JADAVPUR UNIVERSITY

AKSHAR - Automatic Syllabification of Bengali Speech

by

TANUSREE DAS

A Project/Thesis submitted in partial fulfillment for the
degree of Master of Computer Application

in the
Department of Computer Science & Engineering
FACULTY OF ENGINEERING AND TECHNOLOGY

2019

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

To Whom It May Concern

I hereby recommend that the thesis entitled "**AKSHAR - Automatic syllabification of Bengali speech**" has been carried out by **TANUSREE DAS** (Reg. No.: 137322 of 2016-2017 , Exam Roll: MCA196013), under my guidance and supervision and be accepted in partial fulfillment of the requirement for the degree of Master of Computer Application in Department of Computer Science and Engineering, Jadavpur University.

.....

Dr. Sudip Kumar Naskar
Project Supervisor
Department of Computer Science and Engineering
Jadavpur University, Kolkata-700032

.....

Prof. Mahantapas Kundu
Head of Department
Department of Computer Science and Engineering
Jadavpur University, Kolkata-700032

.....

Prof. Chiranjib Bhattacharjee
Dean
Faculty of Engineering and Technology
Jadavpur University, Kolkata-700032

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

Certificate of Approval

This is to certify that the thesis entitled "**AKSHAR - Automatic syllabification of Bengali speech**" is a bonafide record of work carried out by **TANUSREE DAS** in fulfillment of the requirements for the award of the degree of Master of Computer Application in the Department of Computer Science and Engineering, Jadavpur University. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose for which it has been submitted.

.....
Signature of Examiner 1
Date:

.....
Signature of Examiner 2
Date:

Declaration of Originality & Compliance of Academic Ethics

I hereby declare that this thesis contains literature survey and original research work by the undersigned candidate, as a part of his Master of Computer Application studies.

All information in this document has been obtained and present in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: TANUSREE DAS

Registration No: 137322 of 2016-2017

Exam Roll No.: MCA196013

Thesis Title: AKSHAR - Automatic Syllabification of Bengali Speech

.....
Signature with Date

Acknowledgements

The writing of the thesis as well as the related work has been a long journey with input from many individuals, right from the first day till the development of the final project.

With my most sincere and gratitude, I would like to thank **Dr. Sudip Kumar Naskar**, *Assistant Professor, Department of Computer Science & Engineering,* my supervisor, for his overwhelming support throughout the duration of the project. His motivation always gave me the required inputs and momentum to continue with my work, without which the project work would not have taken its current shape. His valuable suggestion and numerous discussions have always inspired new ways of thinking. I feel deeply honored that I got this opportunity to work under him.

I would also like to thanks **Prof. Mahidas Bhattacharya**, *Former Director School of Languages & Linguistics*, for providing his linguistics knowledge.

I would like to thank all the faculty members of the Department of Computer Science and Engineering of Jadavpur University for their continuous support.

I would also like to thank PhD scholars at JU NLP lab, Jadavpur University for their valuable suggestions and discussions to accomplish the work.

Last, but not the least, I would like to thank my batch mates of Master of Computer Application, Jadavpur University for staying by side when I need them.

.....

Name: TANUSREE DAS

Examination Roll No.: MCA196013

University Registration No: 137322 of 2016-2017

ABSTRACT

Syllabification is an essential component of speech and language processing systems. The development of automatic speech recognizers frequently requires working with sub-word units such as syllables so automatic syllabification is important to develop automatic speech recognition system with more proficiency and robustness. This project is on an automatic syllabification model for Bangla language . The syllabification model is a data driven model along with some rules extracted from Bangla syllable analysis. We focus on crucial linguistic concepts regarding syllable structure, consonant clusters in the word medial environment, and the acoustic features.

During feature extraction, Mel Frequency Cepstral Coefficient (MFCC) is used and Convolutions Neural Network (CNN) is used for training. We focus on the correspondence between intensity cues and syllable boundary. Acoustic properties are analyzed for syllable boundary detection using the Praat software tool. The system is designed with two transduces, syllabification algorithm and mapping relation . Firstly the input is translated to Bengali font then the rules are implemented in linear order from left to right, if pattern matches one of the rules apply. We performed evaluation of Speech Recognition on the Shruti Speech Corpus to validate our proposed approach and the system produces a BLEU score of 87.37 and F1 Score of 0.9043 (90.43%) and 0.74 (74%) for IPA syllable and Bengali Syllable, respectively.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The aim of this project is to present an automatic syllabification model (Iacoponi and Savy (2011)) for Bengali (henceforth Bangla) language. Pronouncing an unfamiliar word is a task that is often accomplished by breaking the word down into smaller components. Even small children learning to read are taught to pronounce a word by "sounding out" its parts.The syllable consists of both the segmental and suprasegmental or prosodic properties (Bigi and Klessa (2015)). In order to describe our system, we will briefly sketch some important linguistic issues regarding different syllable structures and syllabification rules of Bangla. We will also focus on the acoustic property (intensity) analysis of audio datasets. The syllabification rules and acoustic analysis is implemented in the proposed system. In speech technology automatic syllabification is important in terms of automatic speech recognition (henceforth ASR) and speech synthesis (Asliyan (2001)).

1.1 Syllable and Syllabification: A Theoretical Perspective

Basically there are number of controversies in terms of defining the syllable. The syllable is defined as a unit of cognitive organization, comprised of a 'sonority peak'. Sonority is traditionally defined as intrinsic loudness (Chomsky and Halle 1968). But there is no controversy in the vowels (and vowel like units) are syllabic (Strazny (2004)). In theoretical linguistics syllable is an essential phonological unit especially in terms of prosody. The phonological structure of syllable is similar across different languages, but the phonological and phonotactic rules of syllabification is language specific (van der Hulst and Ritter (1999)).Therefore, it is required to incorporate such language specific constraints with the automatic syllabification approach to detection of syllable boundary(Bigi and Klessa (2015)).

Now the question is why syllable? A syllable is a subdivision of a word, typically consisting of a vowel, called the nucleus, and the consonants preceding and following

the vowel, called the onset and the coda, respectively. The fact is, as the syllable is a minimum prosodic unit of natural speech(Itô (1986)), if we can able to build a system with the naturalness, the human-machine interface can overcome the problem of unnaturalness. Though, there are number of automatic syllabification model for different languages, such type of tool is not available for Bangla language.(Iacoponi and Savy (2011))

1.2 Syllable Structure in Bangla

Bangla is the national language of Bangladesh and official language of the Indian state of West Bengal. A good number of research work on Bangla syllable has already been done by some great scholars like S.K Chatterjee(Chatterji (2002)), A.M Hai(Ha (1975)). In fact , there is a recent research work on Bangla syllable in Optimality Theory (OT)(Kar (2010)). In this section we will focus on the syllable structure of Bangla. Bangla follows the Sonority Sequencing Principle (henceforth SSP). The SSP requires the sonority in a syllable to be rising towards nucleus and falling afterwards.

But there some cases where this principle is violated(Hayes and Lahiri (1991)). In Bangla, the maximum three consonants can occur in the onset position. In coda position the maximum number of syllable is two(Majumder (2014)). According to the generative phonological convention, the structural notation of syllable in Bangla is:

$C_0^3 V C_0^2$

Words	**Syllable Structures**
spriha	CCCVCV
briʃti	CCVCCV
kal	CVC
O	V
am	VC
dea	CVV
saŋskriti	CVCCCCVCV

Table 1.1: Example of Different Syllable Structures in Bangla

Few examples of Bangla syllable are shown in Table 1.1.

1.3 Ambiguity of Intervocalic Consonants

The consonant in intervocalic position is ambiguous. Ambiguous in respect of the status they can have in that particular position. The intervocalic consonant can be coda of preceding syllable or can be onset of the next one. For example in the word /d͡zari/, the intervocalic consonant is /r/. The possible syllabification can be: /d͡zar.i/ or /d͡za.ri/.

So, the /r/ can be the coda of the first syllable i.e in /d͡zar/ . Again the same /r/ can be the onset of the next syllable i.e in /ri/. But acoustically, detection of syllable boundary can be resolved for this type of cases. But the problem becomes more complicated where the intervocalic consonant is nasal(like m, n) and liquid like l. If there is a nasal sound or l in the intervocalic position, the impact of nasality and liquidity starts from the coda of the preceding syllable and it gets prominent in the onset of the next syllable. For example: **/grame/**, or **/holo/**. Here, the falling of intensity is very low. The acoustic analysis is crucial to determine the syllable boundary (Datta (2017)). In this paper we focused on the intensity cues of speech. There is a clear co-relation between intensity cue and syllable, where falling intensity defines boundary of a syllable.

Chapter 2

Literature Survey

Automatic preprocessing of words is desirable because the productive nature of language ensures that no finite lexicon will contain all words.

Two approaches are mainly considered for automatic syllabification of a language. They are Rule Based approach and Data driven approach.

Rule-based approaches are language dependent. A prior knowledge of the syllable structure and phonotactics of the language is necessary to derive such rules. Further, syllabification purely based on rule-based approached is mostly inadequate due to the presence of various ambiguities. There are various instances where correct syllabification cannot be obtained by a definite rule or may even break the conventional syllabification principle.

Data driven approaches are neural network based syllabification, Decision tree-based syllabification along with self-correction algorithm.

Obligatory Onset Principle (Hooper, 1972), or Principle of Maximum Open Syllabicity (Pulgram, 1970): It is one of the simplest rule based principles, which is based on the assumption that open syllable i.e., a syllable with no coda, is significance. This is in fact very naive approach and it is impractical for Manipuri as it has various legal consonant clusters other that just CV syllable structure.(Singh et al. (2016))

Sonority principles (Selkirk, 1984): In this rule based method, syllabification depends on sonority of each phoneme i.e., its quality of being sonorous. Sonorous is the capability of giving out sound especially resonance, deep sound. This principle assigns numerical values to every phoneme of a syllable depending on its sonority level where vowels have the highest value followed by nasals, fricatives, and plosives. The main disadvantage of sonority information is the position dependency of phonemes. The same phoneme present at a different location of a syllable may have different sonorous property. There exists several instances where phoneme does not fit its typical sonority rules. (Singh et al. (2016))

Daelemans and van den Bosch (1992) present one of the earliest systems on automatic syllabification: a neural network-based implementation for Dutch. Daelemans et al. (1997) also explore the application of exemplar-based generalization (EBG), sometimes called instance-based learning. EBG generally performs a simple database look-up to syllabify a test pattern, choosing the most common syllabification. In cases where the test pattern is not found in the database, the most similar pattern is used to syllabify the test pattern.(Bartlett et al. (2009))

Zhang and Hamilton(1997) suggested Learning English Syllabification Rules system that learn rules using a symbolic pattern recognition approach. (Singh et al. (2016))

Kiraz and Mobius (1998) present a weighted finite-state-based approach to syllabification. Their language-independent method builds an automaton for each of onsets, nuclei, and codas, by counting occurrences in training data. These automatons are then composed into a transducer accepting sequences of one or more syllables. They do not report quantitative results for their method.(Bartlett et al. (2009))

The legality principle (Goslin and Frauenfelder,2001): Syllabification is based on the validation of the syllable structure considering its onsets and codas. It also considers the legality of the consonant clusters to detect its splitting points. It allows consonant clusters to be valid if they appear in some syllables. Legality principle requires a big corpus to study the legality of its constituents structures. Its main problem arises when several valid splitting instances of a consonant cluster are possible resulting on ambiguous syllabification rules. It is a rule based approach.(Singh et al. (2016))

Pearson et al. (2000) compare two rule-based systems (they do not elaborate on the rules employed) with a CART decision tree-based approach and a "global statistics" algorithm. The global statistics method is based on counts of consonant clusters in contexts such as word boundaries, short vowels, or long vowels. Each test word has syllable boundaries placed according to the most likely location given a cluster and its context. In experiments performed with their in-house dataset, their statistics-based method outperforms the decisiontree approach and the two rule-based methods.(Bartlett et al. (2009))

Muller (2001) presents a hybrid of a categorical and data-driven approach. First, she manually constructs a context-free grammar of possible syllables. This grammar is then made probabilistic using counts obtained from training data. Muller (2006) attempts to make her method language-independent. Rather than hand-crafting her context-free grammar, she automatically generates all possible onsets, nuclei, and codas, based on the phonemes existing in the language. The results are somewhat lower than in (Muller, 2001), but the approach can be more easily ported across languages.(Bartlett et al. (2009))

Bouma (2002) explores syllabification in Dutch. He begins with finite state transducers, which essentially implement a general preference for onsets. Subsequently ,he uses transformation based learning to automatically extract rules that improve his system. Bouma's best system, trained on some 250K examples, achieves 98.17% word accuracy.(Bartlett et al. (2008))

Daelemans and van den Bosch (1992) implement a back propagation network for Dutch orthography, but find it is out performed by less complex look-up table approaches.(Bartlett et al. (2009))

Chen (2003) uses an n-gram model and Viterbi decoder as a syllabifier, and then applies it as a preprocessing step in his maximum-entropy-based English L2P system. He finds that the syllabification pre-processing produces no gains over his baseline system.(Bartlett et al. (2008))

Demberg (2006) uses a fourth-order Hidden Markov Model to tackle orthographic syllabification in German. When added to her L2P system, Demberg's orthographic syllabification model effects a one percent absolute improvement in L2P word accuracy.(Bartlett et al. (2008))

Marchand and Damper(2007) investigate the impact of syllabification on the L2P problem in English. Their Syllabification by Analogy (SbA) algorithm is a data-driven, lazy learning approach. For each input word, SbA finds the most similar substrings in a lexicon of syllabified words and then applies these dictionary syllabifications to the input word. Marchandand Damper report 78.1% word accuracy on the NETtalk dataset, which is not good enough to improve their L2P system.(Bartlett et al. (2008))

Marchand etr al.(2007) conduct a more systematic study of existing syllabification approaches. They examine syllabification in both the pronunciation and orthographic domains, comparing their own SbA algorithm with several instance-based learning approaches (Daelemansetal.,1997; vandenBosch, 1997) and rule-based implementations. They find that SbA universally outperforms these other approaches by quite a wide margin.(Bartlett et al. (2008))

Maximum onset principle(MOP) (Kahn, 2015): It is very similar to legality principle. Here, if multiple legal splits are possible, it gives preference to longer onset sir relevant of the legality of the syllable coda.(Singh et al. (2016))

Chapter 3

Methodology

3.1 Mel Frequency Cepstral Coefficient (MFCC)

Mel Frequency Cepstral Coefficients (MFCCs) are a feature widely used in automatic speech and speaker recognition. They were introduced by Davis and Mermelstein in 1980. MFCC takes human perception sensitivity with respect to frequencies into consideration, and therefore are best for speech/speaker recognition.

During the process frame the signal into short frames and Each frame has to be multiplied with a hamming window in order to keep the continuity of the first and the last points in the frame. Each frame was converted to frequency domain using Fast Fourier Transform (FFT). After having the spectral feature from speech signal, frequencies are converted to MEL frequencies. Number of triangular band-pass filter is placed in mel-frequency scale within a band. MFCC features are computed with discrete cosine transform using filter output amplitude. Each frame is specified to feature vector. It consists of base MFCC features along with first and second order time derivatives.

3.1.1 Pre-emphasis

In this step isolated word sample is passed through a filter which emphasizes higher frequencies. It will increase the energy of signal at higher frequency.

$$Y(n) = X(n) - 0.95 \times X(n-1) \tag{3.1}$$

where, Y(n)=Output signal , X(n)=Input signal

It is needed because high frequency components of the speech signal have lower amplitude with respect to low frequency components.

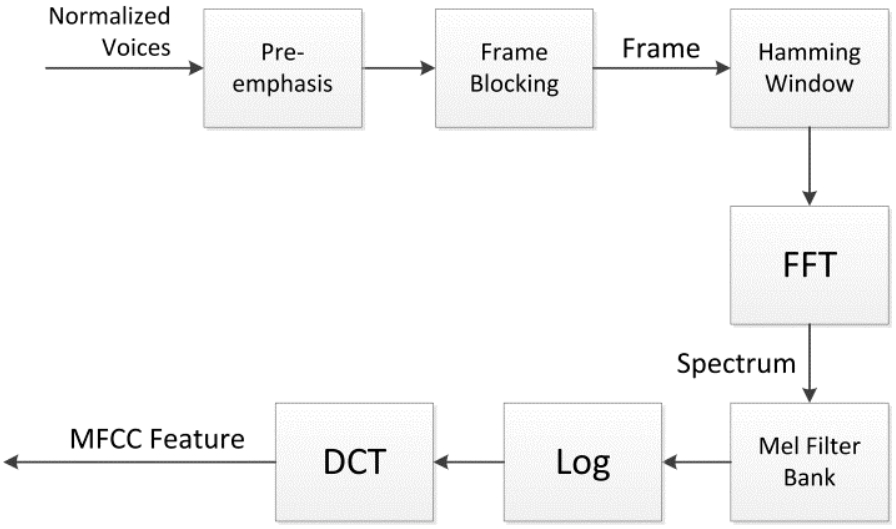

FIGURE 3.1: Block diagram of MFCC processor

3.1.2 Framing

The speech signal is segmented into small duration blocks of 20-30 ms known as frames. It is required as speech is a time varying signal but when it is examined over a sufficiently short period of time, its properties are fairly stationary.

3.1.3 Hamming Windowing

If the input signal in a frame is denoted by $X(n)$, where, $n = 0, 1, \ldots, n-1$, then the signal after Hamming windowing is

$$y(n) = X(n) * w(n) \tag{3.2}$$

where, w(n) is the Hamming window.

$$w(n) = \begin{cases} \text{w(n)} = 0.54 - 0.46\cos(\frac{2\pi n}{N-1}),\ 0 \leq \text{n} \leq \text{N-1} \\ \text{w(n)} = 0, \text{ otherwise} \end{cases} \tag{3.3}$$

where, N = number of samples in each frame

3.1.4 Fast Fourier Transform (FFT)

Then we took the Fourier transform of the signal. The Fourier transform is a mapping function that takes a series of samples in the time domain and maps them into the

frequency domain. By applying FFT the output is a spectrum or periodogram. The equation of Fourier transform is

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx \tag{3.4}$$

where, i = $\sqrt{-1}$
k = any real number
f(x) = time domain
F(k) = frequency domain

FFT reduces the computation time required to compute a discrete Fourier transform and improves the performance also.

3.1.5 Mel filter bank

Human perception of frequency is non-linear. It is Less sensitive at higher frequencies (roughly 1kHz). Mel (melody) is a unit of pitch. Mel-frequency scale is approximately linear up to the frequency of 1KHz and then becomes close to logarithmic for the higher frequencies. Equation to convert f Hz to mel frequency is

$$mel(f) = 1125 \times \ln(1 + \frac{f}{700}) \tag{3.5}$$

Human ear acts as filters called Band-pass filters that concentrate on only certain frequency components. We multiply magnitude frequency response by a set of triangular band pass filters in order to get smooth magnitude spectrum.

Thus, with the help of Filter bank with proper spacing done by Mel scaling it becomes easy to get the estimation about the energies at each spot and once this energies are estimated then the log of these energies also known as Mel spectrum can be used for calculating coefficients using DCT.

3.1.6 Discrete Cosine Transform (DCT)

It is done in order to convert the log Mel spectrum back into the spatial domain. DCT used for data compression as DCT signals have more information concentrated in a small number of coefficients and hence, it is easy and requires less storage to represent Mel spectrum in a relative small number of coefficients.

DCT formula is shown below

$$C_0 = \sum_{k-1}^{N} \cos(\frac{m \times (k-1)\pi}{N}) \times E_k \tag{3.6}$$

where, m=1,2,...,L
N = number of triangular band pass filters
L = number of mel-scale cepstral coefficients
E_k= energy obtained from the triangular band pass filters

The output after applying DCT is known as MFCC (Mel Frequency Cepstrum Coefficient).

Other usages of MFCC are Speaker identification ,Language identification ,Emotion recognition , Musical instrument recognition , Speaker conversion etc.

3.2 Convolutional Neural Network (CNN)

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analysing visual imagery. CNNs use a variation of multilayer perceptron designed to require minimal preprocessing. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. They have applications in image and video recognition, recommender systems image classification, medical image analysis and natural language processing.

3.2.1 Layers used to build ConvNets

As we described above, a simple ConvNet is a sequence of layers, and every layer of a ConvNet transforms one volume of activations to another through a differentiable function. We use three main types of layers to build ConvNet architectures: Convolutional Layer, Pooling Layer, and Fully-Connected Layer.

3.2.1.1 Convolution Layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel

- An image matrix (volume) of dimension **(h x w x d)**
- A filter **(f_h x f_w x d)**
- Outputs a volume dimension **(h - f_h + 1) x (w - f_w + 1) x 1**

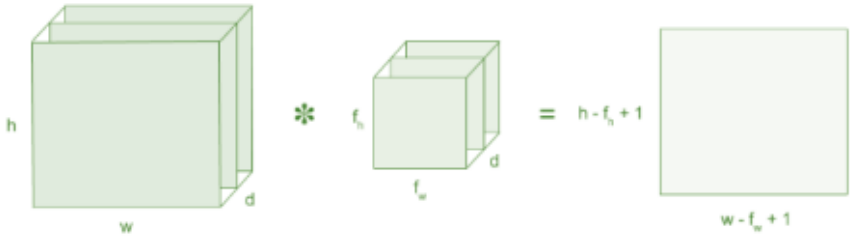

FIGURE 3.2: Convolution Layer

3.2.1.2 Non Linearity (ReLU)

ReLU stands for Rectified Linear Unit for a non-linear operation. It is a activation function. The output is f(x) = max(0,x) . ReLU's purpose is to introduce non-linearity in our ConvNet. Since, the real world data would want our ConvNet to learn would be non-negative linear values.

There are other non linear functions such as tanh or sigmoid can also be used instead of ReLU. Most of the data scientists uses ReLU since performance wise ReLU is better than other two.

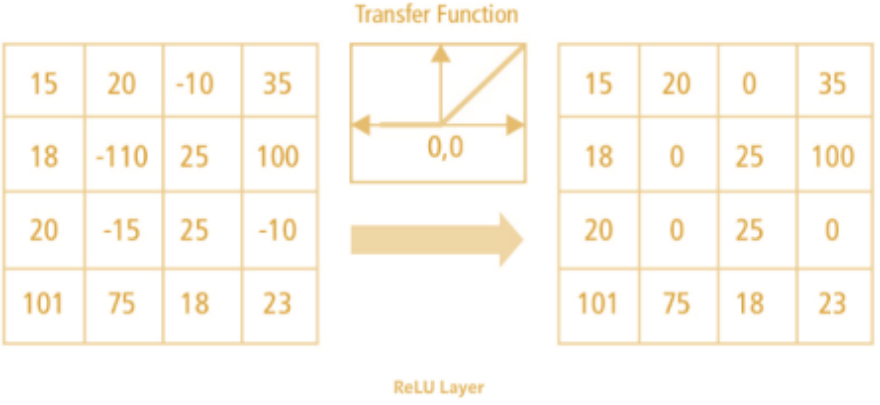

FIGURE 3.3: ReLU operation

3.2.1.3 Pooling Layer

Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or down sampling which reduces the dimensionality of each map but retains the important information. Spatial pooling can be of different types:

- Max Pooling
- Average Pooling
- Sum Pooling

Max pooling take the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling.

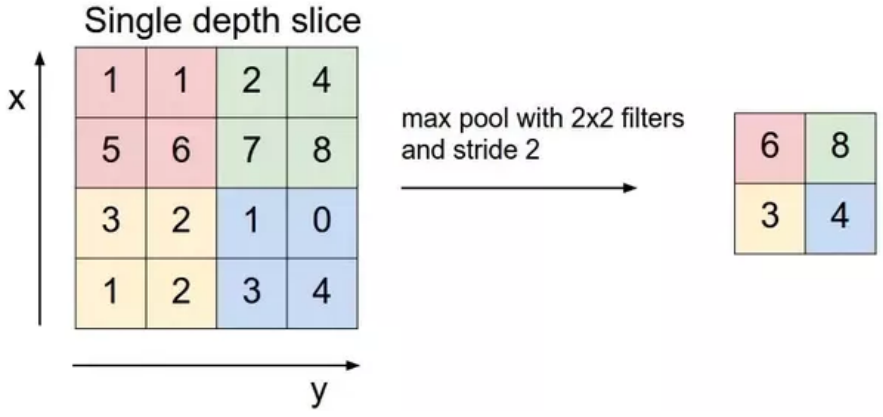

FIGURE 3.4: Max Pooling

3.2.1.4 Fully Connected Layer

The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like neural network.

In the above diagram, feature map matrix will be converted as vector (x1, x2, x3, …). With the fully connected layers, we combined these features together to create a model. Finally, we have an activation function such as softmax or sigmoid to classify the outputs.

In this way, ConvNets transform the original image layer by layer from the original pixel values to the final class scores. Note that some layers contain parameters and other don't. In particular, the CONV/FC layers perform transformations that are a function of not only the activations in the input volume, but also of the parameters (the weights

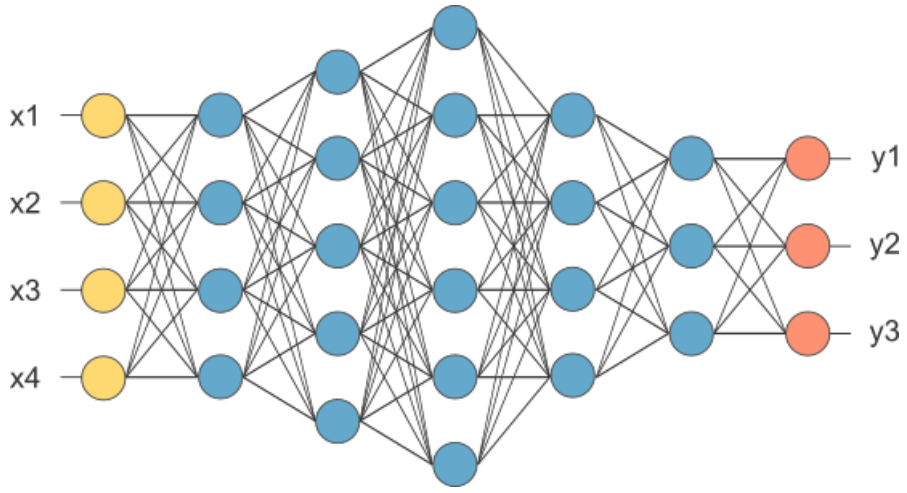

FIGURE 3.5: After pooling layer, flattened as FC layer

and biases of the neurons). On the other hand, the RELU/POOL layers will implement a fixed function. The parameters in the CONV/FC layers will be trained with gradient descent so that the class scores that the ConvNet computes are consistent with the labels in the training set for each image.

3.3 N-gram language model

Here we used N-gram language model for spelling correction. Language model is one of the most important part of Natural Language Processing. Its goal is to compute the probability of a sentence or sequence of words. It is used for spelling correction, speech recognition, machine translation etc.

N-gram is a contiguous (order matters) sequence of words. The n-grams depends on the size of the prefix. The simplest case is the Unigram and then bigram, 3-gram, 4-gram, 5-gram etc.

We can estimate n-gram probabilities by using chain rule, Markov Assumption, Maximum likelihood estimate(MLE) etc.

The Chain Rule is applied to compute joint probability of words in sentence.

$$P(w_1, w_2, \ldots, w_n) = \prod_i P(w_i|w_1, w_2, \ldots, w_{i-1}) \tag{3.7}$$

To estimate each probability a straightforward solution could be to use simple counting.

$$P(w_5|w_1, w_2, w_3, w_4) = \frac{count(w_1, w_2, w_3, w_4, w_5)}{count(w_1, w_2, w_3, w_4)} \tag{3.8}$$

but this gives us to many possible sequences to ever estimate. To cope with this issue we use Markov Assumption,

$$P(w_1, w_2, \ldots, w_n) \approx \prod_i P(w_i | w_{i-k}, \ldots, w_{i-1}) \tag{3.9}$$

For Maximum likelihood estimate(MLE) ,

$$P(w_i | w_{i-1}) = \frac{count(w_{i-1}, w_i)}{count(w_{i-1})} \tag{3.10}$$

So the bigram estimate of sentence probability would be the product of all component tandems ordered as in the sentence.

$$P(w_1, \ldots, w_n) \approx \prod_i P(w_i | w_{i-1}) \tag{3.11}$$

In practice, the outcome should be represented in log form to avoid underflow (if the sentence is long and the probabilities are really small, then such product might end in arithmetic underflow) and adding is faster.

To evaluate Language Model there are two ways, extrinsic evaluation and intrinsic evaluation. For extrinsic evaluation, we first train the language model on one set and test it on completely new dataset. Then we can compare the results of our model on the new dataset and evaluate, how good (how accurate in terms of a task e.g. translation, spell check etc) it works on this new, previously unseen dataset. This way one can compare and decide which language model fits the best to a task. But it requires multiple tests on models which are expensive. The intrinsic evaluation is about testing the Language Model itself not some particular task or application. The popular intrinsic evaluation is perplexity. It is a function of probability of the sentence. Lower perplexity implies better probability. But perplexity is a bad approximation if the test dataset does not look just like the training set. Perplexity metric is the probability of the test set.

$$PP(S) = \sqrt[n]{\frac{1}{P(w_1, w_2, \ldots, w_n)}} \tag{3.12}$$

There is a one problem with the MLE process. It is the problem of sparse data caused by the fact that maximum likelihood estimate was based on a particular set of training data. For any N-gram that occurred a sufficient number of times, we might have a good estimate of its probability. But because any corpus is limited, some perfectly acceptable English word sequences are bound to be missing from it. This missing data means that the N-gram matrix for any given training corpus is bound to have a very large number of cases of putative “zero probability N-grams” that should really have some non-zero probability. Furthermore, the MLE method also produces poor estimates when the counts are non-zero but still small. To avoid this problem there is smoothing

techniques. We simply add 1 to the numerator and the vocabulary size (V = total number of distinct words) to the denominator of our probability estimate.

$$P(w_i\ w_{i-1}) = \frac{count(w_{i-1}, w_i) + 1}{count(w_{i-1}) + V} \tag{3.13}$$

Now our probabilities will approach 0, but never actually reach 0.

3.4 Evaluation Techniques

3.4.1 BLUE Score

The Bilingual Evaluation Understudy (BLEU) Score, is a metric for evaluating a generated sentence to a reference sentence. Although developed for translation, it can be used to evaluate text generated for a suite of natural language processing tasks. A perfect match results in a score of 1.0, whereas a perfect mismatch results in a score of 0.0.The score was developed for evaluating the predictions made by automatic machine translation systems. It is not perfect, but does offer 5 compelling benefits

- It is quick and inexpensive to calculate
- It is easy to understand
- It is language independent
- It correlates highly with human evaluation
- It has been widely adopted

3.4.2 Precision & Recall

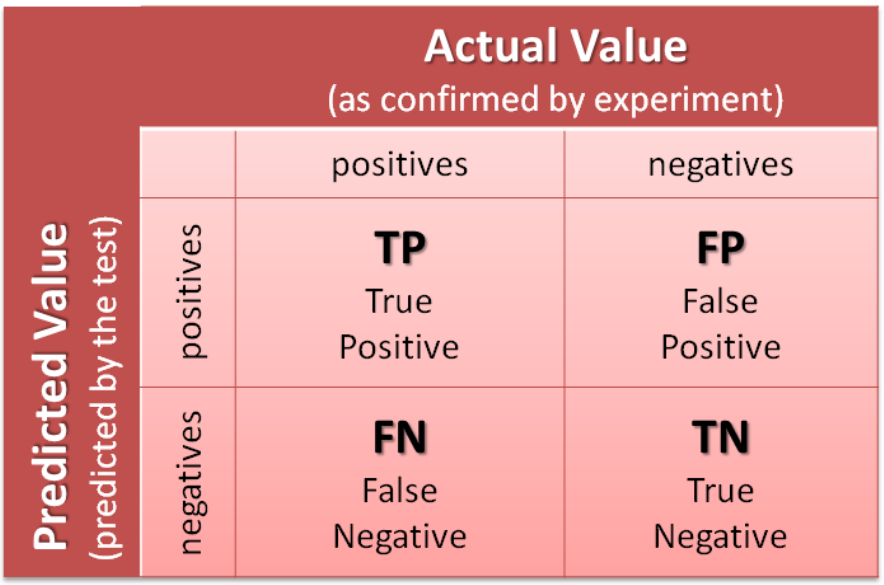

FIGURE 3.6: Precision Recall Confusion Matrix

Precision-Recall is a useful measure of success of prediction when the classes are very imbalanced. In information retrieval, precision is a measure of result relevancy, while recall is a measure of how many truly relevant results are returned.

A system with high recall but low precision returns many results, but most of its predicted labels are incorrect when compared to the training labels. A system with high precision but low recall is just the opposite, returning very few results, but most of its predicted labels are correct when compared to the training labels. An ideal system with high precision and high recall will return many results, with all results labeled correctly.

$$Precision = \frac{\text{True Positive}}{\text{Actual Results}} \tag{3.14}$$

where, Actual Results= True Positive+False Positive

$$Recall = \frac{\text{True Positive}}{\text{Predicted Results}} \tag{3.15}$$

where, Predicted Results= True Positive + False Negative

3.4.3 F1 Score

F1 is a function of weighted average of Precision and Recall. F1 Score is needed when you want to seek a balance between Precision and Recall. Therefore, this score takes both false positives and false negatives into account.

$$F1Score = \frac{2 \times P \times R}{P + R} \tag{3.16}$$

where, P=Precision , R=Recall

3.5 Linguistics Knowledge

Bengali spoken mostly in West Bengal of India and Bangladesh are from language family Indo-European Language. Bengali has 37 phonemes. Of the 37 Bengali phonemes, 7 are vowels *(without considering the nasalized vowels)* and 30 are consonants. Bengali have many common sounds with other languages like English.

TABLE 3.1: Vowels Bengali

Sl. No.	**Phoneme**	**Word** *English*	**Word** *Bengali*	**Meaning**
1	/ɑ/	k**a**n	কান	ear
2	/i/	m**i**l	মিল	similarity
3	/e/	p**e**t	পেট	belly
4	/o/	g**o**l	গোল	round
5	/u/	b**oo**k	বুক	chest
6	/ɔ/	b**o**k	বক	egret
7	/æ/	b**a**ng	ব্যাঙ্	frog

Among vowels, two are common in all respects: /e æ/. If the length of vowel is ignored, /ɑ i u ɔ/ will also be found similar. Among consonants, fifteen are common: / p b t d k ɡ s ʃ n m ŋ r l j h/. From this, one may deduce that there is considerable similarity with English languages in the territory of phonemic. But, yet there are controversy on couple of phonemes. The Bengali phonemes are illustrated in the Tables 3.1 and 3.2.

Sl. No.	**Phoneme**	Word English	Word Bengali	**Meaning**
1	/p/	**p**ul	পুল	bridge
2	/p^{h}/	**ph**ul	ফুল	flower
3	/b/	**b**or	বর	groom
4	/b^{h}/	**bh**or	ওজন	weight
5	/t/	**t**ok	টক	sour
6	/t^{h}/	**th**ok	ঠক	cheat
7	/d/	**d**al	ডাল	pulse
8	/d^{h}/	**dh**al	ঢাল	shield
9	/t̪/	**t**al	তাল	rhythm
10	/t̪^{h}/	**th**ala	থালা	plate
11	/d̪/	**d**an	দান	donation

Table 3.2 continued from previous page

Sl. No.	**Phoneme**	Word English	Word Bengali	**Meaning**
12	/ɖʰ/	**dh**an	ধান	paddy
13	/c/	**ch**al	চাল	rice
14	/cʰ/	**chh**al	ছাল	peel
15	/ɟ/	**j**al	জাল	net
16	/ɟʰ/	**jh**al	ঝাল	solder
17	/k/	**k**al	কাল	yesterday
18	/kʰ/	**kh**al	খাল	canal
19	/ɡ/	**g**oon	গুণ	quality
20	/ɡʰ/	**gh**oon	ঘুণ	termite
21	/s/	**s**aph	সাফ	clean
22	/ʃ/	**sh**ap	সাপ	snake
23	/m/	**m**on	মন	mind
24	/n/	**n**am	নাম	name
25	/ŋ/	ga**ng**	গ্যাং	river
26	/l/	**l**a**l**	লাল	red
27	/r/	**r**ong	রঙ	colour
28	/ɾ/	bha**r**a	ভাড়া	fare
29	/h/	**h**ool	হুল	sting
30	/j/	aa**y**	আয়	income

Table 3.2: Consonants Bengali

Syllabification rules has yet not been proposed till date for Bengali languages.

Chapter 4

Implementation

In this paper a system is proposed with the Syllabification algorithm. It receives Bangla speech as an input and provides its syllable as output in Bangla font. We will discuss about the architecture of the tool and the algorithm in the below sections.

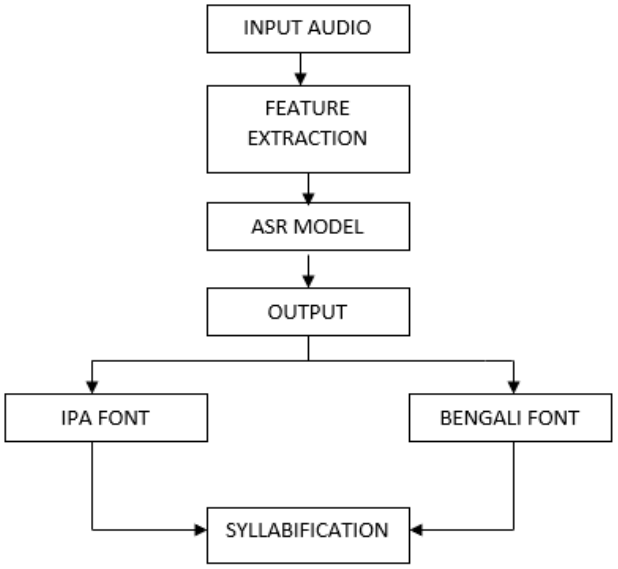

FIGURE 4.1: Block diagram of the system

4.1 The Architecture

The system is designed with two transducers (one for input and one for output)(Jurafsky and Martin (2013)), the Syllabification algorithm and the mapping system from speech to text and then text to speech. It consists of the script of the speech sound. First step is to extract the feature from each signal. Next, acoustic modeling has been done for

each phoneme with the feature list. In the last segment is the language model, which captures the linguistic information from the text (Mitkov (2003)).

4.1.1 Feature Extraction

Speech is a non stationary signal. Mel Frequency Cepstral Coefficient popularly known as MFCC is the most popular feature extraction procedure in cepstral domain. In MFCC, Frequency components are mapped with MEL scale. To propose AKSHAR we have used MFCC feature. We have discussed the MFCC feature in section 4.1.2

4.1.2 MFCC feature

The frames of 10-15 msec (approx) which is the short time period defined by the silence of the audio is extracted. During the process each frame has to be multiplied with a hamming window in order to keep the continuity of the first and the last points in the frame and then each frame was converted to frequency domain using Fast Fourier Transform (FFT). After having the spectral feature from speech signal, frequencies are converted to MEL frequencies. We multiply magnitude frequency response by a set of triangular band pass filters in order to get smooth magnitude spectrum. MFCC (Sultana and Palit (2014)) features are computed with discrete cosine transform using filter output amplitude. Each frame is specified to feature vector. It consists of base MFCC features along with first and second order time derivatives.

4.1.3 ASR Model

Convolutions Neural Network (CNN) is used for training. During the process mentioned in 4.1.2, we get a two dimensional, one-channel representation so it can be treated like an image too. Using the supervised learning each two dimensional is trained to system to learn which audio having the features. The more it is trained it provides more information.

In CNN, non linear activation function for example rectified linear unit (ReLU) or Sigmoid function are used to the result. In neural networks nodes of input layer are connected with the nodes of hidden layer and those hidden layer nodes are fully connected with nodes of output layer. In CNN, convolution are applied on input layer to generate the output. Each part of input are convoluted by different filters and combining them we get the final result. In CNN, there is a pooling layer and the purpose of these layer is sub sampling the input from a specified filter and reduce the size of input . A common pooling technique is max pooling where a maximum value is selected from each filter. Max pooling layer can also perform over a window instead of performing the whole matrix.

We carried this task by labeling the images, the system will start recognizing patterns present in audio that are absent from other ones and will start building its own cognition. We trained the system with 11,417 audio files, in 30 epochs. Loss are shown in the figure 4.2.

FIGURE 4.2: Training Loss with epochs

4.2 Syllabification Algorithm

During this phase the audio was provided in the system in which all the features were extracted using MFCC and tried to map with the nearest word having the same feature using the CNN as mentioned in Section 4.1.3 and will provide the output. The output is translated into Bangla and then it is corrected for the spelling correction using n-gram model. After the procedure using the algorithm 4.3 the Syllabification is done. The algorithm 4.3 is very simple.

Algorithm 1 SYLLABILFIER($text\_output, rules$)

```
array[] ← text_output
size = len(array)
for i = 1; i ≤ size; i++ do
    if i ≡ rules then
        put a syllable boundary
    end if
end for
return syllable output
```

FIGURE 4.3: Algorithm for Syllabification

To be precise, first the transducer recognizes the audio input and stored in text output. After translating into Bangla, it will proceed to the Syllabification algorithm and will be syllabified. Syllabification algorithm will put the syllable boundary based on the three rules. The Syllabification algorithm is given below.

In the algorithm 4.3, the output is stored into an array where each character is stored, and the size of the array is stored in variable named size. We used the algorithmic idea of Knuth Morris-Pratt (KMP) algorithm to search the pattern. If the pattern matches the rules for syllable boundary, it puts syllable boundary. The rules are mentioned in the following table 4.1.

Rule Number	**Pattern**	**Syllabification**
1	VCC	VC.C
2	VCV	V.CV
3	VV	V.V

TABLE 4.1: Rules for Syllabification

Chapter 5

Experimental Results, Comparison and Discussion

5.1 Corpus

Shruti speech corpus has been used which contains total of 7,383 unique sentences. The total number of phoneme is 49 and total number of word is 22,012. The sentences are spoken by 34 speakers of standard colloquial Bangla, in the region of West Bengal, India.

Another data set used was from Centre for Development of Advanced Computing, Kolkata (CDAC-Kolkata) who has developed a large spoken corpus of standard Bangla. The quality of ASR is noise free which is most import while training. In our proposed system only the word set was used. This contains around 40 thousands words, but not the unique word utterances.

90% of this corpus is used for training and remaining 10% is used for test.

5.2 Speech Recognizer Output

In the following table 5.1 output of the system is shown. Left column gives the Gold Data of the Data Corpus whereas Right Column is the System generated output.On evaluating the system the results shows **87.37** BLEU score. Performance check with different benchmark was unable to achieved as no one has yet used SHRUTI dataset for evaluation.

Table 5.1: Speech Recognizer Output

Reference	**Machine Output (English Phonemes)**
isalamabadera china dutabasera eka anushthane ei mantabya karechhena paka bideshamantri shaha mehamuda kureshi	isalamabadera china dutabasera eka anshthana ei mantapya karechhena paka biesamanti shaha mehamuda kureshi
tara kathaya pakistana tara dirghadinera bandhu chinake e byapare blyanka cheka diye rakhachhe	tara kathayai pakistana tara dirgadinera bandhu chinke ebe pare ganachet derachachhe
chinera bishesha duta hi iyapheiyera sange ei niye tara kathao hayechhe	cimera bishsaduta si ia heiyera sange ei niyietara katha hayechhe
tanke tini anurodha karechhena shigagirai dilli yete	take dinie anroha karechhena siatira dilijete
dilli giye tini ya alochana karabena tate purna sammati dite raji isalamabada	diligiye tini ja alosana karabena tate purna samatkite raji isa da mabad
asale mumbai jangihana niye bharata o antarjatika kshetra theke pratidina etarakama chapa asachhe ye karyata dishehara pakistana	asale mumbai jamgihana niye bharata o antarjatika kshetra theke pratidina etarakama chaa asachh ye karyata dishehara pakistana
jamata uda daoyara biruddhe byabastha nebara katha bale labha hayani	chamaha udhadaoyara biruddhea byabatanegara patha balde la bhanani
tara upara shapatha niyei baraka obama bale diyechhena pakistana arthika sahayya pabe ekatai sharte	taraopara sapaiyei barakpababa bale diyechhena pakistana arati sahayapabe ekataei sarate
jangidera nirmula karara abhiyana tadera chaliye yete habe	yangidera nierpauna karara abhiyana ptadera chadi jete habe
yadi ta na kare tahale arthika sahayya niye natuna kare bhababe amerika	yadi tana kare tahale arathik sahajaniye nutuna kare bhakabe yamerika
ei abasthaya dirghadinera bandhu chinakei bipada theke uddhare madhyastha karate chaya jaradari gilanira sarakara	ei abasthaya dirgkhadinera bandhu chinakei bipadatheke uddhare madhyasta karate chaya jaradari gilanira sarakara
gata unatrishe aktobara isalamabada giyechhilena hi iyaphei	gata untrishya aktabara isulamabana jiyeshana hi iyane
sekhana theke e bachhara pancha tarikha dilli asena tini	sikhan theke e bachhara pacsakarikadiniyashena thini

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
bharatake anurodha karena upama-hadeshe uttejana kamate pakistanera sange yena shigagirai kathabarta shuru haya	bharatake nurodha karena upama-hadeshe uttejana kamate pakistanera sange yena igagirai kathabarta shuru haye
se kathabarta ajao chalu hayani	she pathabarta aja januhayani
dupakshai simante jado karechhe bipula senabahini	dupakshasa simante jada karechhen bipanshe nabakani
pakistana chaichhe arao ekabara dilli yana iyaphei	pakistan chaichie araoayeka bara dilijana ieaphei
pakistanera sange alochanaya basara janya chapa srrshti karuna	pakistanera sange alochanaya basara janya chapa srrshti karuna
ye paddhatitei china egoka sabetei tadera sammati achhe janiyechhena kureshi	ye paddhatitei china egoka sabetei tadera sammati achhe janiyechhena kureshi
andale natuna bimananagari hale kebala bardhamana naya ashapashera besha kayekati jelarao unnati habe bale mane kara hachchhe	andale natuna bimananagari hale kebala bardhamana naya ashapashera besha kayekati jelarao unnati habe bale mane kara hachchhe
shudhu tai naya natuna bimanabandarera aparisima subidha pabe durgapura o asanasola shilpanchala	shudhu tai naya natuna bimanabandarera aparisima subidha pabe durgapura o asanasola shilpanchala
karmasansthana habe bahu manushera	karmasansthana habe bahu manushera
gata deda du bachhare rajye shilpayana prakriya jora dhakka kheyechhe	gata deda du bachhare rajye shilpayana prakriya jora dhakka kheyechhe
nandigrame gandagolera jere anishchita haye padechhe bahu prakalpa	nandigrame gandagolera jere anishchita haye padechhe bahu prakalpa
tarai madhye singura theke hata gutiyechhe tata prakalpa	tarai madhye singura theke hata gutiyechhe tata prakalpa
bardhamanera katoyaya tapabidyut prakalpa ghire shuru hayechhe andolana	bardhamanera katoyaya tapabidyut prakalpa ghire shuru hayechhe andolana
e abasthaya sarakara bishesha bhabe takiye rayechhe andala bimananagari prakalpatira dike	e abasthaya sarakara bishesha bhabe takiye rayechhe andala bimananagari prakalpatira dike

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
itimadhyei sarakara samiksha kare dekhechhe ye bhabe damadama bimanabandare yatri sankhya badachhe tate adura bhabishyate dbitiya bimanabandarera prayojana abashyambhabi	itimadhyei sarakara samiksha kare dekhechhe ye bhabe damadama bimanabandare yatri sankhya badachhe tate adura bhabishyate dbitiya bimanabandarera prayojana abashyambhabi
se karanei dakshina chabbisha paraganara baruipurera kachhe bimanabandara tairira bhabana chinta shuru kara hayechhilo	se karanei dakshina chabbisha paraganara baruipurera kachhe bimanabandara tairira bhabana chinta shuru kara hayechhilo
erai madhye bengala erotrapilasa du hajara chhaya sale egiye ase bardhamanera andale bimanabandara o take kendra kare nagarayanera kaje	erai madhye bengala erotrapilasa du hajara chhaya sale egiye ase bardhamanera andale bimanabandara o take kendra kare nagarayanera kaje
mukhyamantri buddhadeba bhattacharya atyanta ashabadi ei prakalpatike niye	mukhyamantri buddhadeba bhattacharya atyanta ashabadi ei prakalpatike niye
puro prakalpe praya dasha hajara koti taka biniyoga habe	puro prakalpe praya dasha hajara koti taka biniyoga habe
pashapashi karmasansthana habe bahu manushera	pashapashi karmasansthana habe bahu manushera
shudhu tai naya andale ei bimananagarira prabhaba padabe bardhamanasaha dakshinabangera kayekati jelaya	shudhu tai naya andale ei bimananagarira prabhaba padabe bardhamanasaha dakshinabangera kayekati jelaya
itimadhyei bengala erotrapilasa samiksha kare dekhechhe pashchimabanga simantabarti jhadakhandera bibhinna elakara manusha ei bimanabandarati byabahara karate parabena	itimadhyei bengala erotrapilasa samiksha kare dekhechhe pashchimabanga simantabarti jhadakhandera bibhinna elakara manusha ei bimanabandarati byabahara karate parabena
phale bimanabandaratira labhera sambhabana yatheshta	phale bimanabandaratira labhera sambhabana yatheshta
durgapura asanasola shilpanchaleo ei bimananagarira prabhaba padabe yatheshta	durgapura asanasola shilpanchaleo ei bimananagarira prabhaba padabe yatheshta
baraka obama markina presidentera dayitba bhara neoyara para bharata markina samparka arao majabuta habe	baraka obama markina presidentera dayitba bhara neoyara para bharata markina samparka arao majabuta habe

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
habu markina rashtrapatira kachha theke paoya chithira bayanera bhittitei jora galaya ekatha janalena pradhana-mantri manamohana sinha	habu markina rashtrapatira kachha theke paoya chithira bayanera bhittitei jora galaya ekatha janalena pradhana-mantri manamohana sinha
nirbachane jaya labhera para ushna abhinandana janiye obamake chithi likhechhilena manamohana	nirbachane jaya labhera para ushna abhinandana janiye obamake chithi likhechhilena manamohana
tara uttare obamao chithi likhechhena tanke	tara uttare obamao chithi likhechhena tanke
sei chithi pade mugdha manamohana	sei chithi pade mugdha manamohana
chithite bharata markina bhabishyat samparka ki bhabe arao nibida kara yaya sei bishaye tara drrshtibhangii chithite likhechhena baraka obama	chithite bharata markina bhabishyat samparka ki bhabe arao nibida kara yaya sei bishaye tara drrshtibhangii chithite likhechhena baraka obama
jayalabhera para manamohanake phona na kare bharatake abahela karechhena obama ei yukti udiye diye pradhanamantri janana	jayalabhera para manamohanake phona na kare bharatake abahela karechhena obama ei yukti udiye diye pradhanamantri janana
upasagariya deshagulite tara sapharera samayasuchigata asubidhara karanei du janera katha bala haye otheni	upasagariya deshagulite tara sapharera samayasuchigata asubidhara karanei du janera katha bala haye otheni
kintu obama tara sange phone katha balara ichchha prakasha karechhilena	kintu obama tara sange phone katha balara ichchha prakasha karechhilena
deshe phirei tara sange obamara katha habe baleo janana pradhanamantri	deshe phirei tara sange obamara katha habe baleo janana pradhanamantri
obamara sange itimadhyei panera ti deshera netadera sange katha hayechhe teliphone	obamara sange itimadhyei panera ti deshera netadera sange katha hayechhe teliphone
edera madhye rayechhena paka presidenta asipha ali jaradario	edera madhye rayechhena paka presidenta asipha ali jaradaro
itimadhyei rajadhanite phire esechhena manamohana	itimadhyei rajadhanite phire esechhena manamohana
ekhana obamara sange tara alapera madhyamei shuru habe bharata markina natuna samparkera suchana	ekhana obamara sange tara alapera madhyamei shuru habe bharata markina natuna samparkera suchana
amerikara prathama krrshnanga presidenta hisabe obamara jayera para besha dushchintaya rayechhe pakistana	amerikara prathama krrshnanga presidenta hisabe obamara jayera para besha dushchintaya rayechhe pakistana

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
anekei ekhana bhabate shuru karechhena pakistanera uttaranchale sena abhiyanera matra na badiye dena obama	anekei ekhana bhabate shuru karechhena pakistanera uttaranchale sena abhiyanera matra na badiye dena obama
ei abasthaya chintabhabana chalachhe obama sarakarake bojhanora ye ei anchalera jangidera mokabilaya pakistana sarakara ekai yatheshta	ei abasthaya chintabhabana chalachhe obama sarakarake bojhanora ye ei anchalera jangidera mokabilaya pakistana sarakara ekai yatheshta
kathata yadi obamake bojhano na yaya tahale du deshera samparkera abanati ghata bichitra naya	kathata yadi obamake bojhano na yaya tahale du deshera samparkera abanati ghata bichitra naya
santrasera biruddhe ladaiyera prashne demokryata ripabalikanera manobhabe bishesha pharaka nei	santrasera biruddhe ladaiyera prashne demokryata ripabalikanera manobhabe bishesha pharaka nei
du dalai chaya santrasake ekebare shikada theka upade phelate	du dalai chaya santrasake ekebare shikada theka upade phelate
pakistanera uttaranchala o aphaganistanera simante durgama girikandarera kothao ga dhaka diye achhena osama bina ladena ei bishbasa anekadinera	pakistanera uttaranchala o aphaganistanera simante durgama girikandarera kothao ga dhaka diye achhena osama bina ladena ei bishbasa anekadinera
jangidera sandhane besha kichhudina dhare ei anchale abhiyana chalachchhe	jangidera sandhane besha kichhudina dhare ei anchale abhiyana chalachchhe
pracharera samaya obama balechhilena ala kayadake kayada karate pakistana yadi byartha haya tahale tara netrrtbe odera nishana karabe amerikai	pracharera samaya obama balechhilena ala kayadake kayada karate pakistana yadi byartha haya tahale tara netrrtbe odera nishana karabe amerikai
tini ingita diyechhilena prayojane iraka theke senabahini sariye niye aphaganistane ladai karate pathabena	tini ingita diyechhilena prayojane iraka theke senabahini sariye niye aphaganistane ladai karate pathabena
presidenta haoyara age obamara eisaba kathabartai apatata chintaya rakhachhe pakistanake	presidenta haoyara age obamara eisaba kathabartai apatata chintaya rakhachhe pakistanake
tahale emana dandabe ye markina phaujera tattbabadhanei jangidera biruddhe yuddha chalate habe tadera	tahale emana dandabe ye markina phaujera tattbabadhanei jangidera biruddhe yuddha chalate habe tadera
phera baghabandi kanda	phera baghabandi kanda

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
somabara rate gosabara rajatajubali grame bana karmidera pata phande dhara padalo rayyala bengala taigara	somabara rate gosabara rajatajubali grame bana karmidera pata phande dhara padalo rayyala bengala taigara
mangalabara take niye kendo dbipera dike raona diyechhena bana daphatarera karmira	mangalabara take niye kendo dbipera dike raona diyechhena bana daphatarera karmira
sekhanei budhabara baghatike jangale chhede deoya habe	sekhanei budhabara baghatike jangale chhede deoya habe
abasheshe somabara gabhira rate dhara padalo sundarabanera gosabara rajatajubali gramera bagha	abasheshe somabara gabhira rate dhara padalo sundarabanera gosabara rajatajubali gramera bagha
brrhaspatibara theke ekati rayyala bengala taigara daphaya daphaya hana dichchhila rajatajubali grame	brrhaspatibara theke ekati rayyala bengala taigara daphaya daphaya hana dichchhila rajatajubali grame
grama theke ekati chhagalao niye yaya se	grama theke ekati chhagalao niye yaya se
shuru haya rata jege gramabasidera pahara	shuru haya rata jege gramabasidera pahara
banakarmidera pata phande somabara rate dhara pade baghati	banakarmidera pata phande somabara rate dhara pade baghati
tabe shudhu eka bagha naya	tabe shudhu eka bagha naya
sara rata khanchara baire apeksha karachhila ekati baghinio	sara rata khanchara baire apeksha karachhila ekati baghinio
mangalabara bhore aguna jbaliye banakarmira baghinitike tadiye dena	mangalabara bhore aguna jbaliye banakarmira baghinitike tadiye dena
erapara khancha bandi baghatike niye prathame yaoya haya sajanekhali taigara syanachuyari	erapara khancha bandi baghatike niye prathame yaoya haya sajanekhali taigara syanachuyari
tarapara lanche kare baghatike niye bana daphatarera karmira raona dena kendo dbipera dike	tarapara lanche kare baghatike niye bana daphatarera karmira raona dena kendo dbipera dike
bane thikamato khabara na paoyatei phera lokalaye dhuke padechhe rayyala bengala taigara	bane thikamato khabara na paoyatei phera lokalaye dhuke padechhe rayyala bengala taigara
mane karachhena bisheshajnara	mane karachhena bisheshajnara
gata phebruyari masa theke besha kayekabara sundarabana sanlagna gramaguli theke dhara padalo bagha	gata phebruyari masa theke besha kayekabara sundarabana sanlagna gramaguli theke dhara padalo bagha

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
gata saterai phebruyari sundarabanera kulatali grama kheke dhara pade rayyala bengala taigara	gata saterai phebruyari sundarabanera kulatali grama kheke dhara pade rayyala bengala taigara
bishe phebruyari rate jhadakhalite baghake khanchabandi karena bana karmira	bishe phebruyari rate jhadakhalite baghake khanchabandi karena bana karmira
erapara naya juna gosabara kumira-mari grame phera dhara pade bagha	erapara naya juna gosabara kumira-mari grame phera dhara pade bagha
adai masa bade trishe agasta gosabara yanapura grame phera banakarmidera phande bandi haya bagha	adai masa bade trishe agasta gosabara yanapura grame phera banakarmidera phande bandi haya bagha
gosabatei pakhiralayera kachhe ekati grame chabbisha aktobara bhore bagha dhara pade	gosabatei pakhiralayera kachhe ekati grame chabbisha aktobara bhore bagha dhara pade
tarapara budhabara gabhira rate arthat teshara nabhembara gosabara rajatajubali grame abara baghabandi kanda	tarapara budhabara gabhira rate arthat teshara nabhembara gosabara rajatajubali grame abara baghabandi kanda
rajatajubali grama theke bagha bandi parbera puro ghatanatira netrrtba diyechhena taigara rijarbha pharestera philda direktara subrata mukhopad-hyaya	rajatajubali grama theke bagha bandi parbera puro ghatanatira netrrtba diyechhena taigara rijarbha pharestera philda direktara subrata mukhopad-hyaya
tini janiyechhena purna bayaska ei baghati sustha	tini janiyechhena purna bayaska ei baghati sustha
budhabara bikelera madhye take niye lancha kendo dbipe paunchhe yabe	budhabara bikelera madhye take niye lancha kendo dbipe paunchhe yabe
sekhanei take jangale chhede debena banadaphatarera karmira	sekhanei take jangale chhede debena banadaphatarera karmira
piesaelabhi si tu raketera sahayye nikhuntabhabe chandrayana eka ke pathano hayechhe mahakashe	piesaelabhi si tu raketera sahayye nikhuntabhabe chandrayana eka ke pathano hayechhe mahakashe
ebara take nikhuntabhabe chandera kakshapathe sthapana karara pala	ebara take nikhuntabhabe chandera kakshapathe sthapana karara pala
itimadhye prrthibira chhabi tule pathiyechhe chandrayana	itimadhye prrthibira chhabi tule pathiyechhe chandrayana
tate bojha giyechhe ashatita bhala kaja karachhe sei mahakashayana	tate bojha giyechhe ashatita bhala kaja karachhe sei mahakashayana

Table 5.1 continued from previous page

Reference	Machine Output (English Phonemes)
sabakichhu thikathaka thakale agami atai nabhembara chandrayana pradakshina shuru karabe chandake	sabakichhu thikathaka thakale agami atai nabhembara chandrayana pradakshina shuru karabe chandake
ara panera nabhembara bijnanira take sthapana karabena chandera kakshapathe	ara panera nabhembara bijnanira take sthapana karabena chandera kakshapathe
mahakashayanake chandera kakshapathe sthapana karara kajati atyanta duruha	mahakashayanake chandera kakshapathe sthapana karara kajati atyanta duruha
era age bahu mahakashayanake kakshapathe sthapana karate giye haya ta achhade padechhe sei upagrahera buke	era age bahu mahakashayanake kakshapathe sthapana karate giye haya ta achhade padechhe sei upagrahera buke
nayato ta hariye giyechhe mahashunye	nayato ta hariye giyechhe mahashunye
gata unatrishe aktobara chandrayana prrthibira ye chhabi tule pathiyechhe ta bishleshana kare bijnanira janate perechhena mahakashayanati ekhana ki abasthaya achhe	gata unatrishe aktobara chandrayana prrthibira ye chhabi tule pathiyechhe ta bishleshana kare bijnanira janate perechhena mahakashayanati ekhana ki abasthaya achhe

5.3 Speech Recognizer Output to Bengali

The following table 5.2 shows the result of the system which was translated to the Bengali characters.

Table 5.2: Speech Recognizer Output to Bengali

Reference	Machine Output (Bengali Phonemes)
isalamabadera china dutabasera eka anushthane ei mantabya karechhena paka bideshamantri shaha mehamuda kureshi	ইসালামাবাদেরা চিনা দুটাবাসেরা একা আনসহথানা এি মানটাপয়া কারেছেনা পাকা বিসামানটি সহাহা মেহামুদা কুরেসহি
tara kathaya pakistana tara dirghadinera bandhu chinake e byapare blyanka cheka diye rakhachhe	টারা কাথায়িা পাকিসটানা টারা দিরগাদিনেরা বানধু চিনকে এবে পারে গানাচেট দেরাচাছে
chinera bishesha duta hi iyapheiyera sange ei niye tara kathao hayechhe	চিমেরা বিসহসাদুটা সি ইা হেিয়েরা সাঙে এি নিয়েটারা কাথা হায়েছে

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
tanke tini anurodha karechhena shigagirai dilli yete	টাকে দিনি আনরহা কারেছেনা সিটিরা দিলিজেটে
dilli giye tini ya alochana karabena tate purna sammati dite raji isalamabada	দিলিগিয়ে টিনি জা আলসানা কারাবেনা টাটে পুরনা সামাটকিটে রাজি ইসা দা মাবাদ
asale mumbai jangihana niye bharata o antarjatika kshetra theke pratidina etarakama chapa asachhe ye karyata dishehara pakistana	আসালে মুমবিা জামগিহানা নিয়ে ভারাটা অ আনটা-রজাটিকা কসহেটেরা থেকে পরাটিদিনা এটারাকামা চাা আসাছ য়ে কারয়াটা দিসহেহারা পাকিসটানা
jamata uda daoyara biruddhe byabastha nebara katha bale labha hayani	চামাহা উধাদায়ারা বিরুদধো বয়াবাটানেগারা পাথা বালদে লা ভানানি
tara upara shapatha niyei baraka obama bale diyechhena pakistana arthika sahayya pabe ekatai sharte	টারাপারা সাপিায়িে বারাকপাবাবা বালে দিয়েছেনা পাকিসটানা আরাটি সাহায়াপাবে একাটেিা সারাটে
jangidera nirmula karara abhiyana tadera chaliye yete habe	য়াঙিদেরা নিরপুানা কারারা আভিয়ানা পটাদেরা চাদি জেটে হাবে
yadi ta na kare tahale arthika sahayya niye natuna kare bhababe amerika	য়াদি টানা কারে টাহালে আরাথিক সাহাজানিয়ে নু-টুনা কারে ভাকাবে য়ামেৃকা
ei abasthaya dirghadinera bandhu chinakei bipada theke uddhare madhyastha karate chaya jaradari gilanira sarakara	এি আবাসথায়া দিরগখাদিনেরা বানধু চিনাকেি বি-পাদাথেকে উদধারে মাধয়াসটা কারাটে চায়া জারাদৃা গিলানিরা সারাকারা
gata unatrishe aktobara isalamabada giyechhilena hi iyaphei	গাটা উনটৃসহয়া আকটাবারা ইসুলামাবানা জিয়েস-হানা হি ইয়ানে
sekhana theke e bachhara pancha tarikha dilli asena tini	সিখান থেকে এ বাছারা পাচসাকৃাকাদিনিয়াসহেনা থিনি
bharatake anurodha karena upamahadeshe uttejana kamate pakistanera sange yena shigagirai kathabarta shuru haya	ভারাটাকে নুরধা কারেনা উপামাহাদেসহে উটটে-জানা কামাটে পাকিসটানেরা সাঙে য়েনা ইগাগিরিা কাথাবারটা সহুরু হায়ে
se kathabarta ajao chalu hayani	সহে পাথাবারটা আজা জানুহায়ানি
dupakshai simante jado karechhe bipula senabahini	দুপাকসহাসা সিমানটে জাদা কারেছেন বিপানসহে নাবাকানি
pakistana chaichhe arao ekabara dilli yana iyaphei	পাকিসটান চিাচিে আরাায়েকা বারা দিলিজানা হৌফে
pakistanera sange alochanaya basara janya chapa srrshti karuna	পাকিসটানেরা সাঙে আলচানায়া বাসারা জানয়া চা-পা সররসহটি কারুনা

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
ye paddhatitei china egoka sabetei tadera sammati achhe janiyechhena kureshi	য়ে পাদধাটিটেি চিনা এগকা সাবেটেি টাদেরা সাম-মাটি আছে জানিয়েছেনা কুরেসহি
andale natuna bimananagari hale kebala bardhamana naya ashapashera besha kayekati jelarao unnati habe bale mane kara hachchhe	আনদালে নাটুনা বিমানানাগ্া হালে কেবালা বারধা-মানা নায়া আসহাপাসহেরা বেসহা কায়েকাটি জে-লারা উননাটি হাবে বালে মানে কারা হাচছে
shudhu tai naya natuna bimanabandarera aparisima subidha pabe durgapura o asanasola shilpanchala	সহুধু টিা নায়া নাটুনা বিমানাবানদারেরা আপ্াসিমা সুবিধা পাবে দুরগাপুরা অ আসানাসলা সহিলপান-চালা
karmasansthana habe bahu manushera	কারমাসানসথানা হাবে বাহু মানুসহেরা
gata deda du bachhare rajye shilpayana prakriya jora dhakka kheyechhe	গাটা দেদা দু বাছারে রাজয়ে সহিলপায়ানা পরাকৃয়া জরা ধাককা খেয়েছে
nandigrame gandagolera jere anishchita haye padechhe bahu prakalpa	নানদিগরামে গানদাগলেরা জেরে আনিসহচিটা হায়ে পাদেছে বাহু পরাকালপা
tarai madhye singura theke hata gutiyechhe tata prakalpa	টারিা মাধয়ে সিঙুরা থেকে হাটা গুটিয়েছে টাটা পরা-কালপা
bardhamanera katoyaya tapabidyut prakalpa ghire shuru hayechhe andolana	বারধামানেরা কাটয়ায়া টাপাবিদয়ুট পরাকালপা ঘি-রে সহুরু হায়েছে আনদলানা
e abasthaya sarakara bishesha bhabe takiye rayechhe andala bimananagari prakalpatira dike	এ আবাসথায়া সারাকারা বিসহেসহা ভাবে টাকিয়ে রায়েছে আনদালা বিমানানাগ্া পরাকালপাটিরা দিকে
itimadhyei sarakara samiksha kare dekhechhe ye bhabe damadama bimanabandare yatri sankhya badachhe tate adura bhabishyate dbitiya bimanabandarera prayojana abashyambhabi	ইটিমাধয়েি সারাকারা সামিকসহা কারে দেখেছে য়ে ভাবে দামাদামা বিমানাবানদারে য়াটৃ সানখয়া বাদা-ছে টাটে আদুরা ভাবিসহয়াটে দবিটিয়া বিমানাবান-দারেরা পরায়জানা আবাসহয়ামভাবি
se karanei dakshina chabbisha paraganara baruipurera kachhe bimanabandara tairira bhabana chinta shuru kara hayechhilo	সে কারানেি দাকসহিনা চাববিসহা পারাগানারা বা-রুিপুরেরা কাছে বিমানাবানদারা ট্িরা ভাবানা চিনটা সহুরু কারা হায়েছিল
erai madhye bengala erotrapilasa du hajara chhaya sale egiye ase bardhamanera andale bimanabandara o take kendra kare nagarayanera kaje	এরিা মাধয়ে বেঙালা এরটরাপিলাসা দু হাজারা ছায়া সালে এগিয়ে আসে বারধামানেরা আনদালে বিমা-নাবানদারা অ টাকে কেনদরা কারে নাগারায়ানেরা কাজে

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
mukhyamantri buddhadeba bhattacharya atyanta ashabadi ei prakalpatike niye	মুখয়ামানট্ বুদধাদেবা ভাটটাচারয়া আটয়ানটা আসহাবাদি এি পরাকালপাটিকে নিয়ে
puro prakalpe praya dasha hajara koti taka biniyoga habe	পুর পরাকালপে পরায়া দাসহা হাজারা কটি টাকা বিনিয়গা হাবে
pashapashi karmasansthana habe bahu manushera	পাসহাপাসহি কারমাসানসথানা হাবে বাহু মানুসহেরা
shudhu tai naya andale ei bimananagarira prabhaba padabe bardhamanasaha dakshinabangera kayekati jelaya	সহুধু টাি নায়া আনদালে এি বিমানানাগ্যারা পরাভাবা পাদাবে বারধামানাসাহা দাকসহিনাবাঙেরা কায়েকাটি জেলায়া
itimadhyei bengala erotrapilasa samiksha kare dekhechhe pashchimabanga simantabarti jhadakhandera bibhinna elakara manusha ei bimanabandarati byabahara karate parabena	ইটিমাধয়েি বেঙালা এরটরাপিলাসা সামিকসহা কারে দেখেছে পাসহচিমাবাঙা সিমানটাবারটি ঝাদাখানদেরা বিভিননা এলাকারা মানুসহা এি বিমানাবানদারাটি বয়াবাহারা কারাটে পারাবেনা
phale bimanabandaratira labhera sambhabana yatheshta	ফালে বিমানাবানদারাটিরা লাভেরা সামভাবানা য়াথেসহটা
durgapura asanasola shilpanchaleo ei bimananagarira prabhaba padabe yatheshta	দুরগাপুরা আসানাসলা সহিলপানচালে এি বিমানানাগ্যারা পরাভাবা পাদাবে য়াথেসহটা
baraka obama markina presidentera dayitba bhara neoyara para bharata markina samparka arao majabuta habe	বারাকা অবামা মারকিনা পরেসিদেনটেরা দায়িটবা ভারা নেয়ারা পারা ভারাটা মারকিনা সামপারকা আরা মাজাবুটা হাবে
habu markina rashtrapatira kachha theke paoya chithira bayanera bhittitei jora galaya ekatha janalena pradhanamantri manamohana sinha	হাবু মারকিনা রাসহটরাপাটিরা কাছা থেকে পায়া চিথিরা বায়ানেরা ভিটটিটেি জরা গালায়া একাথা জানালেনা পরাধানামানট্ মানামহানা সিনহা
nirbachane jaya labhera para ushna abhinandana janiye obamake chithi likhechhilena manamohana	নিরবাচানে জায়া লাভেরা পারা উসহনা আভিনানদানা জানিয়ে অবামাকে চিথি লিখেছিলেনা মানামহানা
tara uttare obamao chithi likhechhena tanke	টারা উটটারে অবামা চিথি লিখেছেনা টানকে
sei chithi pade mugdha manamohana	সেি চিথি পাদে মুগধা মানামহানা
chithite bharata markina bhabishyat samparka ki bhabe arao nibida kara yaya sei bishaye tara drrshtibhangii chithite likhechhena baraka obama	চিথিটে ভারাটা মারকিনা ভাবিসহয়াট সামপারকা কি ভাবে আরা নিবিদা কারা য়ায়া সেি বিসহায়ে টারা দররসহটিভাঙিি চিথিটে লিখেছেনা বারাকা অবামা

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
jayalabhera para manamohanake phona na kare bharatake abahela karechhena obama ei yukti udiye diye pradhanamantri janana	জায়ালাভেরা পারা মানামহানাকে ফনা না কারে ভারাটাকে আবাহেলা কারেছেনা অবামা এি য়ুকটি উদিয়ে দিয়ে পরাধানামানট্ জানানা
upasagariya deshagulite tara sapharera samayasuchigata asubidhara karanei du janera katha bala haye otheni	উপাসাগৃায়া দেসহাগুলিটে টারা সাফারেরা সামায়া-সুচিগাটা আসুবিধারা কারানেি দু জানেরা কাথা বালা হায়ে অথেনি
kintu obama tara sange phone katha balara ichchha prakasha karechhilena	কিনটু অবামা টারা সাঙে ফনে কাথা বালারা ইচছা পরাকাসহা কারেছিলেনা
deshe phirei tara sange obamara katha habe baleo janana pradhanamantri	দেসহে ফিরেি টারা সাঙে অবামারা কাথা হাবে বালে জানানা পরাধানামানট্
obamara sange itimadhyei panera ti deshera netadera sange katha hayechhe teliphone	অবামারা সাঙে ইটিমাধয়েি পানেরা টি দেসহেরা নেটাদেরা সাঙে কাথা হায়েছে টেলিফনে
edera madhye rayechhena paka presidenta asipha ali jaradario	এদেরা মাধয়ে রায়েছেনা পাকা পরেসিদেনটা আসি-ফা আলি জারাদ্া
itimadhyei rajadhanite phire esechhena manamohana	ইটিমাধয়েি রাজাধানিটে ফিরে এসেছেনা মানামহানা
ekhana obamara sange tara alapera madhyamei shuru habe bharata markina natuna samparkera suchana	এখানা অবামারা সাঙে টারা আলাপেরা মাধয়ামেি সহুরু হাবে ভারাটা মারকিনা নাটুনা সামপারকেরা সুচানা
amerikara prathama krrshnanga presidenta hisabe obamara jayera para besha dushchintaya rayechhe pakistana	আমেৃকারা পরাথামা কররসহনাঙা পরেসিদেনটা হি-সাবে অবামারা জায়েরা পারা বেসহা দুসহচিনটায়া রায়েছে পাকিসটানা
anekei ekhana bhabate shuru karechhena pakistanera uttaranchale sena abhiyanera matra na badiye dena obama	আনেকেি এখানা ভাবাটে সহুরু কারেছেনা পাকিস-টানেরা উটটারানচালে সেনা আভিয়ানেরা মাটরা না বাদিয়ে দেনা অবামা
ei abasthaya chintabhabana chalachhe obama sarakarake bojhanora ye ei anchalera jangidera mokabilaya pakistana sarakara ekai yatheshta	এি আবাসথায়া চিনটাভাবানা চালাছে অবামা সা-রাকারাকে বঝানরা য়ে এি আনচালেরা জাঙিদেরা মকাবিলায়া পাকিসটানা সারাকারা একিা য়াথেসহ-টা
kathata yadi obamake bojhano na yaya tahale du deshera samparkera abanati ghata bichitra naya	কাথাটা য়াদি অবামাকে বঝান না য়ায়া টাহালে দু দেসহেরা সামপারকেরা আবানাটি ঘাটা বিচিটরা না-য়া
santrasera biruddhe ladaiyera prashne demokryata ripabalikanera manobhabe bishesha pharaka nei	সানটরাসেরা বিরুদধে লাদিায়েরা পরাসহনে দেম-করয়াটা ঋপাবালিকানেরা মানভাবে বিসহেসহা ফা-রাকা নেি

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
du dalai chaya santrasake ekebare shikada theka upade phelate	দু দালি চায়া সানটরাসাকে একেবারে সহিকাদা থেকা উপাদে ফেলাটে
pakistanera uttaranchala o aphaganistanera simante durgama girikandarera kothao ga dhaka diye achhena osama bina ladena ei bishbasa anekadinera	পাকিসটানেরা উটটারানচালা অ আফাগানিসটানেরা সিমানটে দুরগামা গিৃকানদারেরা কথা গা ধাকা দিয়ে আছেনা অসামা বিনা লাদেনা এি বিসহবাসা আনেকাদিনেরা
jangidera sandhane besha kichhudina dhare ei anchale abhiyana chalachchhe	জাঙিদেরা সানধানে বেসহা কিছুদিনা ধারে এি আনচালে আভিয়ানা চালাচছে
pracharera samaya obama balechhilena ala kayadake kayada karate pakistana yadi byartha haya tahale tara netrrtbe odera nishana karabe amerikai	পরাচারেরা সামায়া অবামা বালেছিলেনা আলা কায়াদাকে কায়াদা কারাটে পাকিসটানা য়াদি বয়ারথা হায়া টাহালে টারা নেটররটবে অদেরা নিসহানা কারাবে আমেৃকিা
tini ingita diyechhilena prayojane iraka theke senabahini sariye niye aphaganistane ladai karate pathabena	টিনি ইঙিটা দিয়েছিলেনা পরায়জানে ইরাকা থেকে সেনাবাহিনি স্যায়ে নিয়ে আফাগানিসটানে লাদি কারাটে পাথাবেনা
presidenta haoyara age obamara eisaba kathabartai apatata chintaya rakhachhe pakistanake	পরেসিদেনটা হায়ারা আগে অবামারা এিসাবা কাথাবারটিা আপাটাটা চিনটায়া রাখাছে পাকিসটানাকে
tahale emana dandabe ye markina phaujera tattbabadhanei jangidera biruddhe yuddha chalate habe tadera	টাহালে এমানা দানদাবে য়ে মারকিনা ফুাজেরা টাটটবাবাধানে জাঙিদেরা বিরুদধে য়ুদধা চালাটে হাবে টাদেরা
phera baghabandi kanda	ফেরা বাঘাবানদি কানদা
somabara rate gosabara rajatajubali grame bana karmidera pata phande dhara padalo rayyala bengala taigara	সমাবারা রাটে গসাবারা রাজাটাজুবালি গরামে বানা কারমিদেরা পাটা ফানদে ধারা পাদাল রায়য়ালা বেঙালা টিাগারা
mangalabara take niye kendo dbipera dike raona diyechhena bana daphatarera karmira	মাঙালাবারা টাকে নিয়ে কেনদ দবিপেরা দিকে রানা দিয়েছেনা বানা দাফাটারেরা কারমিরা
sekhanei budhabara baghatike jangale chhede deoya habe	সেখানে বুধাবারা বাঘাটিকে জাঙালে ছেদে দেয়া হাবে
abasheshe somabara gabhira rate dhara padalo sundarabanera gosabara rajatajubali gramera bagha	আবাসহেসহে সমাবারা গাভিরা রাটে ধারা পাদাল সুনদারাবানেরা গসাবারা রাজাটাজুবালি গরামেরা বাঘা
brrhaspatibara theke ekati rayyala bengala taigara daphaya daphaya hana dichchhila rajatajubali grame	বররহাসপাটিবারা থেকে একাটি রায়য়ালা বেঙালা টিাগারা দাফায়া দাফায়া হানা দিচছিলা রাজাটাজুবালি গরামে

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
grama theke ekati chhagalao niye yaya se	গরামা থেকে একাটি ছাগালা নিয়ে য়ায়া সে
shuru haya rata jege gramabasidera pahara	সহুরু হায়া রাটা জেগে গরামাবাসিদেরা পাহারা
banakarmidera pata phande somabara rate dhara pade baghati	বানাকারমিদেরা পাটা ফানদে সমাবারা রাটে ধারা পাদে বাঘাটি
tabe shudhu eka bagha naya	টাবে সহুধু একা বাঘা নায়া
sara rata khanchara baire apeksha karachhila ekati baghinio	সারা রাটা খানচারা বিারে আপেকসহা কারাছিলা একাটি বাঘিনি
mangalabara bhore aguna jbaliye banakarmira baghinitike tadiye dena	মাঙালাবারা ভরে আগুনা জবালিয়ে বানাকারমিরা বাঘিনিটিকে টাদিয়ে দেনা
erapara khancha bandi baghatike niye prathame yaoya haya sajanekhali taigara syanachuyari	এরাপারা খানচা বানদি বাঘাটিকে নিয়ে পরাথামে য়ায়া হায়া সাজানেখালি টিাগারা সয়ানাচুয়্যা
tarapara lanche kare baghatike niye bana daphatarera karmira raona dena kendo dbipera dike	টারাপারা লানচে কারে বাঘাটিকে নিয়ে বানা দাফাটারেরা কারমিরা রানা দেনা কেনদ দবিপেরা দিকে
bane thikamato khabara na paoyatei phera lokalaye dhuke padechhe rayyala bengala taigara	বানে থিকামাট খাবারা না পায়াটে ফেরা লকালায়ে ধুকে পাদেছে রায়য়ালা বেঙালা টিাগারা
mane karachhena bisheshajnara	মানে কারাছেনা বিসহেসহাজনারা
gata phebruyari masa theke besha kayekabara sundarabana sanlagna gramaguli theke dhara padalo bagha	গাটা ফেবরুয়্যা মাসা থেকে বেসহা কায়েকাবারা সুনদারাবানা সানলাগনা গরামাগুলি থেকে ধারা পাদাল বাঘা
gata saterai phebruyari sundarabanera kulatali grama kheke dhara pade rayyala bengala taigara	গাটা সাটেরিা ফেবরুয়্যা সুনদারাবানেরা কুলাটালি গরামা খেকে ধারা পাদে রায়য়ালা বেঙালা টিাগারা
bishe phebruyari rate jhadakhalite baghake khanchabandi karena bana karmira	বিসহে ফেবরুয়্যা রাটে ঝাদাখালিটে বাঘাকে খানচাবানদি কারেনা বানা কারমিরা
erapara naya juna gosabara kumiramari grame phera dhara pade bagha	এরাপারা নায়া জুনা গসাবারা কুমিরাম্যা গরামে ফেরা ধারা পাদে বাঘা
adai masa bade trishe agasta gosabara yanapura grame phera banakarmidera phande bandi haya bagha	আদিা মাসা বাদে ট্রসহে আগাসটা গসাবারা য়ানাপুরা গরামে ফেরা বানাকারমিদেরা ফানদে বানদি হায়া বাঘা
gosabatei pakhiralayera kachhe ekati grame chabbisha aktobara bhore bagha dhara pade	গসাবাটে পাখিরালায়েরা কাছে একাটি গরামে চাববিসহা আকটবারা ভরে বাঘা ধারা পাদে

Table 5.2 continued from previous page

Reference	Machine Output (Bengali Phonemes)
tarapara budhabara gabhira rate arthat teshara nabhembara gosabara rajatajubali grame abara baghabandi kanda	টারাপারা বুধাবারা গাভিরা রাটে আরথাট টেসহারা নাভেমবারা গসাবারা রাজাটাজুবালি গরামে আবারা বাঘাবানদি কানদা
rajatajubali grama theke bagha bandi parbera puro ghatanatira netrrtba diyechhena taigara rijarbha pharestera phillda direktara subrata mukhopadhyaya	রাজাটাজুবালি গরামা থেকে বাঘা বানদি পারবেরা পুর ঘাটানাটিরা নেটররটবা দিয়েছেনা টিগারা ঋজারভা ফারেসটেরা ফিলদা দিরেকটারা সুবরাটা মুখপাধয়ায়া
tini janiyechhena purna bayaska ei baghati sustha	টিনি জানিয়েছেনা পুরনা বায়াসকা এি বাঘাটি সুসথা
budhabara bikelera madhye take niye lancha kendo dbipe paunchhe yabe	বুধাবারা বিকেলেরা মাধয়ে টাকে নিয়ে লানচা কেনদ দবিপে পুানছে য়াবে
sekhanei take jangale chhede debena banadaphatarera karmira	সেখানেি টাকে জাঙালে ছেদে দেবেনা বানাদাফাটারেরা কারমিরা
piesaelabhi si tu raketera sahayye nikhuntabhabe chandrayana eka ke pathano hayechhe mahakashe	পিেসোলাভি সি টু রাকেটেরা সাহায়য়ে নিখুনটাভাবে চানদরায়ানা একা কে পাথান হায়েছে মাহাকাসহে
ebara take nikhuntabhabe chandera kakshapathe sthapana karara pala	এবারা টাকে নিখুনটাভাবে চানদেরা কাকসহাপাথে সথাপানা কারারা পালা
itimadhye prrthibira chhabi tule pathiyechhe chandrayana	ইটিমাধয়ে পররথিবিরা ছাবি টুলে পাথিয়েছে চানদরায়ানা
tate bojha giyechhe ashatita bhala kaja karachhe sei mahakashayana	টাটে বঝা গিয়েছে আসহাটিটা ভালা কাজা কারাছে সেি মাহাকাসহায়ানা
sabakichhu thikathaka thakale agami atai nabhembara chandrayana pradakshina shuru karabe chandake	সাবাকিছু থিকাথাকা থাকালে আগামি আটিা নাভেমবারা চানদরায়ানা পরাদাকসহিনা সহুরু কারাবে চানদাকে
ara panera nabhembara bijnanira take sthapana karabena chandera kakshapathe	আরা পানেরা নাভেমবারা বিজনানিরা টাকে সথাপানা কারাবেনা চানদেরা কাকসহাপাথে
mahakashayanake chandera kakshapathe sthapana karara kajati atyanta duruha	মাহাকাসহায়ানাকে চানদেরা কাকসহাপাথে সথাপানা কারারা কাজাটি আটয়ানটা দুরুহা
era age bahu mahakashayanake kakshapathe sthapana karate giye haya ta achhade padechhe sei upagrahera buke	এরা আগে বাহু মাহাকাসহায়ানাকে কাকসহাপাথে সথাপানা কারাটে গিয়ে হায়া টা আছাদে পাদেছে সেি উপাগরাহেরা বুকে
nayato ta hariye giyechhe mahashunye	নায়াট টা হায়ে গিয়েছে মাহাসহুনয়ে

Table 5.2 continued from previous page

Reference	**Machine Output (Bengali Phonemes)**
gata unatrishe aktobara chandrayana prrthibira ye chhabi tule pathiyechhe ta bishleshana kare bijnanira janate perechhena mahakashayanati ekhana ki abasthaya achhe	গাটা উনাট্রসহে আকটবারা চানদরায়ানা পররথি-বিরা য়ে ছাবি টুলে পাথিয়েছে টা বিসহলেসহানা কারে বিজনানিরা জানাটে পেরেছেনা মাহাকাসহা-য়ানাটি এখানা কি আবাসথায়া আছে

5.4 Precision Recall for IPA Syllable

Following table 5.3 shows the evaluation with the Human Generated syllable with the Machine Generated using F score.

Table 5.3: Precision Recall Value for IPA Syllable

Human Generated	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
is-la-ma-ba-de-ra	is-la-ma-ba-de-ra	6	6	6
cay-na	cay-na	2	2	2
du-ta-ba-ʃera	du-ta-ba-ʃe-ra	3	5	4
æk	æk	1	1	1
o-nuʃ-t ^{h}an-e	o-nuʃ-t ^{h}a-ne	2	4	4
e-i	e-i	2	2	2
mɔn-tab-byo	mɔn-tab-byo	3	3	3
ko-re-c ^{h}en	ko-re-c ^{h}en	3	3	3
pa-ka	pa-ka	2	2	2
bi-deʃ	bi-deʃ	2	2	2
mɔn-tri	mɔn-tri	2	2	2
ʃa-ha	ʃa-ha	2	2	2
mæh-mud	mæh-mu-da	1	3	2
ku-re-ʃi	ku-re-ʃi	3	3	3
t̪̃ar	t̪	0	1	1
kɔ-t ^{h}aɐ̯	kɔ-t ^{h}aɐ̯	2	2	2
pa-kis-t^{h}an	pa-kis-t^{h}an	3	3	3
t̪ar	t̪ar	1	1	1
dir-g ^{h}o-din-er	dir-g ^{h}o-di-ner	2	4	4
bɔn-d u	bɔn-d u	2	2	2
t͡sa-e-na-ke	t͡sa-e-na-ke	4	4	4

Table 5.3 continued from previous page

Human Generated	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
e	e	1	1	1
bæ-pa-re	bæ-pa-re	3	3	3
blænk	blæn-k	0	2	1
cek	cek	1	1	1
di-e	di-e	2	2	2
rak^{h}-c^{h}e	rak^{h}-c^{h}	1	2	2
t͡si-ner	t͡si-ner	2	2	2
bi-ʃeʃ	bi-ʃeʃ	2	2	2
d̪ut	d̪u	0	1	1
hi	hi	1	1	1
iYApheiYera	i—-	0	1	1
ʃɔŋ-ge	ʃɔŋ-ge	2	2	2
t̪ar	t̪ar	1	1	1
kɔ-t ^{h}a-o	kɔ-t ^{h}a-o	3	3	3
ho-e-ts e	ho-e-ts e	3	3	3
t̪̃a-ke	t̪-	0	2	1
t̪i-ni	t̪i-ni	2	2	2
o-nu-rod	o-nu-rod	3	3	3
kɔ-re-ts^{h}en	kɔ-re-ts^{h}en	3	3	3
ʃig-gi-ri	ʃig-gi-ri	3	3	3
dil-li	dil-li	2	2	2
d͡ze-te	d͡ze-te	2	2	2
dil-li	dil-li	2	2	2
gi-e	gi-e	2	2	2
t̪i-ni	t̪i-ni	2	2	2
d͡ze	d͡ze	1	1	1
a-lo-t͡so-na	a-lo-t͡so-na	4	4	4
kɔr-ben	kɔr-ben	2	2	2
t̪a-t̪e	t̪a-t̪e	2	2	2
pur-no	pur-no	2	2	2
ʃɔm-mo-ti	t̪ɔm-mo-ti	3	3	3
di-t̪e	di-t̪e	2	2	2
ra-d͡zi	ra-d͡zi	2	2	2
is-la-ma-bad	is-la-ma-bad	4	4	4
a-ʃole	a-ʃo-le	1	3	2
mum-ba-i	mum-ba-i	3	3	3

Table 5.3 continued from previous page

Human Generated	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
d͡zɔŋ-gi	d͡zɔŋ-gi	2	2	2
ha-na	ha-na	2	2	2
ni-e	ni-e	2	2	2
b ʰa-rot	b ʰa-rot	2	2	2
o	o	1	1	1
an-ʈor-d͡za-tik	an-ʈor-d͡za-tik	4	4	4
k ʰet-tro	k ʰet-ro	1	2	2
t ʰe-ke	t ʰe-ke	2	2	2
pro-ti-din	pro-ti-din	3	3	3
æ-to-rɔ-kom	æ-to-rɔ-kom	4	4	4
t͡sap	t͡sap	1	1	1
aʃ-ts͡ʰe	aʃ-ts͡ʰe	2	2	2
d͡ze	d͡ze	1	1	1
kar-d͡zo-to	kar-d͡zo-to	3	3	3
di-ʃe-ha-ra	di-ʃe-ha-ra	4	4	4
d͡za-mat	d͡za-mat	2	2	2
ud	ud	1	1	1
da-o-ar	da-o-ar	3	3	3
bi-rud-d ʰe	bi-rud-d ʰe	3	3	3
bæ-bos-t ʰa	bæ-bos-t ʰa	3	3	3
ne-bar	ne-bar	2	2	2
kɔ-t ʰa	kɔ-t ʰa	2	2	2
bo-le	bo-le	2	2	2
lab ʰ	lab	0	1	1
ho-y	hoy	0	1	2
ni	ni	1	1	1
tar	tar	1	1	1
u-po-re	u-po-re	3	3	3
ʃɔ-pot ʰ	ʃɔ-pot ʰ	2	2	2
ni-e-i	ni-e-i	3	3	3
ba-rak	ba-rak	2	2	2
o-ba-ma	o-ba-ma	3	3	3
bɔ-le	bɔ-le	2	2	2
di-e-ts͡ʰen	di-e-ts͡ʰen	3	3	3
ar-t ʰik	ar-t ʰik	2	2	2
ʃa-had͡z-d͡zo	ʃa-ha-d͡zo	1	3	3

Table 5.3 continued from previous page

Human Generated	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
pa-be	pa-be	2	2	2
æk-ʈa-i	aek-ʈa-i	3	3	3
ʃɔr-te	ʃɔr-te	2	2	2
d͡zoŋ-gi-der	d͡zoŋ-gi-der	3	3	3
nir-mul	nir-mul	2	2	2
kɛ-bol	kɛ-bol	2	2	2
bɔr-d ^{h}o-man	bɔr-d ^{h}o-man	3	3	3
		208	232	228

From the table 5.3 shown **208** is the "Number of Syllable Matching the Text", **232** is "Syllable Count in Machine Output" and **228** is "Syllable Count in Human Output".

$$Precision(P) = \frac{208}{232} = \mathbf{0.896551724} \tag{5.1}$$

$$Recall(R) = \frac{208}{228} = \mathbf{0.912280702} \tag{5.2}$$

Then F1 Score is

$$F = 2 \times \frac{(P \times R)}{(P + R)}$$
$$F = 2 \times \frac{(0.896551724 \times 0.912280702)}{(0.896551724 + 0.912280702)} = \mathbf{0.904347826} \tag{5.3}$$

5.5 Precision Recall for Bengali Syllable

Following table 5.4 shows the evaluation with the Human Generated syllable with the Machine Generated using F score.

Table 5.4: Precision Recall Value for Bengali Syllable

Original Word	**Human Generated**	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
বাঁকুড়া	বাঁ-কু-ড়া	বাঁ-কু-ড়া	3	3	3
মানুষদের	মা-নুষ-দের	মা-নুষ-দের	3	3	3
সাদর	সা-দর	সাদ-রর	0	2	2
জানাচ্ছে	জা-নাচ-ছে	জা-নাচ-ছে	3	3	3
পশ্চিমবঙ্গ	পশ-চিম-বঙ-গ	পশ্চিম-বঙ্গ	0	2	4
রাজ্য	রাজ-য	রাজ-য	2	2	2
পর্যটন	পর-য-টন	পর্যটন	0	1	3
জেলা	জে-লা	জে-লা	2	2	2
প্রশাসন	প্র-শা-সন	প্রশাস-ন	0	2	3
থেকে	থে-কে	থে-কে	2	2	2
দুই	দু-ই	দু-ই	2	2	2
মহকুমায়	ম-হ-কু-মায়	মহকু-মায়	1	2	4
শুরু	শু-রু	শু-রু	2	2	2
হয়েছে	হ-য়ে-ছে	হয়ে-ছে	1	2	3
ভিন্ন	ভিন-ন	ভিন-ন	2	2	2
ধরনের	ধ-র-নের	ধরনের	0	1	3
দুই	দু-ই	দু-ই	2	2	2
মেলা	মে-লা	মে-লা	2	2	2
চলবে	চল-বে	চলবে	0	1	2
পর্যন্ত	পর-যন-ত	পর্যন্ত	0	1	3
আর	আর	আর	1	1	1
প্রাকৃতিক	প্রা-কৃ-তিক	প্রাক-ৃতিক	0	2	3
সৌন্দর্য	সৌন-দর-য	সৌন-দর্য	1	2	3
ঘেরা	ঘে-রা	ঘে-রা	2	2	2
রানি	রা-নি	রা-নি	2	2	2
মুকুটমণিপুরে	মু-কুট-ম-ণি-পু-রে	মু-কুট-মণি-পু-রে	4	5	6
শুরু	শু-রু	শু-রু	2	2	2
মেলা	মে-লা	মে-লা	2	2	2
দিকে	দি-কে	দি-কে	2	2	2
নগরী	ন-গ-রী	নগরী	0	1	3

Table 5.4 continued from previous page

Original Word	**Human Generated**	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
বিষ্ণুপুর	বিষ-ণু-পুর	বিষ-ণু-পুর	3	3	3
শুরু	শু-রু	শু-রু	2	2	2
হল	হল	হল	1	1	1
মিউজিক	মি-উ-জিক	মি-উ-জিক	3	3	3
আদিবাসী	আ-দি-বা-সী	আ-দি-বা-সী	4	4	4
ও	ও	ও	1	1	1
শিল্পের	শিল-পের	শিল-পের	2	2	2
তিন	তিন	তিন	1	1	1
দিনের	দি-নের	দি-নের	2	2	2
মেলার	মে-লার	মে-লার	2	2	2
উদ্বোধন	উদ-বো-ধন	উদ-বোধ-ন	1	3	3
করেন	ক-রেন	করেন	0	1	2
জেলা	জে-লা	জে-লা	2	2	2
পরিষদের	প-রি-ষ-দের	পরিষ-দের	1	2	4
উপস্থিত	উ-প-স্থিত	উপ-স্থিত	1	2	3
ছিলেন	ছি-লেন	ছি-লেন	2	2	2
মহকুমা	ম-হ-কু-মা	মহকু-মা	1	2	4
শাসক	শা-সক	শাস-ক	0	2	2
রাজু	রা-জু	রা-জু	2	2	2
মিশ্র	মি-শর	মিশ-র	0	2	2
খাতড়া	খাত-ড়া	খাত-ড়া	2	2	2
সমিতির	স-মি-তির	সমি-তির	1	2	3
আধিকারিক	আ-ধি-কা-রিক	আ-ধি-কা-রিক	4	4	4
রামশঙ্কর	রাম-শন-কর	রাম-শঙ্কর	1	2	3
প্রমুখ	প্র-মুখ	প্রমুখ	0	1	2
এখানে	এ-খা-নে	এ-খা-নে	3	3	3
আসা	আ-সা	আ-সা	2	2	2
সৌন্দর্য	সৌন-দর-য	সৌন-দর্য	1	2	3
উপভোগের	উ-প-ভো-গের	উপ-ভো-গের	2	3	4
পাশাপাশি	পা-শা-পা-শি	পা-শা-পা-শি	4	4	4
আদিবাসী	আ-দি-বা-সী	আ-দি-বা-সী	4	4	4

Table 5.4 continued from previous page

Original Word	**Human Generated**	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
সংগীত	সং-গীত	সংগীত	0	1	2
ও	ও	ও	1	1	1
অনুষ্ঠান	অ-নুষ-ঠান	অ-নুষ-ঠান	3	3	3
দেখার	দে-খার	দে-খার	2	2	2
সুযোগ	সু-যোগ	সু-যোগ	2	2	2
পাবেন	পা-বেন	পা-বেন	2	2	2
জানানো	জা-না-নো	জা-না-নো	3	3	3
হয়েছে	হয়ে-ছে	হয়ে-ছে	2	2	2
আটটি	আট-টি	আট-টি	2	2	2
লোকশিল্পীরা	লোক-শিল-পী-রা	লোক-শিল-পী-রা	4	4	4
তিন	তিন	তিন	1	1	1
দিন	দিন	দিন	1	1	1
তারা	তা-রা	তা-রা	2	2	2
পরিবেশ	পরি-বেশ	পরি-বেশ	2	2	2
করবেন	কর-বেন	করবেন	0	1	2
বিষ্ণুপুর	বিষ-ণু-পুর	বিষ-ণু-পুর	3	3	3
মিউজিক	মি-উ-জিক	মি-উ-জিক	3	3	3
অপরূপ	অ-প-রূপ	অপ-রূপ	1	2	3
আলোকসজ্জায়	আ-লোক-সজ-জায়	আ-লোক-সজ্জায়	2	3	4
সেজে	সে-জে	সে-জে	2	2	2
উঠেছে	উ-ঠে-ছে	উ-ঠে-ছে	3	3	3
প্রাঙ্গণ	প্রাঙ-গণ	প্রাঙ-গণ	2	2	2
এক	এক	এক	1	1	1
দিকে	দি-কে	দি-কে	2	2	2
পোড়া	পো-ড়া	পো-ড়া	2	2	2
জয়	জয়	জয়	1	1	1
এক	এক	এক	1	1	1
বছরের	বছ-রের	বছরের	0	1	2
নির্বাসনের	নির-বা-স-নের	নির-বাস-নের	2	3	4
পালা	পা-লা	পা-লা	2	2	2
বিধ্বংসী	বিধ-ধন-সী	বিধ-বংসী	1	2	3

Table 5.4 continued from previous page

Original Word	**Human Generated**	**Machine Generated**	# Correct Syllables	# Syllables in Machine output	# Syllables in Human output
মেজাজে	মে-জা-জে	মে-জা-জে	3	3	3
এবং	এ-বং	এব-ং	0	2	2
রুটের	রু-টের	রু-টের	2	2	2
ভাবনায়	ভাব-নায়	ভাব-নায়	2	2	2
তিনি	তি-নি	তি-নি	2	2	2
বলেছেন	বলে-ছেন	বলে-ছেন	2	2	2
সেরা	সে-রা	সে-রা	2	2	2
আরও	আ-রও	আ-রও	2	2	2
			168	209	245

From the table 5.4 shown **168** is the "Number of Syllable Matching the Text", **209** is "Syllable Count in Machine Output" and **245** is "Syllable Count in Human Output".

$$Precision(P) = \frac{168}{209} = \mathbf{0.803828} \tag{5.4}$$

$$Recall(R) = \frac{168}{245} = \mathbf{0.685714} \tag{5.5}$$

Then F1 Score is

$$\begin{aligned} F &= 2 \times \frac{(P \times R)}{(P + R)} \\ F &= 2 \times \frac{(0.803828 \times 0.685714)}{(0.803828 + 0.685714)} = \mathbf{0.740088} \end{aligned} \tag{5.6}$$

Chapter 6

Conclusions and Future Scope

6.1 Conclusion

In this work, the task of the system is syllabifying Bangla word automatically. The structure of the program is based on linguistic features of speech sound combined with the computational methods. It has been implemented in Python version 2.7.15. The automatic syllabification can help to improve ASR system of Bangla as the syllabification is an important problem in automatic speech recognition and speech synthesis applications. Keeping this in mind we focused on the acoustic signals and its mapping with syllables for better output. We found an overwhelming results on performing evaluation of Speech Recognition on the Shruti Speech Corpus to validate our proposed approach and the system produces a BLEU score of 87.37. Beside, independently we performed evaluation for IPA syllable and Bengali Syllable of the system generated output with syllable by humans using F1 score and found a result of 0.9043 (90.43%) and 0.74 (74%), respectively. All evaluation sets are shown in chapter 5.

6.2 Future Scope

The work in the paper requires more investigation on the exceptions, in order to get more accuracy of the program. Changing in rules may also lead us to better result as we clearly found that the rules we used as mentioned in table 4.1 is not same for IPA representation 5.3 and Bengali representation 5.4 of Bengali Speech. We will try to implement it beyond word level. As syllable is an essential component of many speech and language processing system, it is also important for prosodic analysis of human speech. So, our future work will try to incorporate the prosodic features with syllabification.

Bibliography

R. Asliyan. *Syllable Based Speech Recognition.* Speech Technologies, Prof. Ivo Ipsic (Ed.), ISBN: 978-953-307-996-7, 2001.

Susan Bartlett, Grzegorz Kondrak, and Colin Cherry. Automatic syllabification with structured svms for letter-to-phoneme conversion. *Proceedings of ACL-08: HLT*, pages 568–576, 2008.

Susan Bartlett, Grzegorz Kondrak, and Colin Cherry. On the syllabification of phonemes. In *Proceedings of human language technologies: The 2009 annual conference of the north american chapter of the association for computational linguistics*, pages 308–316. Association for Computational Linguistics, 2009.

Brigitte Bigi and Katarzyna Klessa. Automatic syllabification of polish. In *7th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, pages 262–266, 2015.

Suniti Kumar Chatterji. *The Origin And Development Of The Bengali Language.* Rupa Co., 2002.

Asoke Kumar Datta. *Acoustics of Bangla Speech Sounds.* Springer; 1st ed. 2018 edition, 2017.

A. M. Ha. *Dhonibiggan O Bangla Dhonitatto.* 1975.

B. Hayes and A. Lahiri. *Bengali Intonational Phonology.* Natural Language Linguistic Theory, 1991.

Luca Iacoponi and Renata Savy. Sylli: Automatic phonological syllabification for italian. In *Twelfth Annual Conference of the International Speech Communication Association*, 2011.

Junko Itô. Syllable theory in prosodic phonology. 01 1986.

D. Jurafsky and James H. Martin. *Speech and Language Processing.* Pearson Education India; 2 edition, 2013.

Somdev Kar. *Syllable Structure of Bangla.* 2010.

A Majumder. *Bhasaprasanga O Bhvanivijnan.* Deys Publishing, 2014.

R. Mitkov. *The Oxford Handbook of Computational Linguistics.* Oxford University Press, 2003.

Loitongbam Gyanendro Singh, Lenin Laitonjam, and Sanasam Ranbir Singh. Automatic syllabification for manipuri language. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 349–357, 2016.

Philipp Strazny. *Encyclopedia of Linguistics(Vol.2).* Routledge, 2004.

R. Sultana and R. Palit. A survey on bengali speech-to-text recognition techniques. In *2014 9th International Forum on Strategic Technology (IFOST)*, pages 26–29, Oct 2014.

Harry van der Hulst and N Ritter. *The syllable: views and facts.* 01 1999.