Sindhi Speech Recognition System

Fida Hussain Khoso[a,b*], Dil Nawaz Hakro[c], Syed Zafar Nasir[a]

[a]Indus University, Karachi, Pakistan [b]Dawood University of Engineering & Technology, Karach, Pakistan.
[c]Institute of Information and Communication Technology, University of Sindh, Jamshoro

Abstract
For languages around the world, one of the most revolutionary and significant technology is the Speech Recognition, which helps in advancement of languages computing. Speech recognition systems for most of the languages around the world have either been already developed or under the process of development. Unlike other languages, the number of characters and the sounds in Sindhi language is more than the other languages, which makes it unique so most of the available approaches cannot be used for Sindhi language. Therefore, a totally different approach is required in correspondence of Sindhi sounds and characters for designing and development of the tools for Sindhi Speech Recognition.
The literature review of Sindhi speech recognition reveals that there is a gap in the quality work and no SR system is available (to the best of our knowledge) with respect to software development for recognition of spoken Sindhi characters and sounds, which may be converted in Sindhi writing by the software. If the input is made through Sindhi speaking, it will not only help the people in reducing effort of typing of very difficult character sounds and characters of Sindhi but also be beneficial for developers who need to map the 52 characters of Sindhi Alphabets against 26 alphabets of English keyboard.

Key-word:
Speech Recognition, Acoustic Model, Artificial Intelligence, Natural Language Processing.

1. Introduction

The latest and most important innovation in the field of computer sciences is the Artificial Intelligence (AI). It is the intelligence demonstrated by the computers just like a human being. Through AI, the powers of thinking and analysis are provided to the computers to make these machines capable of making intelligent decision making as shown in Figure 1. The AI has many branches and
one of the branches, wherein the computers are made to understand, recognize and write human languages
through complex programming is called Natural Language Processing (NLP). The algorithms related to
Speech Recognition are developed and programmed to make the computers understand human spoken words, which are then converted into written text by use of system commands and further processing. The efforts of typing are enormously reduced and the same can even be eliminated, thus saving time that otherwise would be required in typing the sentences[1]. The implementation of algorithms for recognition of different languages requires different approaches and can vary significantly from each other. The literature review related to speech recognition has revealed that significant work has been done on the languages spoken in Latin script languages such as English, Spanish, French and German. Due to the small number of phonemes and non-cursiveness; however the work on languages spoken in the less Arabic script has been negligible, in fact nearly non-existent (to the best of our knowledge), specifically on Sindhi language which possesses rich cultural and literary background. Therefore, a need is felt to work on this area, specifically on the development of Speech to Text Recognition System for Sindhi language. The typing efforts can be reduced for typing of very difficult sounds and characters of Sindhi with the development of speech recognition system for Sindhi language[2].

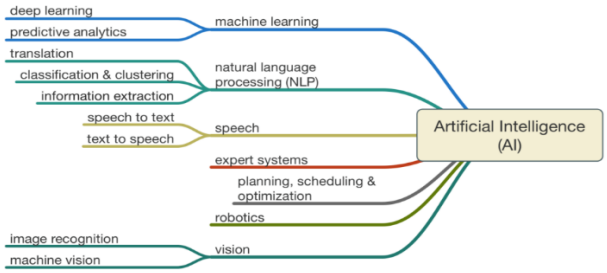

Fig. 1 Branches of AI (Courtesy : google image)

Based on the collection of phoneme library and voice samples, a speech recognition system for the specified language can be developed. The samples can be collected from various subjects of the language to create an exhaustive database. From these databases, spoken words or speech can be recognized, which by converting into text or to control computer system, mobile system, machines, cars and other devices. Multipurpose device controlling can also be done by using speech library [2].

1.1 The History of the Sindhi Language

Sindhi Language is one of the oldest languages in the world with its history spread from 600 BC to 500 AD i.e. over 1100 years. According to this research, it is not right that the Sindhi language was extracted directly from Sansakrit,

Manuscript received November 5, 2019
Manuscript revised November 20, 2019

which is a general impression. This mis-understanding that Sindhi was directly extracted from Sansakrit is based on reasons. One of the reasons is that some Hindu scholars as a preaching instead of facts or researching have seconded this impression in their writings. In addition, the research of European scholars on languages was influenced by the impression that all the languages in Hindustan (Greater India pre-partition) were extracted from Sansakrit due to their initial trainings. Moreover, as the Arabic language evolved from Sansakrit and most Muslim scholars are of the view that Sindhi has been extracted from Arabic language, therefore it is a general impression that Sindhi must also have been extracted from Sansakrit.

At a point when networks meet vorious social foundations start to impact each orther.The networks which are monitored accesssively to pickup speakers from the poorer networks. As far the Sindhi language is conserned the potential risks involve in decreasing no of speakers, loss of renowned, the extended Arabanization of the populace where the guardians are reluctant to bestow Sindhi as the primary language to the kids. It is hard to lean due to substantial no of phonetic components and the poor proficiency rate of the dialects. By the 1998 evalution by legislature of Pakistan the education rate in the area of sindh is 45.29%[4].

Sindhi voice had been advanced for so many centuries. The voice of general public of Sindh is interacted with Aryan and it has evolved to become Indo Arayan(Prakrit). The strong establishment of Sansakrit and Prakrit is Sindhi dialects, and this is called voice of India it has some lexis from Dravidian from Mediterranean, persian and Arabic and therefore it is called Moen-Jo-Daro human advancement(Allna 2002). In India about 1.2 million individuals lions shares of whom relocated from the region of Sindh so Sindhi is likewise one of the perceived authority voices of India. Sindhi language is spoken as there first voice by appromaxiately 4000000.

1.2 Speech Recognition (SR)

The sub-field of computational linguistics that develops methodologies and technologies aimed at enabling the recognition and conversion of spoken language into text by the computers is called as speech recognition.

1.3 Sindhi Alphabets

The total number of Sindhi letters is 50 which is a supper arrangement of Arabic,Persion and urdu. Which appear all to gether in fig1. It has also some special characters belonging to a section from the essential accentuation characters and numbers just like ؕ‘and’and ؔ‘in’ (Dow, 1976). Every alphabet which is inorder posses more than one frmae contingent upon its position which is represented by actual written work. As per rule the last, centre, starting, and independent letter has four structures.

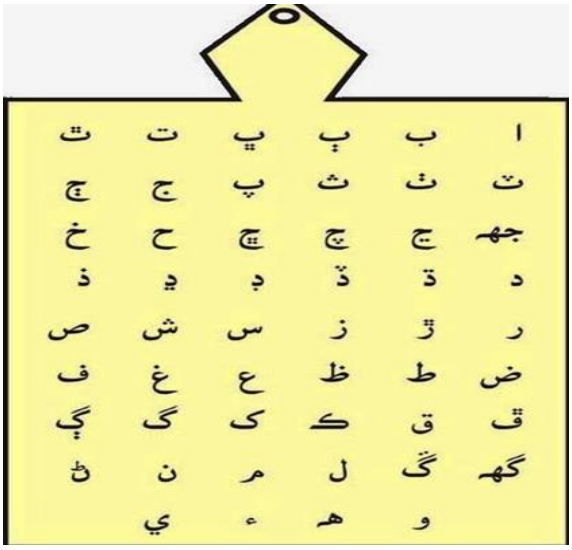

Fig. 2 Sindhi Language Alphabet(Courtesy : google image)

1.4 Sindhi Noun

In Sindhi language every noun is solitary or plural reapteadly associated conclusion(sounds of vowels indistint structures).

i.	چڪور	ساوڻ	رمضان
ii.	حيدرآباد	گمبٽ	دادو
iii.	پاڪستان	هندستان	سريلنڪا
iv.	مچي	اجرڪ	وانگڻ

Fig.3: Sindhi Nouns

2. Literature Review

The probabilistic lexical modeling approach has been presented by the (Rasipuramet,al,2016), which is based on articulatory features for the construction of continuous speech recognition system. The approach is based onintegrating articulatory features inside an automated speech recognition system. The relationships between observations of acoustic features and various types of features (articulatory) has been presented. The author has also proposed graphemes and phonemes as a subword unit. Further, the KL-HMM approaches been used for the estimation of lexical models parameters. The paper has also presented the subword and word level analysis.

As natural languagee processing(NLP) there are modest number definiton acknowledgement methods which are outstanding. Large number of calculations were created as clarifications. For this procedure the common language processing(NLP) systems and they chop back and be

rummage. With respect to the synthestic names these methods do not work productively. For clarification weight on the synthestic substances and in this manner it is complimentary[1].

The paper presents a framework that is based on Carnegie Mellon University's open source CMU Sphinx-4. It is a framework that consists of exhaustive vocabulary, is speaker-autonomous, and constant discourse acknowledgments that is based on the view of discrete Hidden Markov Models (HMMs). The conceivable flexibility of this framework has been used to demonstrate Arabic voice acknowledgment with the help of Hello_Arabic_Digit, which is an application to exhibit the conceivable flexibility of this framework to Arabic discourse. The researchers have tried to reach out to wide Arabic dialect acknowledgment through this application, specifically in the context of Moroccan vernacular dialect [5].

In another study related to Arabic style Sindhi consonants and vowel segmentation, the author has claimed that the approach used by them can be applied on any language of the world especially Arabic script Sindhi language and its neighbour Urdu. The study has used Sindhi language as an example, because of the fact that Sindhi possesses the superset of characters and contains all characters of the other languages. The study is aimed at refinement of theSindhi Language standard phonemes (set by international Phonetic Associations (IPA)) that contain consonants, diphthongs, semivowels and vowel by redefining the same. It is claimed by the study that it has provided a general guideline to create speech recognition not only for Sindhi but also for other Arabic script adapting languages in general[6].

3. Complance and Difficulties of Sindhi Speech Recognition System

3.1 Sindhi Language Features

The back history of sindhi language can be taken from the 5000 years old indus civilazation which is closed to larkana area of sindh (about indus 2014) sindhi alphabets are 24 more than the arabic of 28 and they are total 52. From these altered letters have been added with four dots to suit different sounds. Sindhi language consists of large no of vowels and consonents than arabic and its sister language urdu. The framework of the composition takes after a similar shape of arabic content in which alphabets are written from apporiate to left where as the numbers are written from left to right[8].

Isolated	Start	Middle	End
ع	عـ	ـعـ	ـع
ک	کـ	ـکـ	ـک
ق	قـ	ـقـ	ـق

Fig. 4 Places of Characters with Diverse

As it is shown in Fig3 the Sindhi language comprises of 52 letters in such a way each and every letter has its own position in phrase. To complete a segement these letters have two or four shapes. Every letter depends on different number and same base of class which has a diverse classes which is given below in Fig4. In (Hakro et al 2014) the identification and segmentation of Sindhi letters have been discussed.

Fig. 5 changed number and same shape of dotand their places (Hakro, et al,2024)

3.2 Classification of Various Sounds of Sindhi

By utilizing articulatry data, place and way of explaination the consonents are tending to be less demonding to depict and characterised in bunches than vowels[10].

3.3 Implosive Stops

In a couple of the world dialects sound stock, the implosive stops exist. Among them the remarkable one's are the Hausa and Swahili languages talked in Aferica where as Sindhi is talked and spoken in the Indian Subcontent (Nihalani 1986). By bringing down the larynx with vibrating vocal ropes the implosive of Sindhi which are simply non suctioned glottal in gressive sounds are created (Raza el al 2004). The implosive of Sindhi are created with the vibrating vocal lines so in this way they are alluded to voiced consonents of Sindhi.

To fulfil and complete the accoustic examination of these sounds is one of the best inspirations of this work.In very few of the world languages sound inventory, implosive stops do exist. The noteable examples of these languages are the Sindhi which is spoken in Indian sub-continent and the Hausa and Swahili spoken in Africa. The Sindhi language among all these languages posses the big number of contrastive implosive stops. By lowering the larynx with vibrating vocal cards, the implosive of Sindhi are absolutely non aspirated Glottal in grassive sounds. The vibrating vocal cards which produce implosives of Sindhi which are referred to as voice consonants of Sindhi. The

greatest motivations of this research work are to carry out the acoustic analysis of the sound[12].

Voiced		Place of articulation	Voiced		Place of articulation
/ɓ/	ٻ	Bilabial	/ʃ/	ڇ	Palato-Alveolar
/ɖ/	ڈ	Retroflex	/ɠ/	ڳ	Velar

Fig. 6 Stop consonants of Sindhi

3.4 Nasal Consonants

On lowering the velum in the mouth, Nasals are produced. They allow the air passage through the nasal cavity (Olive 1993). Produced with vibrating vocal cards there are five nasals in Sindhi and they are all voiced[13]. The nasal consonants of Sindhi are shown bellow Table 7.

Voiced		Place of articulation	Voiced		Place of articulation
/m/	م	Bilabial	/ɳ/	ڻ	Retroflex
/n/	ن	Alveolar	/ɲ/	ڃ	Palatal
/ŋ/	ڱ	Velar	--	--	--

Fig. 7 Nasal Consonants of Sindhi

3.5 Stop Consonants

The vocal tract releases the airstream suddenly when due to any blockage of airstream totally the stops are produced. On closer of lips and the nasal cavity which produces stops we can achieve the complete closer of air passage through the vocal tract. In Sindhi language one of the unique prosperities is linked with the stops. For example if a stop consonants is followed in an utterance by a vowel sounds it is referred to as a plosive stop consonants. In other classes of consonants sounds the phonemic inventory of Sindhi comprises of the biggest set.

3.6 Articulatory Synthesis

This synthesis utilizes for computational biomechanical models of speech production. For example the models for the Glottis (the expression and periodic excitation is generated by it and the moving vocal tract. The simulated muscle actions of the articulators controlled the articulatory synthesis ideally. It is like this that the tongue, the lips and the Glottis act.

3.7 Acoustic Phonetics

By utilizing the fundamental frequency, Amplitude and durational measurement and computer generated spectrogram, the instrumental study of the speech sounds is referred to acoustic phonetic. Although the incessant research and development Endeavour's in the past six deceits still lacking the design of a state of the art, efficient and speech recognizer instead of remarkable technological advancement in the are of speech production and recognition systems[14].

Many of the languages including sindhi can be denoted as the arrangement of sounds,phonems. There are about 50 phonems including 38 consonents, 3 semi vowels, 8 vowels are speciallicaly in sindhi language and it is appeared (in) fig 8.The vowels diphthongs,semi vowels and consonents are four categories of souonds. Each one of these classifications can be split into sub categories which are recognized with way, and place of vervalization of the sound inside the vocal taract.

Fig. 8 Sindhi language Phonemes

3.8 Classification of constant phonemes

The sindhi consonents framework which consists of 12 stop are plosives have been classified in phonetic images by international phonetic association (4 implosive stops counting),5 nasals, 2 affricates, 8 suctions, 6 fricatives, 2 retroflex, 1 parallel, and 2 semivowels[16].

3.9 Formant Synthesis

It is assumed that the (glottal) source is totally free from the filter with the use of set of rules to control highly simplified source filter model. The formant frequencies and bandwidth which are control parameters determine the filter. A specific resonance of the vocal tract link to each formant. Stylized glottal or other pulses for periodic noise and sounds are generated by the source. Incomplete natural sounding speech but highly intelligible can be generated by

formant synthesis. Any how it posses the advantage of only moderate computational requirements and a low memory foot print.

3.10 value of Dots in Sindhi Script

To broaden the first Arabic content the dots are very important in Sindhi content. Arrangement and introduction of spots can make an other character, the changing of quantity of dabs is noteable in Sindhi. There are utilization of two accents in Sindhi content one is spot and other is littel"ط" as in "ٹ". To make a verb in Sindhi language, the letter"ٹ" is composed before a basic. The easy elements in Sindhi characters as specks and accents are given in figure 9. Without spots than with specks,there are a greater number of characters[8].

Number of Dots	=	Characters
With single dot	12	ب, ج, جھ, خ, ڈ, ذ, ز, ض, ظ, غ, ف, ن
With Two dots	11	ت, ٹ, چ, ڇ, ڈ, ڊ, ق, گ, گ, ي
With Three dots	06	پ, ٹ, ٹ, ڇ, ڈ, ش
With Four dots	05	پ, ٹ, ڇ, ڑ, ڤ
With small (ط)	01	ٹ
Without dot	17	س, ص, ط, ع, ک, ک, گ, گھ, ل, م, و, ه, ء, ا, ح, د, ر,
Total number characters	**52**	

Fig. 9 Dot Representation in Sindhi Alphabet

3.11 In Sindhi Discretixation Problem

The sentiments of words are called discretization. There are two classes of images in sindhi content and they are letters and diacritics. Although the content is composed without diacritics applications, the letters are constantly composed. The lexical and morphological equivocalness is made of non attendance whereas diacritics are of a big value for comprehensiblity and comprehention. The most basic issues confront computational preparing are the non appearence of diacritics in Sindhi content. Since the missing images dont speak to just the vowels, term discretization is used in this exploration and it is additionaly spoken to some different images. For the most part along these lines the diacritics are not composed as words are vogue to comprehened the right significant of the word.

Sindhi Word	Transliteration	POS	Meaning
ڪُنُ	Kanu	Noun	Ear
ڪَنَ	Kana	Noun	Ears
ڪِنُ	Kinu	Noun	Dirt
ڪِنَ	Kina	Noun	Dirt's
ڪُنُ	Kunu	Noun	Whirlpool
ڪُنَ	Kuna	Noun	Whirlpools
ڪِن	Kini	Pronoun	Many
ڪَن	Kani	Pronoun	
ڪُن	Kuni	Noun	Cesspool

Fig. 10 Orthographic Presentation of ڪن

For such sort of words labeling is troublesome(Mahar, and Memon 2010). For example a word which is comprised of two scripts like(ڪن,(i,e 'k'and 'n') has uncertainty since it is composed without diacritics. There are in Sindhi lexcons nine kinds of homonymy words which are made accessible. Because of the arrangement of diacriticts images, these nine words are not the same to each other grammetatically and semantically. So it is very hard and adjust syntactic classes to these words[17].

3.12 Rural and Urban Sindhi

There is much difference in Arabian and Rural spoken Sindhi. There is further division between rural spoken Sindhi mean the educated and uneducated speakers. The refined and sophisticated modern form of Sindhi is the urban Sindhi and the language of the educated people in rural areas. These people mix the words of other languages such as Urdu, English and Arabic languages without intension. The symmetry in the spoken language of the literate people can be monitored across all the dialects because they speak the taught version of the language which is standard Sindhi. The Sindhi of illiterate persons living in rural areas is non refined version of language in their natural style.

4. General Problem of Automatic Speech Recognition

As for the programmed discourse acaknowlegement framework is conserned they possess a big no of parameters which include and exert the influence on pericision of the acknowledgment framework. For instance the dependent and free speaker who is separated or linked with work acknoledgement, individual voculabary, transducers, prplexity,dialect, demonstarting and voice displaying and such issues of programmed discourse acknowledgement framwork which include given info discourse elocution by one individual for some time clash amongst preparing and testing without finsh acknowledgement[19].

4.1 Speech Comparsion ASR by Human comprehension

To get and learn more information, it is mention by the speaker recognition that the speaker achieves the chance to learn. This process is done by hearing a sound not because of perusing of papers, books, magzines which abstain the speaker from composing a few paper on a solitary theme. It is the tendency of each person that they need to do work with voice or hand free frame work. For the development of a measurable model are syntactic structure it was easier to take more and get more learning in to conquer this issue and get result. To demonstrate the information of the speaker we do confront the issues and in addition to world learning we can not build a model of world learning. So an inquary is arrising that how we might defeat from this issue and contify the husman appreciation in the automatic recognition system[19].

4.2 Noise Problem

To make an application of framework with high accuracy their is programmed this course recognition (ASR) and behind it principle motivation. To expel the all foundation commotion of talked word which do not satified this condition, the common after effect of this course recogiziers or excssively successful. Discourse is vocalized in sound conditions. We can give the example of ticking clock which plays some tunes in a particular place in an other room. The communication of people outside, the vechiles horns and all these distrubing kinds of sounds are generally called clamor.
We should indentify all voices and sounds which are undesrible and irritate clamor from the info flag, we have to increase the pecision of discourse acknowledgement in ASR. In the mouth piece a couple of mili seconds the elective sorts of unliking sounds which are called reverbrate impact, they appear in the mouth keys at the season of talking. For increasing acknoledgment discourse strength, there are a big no of common straiges which are accessiable. For example we can apply measurable techinques for achieved discourse units in loud condition just viva preparing in different uproarious conditions.

4.3 Communication Style

It is obvious and natural that every body as its own alternate style of talking. By substitute method each human being shows his identity. All the human being do utilize their individual way to deal with accentuation and articulate words without doubt and style. For this purpose they don't use their own vocabulary. It is a common practice that while talking to each other in different occusion we use different styles and ways for example in banks, chatting with colleagues, talking to folks and speaking with highly educated persons. It means while talking to different class of people, we change the style of talking in different circumantance. It is obvious that by means of discourse every person has his own alternate method to express his feelings. We use various styles in different atmosphere and environment in expressing our sentiments such as when we are disappoined,foucsed,baffled, depleted, chairful,pitiful and energized. For example when some one (he,or she) is not happy and glad he will talk in moderate and some what in high voice and on the other hand when he is angry he will use his talk in shotting way[20].

4.4 Body Language

Body language is generally called and perceived by the movement of hands, eyes, experssion of feelings body motion had movement so end and so fouth. It means only the voice is not the mean of communication but the body language also communicate the programmed discourse acknowledgement framework.

4.5 Variability of Channel

As indicated by requisitions and indentification the all speakers have their own tone of voice because every single individual has is own characteristics of indentity. Each speaker is differenciated with other not only by the single factor of voice but they are separated by different attributes and likewise matters.

4.6 Continue Speech

Dispite the word limit between the discourse there is no normal respite. It is commonly seen that the properties and characteristics stop on a manufactured level like a short time and later on an articulation or a sentence. The potential issue is the transition of words from singnals and there is a lone strategy to confront this test. This task is done by applying a settled hole or a delay in the middle of discourses[20].

5. Speaker Dependent System

This kind of system is dependent on the training session before the spoken word is identified. In this system the example of every speaker is perceived and the software program of PC must be ready. When a user some limit of words spoken the computer learns the basic feature of voice. When a user speaks dissimilar word it is stored in database and that context is included in the chart of speaker voice patterns. For this purpose the spoken words can be recognized by the computer machine which are related to the different and same speakers and then it is matched with the database. On the basis of training of large no of

representative sentences this kind of system is developed [31].

5.1 Speaker Independent System

When the training of machine is mandatory this system takes place. With the use of minimum no of task oriented system and spoken word lexicon these devices are utilized. With the use of microphone the speaker speaks words in front of computer and they are then identifying as spoken words. In this system the accuracy rate is less because of the hardness of this type of system as compared to the speaker independent system. To control flexibility these types of system in home based application are used for some times. For the purpose of providing the free platform to user without any training this system is developed. When a person can interact with style or accent this technology permits him to use continuous voice and progressively verbal patterns. This system does not provide any necessity of prerequisite training [34].

6. Training and Testing

6.1 Training Phase

For the acquisition of exit an accurate results we do provide and furnish various repeated words. With the use of Microphone we have provided five Sindhi words in the preparation and starting phase. For identification and testification of pronunciation and accent we have repeated them for so many times. Meanwhile the model has adopted three words out of five and two words have been left behind. Similarly by some models or common process each and every talked outlines alongwith configurations' are achieved which is reflected by true example characteristics [32].

6.2 Testing Phase

The testing phase is started after passing through different phases when we make and prepared any model. The perfectness, reliability accuracy and operativeness which is tested and investigated by us. The errors and bugs which occur as mistakes during the execution phase of programming. It is usual that whenever we hand over any software to the consumer then some defects expose as mistakes but they are not permanently removed out. If we wish that the software may continue life time then as the deployment phase to the consumer begins and the testing is complete and similarly the phase of maintenance starts.

7. User Interface or Prototype

The platform of the any programming are called user Interface and it is also intuitive as nature of project demand. Over there the interface of customer graphic is used. Utilizing the Notepad++ and C++ are used for initially as the interface design of user. In this case some changes were also made. Additionally the clock functionality has been added by us in the program to every screen so that no matter the user is viewing screen.

Fig. 10 User Interface or Prototype

8. Conclusion

Speech recognition process is the process of listening and understanding human voices. Computer or smartphone understand the signals or voices of human being and they are used for various task from writting text on screen to perform different activities. For many languages accept Sindhi there are available speech recognition system but there is still need for Sindhi an automatic speech recognition system. To establish a speech recognition system there have been research works and the details of issues and challenges for the development of a workable speech recognition system.

Here in this research paper the for Sindhi speech recognition, general complexities and complicacies have been presented. This current research study will provide and give helpful for the guidence of empirical approach of Sindhi speech recognition system.

References

[1] Rasipuram, R., & Doss, M. M. (2016), 'Articulatory feature based continuous speech recognition using probabilistic lexical modeling' , Computer speech & language 36,233 – 259.

[2] Li, X.; Yang, Y.; Pang, Z. & Wu, X. (2015), 'A comparative study on selecting acoustic modeling units in deep neural networks based large vocabulary Chinese speech recognition ',Neuro computing 170, 251 - 256.

[3] Anusuya, M. A. (2009). Speech Recognition by Machine: A Review. International Journal of Computer Science and Information Security (IJCSIS) Vol. 6 (3), 181-205.
[4] Jennifer, S. C. (2006). The Sindhi language. In Brown (ed.) Encyclopedia of Language and. Linguistics 2ed., Vol. 11, 384-386). Oxford: Elsevier
[5] Satori, H., Harti, M., &Chenfour, N. (2007).Introduction to Arabic speech recognition using CMUSphinx system.
[6] Muhammad Asif,;DrNajmi G.(2004),'Acoustic Analysis of Phonetics of Arabic Script Sindhi Language to evaluate Voerl-Consonant Segmentation'.
[7] Acoustics, Speech, and Signal Processing, ICASSP-95., 1995 International Conference Vol. 1, 153-156. IEEE
[8] Hakro, D. N. I. A. Ismaili, A. Z. Talib. Z. Bhatti. AndG. N. Mojai. (2014), Issues and Challenges in Sindhi OCR, Sindh University Research Journal (Science Series) 46(2), 143-152.
[9] Ahmed H Makhdoom, International Management Centre (MIMC), Singapore,
[10] Kent, R. D., and R.Charles, (2002). The acoustic analysis of speech (2 ed.). Singular Publishing Group.
[11] Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In Acoustics, speechand signal processing (icassp), 2013 international conference on (pp. 6645-6649). IEEE.
[12] Raza, S., F. Z. Agha, and R.Usman, (2004).Phonemic inventory of Sindhi and acoustic analysis of voiced implosives.Center for Research in Urdu language Processing (CRULP).[13] Olive, J. P., G. Alice, and C. John. (1993). Acoustics of American English Speech: a dynamic approach. New York: Springer-Verlag.
[13] Furui, S. (2005). 50 years of Progress in speech and Speaker Recognition Research. ECTI Transactions on Computerand Information Technology, Vol1 (2)64-74.
[14] Cole, J., (2005). "Sindhi", In Strazny, Philipp(ed) Encyclopedia of Linguistic. New Yark: Routledge
[15] Handbook of the International Phonetic Association, (2002) A Guide to the Use of the International Phonetic Alphabet, Cambridge University Press.
[16] Mahar, J. A., and G. Q. Memon, (2010). Sindhi part of speech tagging system usingwordnet. International Journal of Computer Theory and Eng. 2(4), 538Pp.
[17] Dow, H. (1976). A note on the Sindhi alphabet. Asian Affairs, 7(1), 54-56.
[18] Gong, Y(1995) Speech recognition environments: A survey.Speech communication, 16(3), 261-291.
[19] Furui, S. (1997). Recent advances in robust speech recognition. In ESCA-NATO Workshop on Robust speech recognition for unknown communication channels 11-20.
[20] Forsberg, M. (2003). Why is speech recognition difficult.The Chalmers University of Technology.
[21] Jennifer, S. C. (2006). The Sindhi language.In K. Brown (ed.) Encyclopedia of Language and.
[22] Khawaja, M. A., and G. H. Najmi, (2004). Acoustic Analysis of Phonetics of Arabic Script Sindhi Language to evaluate Vowel-Consonant Segmentation. Journal of Independent Studies and Research (JISR), 2(2), 15-26.
[23] About Indus (2014). [Online] Available: http://en.wikipedia.org/wiki/Indus_civilization
[24] Kaur, R., P. Bhatia, (2010). Design and Implementation of SUDHAAR-Punjabi Spell Checker. International J. of Information and Telecommunication Technology (IJITT), 1(1) (ISSN: 0976-5972), North American
[25] Wikipedia, The Free Encyclopedia, Phonetics, http://wikipedia.org
[26] Kashif K., (2004) Audio Visual Officer, Sindhi Language Authority (SLA), Hyderabad, Sindh, Pakistan,
[27] Ziegenhain, U., Hoge, H., Arranz, V., Bisani, M., Bonafonte, A., Castell, N.,&Rabie, A. (2003).Specification of corpora and word lists in 12 languages. LC-STAR Deliverable D, 1.
[28] Javed Ahmed Mahar& Ghulam Qadir Memon.(2009)Phonology for Sindhi Letter-to-Sound Conversion,Vol. 3,
[29] Chourasia, Vishal, K. Samudravijaya, and Manohar Chandwani (2005). "Phonetically rich hindi sentence corpus for creation of speech database." Proc. O-Cocosda : 132-137.
[30] Punitha, P., and G. Hemakumar (2014). Speaker dependent continuous Kannada speech recognition using HMM. In Intelligent Computing Applications (ICICA), 2014 International Conference on, pp. 402-405. IEEE.
[31] Park, J., Diehl, F., Gales, M. J. F., Tomalin, M., & Woodland, P. C. (2009, April). Training and adapting MLP features for Arabic speech recognition. In Acoustics, Speech and Signal Processing, 2009.ICASSP 2009. International Conference on (pp. 4461-4464). IEEE.
[32] Ali, A., Zhang, Y., Cardinal, P., Dahak, N., Vogel, S., & Glass, J. (2014). A Complete KALDI Recipe for building Arabic Speech Recognition Systems, 525–529.
[33] Mohammad A.M,;Raja.N.; Roziati. Moustafa,; Othman O.Khalifa. (2010), Natural Speaker-Independent Arabic Speech Recognition System Based on Hidden Markov Models Using Sphinx Tools
[34] Këpuska, V., & Bohouta, G. (2017). Comparing speech recognition systems (Microsoft API, Google API and CMU Sphinx). Int. J. Eng. Res. Appl, 7(03), 20-24.