
The design of a reconstructive hand surgery text database based on a speech recognition system in Iran

Marjan Ghazisaeedi¹, Reza Safdari², Abdoljalil Kalantarhormozi³, Leila Shahmoradi¹,
Fateme Sadeghi^{4,*}

¹Faculty Member, Health Information Management, School of Allied medicine, Tehran University of Medical Sciences, Tehran, Iran

²Health Information Management, School of Allied medicine, Tehran University of Medical Sciences, Tehran, Iran

³Professor of Medicine, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁴Health Information Technology, School of Allied Health Sciences, Tehran University of Medical Sciences, Tehran, Iran

Email address:

Ghazimar@tums.ac.ir (M. GHazisaeedi), RSAFDARI@tums.ac.ir (R. Safdari),

Abdoljalil_kalantarhormozi@yahoo.com (A. Kalantarhormozi), leilashahmoradi1@gmail.com (L. Shahmoradi),

Fateme_S2860@yahoo.com (F. Sadeghi)

To cite this article:

Marjan Ghazisaeedi, Reza Safdari, Abdoljalil Kalantarhormozi, Leila Shahmoradi, Fateme Sadeghi. The Design of a Reconstructive Hand Surgery Text Database based on a Speech Recognition System in Iran. *International Journal of Biomedical Science and Engineering*. Vol. 2, No. 3, 2014, pp. 17-22. doi: 10.11648/j.ijbse.20140203.11

Abstract: Speech recognition system (SR) is a tool to create comprehensive, accurate and legible information in the patient's medical records. The present study is a part of a project to create a specialty text database for SR software in the field of reconstructive hand surgery. This project is considered a fundamental work for practical application of the system to enter the healthcare field in Iran. An interventional, quasi-experimental study and fundamental work to create surgical text database in the existing Persian software of "Nevisa". 1863 reconstructive hand surgical descriptions during a 3-month period (March 20, 2012 through June 20, 2012) in 15-Khordad Subspecialty Hospital were used to create database. The statistical population of the study consisted of all patients admitted with hand disorders within this time interval. About 108 dictated voices of physicians were collected after surgery. After the type and required analysis, the reconstructive hand surgery text database was created and tested in a real operating room. In this paper, a reconstructive hand surgery text database was created for Persian SR software. The database contains capital vocabulary and more than 1200 words for use in the hand surgical unit in specialty hospitals and clinics. The record of findings and results of surgery in the patients' medical records is very important. The lack of computerized data recording is one of the reasons for failure and illegibility of medical records in health centers. To solve these problems, software was used for data entry in the medical records and creating comprehensive electronic records. The designed software was tested in the operating room. The obtained accuracy demonstrated the potential for practical applicability of the developed software in Iran.

Keywords: Speech Recognition Software, Reconstructive Hand Surgery, Database

1. Introduction

The accurate and timely record of medical records is considered as a problem in healthcare which still exists despite computer developments. The entrance of technology in the specialty field of medicine was very gradual and slow due to the complexity of the medical profession [1]. With the advancement of technology, the need for computer knowledge and the required time for data entry into these devices obliged human to think how to use comfortably this technology spending less time [2]. In this regard, human interested in the development of

computer technology and tried to control and communicate with the computer via speech [3]. This idea has received much attention by many researchers [4, 5]. In this regard, some software (e.g. SR) have been designed with the capability of converting speech into typed text [6].

The data recorded in the medical records has a significant impact on the quality of patients' care [7]. The data must be recorded timely, accurately and completely in all pages of the medical record including surgical specific pages for patients who underwent surgery. Given the

importance of recording information in medical files and the significant role of these files in research works, the correct record of data in all papers, especially the pages related to surgeries is of remarkable importance [8]. One of the obvious benefits for speech communication with computers is to enter data when the hands are busy with other works, for instance, a surgeon who involved in doing surgery. In such a situation, we first have to finish one job and then be able to enter data. After surgery, due to fatigue and excessive workload because of high accuracy in work, the use of speech interface is highly desirable in terms of completeness and accuracy of the records [4, 9]. It also provides more timely documentation and maintenance of data quality [1].

Of problems which involved the health centers in Iran is unwillingness of physicians to record the findings of patients' examination in the medical records [10]. The other problem facing the medical staff is illegibility of handwritings related to other specialists and surgeons. This problem may make the vital information useless and exert irreparable damage on the status and health of society [11].

In medicine, hand is the most common organ that is damaged during the industrial works or agricultural activities. The incidence of hand injuries has been reported in the range of 6.6 to 28.6%. Hand injuries are observed in 20% of patients with multiple trauma. The treatments of hand injuries are often long-term and expensive process and a lot of good time is lost as well. There is the likelihood of lifelong disability, while the previous studies reveal that the hand injuries are dominant [12].

On the other hand, hand injuries are the most common reconstructive hand surgery emergency cases. The most common hand surgical emergency is hand *lacerations* which include various causes such as work accidents, street battles, suicide, car accidents, inattention cuts while working at home and self-mutilation. Most patients with hand injuries refer to the physician with acute injuries [13].

The existing system in Iran is able to meet the needs of users in Persian language. However, due to lack of coverage of the database of this system in different specialty sectors, it is necessary to define the database for each specialty to employ the system in various scientific fields. The present study provides the context for practical use and the entry of technology into the field of healthcare by creating a text database for SR system and optimization of specialty database in the field of reconstructive hand surgery.

2. Methods & Material

2.1. Text Data Collection

In this study, all surgical descriptions in the field of reconstructive hand surgery were collected during a 3-month period (March 20, 2012 through June 20, 2012) in 15-Khordad Subspecialty Hospital affiliated to Shahid Beheshti University of Medical Sciences. The relevant

descriptions were typed and turned into electronic texts. The total number of patients admitted within the selected interval was 2407. The number of patients in the selected field was 1737 patients. 1863 descriptions were collected. The difference between the number of patients and descriptions (126 cases) was due to more than one surgery for each hospitalized patient which included about 77.4% of all patients referred within the selected interval.

2.2. Audio Data Collection

Along (simultaneous) with the first stage, 108 surgical descriptions were randomly selected among 1737 descriptions. The selected descriptions were dictated by physicians (5 first-year Plastic Surgery Residents and one Hand Surgery Fellowship) to achieve various pronunciations of words and abbreviations used by specialists and comparative investigation of dictated surgical descriptions and recorded formats. After writing the surgical description by physician, the audio data of surgery was dictated and saved using a Zen Stonplus MP3 Player. The audio files were converted into electronic documents in a separate file and compared with files typed of paper surgical descriptions. After comparing the files, a large difference was not observed between the speech and texts written by physicians. However, it was very helpful to access pronunciation varieties. To use the surgical descriptions in the next steps, the data should be converted into a format such that the extraction of statistical data would be easily feasible.

2.3. Uniformization Section

The lack of following a single standard in Persian writing (especially in the computer environment) and writing words with two or more different spellings causes the same two words with different characters. For calculating the frequency, the same words with different spellings are considered and counted as two different words. Attempt was done to find different spellings of the words to convert different spelling states to a single spelling. On the other hand, the text contained abbreviations, non-standard words, individual dialects, various forms of writing and expression. The accuracy of the system was enhanced to use it in real situations by uniformization of texts. Also, the numbers were written in Persian, English, Latin and Roman in some cases which converted only into numerical style. In pronunciation section, all items were defined for the software as pronunciation species.

2.4. Vocabulary Frequency

To create the database, an important part of the work is to obtain a list of all words and the frequency of the word in the text. This is used to assess the accuracy and efficiency of this system in the experimental environment. In this case, whole electronic file was separated as words, and the frequency of occurrence of each word was obtained using Microsoft Excel.

2.5. Pronunciation Writing

In this step, the pronunciation of existing words in the text were written using IPA standard and 29 phonemes of Persian language to complete the phonetic or pronunciation section of the system.

Table 1. 29 phonemes of the Iranian phonetic alphabet (IPA)

E	A	AA	I	U
R	S	SH	Y	O
K	L	Z	ZH	X
CH	V	B	N	M
T	AH	D	F	P
	Q	G	H	JE

To write the pronunciation of a word, the sequence of phonemes must obey the pronunciation pattern in Persian language. The rule is as follows:

If C and V represent the consonant and vowel phonemes, respectively, then the resulting sequence, from left to right, must be consistent with an arbitrary sequence of CVC, CV and CVCC patterns.

2.6. The Linguistic Model

The linguistic model trainer was used to create the linguistic model. The trainer creates the grammar and vocabulary which are used by recognition module.

2.7. Testing the Software

The software was installed on ASUS COREi5 laptop after designing and incorporating the specialty text database to "Nevisa" software. The ANDREA professional headset microphone was used to collect voices. To test the system in the operating room, physicians were divided into two categories of trained (voice training course approximately for 20 minutes to learn the dialect and voice of the physician by the system) and untrained groups. Each physician was asked to dictate 10 surgical descriptions to the system. Since the software was able to type in Microsoft Word, dictating and typing were performed in the standard surgical description by designing the surgical description page and placement in the type environment of the system. After analyzing the findings, the system accuracy was obtained for both groups.

2.8. Hardware and Software

The first section contains the hardware and software requirements for a computerized system to support the implementation of Nevisa software. The system must be running Windows XP and above, Persian language and keyboard setup, Dot Net 2.0 environment, 2 GHz P-IV CPU and higher for PC (and its equivalent for other systems) equivalent to 512 MB RAM and 100 MB of disk space and soundcard with headset microphone input (with a

quality suitable for speech recognition).

The second part includes the main components of the SR software. The software follows the pattern-based approach. This approach consists of two phases: testing and training. The testing phase includes comparison of the input speech patterns with trained patterns to identify the phonetic units of the input speech which consisted of the search and feature extraction. The search is performed through linguistic models, the phonetic models and terminology in the training phase.

In the training phase, patterns associated with each class with same phonetic units are being modeled. In this phase, there are usually phonetic and linguistic models which are used in the testing phase. The extraction of phonetic models is carried out using different methods employing speech database, while linguistic models and vocabulary are extracted using text database. The need for both parts has been met in the design of reconstructive hand surgery database to employ the system in the specialty field of reconstructive hand surgery [5, 14].

3. Results

The first step for designing a database is to create a text configuration of relevant specialty [15]. The text configuration must cover the most electronic texts of the given specialty. The electronic texts obtained from paper surgical descriptions contained vague and incomplete words. The problem was resolved with the help of documentary-makers. The words were uniformed after typing the texts.

The dictated audio data were compared with corresponding text file. There was no significant difference between speech and writing of both files. For comparison, the total number of the words in both files and their varieties were obtained. The comparative results are presented in Table 2.

Table 2. Comparison of the number and variety of words in written and recorded surgical descriptions

Word	Word Repetition	
	Audio data	Textual data
Total word	5226	5261
Variety of words	520	569

According to Table 2, the lengths of surgical descriptions in audio files do not differ much with text files in terms of word count and the length is almost identical in both files.

Table 3 represents the comparison of writing methods of various flaps and their frequencies in both files. According to the results, there are different writing methods for surgical descriptions which well indicates the need for text uniformization. For example, a word such as "flap" was written in three forms. All three written forms have a same pronunciation. Such cases result in acoustics error in the software and significantly reduce the accuracy of the system.

Table 3. Comparison of word count in both texts and matching different samples

Word within files	Number of words in audio file	The number of words in a text file
Advance	-	2
Advancement	-	5
Flap	-	1
v-y advance	-	3
v-y advancement	-	2
F E L A A P	15	12
F E L A A P H A A	1	1
F E L A A P H A A Y E	6	1
F E L A P	18	22
F E L A P H A A Y E	5	2
V I V A A Y A D V A N S	3	1
V I V A A Y A D V A N S M E N T	4	-
Sum	52	52

Table 4 shows the comparison of various anesthesia in the paper and dictated surgical descriptions. The results indicate that there are 11 writing and spelling models for two types of surgical anesthesia.

Table 4. Comparison and matching of the type of anesthesia used during surgery for each patient in dictated and text files

Word within files	Number of words in audio file	The number of words in a text file
GA(General Anesthesia)	-	24
LA(Local Anesthesia)	-	41
AH A A N E S T E Z I Y A A Y	5	-
AH A A N E S T E Z I Y A A Y E J E E N E R A A L	8	-
AH A A N E S T E Z I Y A A Y E L O K A A L	4	-
AH A A N E S T E Z I	3	34
AH A A N E S T E Z I Y A A	19	-
J E E N E R A A L	8	4
J E E N E R A A L A H A A N E S T E Z I Y A A	25	-
L O K A A L	28	4
L O K A A L A H A A N E S T E Z I Y A A	6	-
Sum	106	107

For this reason, after uniforming all words and converting them into a single word in the text, in addition to the pronunciations of the word, the written forms have been defined in the pronunciation writing section of the system as pronunciation. The total number of words in the text was 1324 words which should be pronounced. The number of pronunciations consisted of one to 18 cases in

the various types of words. Within the whole text, there were only two words with the same spelling and different meanings. The first word was PIN (with two meanings: a tool for bone fracture and crash fixation and the abbreviation corresponding to Posterior Interosseous Nerve). The second case includes two words; "In in Persian" and "In in English" which have different written forms and same pronunciation.

In the next step, to build a linguistic model, the training set was obtained using the words in the electronic texts of surgical descriptions. The vocabulary file includes the vocabulary of training text and the phonetic versions of each word. The phonetic versions are used to relate each word with the corresponding sequence of phonetics. The grammar file contained the sum of the probabilities of occurrence of each word after the other words in the training set. The following data were extracted from the content for the linguistic model:

1. Number of occurrences of each word in the context (monogram of words)
2. Number of occurrences of each pair of words in the context (bigram of words)
3. The number of occurrences for each triple of words in the context (trigram of words)

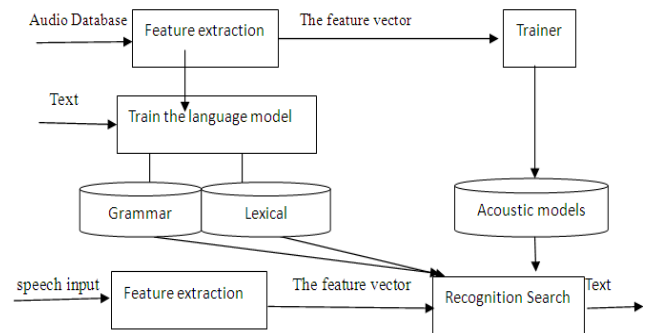


Figure 1. The stages in the development of the linguistic model

The linguistic model trainer consists of several parts. The data entered into this section of the text are different with audio files. With the help of this feature, the linguistic model can be trained independent of the audio model. The number of occurrences of the words and pair of words are derived from intermediate section of the database. If the size of training set is limited, the estimation of the probability of words may be inaccurate. In this case, the probability can be calculated for a set of words instead of words.

The database design consists of three stages. The final step is to add the obtained items to the software. The first stage of software design is to create electronic texts. The other relevant items are extracted from electronic texts. Then, the extracted items are added to the speech database to prepare the software to use in the desired specialty. The developed software can be used in hospitals and sectors with reconstructive hand surgery specialty. This process is shown in Figure 2.

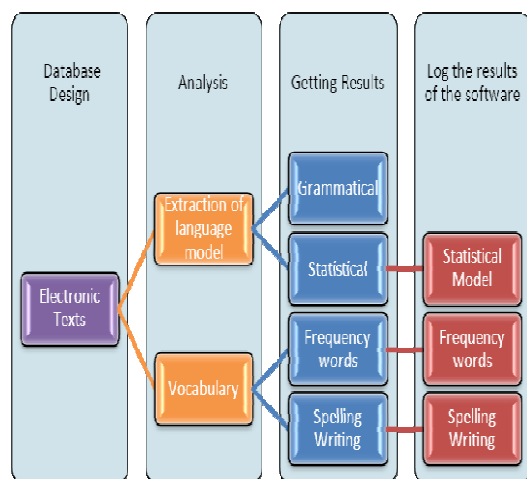


Figure 2. The steps and components of text database design in the speech recognition system

Errors were divided into several groups for analyzing the items in the testing phase. This information is completely presented in Table 5.

Table 5. The accuracy and the type of errors in the software after testing in operating room

Errors	User with training	user without training	Total
out of the database	10	32	3%
Substitution	3	9	0.9%
Spelling	3	6	0.6%
Insertion(noise)	69	124	13.7%
Deletion	5	10	1.1%
Total	90	181	271
Total words dictated	557	848	1405
System accuracy	83.8%	78.7%	80.7%
Error rate	16.2%	21.3%	19.3%

The most error (13.7%) is related to environmental noise. This had a significant impact in reducing system accuracy in the testing phase.

4. Discussion

The present paper examines a new database in the field of plastic surgery and reconstructive hand surgery for the SR software. This database contains a large vocabulary and more than 1200 words in Persian language and continuous speech recognition software which has the potential for use in plastic and reconstructive hand surgery hospitals. This software is the first attempt to introduce a SR system for the medical field in Iran. The current software is a real-time, speaker-independent, word-based software with the linguistic model of n-gram. In this project, the basic system of Nevisa and the speech database of this software were

used to employ the proposed text database.

After completing the design phase, the system was used experimentally in real environment to evaluate the performance of the software. The accuracy of the system for trained and untrained users was obtained equal to 83.8% and 78.7% respectively after testing the software in the operating room and collecting 60 dictated surgical descriptions.

According to the results of [16], 5% improvement in accuracy among users with different levels of training indicates this fact that the training of users leads to enhanced system accuracy. About 3% of the error is related to the non-coverage of the vocabulary of the database which can be added to the system easily.

The most error is related to environmental noise (13.7%). Given the use of professional headset microphone, the environmental noise is due to the lack of a separate room for the dictating the system. According to [16-18], data collection in a separate environment or independent room for dictation can eliminate this problem and enhance the accuracy of the system up to 95%.

5. Conclusion

The database of SR systems as the heart of the system must have sufficient integrity and a maximum vocabulary with the capability of adding new terms [11]. Accuracy, the detection power of the system, speed, ease of application for the user, cost-effectiveness [19] and communicating with hospital information systems [15, 20-22] are important parameters in evaluating the success of the system.

The record of the findings and results of surgery in the patients' recordings is very important. In fact, the lack of computerized recording of data is one of the reasons for failure and illegibility of medical records in health centers. For this reason and to solve these problems, a software was designed for data entry in the medical records and creating comprehensive electronic records. Since the SR technology makes it easier to create electronic medical records [23] and due to increased required time for data entry into the EHR which has reduced its relevance and application, the use of SR software could potentially facilitate the data entry into EHR systems and enhance the EHR relevance [24]. For this reason, some users consider this software as a national solution for EHR production [23].

Acknowledgements

Thanks to the cooperation of 15 Khordad Hospital authorities and physicians to assist in the design and testing of software. Thanks management and engineering company Guesh prdaz time to work on the project and the use of database software, Nevisa. Also thank of Mr. K. Hosseinzadeh and B. Baba Ali due to the sincere and selfless assistance during project.

References

- [1] Fournier LC, "Towards a natural spoken language order-entry system for the ICU: developing a language model from handwritten physician orders," Portland: Oregon Health & Science University; 2001. (dissertation)
- [2] Safari M, "Automatic Speech Recognition," Iran: Islamic Azad University Lahijan; 2007. (dissertation)
- [3] Ashrie j, "Design and manufacturing of critical control systems using voice," Iran: Science and Technology Tehran; 1998. (dissertation)
- [4] Saeedi M, "voice recognition," Shabakeh Magazine [Internet]. 2003 Dec [cited 2003 Dec];1(38):59-5. Available from: <http://shabakeh-mag.com/article.aspx?id=1000206>.
- [5] Babaali B, "Pruning techniques to improve the performance of a Continuous Speech Recognition system based on hidden Markov models and their application," Iran: Sharif University of Technology; 2003. (dissertation)
- [6] Aziz Khani Z, Naghshine N, "Overview of the speech recognition system and its application," *IranDoc Scientific Communication monthly Journal* [Internet]. 2007 May 25 [cited 2007 Jun 25];7(1):64-9. Available from: http://ejournal.irandoc.ac.ir/browse.php?a_code=A-10-2-99&slc_lang=fa&sid=1.
- [7] Zafar A, Mamlin B, Perkins S, Belsito AM, Overhage JM, McDonald CJ, "A simple error classification system for understanding sources of error in automatic speech recognition and human transcription," *Int J Med Inform.* 2004;73(9-10):719-30. Epub 2004/08/25.
- [8] Sharifian R, Ghazisaeedi M, "Evaluation of the data sheets for specific surgical patients discharged from hospitals in Tehran University of Medical Sciences in 2005," *Payavard Salamat.* 2008;2(1-2):31-9. Persian.
- [9] Edgerton M, "The Art of Surgical Technique," 1nd ed. Tehran: Institute publications puiandeghane teb; 2002. Persian.
- [10] Safdari R, Meidani Z, "Health services accreditation standards for information management in Canada, New Zealand and USA: a comparative study," *The Journal of Qazvin University of Medical Sciences.* 2007;11(1):73-6. Persian.
- [11] Sheikhtaiefe M, "A comparative study of speech recognition systems Clinical data documenting the process in selected countries in 2010," Iran: Tehran University of Medical Sciences; 2010. (dissertation)
- [12] Fatemi M, Elmirad J, Rajabi F, "Distribution of Hand Injuries from a Referral Teaching Hospital," *Iranian Journal of Orthopaedic Surgery.* 2008 Autumn;6(4-A):191-6. Persian.
- [13] Hassanpour E, Loghmani S, Attariyan S, Fadaee A, "Evaluation of accuracy of preoperative examination of flexor tendons in zone 4 and 5 of hand injuries," *Journal of Birjand University of Medical Sciences.* 2009;16(3):16-6. Persian.
- [14] Groc B, Tuffelli D, "A continuous speech recognition system for data base consultation in Acoustics. Acoustics, Speech, and Signal Processing," *IEEE International Conference on ICASSP '80; 1980 Apr 9-11; New York: IEEE; 1980.* (Conference)
- [15] Kanal KM, Hangiandreou NJ, Sykes AM, Eklund HE, Araoz PA, Leon JA, et al, "Initial evaluation of a continuous speech recognition program for radiology," *J Digit Imaging.* 2001;14(1):30-7. Epub 2001/04/20.
- [16] Houston JD, Rupp FW, "Experience with implementation of a radiology speech recognition system," *J Digit Imaging.* 2000;13(3):124-8. Epub 2004/09/14.
- [17] Alapetite A, "On speech recognition during anesthesia," Denmark: Department of Communication, Business and Informations Technologies Roskilde; 2007. (dissertation)
- [18] Molnar B, Gergely J, Toth G, Pronai L, Zagoni T, Papik K, et al, "Development of a speech-based dialogue system for report dictation and machine control in the endoscopic laboratory," *Endoscopy.* 2000;32(1):58-61. Epub 2000/02/26.
- [19] O'Hara SP, Bryant TN, Oji EC, Rowe DJ, "Speech recognition and the clinical microbiology laboratory," *Med Lab Sci.* 1992;49(1):20-6. Epub 1992/03/01.
- [20] Langer SG, "Impact of speech recognition on radiologist productivity," *J Digit Imaging.* 2002;15(4):203-9. Epub 2002/11/05.
- [21] White KS, "Speech recognition implementation in radiology," *Pediatr Radiol.* 2005;35(9):841-6. Epub 2005/05/19.
- [22] Hayt DB, Alexander S, "The pros and cons of implementing PACS and speech recognition systems," *J Digit Imaging.* 2001;14(3):149-57. Epub 2001/11/27.
- [23] Bunschoten B, "What role will speech recognition play in health care?," *Health Data Manag.* 1996 Jan;4(1):38-6.
- [24] Hoyt R, Yoshihashi A, "Lessons learned from implementation of voice recognition for documentation in the military electronic health record system," *Perspect Health Inf Manag.* 2010;7:1e. Epub 2010/08/11.