their secretaries when they start using speech recognition, which leads to a decrease in the report turnaround time (RTT) (Koivikko, Kauppinen, and Ahovuo, 2008). Other departments started working with the Electronic Health Record (EHR) before speech technology, and already cut down on their secretaries. Because of this, speech technology lacks this benefit for departments other than radiology and pathology, including the decrease in RTT and the financial benefits of the staffing costs (M1).

Before doctors can start using speech technology, a profile must be prepared whereby the system gets familiar with the user's speech and vocabulary. This can be done by reading a text aloud (Bosch, 2005). This is beneficial for the accuracy of the system (Vervoort, 2017), but takes time (Ajami, 2016; Johnson, et al., 2014). Speech technology uses a lexicon, as described in paragraph 2.1. For medical staff, medical terminology is added, but not terminology that is used in daily life (S2). A disadvantage of this dictionary holds that words that are not included, cannot be recognized by the system (S2). Patient friendliness increases (Ajami, 2016). When a doctor types during a conversation, he or she has less attention for the patient. Using speech technology, he or she can listen to the patient without this distraction (U1). The doctor has to dictate during the conversation, or afterwards, since it is not (yet) possible for software to recognize two voices at once, *i.e.* Advanced Voice Technology (Tuin, 2016). Besides, reports are available faster (Ajami, 2016) (Johnson, et al., 2014; Koivikko, Kauppinen, and Ahovuo, 2008), therefore patients can be cured faster, which leads to an increased quality of patient care (Koivikko, Kauppinen, and Ahovuo, 2008; Parente, Kock, and Sonsini, 2004). A challenge for implementing speech technology is the human factor (Ajami, 2016; Dawson *et al.*, 2014; Parente, Kock, and Sonsini, 2004). Doctors need to adapt their way of working and this often leads to problems (Dawson, et al., 2014). To avoid this, intensive support is needed (S2; Ajami, 2016). An overview of all found pros and cons in literature is represented in Table 1.

3 METHODS

For this study we performed a literature review and a qualitative study. We searched PubMed, Springerlink and Elsevier for finding the relevant articles. The following key words and/or their combinations are used: *speech recognition, health care, spraaktechnologie, spraakherkenning, zorg, medisch, pros, advantages, cons, werking, neural network, acoustic model, akoestisch model* and *Hidden Markov Model*. While selecting articles, we focused on the publication date and Citation index.

The data for the qualitative study were gathered by performing ten semi-structured interviews. We used a structured topic list and an operational model to establish the topics of the interviews and corresponding questions. The participants consisted of four managers working at two different hospitals, four suppliers of speech technology working at different companies, and two users of speech technology with different professions. An overview of the participants can be found in Table 2.

Table 1: Pros and cons of speech technology according to other studies.

Pros	Cons
- Improved and time-saving workflow in provision of health care services (Ajami, 2016) (Johnson, et al., 2014) (Koivikko, Kauppinen, & Ahovuo, 2008) - Online registration (coordination of activity and documentation) (Ajami, 2016) - Quick access to documents (Ajami, 2016) - Reduction in cost and time in documenting information (Ajami, 2016) (Johnson, et al., 2014) (Koivikko, Kauppinen, & Ahovuo, 2008) - Accurate (99%), legible and complete information (Ajami, 2016) (Parente, Kock, & Sonsini, 2004) (Dawson, et al., 2014) (Koivikko, Kauppinen, & Ahovuo, 2008) - Reduction in patients' duration of hospital stay (Ajami, 2016) - Quality of patient care increases (Koivikko, Kauppinen, & Ahovuo, 2008) (Parente, Kock, & Sonsini, 2004) - Patient friendliness (Ajami, 2016) - Patient security (Ajami, 2016)	- Quality of the system depends on the dictionary (Ajami, 2016) - Initial cost of implementation and maintainance of technology (Ajami, 2016) - Sound pollution (noise) (Ajami, 2016) - Variety of accents (Ajami, 2016) (Johnson, et al., 2014) - Adaption of work and acceptance by doctors and other health care providers (Ajami, 2016) (Dawson, et al., 2014) (Parente, Kock, & Sonsini, 2004) - Correcting mistakes is time-consuming and needs to be done dedicated (Ajami, 2016) - Hardware requirement (microphones, sound cards and speech engine software) (Ajami, 2016) (Johnson, et al., 2014) - User training (and the time needed for this) (Ajami, 2016) (Johnson, et al., 2014)

Table 2: Background overview and IDs of the participants.

Managers (of departments)	NLP Suppliers	Users/ Doctors
M1 Martini Ziekenhuis	**S1** Cedere	**U1** Orthopaedist
M2 UMC Utrecht	**S2** G2Speech	**U2** Radiologist
M3 UMC Utrecht	**S3** G2Speech	
M4 UMC Utrecht	**S4** Nuance	

Participants were approached when they had experience with using speech technology, the implementation of speech technology, or facilitated speech technology. They were approached via LinkedIn, or participants referred to other interviewees with experience with speech technology, who were approached next.

The respondents had to sign an informed consent to give permission for using the information of the conversation. With their permission the conversation was recorded. Next, we transcribed the interviews. After the transcription, we validated our transcriptions by sending it back to the participant for their final approval. The document was then added to the report. Next, we performed a content analysis for finding the potentials and barriers as mentioned by our participants. The analysis was done using Nvivo, version 11.

4 RESULTS

4.1 Speech Technology in Practice

The orthopaedist who was interviewed as one of the users of speech technology explained that documentation became a task for the doctor, due to implementation of the EHR. Besides, documenting transferred from speaking to typing according to him, confirming what was explained in paragraph 2.1 (U1).

After completing their profile, speech technology is functional. However, time is still needed to optimize the profile by correcting mistakes, as speech technology is a learning system (S3).

Only 1% of medical staff uses speech technology. According to one of the suppliers (S4), this is relatively low compared to other countries in the Benelux. In the U.S., this technology is extensively used (Shagoury, 2010). In the next paragraph we explain our findings about the limitations of this technology in the Netherlands.

4.2 Limited Use of Speech Technology

An overview of our results is represented in Table 3. It shows which aspect is named by which group. N states how many participants mentioned a particular aspect. NU stands for the group of users, NS for the group of suppliers and NM for the group of managers. The last column contains the totals.

Our results show that, in almost every hospital in the Netherlands, pathology and radiology are using speech technology, as explained by the participating managers. However, these are only two departments which use speech technology often. The remaining departments do not use it, or it is occassionally used by a few doctors from the department, as is the case with the orthopaedist we interviewed. This is one of the clarifications for the limited use of this technology in the Netherlands. One of the suppliers stated that radiologists and pathologists do not see any patients, but only investigate the radiological photos or bodyparts of the patient for diagnosis. Furthermore, these specialisms document the most by far. This is confirmed by a manager at the radiology department of UMC Utrecht. One of the users stated that the EHR is a barrier to use speech technology, since the EHR is not adapted for the use of speech recognition.

Another reason of the limited use of speech technology could be false expectations that people have of speech technology, mentioned by two suppliers. Since this technology was not very functional at its introduction, many users kept this opinion and therefore are not willing to use or obtain it now. Users expect the technology to work immediately, however, this is not the case in practice, since usage of this technology demands a developed profile. In addition, it was mentioned that accuracy actually decreases after using speech technology for a long period.

Besides the fact that the EHR is not adapted to speech technology, it has *smarttexts* and *smartphrases.* Managers tell that these are shortcuts in the EHR. Doctors just have to type the abbreviation of a commonly used word and the complete word shows up on the screen. When doctors are used to working with these shortcuts, speech technology is less beneficial for them. The fields of the EHR do not require much information. Polyclinic letters are therefore set up easily. When the doctor has filled in all information, he just has to click on the information he wants to state in the letter and an automatic text with previously filled in values shows up on the screen. This auto-complete technology competes with speech technology, which could be one of the reasons for the limited use of speech technology.

To use speech technology, various adaptations need to be done. First of all, doctors need to adapt their speech. Users have to speak calmly and articulate well. Besides, the work environment has to be adapted. A place needs to be created where the hardware and software is available, as well as a silent environment.

The human factor is the most important problem when implementing speech technology. Using speech technology, physicians need to do secretarial work (documenting) which evokes resistance, according to one of the managers.

Hospitals have many projects that are legitimately obliged to or that need to be done in terms of patient safety. These projects often have priority over speech technology. Furthermore, hospitals have limited staff and financial capacity for speech technology, and therefore are not able to support these projects. This support is one of the most important aspects when implementing speech recognition software.

The authorities in hospitals and the management structures are slightly different in the Netherlands compared to some other counteries. There is no central authority that makes decisions like working with speech technology or not. This is done departmental based or even individually, making the adoption of speech technology slow.

The extensive use of speech technology in the U.S. can be clarified by the fact that it was initially available for the English language only. English speech technology works better because it exists longer and had more time to mature. This is different for the Dutch recognition rate, due to the relatively small number of Dutch speakers in the world.

To sum up, there are plenty of factors influencing the implementation of speech technology. They are summarized in table 3.

4.3 Potential of Speech Technology

We asked participants their opinion about the potential of speech technology as a tool for documentation. An overview of the answers is showed in table 4. Indecisive answers were excluded from the table.

Participants who thought that speech technology as a tool for documentation has high potential, do not understand why the use is limited to 1%. One of the users compared speech technology to automatically driving cars. "It is already possible, but just a few people bought it. It will become cheaper, easier and people will get used to it, and next, adopt it. This is the same for speech technology" (U1). One of the people who did not expect a large potential worked at the central ICT department of UMC Utrecht, and experienced a booming period around 2005, but stated that the hype of speech technology is over nowadays. He explained that doctors who can work with smarttexts and smartphrases in the EHR do not benefit enough from speech technology.

Table 3: Overview of the different aspects mentioned by participants for the limited use.

Aspect	NU (20%)	NS (40%)	NM (40%)	Total (100%)
▪ Speech technology is only implemented at radiology and pathology departments	10%	20%	20%	50%
▪ Wrong expectations				40%
● It does not work properly		20%		20%
● It works immediately			10%	10%
● Accuracy keeps increasing	10%			10%
▪ Smarttexts and Smartphrases compete with the benefits of speech technology			30%	30%
▪ Adaptations need to be done when implementing speech technology				60%
● Way of speaking	10%			10%
● Work environment	10%	10%		20%
● Doctors need to adapt their way of working	10%	10%	40%	60%
▪ Hospitals have many projects with priority		10%	20%	30%
● No capacity left for support, while support is very important	10%	20%		30%
▪ Financial barrier	20%	20%	10%	50%
▪ No central authority in Dutch hospitals	10%	30%	10%	50%
▪ Extensive use in America due to English language		10%		10%

Table 4: Expected potential of speech technology as tool for documentation consistent with participants.

Speech technology as tool for documentation has potential	Speech technology as tool for documentation has no potential
1 supplier	1 manager
2 users	1 supplier
3 managers	

80% of all participants (all users, half of the suppliers, and all managers) think high potential can be expected in other applications than documentation, such as structured reporting. When data is entered fragmentedly, one can do analyses on these data. This way, more information is obtained from the enormous amount of data. This information can be used internationally by using codes such as ICD10 and ATC. Languages are not understandable by everyone, but these codes are the same for every language, a supplier explained.

In addition to structured reporting, speech commands are mentioned. Computers can be commanded by speech. This is useful when human hands and eyes are busy (Ajami, 2016), for example when operating.

Furthermore, decision support is mentioned as a potential application field by various participants.

This works as follows: the computer suggests a possible diagnosis based on the information entered by the physician. An overview of the particpants' answers on the pros and cons of speech technology can be found in Table 5. Column N states how many participants mentioned that particular aspect in percentage of all participants. The major advantages are the shorter RTT mentioned by 40% of the participants, and the decrease in time needed for administration, mentioned by 50% of the participants. Finally, the major disadvantage is the financial aspect, mentioned by 40% of the participants.

5 DISCUSSION

5.1 Conformity Literature and Results

In this study we investigated the reasons for the limited use of speech technology in Dutch health care. Our main findings to clarify the limited use were: speech technology is only implemented in radiology and pathology departments, doctors need to adapt their way of working, no central authority for Dutch hospitals, and finally the financial barrier.

Our main findings concerning the potentials and barriers of speech technology were the decreased RTT and the decreased time needed for administra-

Table 5: The pros and cons according to participants in percentage of all participants.

Pros	NU	NS	NM	Cons	NU	NS	NM
• Higher quality of documentation		20%		• Initial costs of speech technology are high	10%	10%	20%
• Complete and more adequate because there is more time				• Recognition decreases	10%		
• Doctors do not have to remember the information till they have time for documentation				• System makes a lot of errors	10%		
• More patient friendly	10%			• Users need to adapt their way of speaking	10%		
• In short time, better than typing	10%			• Transformation to speech technology costs money and time	10%		
• Decreased report turnaround time (RTT)	10%	10%	20%	• Doctors and other health care providers need to dictate in the context of the dictionary		20%	
• Less time needed for administration	10%	20%	20%	• Doctors and other health care providers need to correct mistakes by themselves		10%	
• Accuracy 99%		20%		• Doctors and other health care providers need to control the report by themselves		10%	
• Costs earned back by faster workflows		10%	10%	• The system has a long learning curve		10%	10%
• Useful for reports without standard template			10%				

tion as advantages, and the financial aspect as major disadvantage. We compared the mentioned pros and cons by participants with the pros and cons found in literature, by comparing Table 1 and Table 5.

In our study, patient safety was not mentioned by the participants as an advantage, but the factors that lead to an increased patient safety were mentioned. These factors are an increased quality of documentation and a shorter RTT. Ajami (2016) states that the duration of patients' stay is reduced by speech technology. This is not mentioned by our participants. Moreover, Ajami (2016) states that different accents are a disadvantage of speech technology because the technology cannot cope with this. In contrast, our participants explained that the technology can handle different accents because of the profile that needs to be made to get used to different accents and a users' vocabulary.

Furthermore, the study of Parente et al (2004) found that users can speak to the computer like they normally do to other people. This is contradicted with our findings. Our findings show that users need to speak slowly and articulate well. The recognition rate was found to actually decrease after a while. This is not found in literature, but can be explained by a finding of Ajami (2016). Since speech technology is a learning system, the system saves new words and

new pronunciations per word, as an individual does not have the same pronunciation every time. The vocabulary in the dictionary increases, and therefore the system will confuse words more often since dictionaries consisting of a lot of words tend to confuse words with each other more often (Ajami, 2016). Our comparison showed that the remaining aspects from Tables 1 and 5 correspond to each other.

According to the manager from the central ICT department at UMC Utrecht, the hype of speech technology happened around 2005. This statement is in accordance with the predictions of the Gartner hype cycle. This cycle characterizes a typical progression of a new technology (Linden and Fenn, 2003). In 2014, speech technology was already placed at the end of the cycle (Gartner, 2014). In 2015 and 2016 (the most recent one) the technology is not included anymore in the models (Gartner, 2015-2016). This suggests that the adoption of speech technology is already over. However, one of our interviewed users of speech technology refuted this trend. He stated that people will get used to the technology, and the technology will become easier and cheaper. Speech technology will reach the majority of the medical staff after this phase. This is consistent with Rogers's theory of innovations. First, the most progressive 2,5% (the innovators) will adopt

the innovation, and after this, the remaining four groups will follow (Rogers, 1995). Accordingly, speech technology is now only adopted by a part of the innovators which would indicate that the adoption of speech technology has yet to start.

The majority of the participants recognizes the potential for speech technology as a tool for documentation, but most potential is expected in other applications of the technology. The study of Parente, Kock and Sonsini (2004) expects a lot of potential for speech technology as a tool for documentation. The study of Johnson et al. (2014) is more cautious. They state that speech technology can have benefits, but there are many factors that need to be taken into account, such as financial problems and resistance of doctors (Johnson et al, 2014). The study of Ajami (2016) is less positive and states that the use of speech recognition is time-consuming, awkward and not userfriendly. However, they state that the technology will become reality in the end. Nevertheless, the more recent the studies, the more negative they tend to report on speech technology.

Finally, the human factor was mentioned by all groups of participants. Moreover, this is mentioned in many previous studies, and the study of Dawson (2014) is fully committed to this factor. This highlights the human factor as a major problem when implementing speech technology. However, it is possible that this is only the case in the beginning of the implementation process, because of habituation (Groves and Thompson, 1970).

5.2 Strengths and Limitations

To the best of our knowledge, this was the first study to explore the advantages and disadvantages of speech technology and to find the limitations of this technology for Dutch health care. Another strength of our study is the selection procedure of the participants. All participants had relevant experience in using, facilitating and/or implementing speech technology, and all different types of stakeholders were taken into account.

Nevertheless, our study had some limitations. Our findings are based on only ten participants. We would have wanted to increase the group of users, since we interviewed only two and the remaining groups all had four participants. Unfortunately, in the context of this research, all users were doctors, and no nurses were included. Because doctors are busy and hard to reach, we did not succeed in expanding the group of users.

5.3 Recommendations and Future Studies

Our findings have implications for different parties. Our results showed that managers, directors, ICT department staff and other people need to know better what to expect with the implementation of speech technology to get familiar with all the different factors that influence this process and its implementations. This way a well-considered decision can be made. Besides, users should know better what to expect of the technology to be better prepared to possible problems or obstacles. Also, when the decision to implement speech technology is made, we recommend intensive user support. This was rated as very important according to our findings.

More research is needed, preferably with a more extensive study design, to further confirm our findings. For future research, new target groups should be investigated, such as nurses, who document on average 8.7 hours per week (Nuance, 2015). According to Bosch (2005), speech technology could be very useful for the General Practice (GP) as well. It should be investigated how feasible implementing speech technology is for the GP. Furthermore, the other application potentials of speech technology as named in paragraph 4.3 could be investigated. Future studies are needed to investigate structured reporting, speech commands and decision support in practice.

Finally, we propose to design and evaluate such analytical applications of speech technology to improve the daily practices of domain experts from an Applied Data Science context (Spruit & Jagesar, 2016).

6 CONCLUSIONS

We discovered various barriers influencing the adoption of speech technology. However, the majority of participants in our study still thought there is high potential for this technology. They acknowledged that other applications of this technology may be more beneficial than documentation. Our results showed that speech technology is useful as tool for documentation at the radiology and pathology departments, but is less useful as tool for documentation at other departments. For those other departments, higher potential of speech technology is expected in other applications such structured reporting, speech commands and decision support.

REFERENCES

Advanced Voice Technology. (sd). *Is de transcriptie van een interview of vergadering middels spraakherkenning mogelijk?* Retrieved from Advanced Voice Technology: https://spraakherkenning.nl/over-spraak herkenning/interviews-en-vergarderingen-uitwerken/

Ajami, S. (2016). Use of speech-to-text technology for documentation by healthcare providers. *The National Medical Journal of India*, 148-152.

Bosch, B. (2005). Brian Bosch over spraakherkenning op de computer: Voor de huisarts met twee vingers ... *Huisarts en Wetenschap*, 891-892.

Dawson, L., Johnson, M., Suominen, H., Basilakis, J., Sanchez, P., Estival, D., Hanlen, L. (2014). A usability framework for speech recognition technologies in clinical handover: A pre-implementation study. *Journal of Medical Systems*.

Gartner (2014, 2015, 2016). *Gartner Hype Cycle.*

Geitgey, A. (2016, December 23). *Machine Learning is Fun Part 6: How to do Speech Recognition with Deep Learning*. Retrieved May 2, 2017, from Medium.

Groves, P. M., and Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, 419-450.

Johnson, M., Lapkin, S., Long, V., Sanchez, P., Suominen, H., Basilakis, J., and Dawson, L. (2014). A systematic review of speech recognition technology in healthcare. *BMC Medical Informatics and Decision Making*, 14.

Koivikko, M. P., Kauppinen, T., and Ahovuo, J. (2008). Improvement of Report Workflow and Productivity Using Speech Recognition—A Follow-up Study. *Journal of Digital Imaging*, 378-382.

Linden, A., and Fenn, J. (2003, May 30). Understanding Gartner's Hype Cycles.

NOS. (2017, April 3). *Koks, onderwijzers en artsen ervaren hoogste werkdruk*. Retrieved June 3, 2017, from NOS: http://nos.nl/artikel/2166294-koks-onderwijzers-en-artsen-ervaren-hoogste-werkdruk.html

Nuance. (2008). *Dragon NaturallySpeaking 10 Preferred.* : Nuance Communications.

Nuance. (2015, June 24). *Nuance Healthcare Partner event 2015*. Retrieved April 12, 2017, from Cedere: https://www.cedere.nl/2015/06/24/nuance-healthcare-partner-event-2015/

Parente, R., Kock, N., and Sonsini, J. (2004). An Analysis of the Implementation and Impact of Speech-Recognition Technology in the Healtchare Sector. *Perspectives in Health Information Management*.

Renckens, E. (2009, May 7). *Het luisterend oor van de computer*. Retrieved April 28, 2017, from NEMO Kennislink: https://www.nemokennislink.nl/publica ties/het-luisterend-oor-van-de-computer

Rogers, E. M. (1995). *Diffusion of Innovations.* New York: The Free Press.

Schumacher, J. (2017, May 22). *Cijfers: vergrijzing en toenemende zorg*. Retrieved June 8, 2017, from Zorgvoorbeter: http://www.zorgvoorbeter.nl/ ouderenzorg/hervorming-zorg-cijfers-vergrijzing.html

Shagoury, J. (2010). Chapter 11 Dr. "Multi-Task": Using Speech to Build. In A. Neustein, *Advances in Speech Recognition* (pp. 247-273). New York: Springer US.

Spruit, M., & Jagesar, R. (2016). Power to the People! Meta-algorithmic modelling in applied data science. In Fred,A. et al. (Ed.), Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (pp. 400–406). KDIR 2016, November 11-13, 2016, Porto, Portugal: ScitePress.

Tuin, C. (2016, October 20). *Microsoft claimt doorbraak spraakherkenning*. Retrieved from Computable: https://www.computable.be/artikel/nieuws/developme nt/5860772/5440850/microsoft-claimt-doorbraak-spraakherkenning.html

Vervoort, D. (2017, April 21). *Zo werkt spraaktechnologie*. Retrieved April 26, 2017, from Clickx: http://www.clickx.be/achtergrond/169324/zo-werkt-spraaktechnologie/