

# Punjabi Speech Generation System based on Phonemes

Simmi Luthra  
Department of Computer Science  
Guru Nanak Dev Engineering College  
Ludhiana (Punjab)-INDIA

Parminder Singh  
Associate Professor  
Department of Computer Science  
Guru Nanak Dev Engineering College  
Ludhiana (Punjab)-INDIA

## ABSTRACT

This paper discusses the development of a concatenated speech synthesis system for the Punjabi text written in Gurmukhi script based on Phoneme as basic unit. The research work was carried out with the aim that the developed system must be able to produce synthetic speech, corresponding to two qualities of Speech Synthesis: intelligibility and naturalness. In order to get the naturalness in the synthetic speech, Concatenative speech synthesis techniques have been used, with Punjabi Phonemes as a basic unit. In this work, phonemes have been selected as the speech unit for the development of the Punjabi speech database. After analyzing of Phonemes, a carefully selection of unbiased Punjabi Corpus was made, having nearly, 2,33,009 unique and more than four million total words. The two types of Phonemes (V and CV) gave rise to 380 phonemes with non nasalized vowels and 380 phonemes with nasalized vowels, resulting total 722 phonemes. The system is based on a Punjabi speech database that contains the starting and ending positions of phoneme-sounds labeled carefully in a wave file of recorded words. The input text is first processed then these phonemes are searched in the database for corresponding phoneme-sound positions in recorded wave file.

**Keywords:** Speech database, Punjabi Phoneme, Phoneme sounds.

## 1. INTRODUCTION

In recent years, the use of computers in speech synthesis and speech recognition has become an important area of study among speech and computer scientists. The primary motivations are to provide users with a friendly vocal interface with the computer and to allow people with certain handicaps (such as blindness) to use the computer.

Speech is the primary means of communication between people. Punjabi language is being spoken by about 104 million peoples in India, Pakistan and other countries with Punjabi migrants. The language is being written in Gurmukhi script in Indian Punjab, whereas in Shahmukhi script in Pakistani Punjab. In this paper we describe the approach used to develop a system for Punjabi text written in Gurmukhi script. Because Punjabi is a syllabic language, so output waveform is generated by concatenating the phoneme sounds marked in recorded sound file. Phoneme sounds in different contexts have been marked and stored in the speech database to get natural sounding synthesized speech.

## 2. PUNJABI LANGUAGE

Punjabi is a language spoken by inhabitants of the historical Punjab region (north western India and eastern Pakistan). In Pakistan, Punjabi is the most widely spoken language. Punjabi can be subdivided into two varieties, known as Eastern Punjabi and Western Punjabi. Punjabi is called also Gurmukhi (in India) or Shahmukhi (in Pakistan). Gurmukhi means "from the mouth of the Guru". It was devised during the 16th century by Guru Nanak, the first Sikh guru, and popularized by Guru Angad, the second Sikh guru. It is written from left to right using the Gurumukhi (an abugida derived from the *Landā* script and ultimately descended from Brahmi script) as well as Shahmukhi (a version of the Arabic script) scripts. In Gurumukhi script, which follows the "one sound-one symbol" principle, the Punjabi language has thirty eight consonants, ten non-nasal vowels and same numbers of nasal vowels.

## 3. PUNJABI PHONEMES

A Punjabi phoneme can be defined as a minimum sound unit of a language by which the meaning may be differentiated. The phonemes are the elements which stand in contrast with each other in the phonological system of that language. In other words, a phoneme in a language is defined only in terms of its difference from the other phonemes of the same language. Phonemes have the following characteristics:-

- Phoneme is a representative of a class of sounds, which speaker accepts as a single unit, regardless of positional variants.
- Phoneme is a minimum distinctive unit in the sound system of a language, by which the meaning may be differentiated.
- Phonemes of a particular language exist in contrast to the other phonemes of the same language.
- Phoneme is a functional unit and is represented within phonemic notation of two slanted lines (//)

Punjabi phonemes can be classified as: segmental phonemes and supra-segmental phonemes. Punjabi language, like other Indian languages includes segmental phonemes, but not supra-segmental phonemes in its alphabet.

## 4. CATEGORIES OF PUNJABI PHONEMES

Punjabi phonemes are classified into two categories: Segmental phonemes and non Segmental phonemes.

### 4.1 Segmental Phoneme.

Segmental phonemes are the discrete units that can be identified either physically or auditory in the stream of speech. These have their independent existence. Like other languages, in Punjabi vowels and consonants are the segmental phonemes. Segmental phonemes in Punjabi include twenty vowels and thirty eight consonants.

#### 4.1.1 Punjabi Vowels

Vowels are the sounds produced with vibrations of the vocal chords, by the unobstructed passage of air through the oral cavity. In Punjabi language, there are total 20 vowels. Out of twenty vowels ten are non-nasalized and Table1 shows non-nasalized vowels. And ten are nasalized and Table2 shows nasalized vowels

Table1. Non-Nasalized Vowels

Table2. Nasalized vowel

All these vowels are distinctive and significant having characteristic that replacing one vowel with another changes the meaning and so are contrastive.

#### 4.1.2 Punjabi consonants

Consonants are the sounds produced by an obstruction of blocking or some other restriction to the free passage of the air, exhaled from the lungs, through the oral cavity. And out of thirty eight consonants in Punjabi language five (ੳ, ਙ, ਞ, ਣ, ਮ) are nasalized and the remaining consonants are non-nasalized. Punjabi vowels can be classified based on: opening of mouth, position of tongue tip and rounding of the tongue, whereas Punjabi consonants can be classified based upon: place of co-articulation and manner of articulation.

ੲ	ਈ
ਏ	ਐ
ਅ	ਆ
ਐ	ਉ
ਊ	ਓ
ੲ	ਈ
ਏ	ਐ
ਅ	ਆ
ਐ	ਉ
ਊ	ਓ

- ✓ Place of articulation: The classification is based upon identifying the sounds produced at different places of articulation. These places are: lips, teeth, alveolar, hard palate, soft palate, uvula, larynx, glottis etc. These articulation places have been internationally accepted and have been included in International Phonetic Alphabet.
- ✓ Manner of articulation: The second generally accepted parameter for the classification of the consonants is manner of articulation, which classifies the consonants based on “how” consonants are uttered. The consonants are classified based upon the degree of obstruction to the emission of air stream. The various categories are: stops, affricates, nasals, trill or rolled, flap, semi vowels and laterals

### 4.2 Supra-segmental phonemes

Supra-segmental phonemes extend over more than one segment and so these cannot be segmented. However, their presence changes the meaning, hence these are phonemes. Supra-segmental phonemes include stress, nasality, juncture, tone and intonation. In case of mono-syllabic words, stress is given on the whole word during the utterance of the sentence, whereas in case of poly-syllabic words, stress is given on long, middle syllables of the word, during utterance of word in a sentence. The utterance of supra segmental phonemes is possible only with segmental phonemes.

#### 4.2.1 Stress

Stress is a supra-segmental phoneme, because its presence in a speech unit of the sentence, changes the meaning. The stress can cover the phone, syllables or word. The stressed unit is spoken louder than other units, and so the stressed unit is distinguished from others. In Punjabi stress is represented by / ˈ / called ‘*addak*’. Presence of ‘*addak*’ in words changes word meaning and so is phonemic. Stress is multi-dimensional supra-segmental phoneme in Punjabi.

#### 4.2.2 Nasality

Punjabi sounds have nasalized vowels and nasalized consonants. In Punjabi nasality is distinctive and its presence changes the meaning of the word. Unlike nasalized consonants, nasalized vowels are not segmental. The nasality depends upon vowels and so is supra-segmental. In Gurumukhi nasality is represented by / ̃ / (called ‘*tippi*’) and / ̄ / (called ‘*bindi*’).

#### 4.2.3 Tone

The languages, where word meaning or grammatical categories are dependent on pitch level are known as tone languages. Among all Indo- Aryan languages, only Punjabi is a tonal language in which tone has been selected as a supra-segmental phoneme.

There are three types of tones in Punjabi – High tone, Level tone and Low tone. Stress and tone are used with each other and cannot be separated. Phoneme having tone is also stressed and one phoneme has only one tone in Punjabi. Word level tone sequences produce intonation that is at sentence level.

## 5. GENERATION OF PUNJABI SPEECH SYSTEM

There are some steps that have been followed for the generation of Punjabi speech system:

### 5.1 Pre-processing

In this Pre-processing, the user can enter the text. The text can have some abbreviations, numeric values and special symbols etc .which must be processed before passing the text for the TTS conversion. There may be the abbreviations in the input text, which are first searched and then replaced with expanded form, so that written abbreviations be spoken in full word form. Also the numeric values are first analyzed and then expanded to the form required for speaking out that numeric value. The entered text is then segmented into words.

### 5.2 Segment Words into Phoneme

As Phoneme is the basic unit of concatenation, so it is necessary to segment the words of text into phonemes. The word is segmented into Phonemes, upon concatenation the word will be spoken as per Phonemic sounds.

### 5.3 Searching

These components deals to find the values of phonemes in the recorded sound file, from the database. First the Punjabi text is segmented into words, and then the words are segmented into phonemes which are stored in a linear array. Then it is passed to the database for search. Initially the search is made according to the position (Starting, end) of the phoneme in the word. . If search is successful for that particular position of Phoneme, starting and end positions of Phoneme sound in recorded sound file are returned from database. If there is no entry for that particular position of Phoneme in the database, then this will be skipped as invalid phoneme as typing error (Mistake).

### 5.4. Database Development

The following steps have been followed to develop Punjabi speech database:-

#### a).Selection of Phoneme:

For the development of Punjabi system, Phoneme was selected as the basic units of concatenation. Phoneme is a unit of speech made up of vowels and consonants. A phoneme can be defined as a minimum sound unit of a language by which the meaning may be differentiated. The reason for selecting phoneme basic speech units is that, being relatively smaller than words and syllables, these preserve within co-articulation effects. So, phoneme give lesser number of total speech units than words and syllables .Hence it provides lesser storage space.

#### b).Punjabi Corpus used:-

The role of corpus for the study of different aspects of a language is very important.: For the analysis of phonemes, a carefully selection of unbiased Punjabi Corpus was made, having nearly, 2, 33,009 unique and more than four million total words The two types of Phonemes (V and CV) gave rise to 380 phonemes with non nasalized vowels and 380 phonemes with nasalized vowels, resulting total 722 phonemes.

### c) Analysis of Punjabi phonemes:-

In order to select minimum require phonemes for the system, the total phoneme set of Punjabi language was analyzed over the carefully selected Punjabi corpus. There are total two types of phonemes in Punjabi (V and C where V and C represent vowel and consonant respectively).

The Phonemes (V and CV) gave rise to 380 phonemes with non nasalized vowels and 380 phonemes with nasalized vowels, resulting total 722 phonemes.

**Table3. Valid and Invalid Phonemes**

Phoneme Type	No. of Phonemes	Invalid Phonemes	Valid Phonemes
V	20	0	20
CV(Non Nasalized)	380	7	373
CV(Nasalized)	380	53	329
		Total	722

### d) Word Recording:-

The finally selected words were recorded by the native female speaker of Punjabi. The speech quality depends upon the quality of the recorded sound and hence, sound quality of extracted speech units from this recorded sound. So, a professional female speaker of Punjabi is selected for recording. The recording has been done in the studio with the following characteristics, set for recording:  
Sampling Rate: 44,100Hz, BitDepth: 16 bit, Channels: Mono

### e) Labeling of Phoneme sound:

The next phase was to label the Phoneme sound in the recorded sound file. Labeling or extracting the phoneme sounds in the recorded sound file is one of the most important and time consuming tasks that needs to be done very carefully, because the naturalness of the synthetic speech produced by TTS system depends upon how exactly phoneme boundaries have been marked. For this purpose we have used sound editing software - *Sonic Foundry Sound Forge 5.0B* and the phoneme sounds have been labeled manually one by one, after carefully listening and analyzing the word sounds. The phoneme boundaries have been marked by noting down the starting and end position of each phoneme in thee recorded sound file.

**f) Punjabi Speech Database:** The speech database is an important part of a Punjabi system based upon the concatenation technique. For the development of this Punjabi speech generation system database has been implemented. The database that is designed for this system includes three fields: Phoneme, Starting Sample position and end Sample position.

## **6. APPLICATION OF PUNJABI SPEECH GENERATION SYSTEM FOR COGNITIVE DISABLE PERSON**

The main objective for developing a Speech generation system for Punjabi language is to help the persons with cognitive disabilities like dyslexia, visual comprehension and other learning disabilities. The concept of cognitive disabilities is extremely broad, and not always well-defined. When someone has a cognitive disability, this means that her or his brain works differently. That person may struggle with things like learning, thinking, and reasoning, problem solving, understanding information, making decisions or remembering. In other words, a person with a cognitive disability has greater difficulty with one or more types of mental tasks than the average person. There are too many types of cognitive disabilities.

Most cognitive disabilities have some sort of basis in the biology or physiology of the individual. The connection between a person's biology and mental processes is most obvious in the case of traumatic brain injury and genetic disorders, but even the more subtle cognitive disabilities often have a basis in the structure or chemistry of the brain. A person with profound cognitive disabilities will need assistance with nearly every aspect of daily living.

This Punjabi speech generation system will be used as an add-on tool embedded with web browsers that will enable the browser to read aloud a website in Punjabi language. With more and more electronic data becoming available online, software's with this system as add-on tool will be helpful for information dissemination, as the user who cannot read Punjabi but can understand it will then be able to get the information contained in a document/webpage by listening to it. This type of assistive technology can be particularly helpful to individuals with cognitive disabilities, visually impaired persons and old people who find it difficult to read from the computer screen. There are at least two ways to classify cognitive disabilities: by functional disability or by clinical disability

### **1).Clinical Diagnosis of Cognitive Disability:**

Clinical diagnoses may be useful from a medical perspective for treatment. Clinical diagnosis of cognitive disability can include Down syndrome, Traumatic Brain Injury (TBI), Autism, or Dementia. Clinical diagnosis may also include less severe cognitive conditions such as Dyslexia (difficulty reading), Attention Deficit Disorder, Dyscalculia (difficulty with math), and other learning disabilities.

### **✓ Dyslexia**

Dyslexia is the most common form of language-based learning disability. Approximately fifteen to twenty percent of the population has some form of language-based learning disability. Dyslexia is primarily a reading disability, and there is evidence suggesting that Dyslexia is a condition that is inherited. Dyslexia is a condition that is found in both females and males from all ethnic backgrounds. The person may have trouble with different forms of language, reading, and difficulty with spelling and writing as well.

### **✓ Attention Deficit Hyperactivity Disorder (ADHD)**

ADHD is a medical condition affecting a person's ability to focus, sit still, and pay attention. They may have difficulty in focusing on tasks or subjects, or act impulsively; they may also get into trouble. ADHD begins in childhood, but may not be diagnosed until the person reaches adolescence or even adulthood. Persons with ADHD may have difficulty with finishing assignments from school or tasks from home, jumping from one activity to another. They may lose things; forget things like homework or something they were supposed to do. They may have difficulty with following instructions, or following through with tasks they have been assigned. The person may make careless mistakes, or have difficulty paying attention to details. Persons with ADHD may have trouble organizing activities, or tasks, and may interrupt other people. They may fidget, feel restless, or talk excessively.

### **2). Functional Diagnosis of Cognitive Disability:**

Sometimes it is more useful to avoid the medical perspective of cognitive disability and view them from a functional perspective instead. A Functional disability perspective ignores the medical and behavioral causes of cognitive disability and focuses on the abilities and challenges the person with a cognitive disability faces. Functional cognitive disabilities may involve difficulties or deficits involving problem-solving, attention, memory, math comprehension, visual comprehension, reading, linguistic, and verbal comprehension.

## **7. CONCLUSION**

By using the above approach, a good quality Punjabi Text -To-Speech System has been developed. During the development of this system, it has been found that the labeling of phoneme sounds from the recorded sound file should be done very carefully. There are some more important features that have been taken care of:

- Selection of the basic unit
- Selection of corpus
- Labeling of the of speech unit
- The quality of the speech unit database.

## **8. REFERENCES**

- Word: An Implementation of SVM & Concatenative TTS, Volume 26– No.2.
- [1] Harjit Singh Gill and Henry A. Gleason Jr. 1969, "A Reference Grammar to Punjabi", Revised Edition, Dept. of Linguistics, Punjabi University, Patiala, pp. 25-40
- [2] Parminder Singh and Gurpreet Singh Lehal,(2010), "Syllables Selection for the Development of Speech Database for Punjabi TTS System", Punjabi University, Patiala. IJCSI International Journal of Computer Science Issues, Vol.7, Issue 6, pp 164-167.
- [3] Pardeep Gera. (2006), "Text-To-Speech Synthesis for Punjabi Language", M.Tech, Thesis, Thapar Institute of Engineering and Technology, Patiala, pp .10-18.
- [4] Singh, P.and Lehal, G.S. (2005), "Text-To-Speech Synthesis System for Punjabi Language," Punjabi University, Patiala, pp.11-15.
- [5] Singh.(2005), "Development of a Punjabi Text-To-Speech Synthesis System", M.Tech.Thesis Punjabi University, Patiala, pp.1-18.
- [6] Daniel Jurafsky and James H. Martin (2003), "Speech and Language Processing - An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", 2nd Edition, Pearson Education, pp, 5-15.
- [7] Nayyara Karamat, "Phonemic Inventory of Punjabi",pp. 179-18.
- [8] Dutoit,T.(1996), "High Quality Text-To-Speech Synthesis:an Overview",Journal of Electrical and Electronics Engineering, Australia,Special issue on Speech Recognition and synthesis,Vol.17,No.1,pp 1-5.
- [9] Sami Lemmetty.(1999), "Review of Speech Synthesis Technology",M.Sc.Thesis, Helsinki University of Technology,Finland,pp.7-12.
- [10] Dinesh Kumar and Neeta Rana(2011), "Speech Synthesis System for Online and written Punjabi
- [11] Dinesh kumar and Neeta Rana. (2011),"Speech Synthesis System for online Handwritten Punjabi Word:An Implementation of SVM&Concatenative TTS,pp 13-17
- [12] S.S.Joshi 1989,"Phonology of the Punjabi Verb",Classical Publishing Company,New Delhi,pp.49-88.
- [13] N.I.Tolstaya1981,"The Punjabi Language-A descriptive Grammer",Boston,USA,pp10-12
- [14] Hunt,A.J. AND Black,A.W.(1996),"Unit Selection in a Conacatenative Speech Synthesis System using Large Speech Database", ATR Interpreting Telecommunications Research Labs,IEEE,pp 32-34
- [15] Carlson,R.(1993),"Models Of Speech Synthesis", presented at Colloquium on Human Machine Communication by Voice,Irvine,California,pp7-8
- [16] Dulai,N.K., "A pedagogical grammar of Punjabi", Indian Institute of Language Studies, Patiala.
- [17] Macchi,M.(1998), "Issues in Text-to-Speech Synthesis", Bellcore Intelligence and Systems, Proceedings, IEEE International Joint on Symposia, pp6-8
- [18] Sandhu,B.S.(1986), "The Articulatory and Acoustic Structure of the Punjabi consonants", 1<sup>st</sup> Edition, Publication Bureau,Punjabi University, Patiala, pp. 13-15.
- [19] Shih,C.and Kochanski,G.(2001), "Prosody Control for Speaking and Singing Styles", Bell Labortories, Lucent Technologies, pp.3-6
- [20] Wolter,M.(1997), "A diphone based text to speech system for Scottish gaelic", Thesis for Degree of Diploma in Infomatik,University of Bonn,pp. 3-5