DEVELOPMENT OF A BANGLA 'TEXT TO SPEECH CONVERTER'

by

MD. NAZRUL ISLAM

A dissertation submitted to
the Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology, Dhaka,
in partial fulfilment of the requirements for the degree of
Master of Science in Engineering
(Electrical and Electronic Engineering).

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
DHAKA, BANGLADESH

April 1995

DECLARATION

This is to declared that this work has been done by me under the guidance of Dr. Md. Saifur Rahman and this has not been submitted elsewhere for the award of any degree or Diploma or for publication.

Countersigned

Signature of the candidate

20/04/95
(Dr. Md. Saifur Rahman)
Supervisor

Md. Nazrul Islam 20/4/95
(Md. Nazrul Islam)

Approval

The thesis title **"Development of a Bangla text to speech converter"** accepted as satisfactory for partial fulfilment of the requirements for the degree of Master of Science in Engineering (Electrical and Electronic Engineering) of **Md. Nazrul Islam**, Roll No - 911304F, Session 1989-90.

Board of Examiners

1.	Dr. Md. Saifur Rahman Assistant Professor, Department of Electrical & Electronic Emgineering, BUET, Dhaka.	Chairman and Supervisor
2.	Dr. A.B.M. Siddique Hossain Professor and Head, Department of EEE BUET, Dhaka.	Member
3.	Dr. Joarder Kamruzzaman Assistant Professor Department of EEE BUET, Dhaka.	Member
4.	Dr. Mohammad Kaykobad Associate Professor Department of CSE BUET, Dhaka.	Member (External)

ABSTRACT

A text to speech converter is a device that accepts a text input and outputs the corresponding audible speech. In this work, an algorithm has been developed for the Bangla text to speech converter, which has been implemented using the PASCAL programming language. The Bangla text to speech converter has been developed using the method that stores speech sounds of part of the words (i.e. of speech segments). To develop the text to speech converter, several basic sound units are defined corresponding to each letter of the Bangla alphabet. These sound units are recorded from the speech sounds of a particular person using the Sound Galaxy Nx Pro 16 sound card. The pronunciation (or letter to sound) rules for Bangla words have been studied extensively as part of the current research. These text to speech rules for Bangla speech are used to identify specific words from the arbitrary text input. First, the text to speech converter algorithm reads the ASCII values from the input text file, identifies each letter of a word from its ASCII value, and then assigns sound units corresponding to the speech segments comprising a word. When identification of the sound units for a particular Bangla word is completed, all the sound units are combined together into a speech file to form the complete speech signal for that Bangla word. The file merging technique has been developed to accomplish this task. The process of identifying the words and forming their corresponding speech files is continued until a sentence is completed. Finally, the files corresponding to the words are sequentially played back to produce the audio sound of the Bangla sentence. The play back facility of the sound card has been used for this purpose. Some of the possible applications of the 'Bangla text to speech converter' algorithm are also explored.

ACKNOWLEDGEMENT

The author owes the initiation to write the dissertation on **Development of a Bangla 'Text To Speech converter'**, to the inspiring suggestions of his supervisor Dr. Md. Saifur Rahman, Assistant Professor, Department of Electrical and Electronic engineering and expresses his great admiration for his advice, valuable guidance, the keen interest taken and patience shown to him.

The author is indebted to Dr. A.B.M. Siddique Hossain, Professor and Head, Department of Electrical and Electronic Engineering , BUET, for his all out support and cooperation to complete this work.

Acknowledgements are also made to Professor M. A. Matin, Dean, Faculty of Electrical and Electronic Engineering for his admiration and enthusiasm shown during the course of the work.

The author is thankful to Dr. S. P. Majumdar, who is in charge of the Telecommunication Laboratory, for giving his permission to use the PC of the laboratory.

The author is grateful to Dr. M.A. Choudhury, Associate professor of EEE, BUET for his spontaneous help and valuable advice towards successful completion of this work.

The author is also thankful to Engr. M. Soharab Uddin Bhuiyan, Executive Engineer, Sonali Bank, Head Office, Dhaka, for inspiring the author to carry out this research.

Last but not the least, the author would like to register his thanks to his friends, specially Mr. Md. Foezur Rahman Chowdhury, and to all those people who have helped directly and indirectly during the course of this work.

CONTENTS

CHAPTER 1
INTRODUCTION

This chapter describes the definition of a text-to-speech converter, a brief literature survey in this field, the objective of the current research and the organization of this dissertation.

1.1 Introduction

A text to speech converter or synthesizer is a device that accepts a text input and outputs the corresponding audible speech [1, 2, 18, 19, 20].

Ability of man to communicate through speech sets him apart from other earthly species and is often regarded as a sign of spirituality. As speech is the most natural form of communication between humans, speech is a subject which has attracted much interest and attention over many years. The structure of speech, its production and perception mechanisms have long occupied linguists, psychologists and physiologists. Scientists and engineers have endeavoured to construct machines which can convert and recognise human speech. Current speech synthesis systems are capable of producing reasonably intelligible, though not natural-sounding, speech. The performance of speech conversion system is improving slowly and steadily with time, and speech systems are now being used in certain commercial applications [2, 4, 5, 15, 21].

There are three main areas in speech technology. These are:

(i) speech conversion,
(ii) speech recognition, and
(iii) speech coding.

The ultimate goal of the speech conversion is to develop a machine which can accept as input a piece of text and convert it to natural sounding speech, which would be as intelligible and as natural sounding as if spoken by a person. This text to speech converter can be used as speech output from a computer, reading machine for the blind and public message system [8, 12]. The text to speech converter encodes a set of sounds from which a speech sound corresponding to a word could be constructed.

1.2 Brief literature survey and present state of art of the project

Research on text-to speech synthesis began in 1947 [1], and a text to speech synthesizer was first developed for the English Language in 1960 [1, 2, 4, 10]. Following English, text-to-speech synthesizers of Telegu, and Hindi Languages have also been developed in our sub-continent [21]. However, Bangla text-to-speech converter has not yet been developed anywhere. Although a group of researchers of the applied physics department of the University of Rajshahi are trying to find the formant frequencies of Bangla phonemes [6], they are doing this in a conventional way. However, the conventional way of doing this requires the development of complex software and hardware to make a Bangla text-to-speech converter. At the moment, they are confined to the analysis of Bangla phonemes, and have not yet developed any Bangla text-to-speech converter. The text-to-speech converter algorithm and the corresponding software developed and described in this dissertation is the first of its kind.

This does not follow the traditional phoneme method, and this has been implemented using a commercially available sound card.

1.3 Objective of the research

The main objective of the current research is to develop a software for Bangla 'text to speech converter' algorithm.

1.4 Organization of the dissertation

This dissertation has been organized in the following way. Chapter 2 describes the introductory information about text-to-speech synthesis. Chapter 3 defines the basic sound units for Bangla speech and describes the recording and storage procedure of the sound units. Chapter 4 enumerates the rules of assigning sound units from the textual representation of a Bangla text input from the keyboard of a computer. Chapter 5 describes the application of the developed Bangla text-to-speech converter. Chapter 6 concludes the results of the research and comments on the topics of further research in this field. References and appendices are included at the end of this dissertation.

CHAPTER 2

TEXT TO SPEECH CONVERSION

This chapter describes the introductory information about text to speech conversion, methods used for text to speech conversion, and the method that can be used to develop a Bangla 'text-to-speech converter'.

2.1 Introduction

Language conveys meaning by stringing together discrete symbolic units at several concurrent levels. Sequences of sounds form words, and when words are assembled together they form sentences. The combination of these units is governed at each level by a set of principles. The science of linguistic provides rigorous methods for identifying these principles and represents them in a formal manner. The multi-level linguistic message supports a number of possible strategies for their storage and retrieval regardless of which voice output technology is employed in a particular system. There are several options to analyze a message, these are [1, 15]:

(i) Analyze the message in its entirety and reference it as a single block of data,

(ii) The message could also be broken apart into two or more phrases,

(iii) It can also be broken into individual words which would then be called up one after another to speak the full sentence, and

(iv) The most radial alternative would be to encode a set of sounds from which words or parts of words may be patched together as needed.

The last option to a vocabulary of sufficient size would eventually yield a complete inventory control from which any word could be constructed. Such an inventory when combined with algorithms for selecting sounds and stringing them together on the basis of ordinary spelling, will result in a system capable of converting words or phrases from arbitrary text input. This will be known as a 'text-to-speech converter'.

The speech which can be produced by analysing the real speech is refered to as synthetic speech. For the generation of synthetic speech a set of control parameters and a model are required to produce a particular utterance. Text to speech synthesizer is a device which takes text as an input and generate the synthetic speech by controlling the parameters and a model. If the model and parameters are sufficiently accurate then the production of intelligible synthetic speech would be possible. The basic goal in text to speech synthesis is to convert unrestricted text input into natural sounding speech [2, 4, 18].

2.2 Phoneme analysis for text-to-speech synthesis

Phoneme analysis of a particular language is important for synthetic voice generation. Phoneme may be viewed as a sequence of segmental units at a linguistic level [2, 11, 12]. These phonemes are abstract linguistic units and may not be directly observed in the speech signal. Phonetics is concerned with sound of human speech. Phonetic analysis means to derive the phonetic structure of an utterance directly from the speech signal. The sequence of phonemes corresponds to a sequence of articulatory gestures, which have certain well defined acoustic equivalent. However, the acoustic phonetic mapping is exceedingly complex. The most popular way to perform a phonetic analysis is to attempt to segment the speech signals into phonemic like units, and to assign an appropriate label to each unit. Phonetics is the study and science of speech sound, their production and sign used to represent them [1, 2, 15].

Phoneme synthesis is the way of obtaining synthesizer control parameters to generate them from a phonetic transcription of the utterance. In other words, the utterance to be synthesised, represented by a string of phonemes, is input to a computer program, which outputs the sequence of synthesiser control parameters. The computer program is generally based on rules for converting phonetic information to acoustic information. For this reason the synthesis process is often referred to as 'speech synthesis by rule'.

Speech, viewed as a sequence of smeared articulatory gestures, still does not give the complete picture. Superimposed on the basic sequence of gestures are variations in intonation or pitch (the fundamental frequency of variation of the vocal cords), rhythm or timing and intensity or loudness. These variations in an utterance are collectively known as prosody. A speech synthesis system must also have some method of assigning prosodic information (intonation and stress).

Conversion of arbitrary text input to voice output actually involves two separate tasks. The first consists of accepting a sequence of characters, identifying the phonetic components of the required message and extracting information about its syntactic structure. The output of this stage, will then be an intermediate string of symbols representing sound units and boundaries between words, phrases and sentences. The second part of the process is to match the symbols up with items stored in the phonetic inventory, link them together and send the resulting coded waveform to a voice output device [1, 14, 16, 17].

Text-to-speech synthesis requires two major tasks to be performed in addition to that required by phoneme synthesis. These are: (i) to translate the text into phonetics, and then (ii) the prosody of the speech must be determined directly from the textual representation. The first task, that is, text interpretation can be done by choosing appropriate phonetic symbols.

By using the phonetic symbol natural sound can be produced. However, convention for phonetic symbol has failed to keep pace with the rate of phonetic change.

2.3 The other traditional methods used for text to speech conversion

Text to speech synthesis of a language may adopt several methods to segment storage strategies. They are:

(i) sentence storage,
(ii) word and phrase storage, and
(iii) storing parts of words.

In the sentence storage system the voice of a live speaker is recorded with the desired inflection and personality, and then processed to derive a matrix of parameters from which the original utterance may be reconstructed. This system is not flexible because it requires to store repeated words or phrases.

In words or phrase storage system, it stores the similar words or phrases just for one time which reduces the number of words to be stored to a greater extent. Besides these advantages the word and phrase storage system has contextual variation problem, variation in intonation and duration which provide important clues about syntactic structure. Also slapping together words and phrases from different sentences recorded at different times can also result in choppy 'sing song' effect since the pitch (the fundamental frequency of variation on the vocal cords) and duration (existing time) of words will not necessarily match.

Contextual variation problem can be solved by storing parts of words, but in this system if correct variant of a prefix or suffix has been selected there remains the problem of matching its intonation (the rise and fall of the pitch of the voice in speaking, which is an element of meaning in language) to that of the words to which it is attached.

The ultimate way of doing text-to-speech conversion would be to store every word in the language together with its phonetic equivalent and pronunciation could then be effected using a simple dictionary look-up table. The vast number of words in the language prohibit this approach because of the enormous size of the resulting dictionary.

Practical systems tend to use pronunciation rules or letter to sound rules, which specify the phonetic equivalent of single letters or group of letters taking into account the context. This process is often assisted by the use of small 'exceptions' dictionary that stores those words and pronunciations which are constantly recurring and on which the letter to sound rules would fail. Thus, an input is first checked against the entries in the 'exceptions' dictionary and if it is not found then the letter to sound rules are applied.

In English language a very comprehensive system for English pronunciation has been developed by Allen [2, 4] which consists of morpho dictionary, supplemented by letter-to-sound rules. A morpho is either a prefix, a root or suffix. Estimates of the number of morpho varies between 10000 to 30000. The number of morpho used by Allen is 10000. The morpho dictionary contains for each morpho its pronunciation, spelling, parts of speech etc. When a word is input to the system a check is made to see if it exists in the dictionary. If not, it is input to an affix stripping algorithm which attempts to decompose the words into a prefix, a root and a suffix. If an affix can be stripped away, the reduced word is checked against the morpho dictionary. If it still cannot be found in the dictionary, it is presented again to the affix-stripping algorithm. This process is repeated until the reduced word can be found in the morpho dictionary or until the affix stripping algorithm fails. If the morfic decomposition fails, the letter to sound rules are invoked in order to obtain the pronunciation.

2.4 A method of developing Bangla text-to-speech converter

A Bangla text to speech synthesizer can be developed by using the phoneme method. However, to develop a text-to-speech converter using this method requires phoneme synthesis, translation of the text into phonetic symbols, and determination of the prosody from the text. This is a very complex task. It has already been described in section 2.2 that text interpretation can be done by choosing phonetic symbols. However, convention for phonetic symbols has failed to keep pace with the rate of phonetic change. Therefore, we have discarded this method to develop our text-to-speech converter. A text-to-speech converter can also be developed by using any of the methods already described in section 2.3. However, the third method, which stores speech sounds of part of the words (i.e. of syllables), is more appropriate for the development of a Bangla text to speech converter. This method is more appropriate because there is a fixed pronunciation rule for a specific arrangement of Bangla letters in a word. On the contrary, there is no such definite rule for the English language. The particular arrangement of English letters has a different pronunciation rule according to the position of the letter in the word, and therefore, are pronounced differently according to the use of words. A particular arrangement of Bangla letters obey the same pronunciation rule despite its position in the word.

This method also requires to store a small number of sound units compared to word and phrase, or sentence storage system. To develop a Bangla text to speech converter, several basic sound units are defined corresponding to each letter of the Bangla alphabet. These are termed as basic sound units for Bangla speech. The proper combination of the sound units will form a Bangla word. The basic sound units have to be recorded for a specific male or female speaker. In this method, the sound unit will be detected from the text input and will be joined together as per requirement to form a complete word. The identification process and formation of words are repeated until the sentence is completed. Once the sentence is completed, the whole sentence will be played back in the order of the words in the sentence. A block diagram of the proposed text-to-speech converter is shown in figure 2.1.

It was indicated in section 2.2 that although speech may be viewed simply as a sequence of segmental units (phonemes) at a linguistic level, the manifestation of these units at the acoustic level is much more complex. The mechanism of co-articulation, which results in sounds being smeared together, is one of the major influences on speech naturalness. Co-articulation results in complex movements of the formant frequency and amplitude values between one sound and the next. There are also many acoustic variations of each phoneme, depending on its context (place) in an utterance [2, 15].

To avoid such complexity, we have used the third method, the method of basic sound units instead of the phoneme

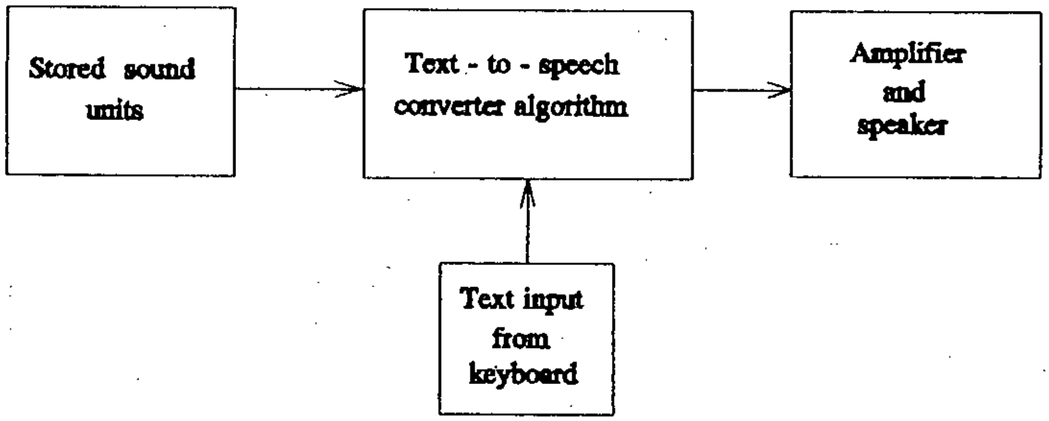


Figure 2.1: Block diagram of the proposed text-to-speech converter.

method. Our method has less problem of contextual variation (the problem of 'sing song' effect described in section 2.3) as our system stores parts of words i.e. speech segments. However, in this system, there remains the problem of matching its intonation (the rise and fall of the pitch of the voice in speaking, which is an element of meaning in language) to that of the words to which it is attached.

2.5 Necessity of using a sound card

In addition to developing the necessary software, a text-to-speech converter requires the recording and storage facility of the sound units to be used, and also the playback facility to play the sound units back sequentially. Except the Apple Macintosh Computer, the IBM PC or its compatibles do not have the built-in recording facility to record audio signals. Moreover, the speaker provided in a PC is not configured to play arbitrary audio signal back. It can only produce beeps of a single frequency at a time from the on-board clock of the PC. However, commercially available sound cards may be inserted in an empty slot of the mother board of a PC, and can be used for the recording and playback of the audio signals. As a text-to-speech converter is mostly PC-based, a sound card is needed to record the sound units, store them in the hard disk surface of the PC, and play them back later.

To develop a Bangla text to speech synthesizer, a sound card is needed to record the basic sound units for a specific male, or a female, or a child speaker (with desired inflection and personality). This is also required to play the sound units back one after another when the specific sound units are detected by the text-to-speech converter algorithm in a full sentence of the Bangla text.

In the next chapter we will define the various Bangla Sound units that will be used to form Bangla words, and their recording and storage procedure using a sound card.

CHAPTER 3

BASIC SOUND UNITS FOR BANGLA SPEECH AND THEIR STORAGE PROCEDURE

In this chapter a description is given on the basic sound units for Bangla speech and their recording and storage procedure using a sound card.

3.1 Assignment of sound units

Like in English alphabet, there are two types of letters in Bangla alphabet. These are: (1) Vowels and (2) Consonants. There are 11 vowels and 39 consonants in Bangla alphabet.

Bangla has a very rich vocabulary. The pronunciation of each letter of Bangla alphabet (vowels and consonants) as one reads them produces unique sounds. However, as we shall see while we proceed further in our discussion, the sounds produced by the utterances of 11 vowels and 39 consonants are not sufficient to pronounce all the Bangla words.

In Bangla language, vowels may sometime occur as a first letter of a word. The vowels এ (pronounced as ´A´) and ঐ (pronounced as ´Oi´) are sometimes used singly as a Bangla word. There are instances when two or more consonants may form a Bangla word (examples বক, কলম etc.). However, in most of the cases, vowels and consonants combine together to form a Bangla word.

A vowel may sometimes occur as a first letter, and it may also appear at the middle or at the end of a word. However, vowels usually appear before and/or after consonants to form a complete word. Here they are used as auxiliary letters and have different symbols (ে ('A'-kar), া ('Aa'-kar) etc.) than the usual symbols of the vowels. We may call these symbols 'vowel auxiliaries'. A list of Bangla vowels, consonants, vowel auxiliaries, and other necessary symbols used to form a Bangla word is given in Appendix A.

Like a vowel, a consonant may take its position at the beginning, at the middle or at the end of a Bangla word. There are some peculiar consonants which never appear as first letters of words. They are ঙ (Uma), ঞ (Yian), ং (Anuswar), ঁ (Chandrabindu) and ঃ (Bisarga). The first two of them sometimes combine with the vowel auxiliaries.

Occasionally, a consonant appearing in a word may not use any vowel auxiliary. However, in most of the cases a consonant has a vowel auxiliary before and/or after it. The pronunciation of the combination of a consonant and vowel auxiliary depends on the particular consonant and the particular vowel auxiliary being used. If we take the Bangla word কে as an example, where the vowel auxiliary ('A'-kar) is placed before ক (Ka). This combination is pronounced as the English word 'K'. However, when ক is followed by an া ('Aa'-kar), the pronunciation becomes as কা 'Kaa'. Similarly, an ি ('E'-kar) followed by ক (ka) will produce the sound like কি 'Kii' etc. Therefore, the sounds corresponding to the utterances of only 11 vowels and 39 consonants alone is not capable of pronouncing the variety of Bangla words.

There is another instance of varied pronunciation, when a vowel or a consonant appears at the middle or end of a word without any vowel auxiliary. Its pronunciation depends on the ending sound of the utterance of the letter immediately preceding it. The following example will clear the above statement.

Let us take the Bangla word কাক (Kaak) as an example. Here the first sound unit is কা ('Kaa') and the second sound unit, if pronounced in isolation, is ক ('Kau'). If these two sound units are now combined together, the word will be pronounced as 'Kaakau'. However, as we know, the actual pronunciation of this word is 'kaak', that is, the utterances কা ('Kaa') + আক ('ȧak') actually produce the sound. If we look at the combination কেক ('Kek'), here as well, the utterances কে ('Kay') + এক্ ('ėk') actually produce the sound. Therefore, if a particular consonant (without any vowel auxiliary with it) appears at the middle or end of a word, its pronunciation depends on the ending sound of the utterance of the letter preceding it. The pronunciation may take roughly one of 5 forms depending on the ending sound corresponding to an 'Aa'-kar(া), 'A'-kar(ে), 'O'-kar(ো), 'E'-kar(ি), or 'U'-kar(ু). The corresponding sounds of the consonant ক (Ka) appearing at the end are: আক ('ȧak'), এক ('ėk'), অক ('ȧuk'), ইক ('i̇k'), উক ('ȯok') etc.. A dot on the letters আ, এ, অ, ই and উ means that they will be pronounced much shorter than they are pronounced normally. Therefore, to identify the ending sound of the previous utterance, a provision of an 'end code' is required to be incorporated in the 'text-to-speech' converter algorithm. If the ending sound of the previous utterance is like 'Au' (অ), ec=0. If the ending sounds of the previous utterances are like 'A' (ে), 'Aa'(া), 'Au' (ো), 'E' (ি or ী), and 'U' (ু or ূ), then the corresponding end codes are ec=1, 2, 3, 4, and 5 respectively. With the help of these end codes the correct sound of a word can be produced.

In Bangla there are some joined-letters as well. When a consonant combines with another consonant or a vowel, in Bangla, this combined letter is known as a joined-letter or Jukta-Aukkhar. These letters are also used widely in forming Bangla words, and they create more complexity in pronouncing a Bangla word. Depending on the position of a joined-letter in a word, the pronunciation of the joined-letter differs. For example, let us take two Bangla words স্তব্ধ ('stabdha') and অস্ত ('austa'). Here, we can see that to pronounce স্তব্ধ we require স্ত 'sta' as the first sound unit corresponding to the joined-letter স্ত , and in অস্ত ('au'+ 'ásta'), we require অস্ত 'ásta' as the second sound unit corresponding to the same joined-letter স্ত . Therefore, if a joined-letter appears at the beginning of a word an additional sound unit is required to pronounce the joined-letter for the consonant which forms the first sound unit of the joined-letter.

To develop a Bangla text-to-speech converter, first, we have considered the utterance of each vowel and consonant (excepting a few peculiar consonants mentioned earlier in this section) as a separate sound unit.

Since the sounds produced by the utterances of 11 vowels and 39 consonants are not sufficient to pronounce the words of the Bangla language, we have also considered several possible sound units for each consonant, corresponding to the auxiliary letters that can be used along with the consonant.

We have also considered six additional sound units for each consonant when appearing at the middle or end of a word, without any vowel auxiliary. The corresponding sound units are assigned only after a preceding end code is checked against the consonant.

There is also another instance, where, ঁ ('chandra bindu') is used with a vowel or a consonant (with or without a vowel auxiliary), the sound corresponding to the combination becomes a nasal sound. To incorporate this in the algorithm, we have also assigned sound units corresponding to some valid nasal sounds for some vowels and consonants (with or without using a vowel auxiliary).

Taking all these factors into account, we have chosen approximately 1500 basic sound units for the Bangla text-to-speech converter, to produce speech from arbitrary text input from the keyboard of a PC. These sound units may be termed 'basic sound units for Bangla speech'. A list of these sound units are included in Appendix B. A particular combination of these sound units will form the pronunciation of a Bangla word. All the words of Bangla language can be generated from these sound units.

There are some other symbols in Bangla, which are র্ ('ref'), ্ ('hasanta'), ্য ('ja-fala'), ৃ ('Ri-fala'), ৌ ('Ou'-kar) etc. However, the sound units that are described earlier can be used to form the correct pronunciation of the words using these symbols. The proposed Bangla text-to-speech converter algorithm will take care of all possible combination of sound units corresponding to the special symbols mentioned above.

3.2 Recording of the basic sound units by using a sound card

The sound units described in section 3.1 has to be recorded for a particular person, and stored in a particular directory of the hard disk of a PC, assigning specific filenames to each of the sound units. A list showing the names of the sound units and the corresponding file names is included in Appendix B.

Suppose that the basic sound units of Bangla speech are stored in a particular directory of the hard disk of a

PC. The next idea is that, if a text is input from the keyboard of a PC, the Bangla text-to-speech converter algorithm (which will be described in chapter 4), will identify which sound units are to be combined together to form the sound corresponding to a word. These sound files will then be merged together to form a play file with a specific name to it. When a sentence will be transformed into such play files they will be played back sequentially and the speaker connected with the sound card will produce the corresponding sounds.

A sound card is necessary to record the sound units, and play the reconstructed sound files back for listening. There are a variety of sound cards commercially available in the market. Some of these are:

(1) Sound Blaster Card

(2) Laser Wave supra 16 Sound Card

(3) Sound Galaxy NX PRO Sound Card (available in both 8-bits and 16-bits).

Of the above three, we have chosen the Sound Galaxy NXPRO 16 sound card. This is a 16-bit sound card and is the most popular one to the users. The recording of the sound units should be done for a particular male or a female speaker with desired inflection and personality. To do that we have chosen a particular male speaker who has a quality voice to pronounce the Bangla words. We shall now briefly describe the recording and editing procedure of the basic sound units of Bangla speech.

For recording the basic sound units for the Bangla text-to-speech converter, we have selected specific sentences. Then the selected person was asked to pronounce the words of the sentence before the microphone of the sound card. The speech sounds of the sentence were simultaneously recorded. The recording was performed using the recording and editing facility of the Sound Galaxy Software in the windows environment [22]. From a particular word thus recorded, the different speech segments have been separated, and specific file names were assigned to store them in the hard disk of a PC. For example, to record the sound units কা ('Kaa') and আঁক ('Ȧak'), we have first recorded the Bangla word কাক ('Kaak') and then separated it to obtain the sound units কা ('Kaa') and আঁক ('Aak') . Similarly, from the Bangla word অহংকার ('Auhangkar'), we may record five sound units অ ('Au'), হ ('Hau'), অং ('Ȧung'), কা ('Kaa'), and আঁর ('Aar'). Using a similar procedure all the required sound units have been recorded. The sound units were not recorded from the speech segments uttered in isolation, because those will not produce natural sound when merged together to form a Bangla word.

In the next chapter we shall describe how the pronunciation information of the Bangla words are derived from its textual representation.

CHAPTER 4

FORMATION OF BANGLA SPEECH FROM THE TEXT INPUT

In the previous chapter specific file names have been assigned to all the possible sound units required to produce Bangla speech, and the recording and storage procedure of those sound units have been described for a specific male (or female) speaker. In this chapter we are going to describe the formation of Bangla speech from its textual representation.

4.1 Introduction

It has already been described in sections 2.3 and 2.4 that a text-to- speech converter requires the texts to be input to it, then the conversion algorithm will operate, and finally produce audible speech sounds corresponding to the text input (refer to figure 2.1). Thus, conversion of arbitrary text input to voice output actually involves three separate tasks:

(i) The first one consists of accepting a sequence of characters.

(ii) The second task is to detect the characters and identify the sound units required to form the message, and

(iii) Finally, merge the sound units together to form a message file (play file) and play these files sequentially to a listening device (a speaker) to produce the audible speech sound.

The first step of text-to-speech conversion is to select a Bangla word processor to enter the Bangla texts, which are to be converted to audible speech sounds. Many Bangla word processors like Bijoy, Bashundhara, Shahid Lipi etc., have been developed and are being used to enter Bangla texts from the Keyboard of a PC.

To incorporate the text-to-speech converter in a Bangla word processor, it is required to know the formet of source file for a particular Bangla word processor. However, the copyright owners of the Bangla word processors mentioned above would not let their secret out to somebody else. Therefore, we had to select a particular Bangla word processor that is currently being developed in the Department of Electrical and Electronic Engineering, B.U.E.T., to be named as 'Amar Bangla'.

We know that in a word processor, when a key is pressed, a particular letter is displayed on the monitor of a PC. However, when it is stored as a text file, the corresponding ASCII (American Standard Code for Information Interchange) codes are actually stored in the file.

To initiate the process of text-to-speech conversion, it is required to know the ASCII value for each Bangla character, vowel auxiliary and symbol. The Keyboard layout and the ASCII value used for each Bangla letter, vowel auxiliary and symbol for the particular Bangla word processor to be used are given in Appendix C.

As the proposed Bangla word processor 'Amar Bangla' has not been completely developed yet, we shall assume that ASCII codes corresponding to a particular text input are already stored in a text file. Considering this text file as the input to the text-to-speech converter, corresponding audible speech sounds will be produced.

4.2 Syntactic structure of Bangla language and the Bangla 'text-to-speech converter' Algorithm

The first step of the text-to-speech converter algorithm is to read the ASCII values from the text file and identify each letter of a word from its ASCII value, and then assign particular sound units corresponding to the speech segments comprising a word. When identification of the sound units for a particular Bangla word is completed, all the sound units have to be combined together into a speech file to form the complete speech signal for that Bangla word. This process of identifying the words and forming their corresponding speech files is continued until a sentence is completed. Finally, the speech files corresponding to the words are sequentially played back to produce the audio sound of the Bangla sentence.

The process of identifying the sound units from Bangla text input is a complex task, because in Bangla each consonant can combine with any one of the vowel auxiliaries ৈ ('Oi'-kar), া ('Aa'-kar), ('A'-kar), ু ('U'-kar), ে ('E'-kar) and other symbols ৃ ('Ri'-fala), র্ ('Ref'), ্র ('Ra'-fala), ্য ('Ja'-fala) etc.. A consonant can also combine with any consonant, and in some cases a consonant may be used as a single character. On the other hand, vowels are usually used as a single character in forming a Bangla word. Due consideration has been given to all these cases to develop the proposed algorithm for a Bangla text-to-speech converter.

Now, the process of translating the Bangla text into its speech equivalent is described with the help of the flow diagram shown in figure 4.1. The algorithm has been implemented in PASCAL programming language. As the size of the program is very large, the algorithm cannot be implemented using a single program. Therefore, it has

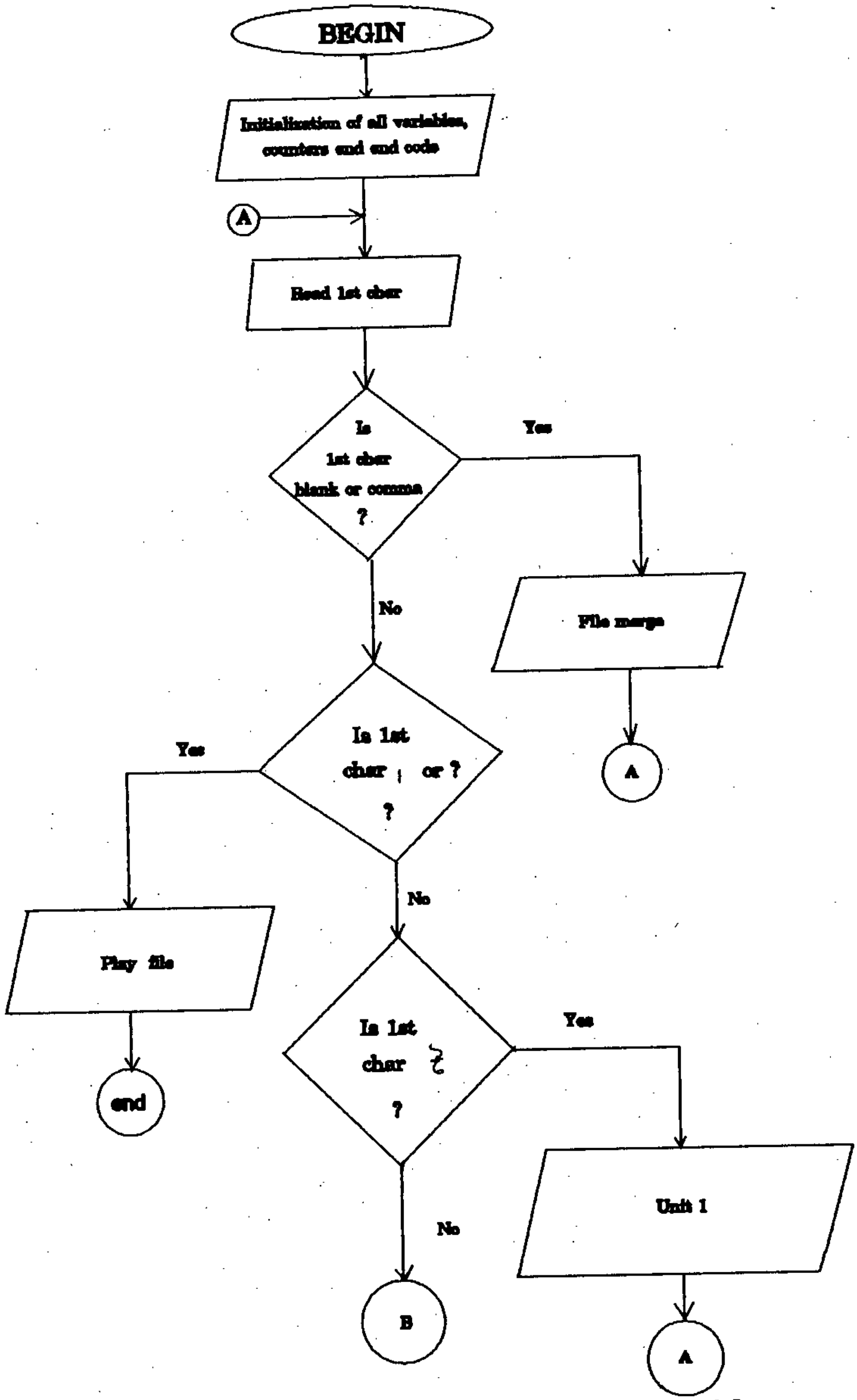


Fig 4.1: Flow diagram of the main algorithm

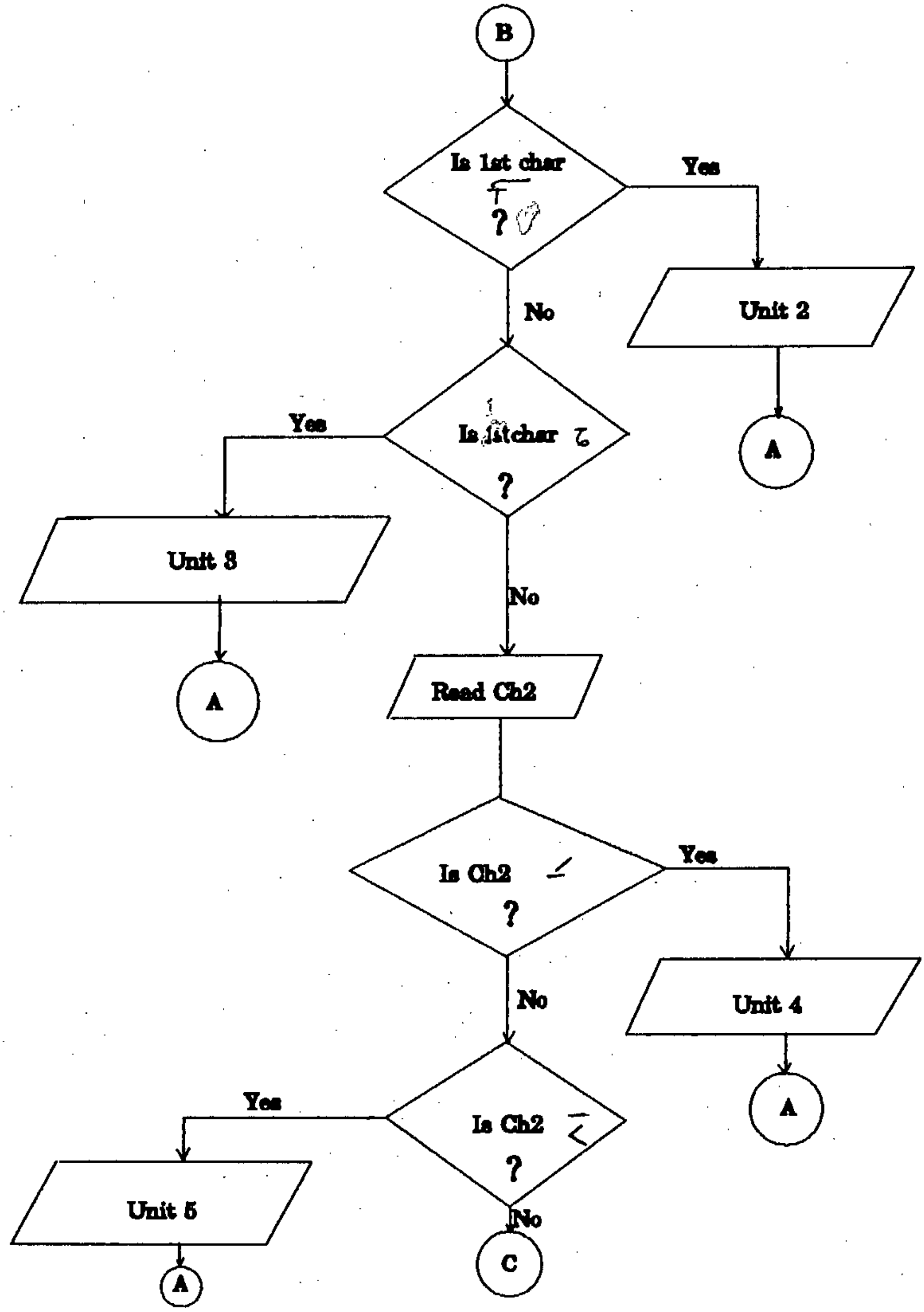


Fig 4.1: Flow diagram of the main algorithm (continued)

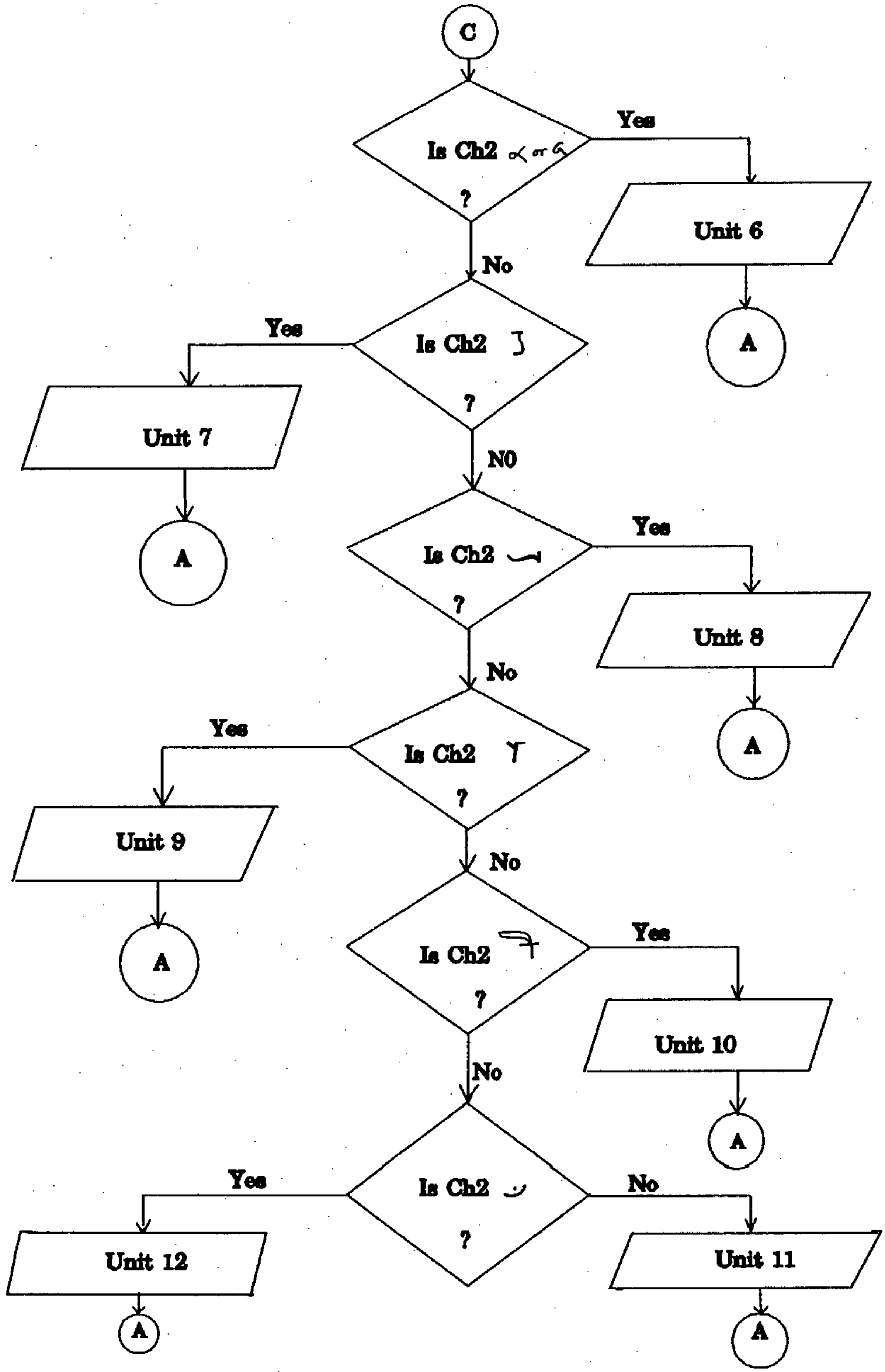


Fig 4.1: Flow diagram of the main algorithm (continued)

been subdivided into 13 units (units 0 to 12), and these units are used by the main program as and when required. The main algorithm is now briefly described below.

The main program first initializes all the character variables, counters and the variable 'ec' for the end code.

Next, it reads the first character in the character variable 'ch1' from the Bangla text file and it checks whether it is a blank ( ) or a comma (,). If the answer is 'yes', it means that an end of a word has been reached. If so, it goes to merge the files corresponding to the sound units already assigned in the earlier stages, then it initializes the variables, if necessary, and returns to point (A) of the main program.

If the answer is 'no' it checks whether the first character is a full stop (।) or a question mark (?). If the answer is 'yes', it means that an end of a sentence has been reached. If so, the program plays back the speech files corresponding to the words of a sentence in a sequential manner.

However, if the answer is 'no' it goes to check whether the first character read is an 'oi'-kar (ৈ). If the answer is 'yes' unit 1 is called on to take necessary steps to assign the sound units, initialize the variables and return to point (A) of the main program.

If the answer is 'no' it checks whether the first character read is an 'E'-kar(ে). If the answer is 'yes', it calls unit 2 to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', it goes to check whether the first character read is an 'A'-kar (া). If the answer is 'yes', unit 3 is called on to take the necessary steps to assign the sound units, initialize the variables if necessary, and return to point (A) of the main program.

If the answer is 'no', it reads the second character in the character variable 'ch2' from the text file, and goes to check whether it is a Ref (র্). If the answer is 'yes', it calls unit 4 to take necessary steps to assign the sound units, initialize the variables and return to point (A) of the main program.

If the answer is 'no', it goes to check whether the second character read from the file is a 'Ri'-fala (ৃ). If the answer is 'yes', unit 5 is called on to take [illegible] necessary steps to assign the required sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', it checks whether the second character read is an 'U'-kar(ু) or Dirgha 'U'-kar (ূ). If the answer is 'yes', unit 6 is called on to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', it goes to check whether the second character read is a Ja-fala (্য). If the answer is 'yes', it calls unit 7 to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', it checks whether the second character read is a Ra-fala (্র). If the answer is 'yes', unit 8 is called on to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no',it goes to check whether the second character read is an 'Aa'-kar (া). If the answer is 'yes', it calls unit 9 to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', it checks whether the second character read is a Dirgha 'E'-kar (ী). If the answer is 'yes', unit 10 is called on to take necessary steps to assign the sound units, initialize the variables if necessary, and return to point (A) of the main program.

If the answer is 'no', it checks whether the second character read is a chandra bindu (ঁ). If the answer is 'yes', it calls unit 12 to take necessary steps to assign the sound units, initialize the variables, if necessary, and return to point (A) of the main program.

If the answer is 'no', unit 11 is called on to take necessary steps to assign the sound units, initialize the variables if necessary, and return to point (A) of the main program.

The flow diagrams of the units from unit 0 to 12 are shown in figures 4.2 to 4.14. Unit 0 defines the global variables and also initializes them. These flow diagrams can be easily understood following the same procedure as used to describe the main algorithm (as shown in figure 4.1) in the preceding paragraphs of this section. Unit 11 is a special unit. This unit assigns all the sound units corresponding to joined-letters, and other exceptions.

One important aspect is worth mentioning here, which is the use of an "end code" to form the correct pronunciation of a word when a single vowel or a single consonant appears at the end of the word (without any vowel auxiliary). This has already been discussed with a

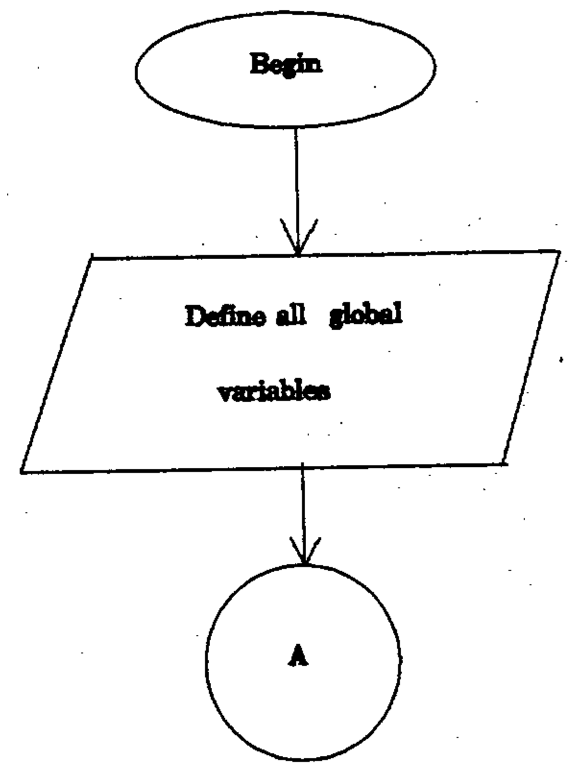


Fig 4.2: Flow diagram for unit 0

The blank box ◘ represents a consonant listed in appendix A (except

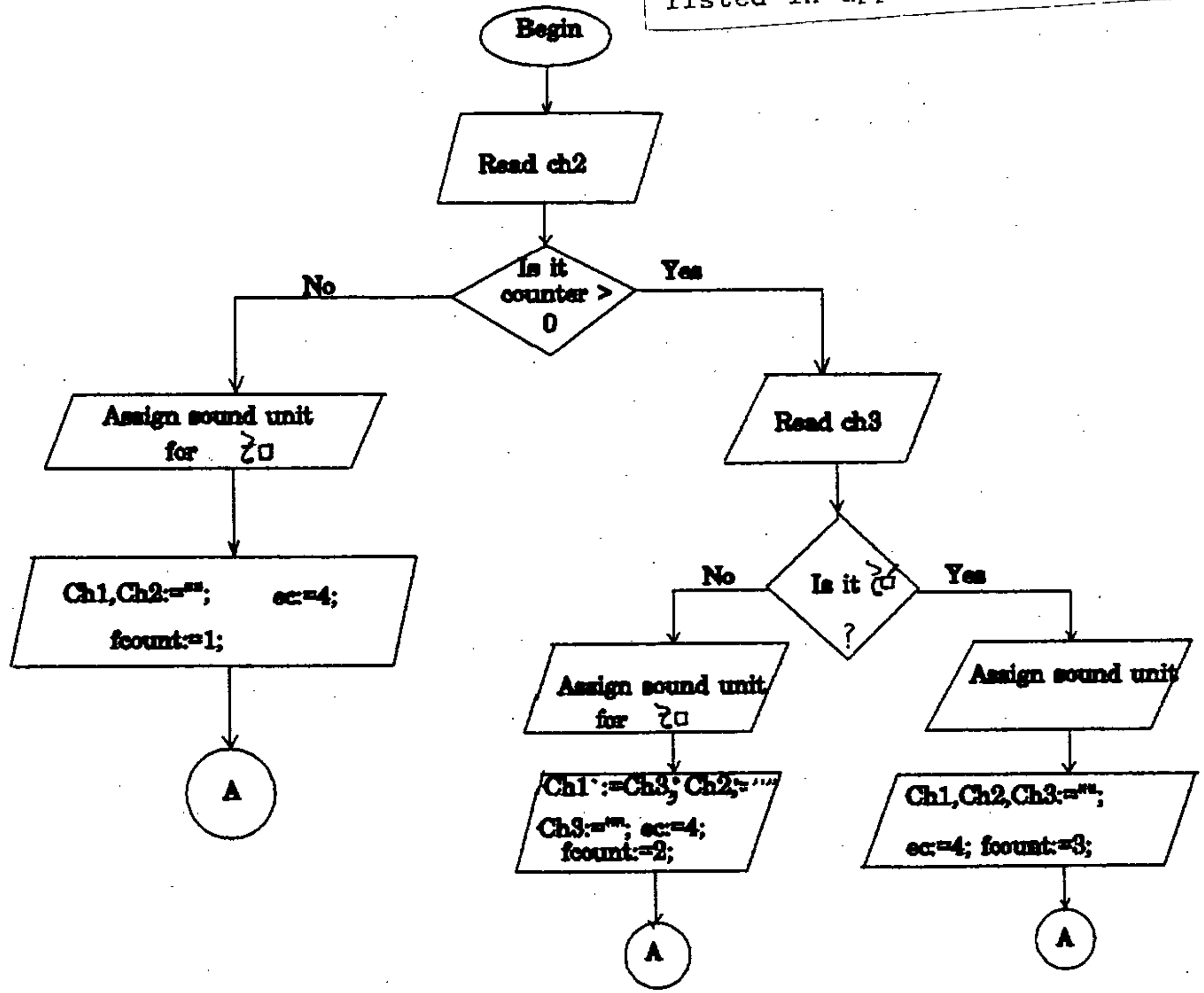


Fig 4.3 flow diagram for unit 1

The blank box ◻ represents a consonant listed in appendix A (except ⌄,⁝ and ◡)

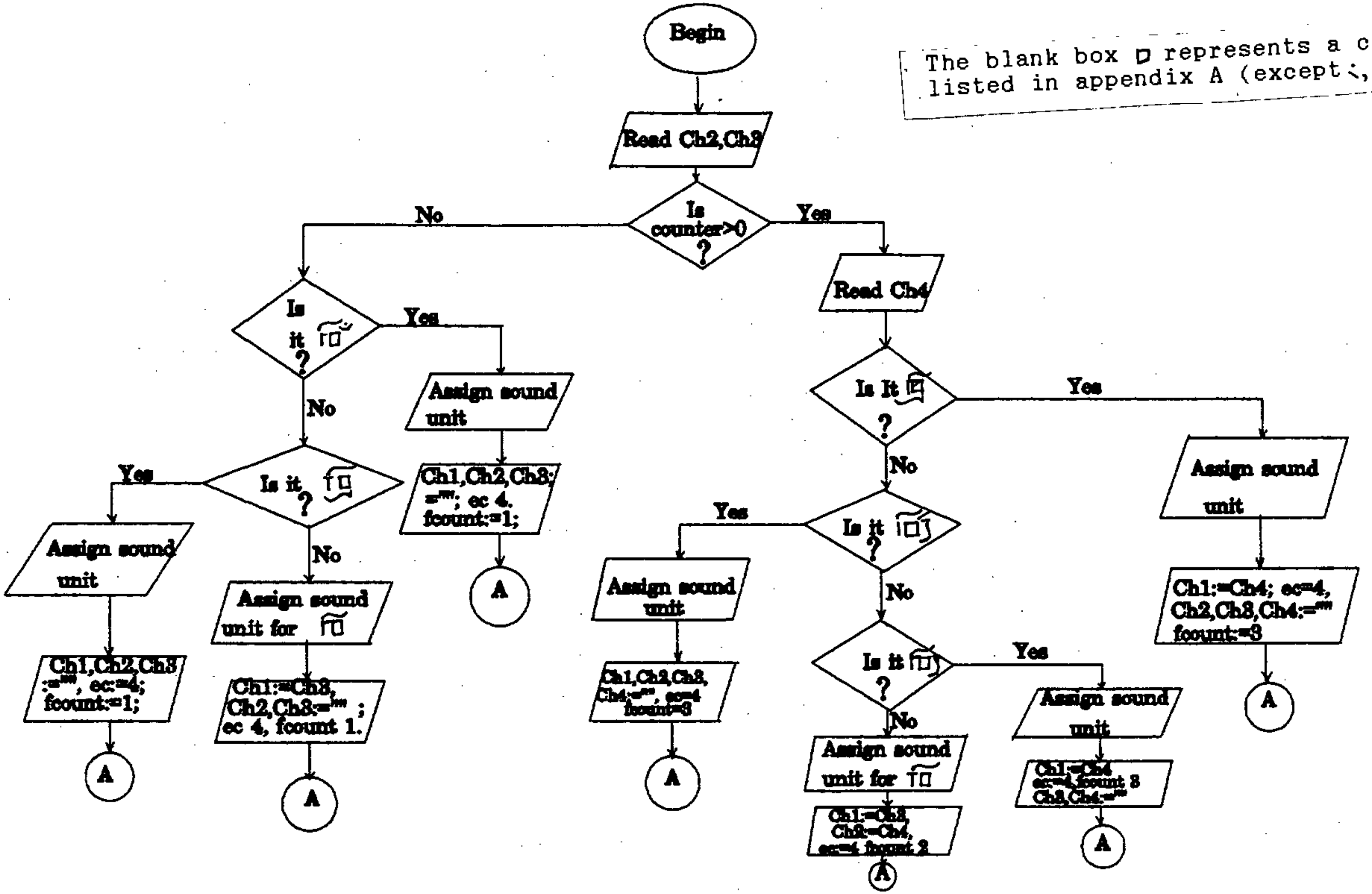


Fig 4.4: Flow diagram for unit 2

The blank box □ represents a consonant listed in appendix A (except ... and ...)

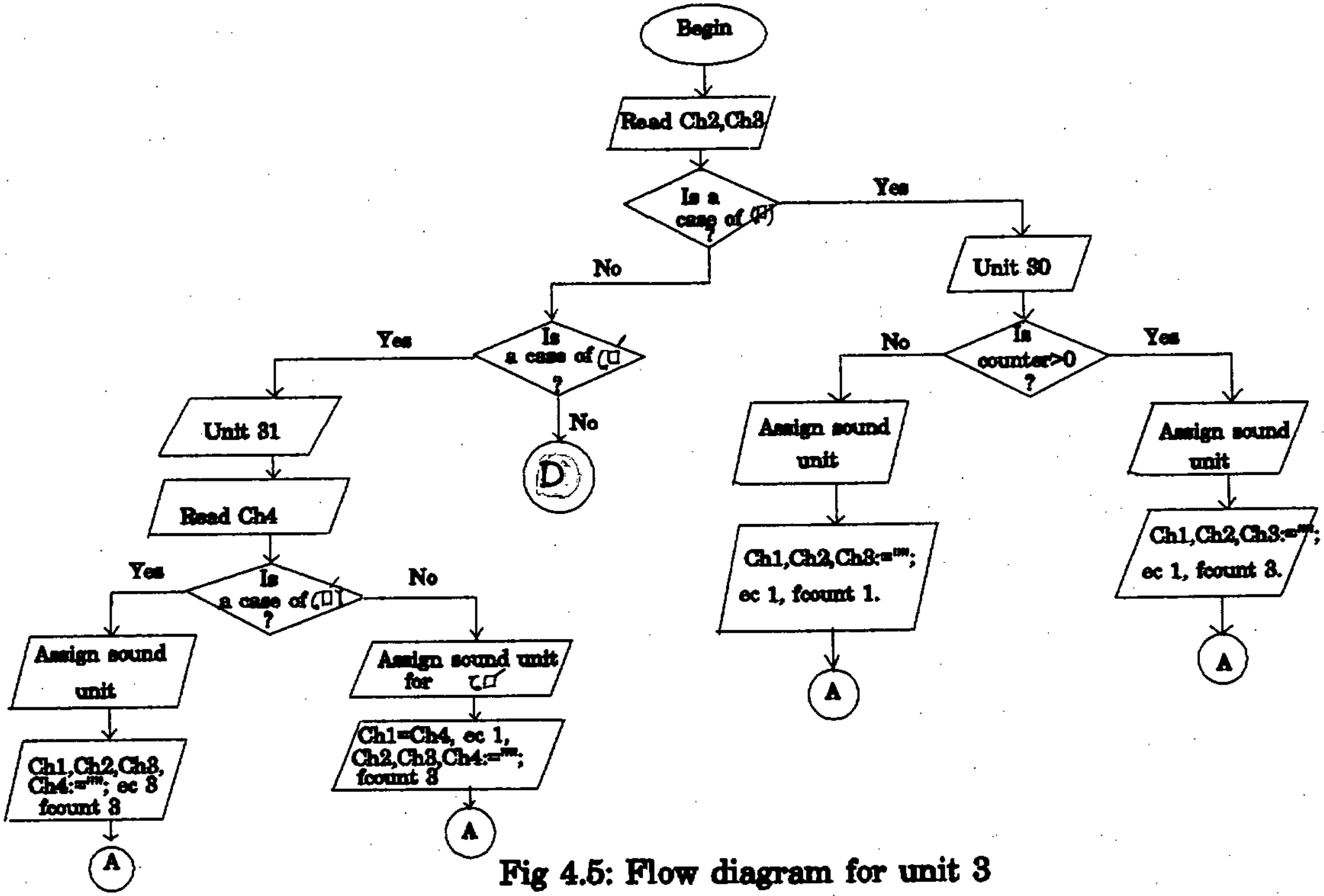


Fig 4.5: Flow diagram for unit 3

The blank box □ represents a consonant listed in appendix A (except ⌐,⌐ and ⌐)

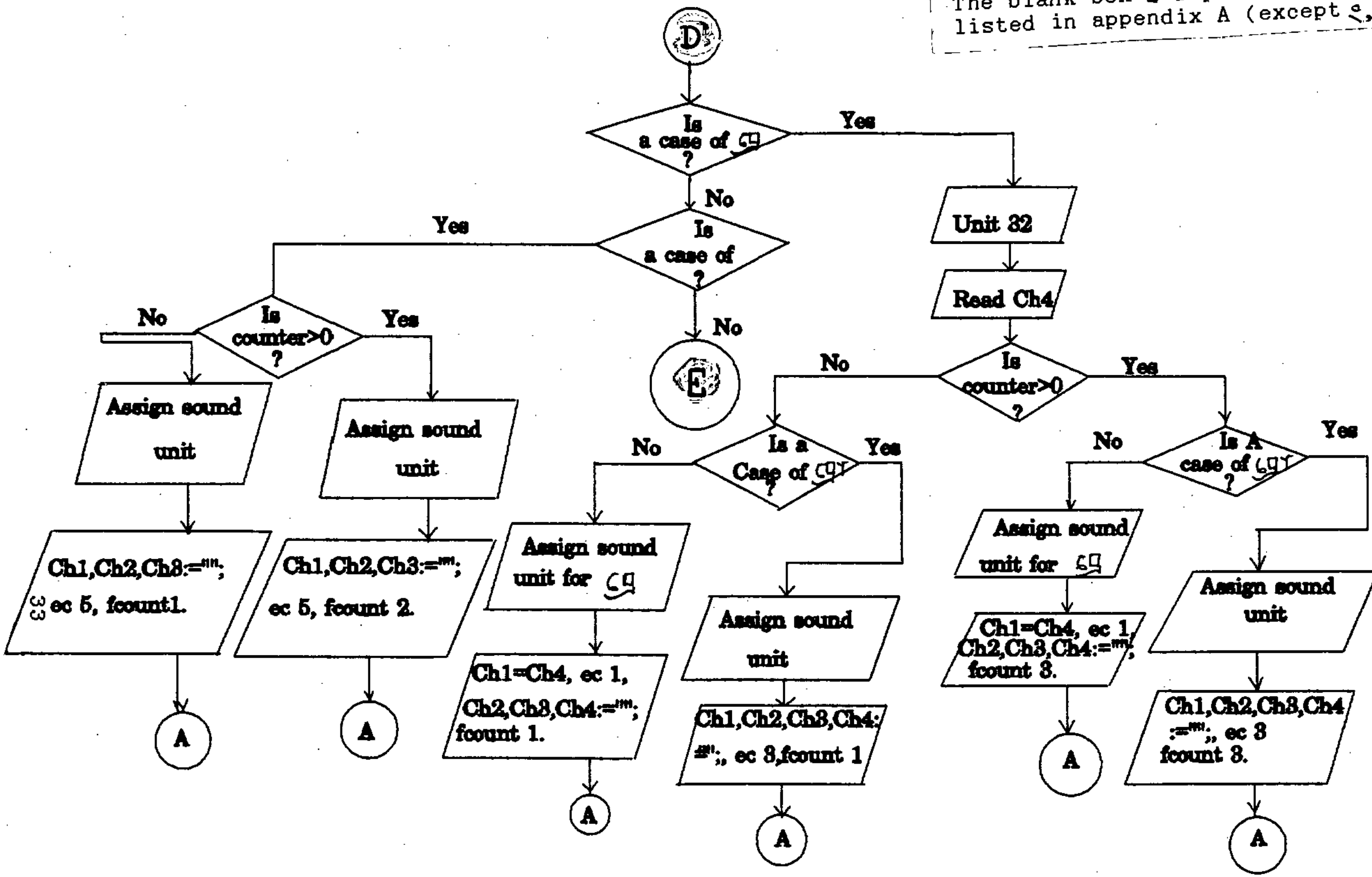


Fig4.5: Flow diagram for unit 3 (continued)

The blank box □ represents a consonant listed in appendix A (except

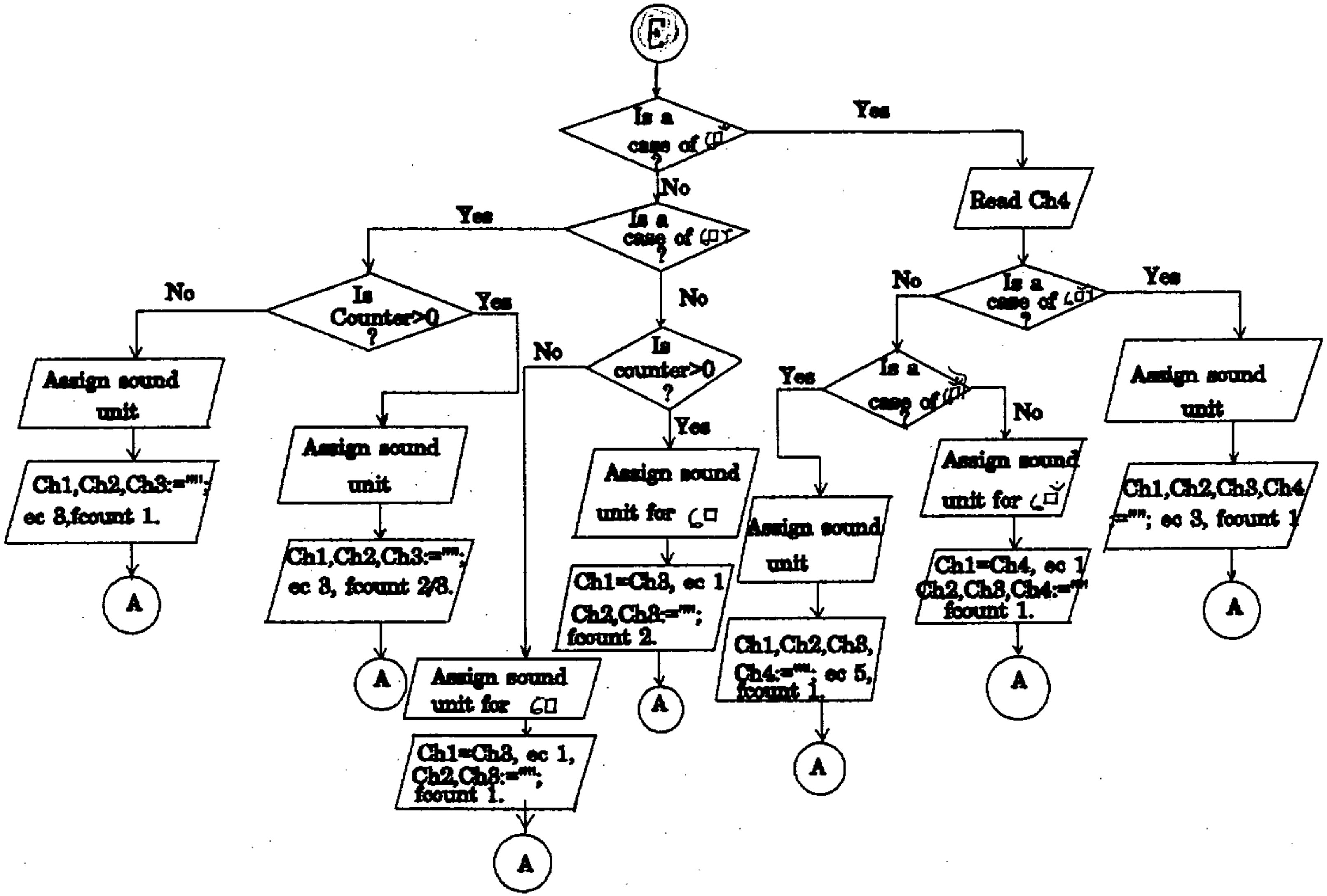


Fig 4.5: Flow diagram for unit 3(continued)

The blank box □ represents a consonant listed in appendix A (except ...)

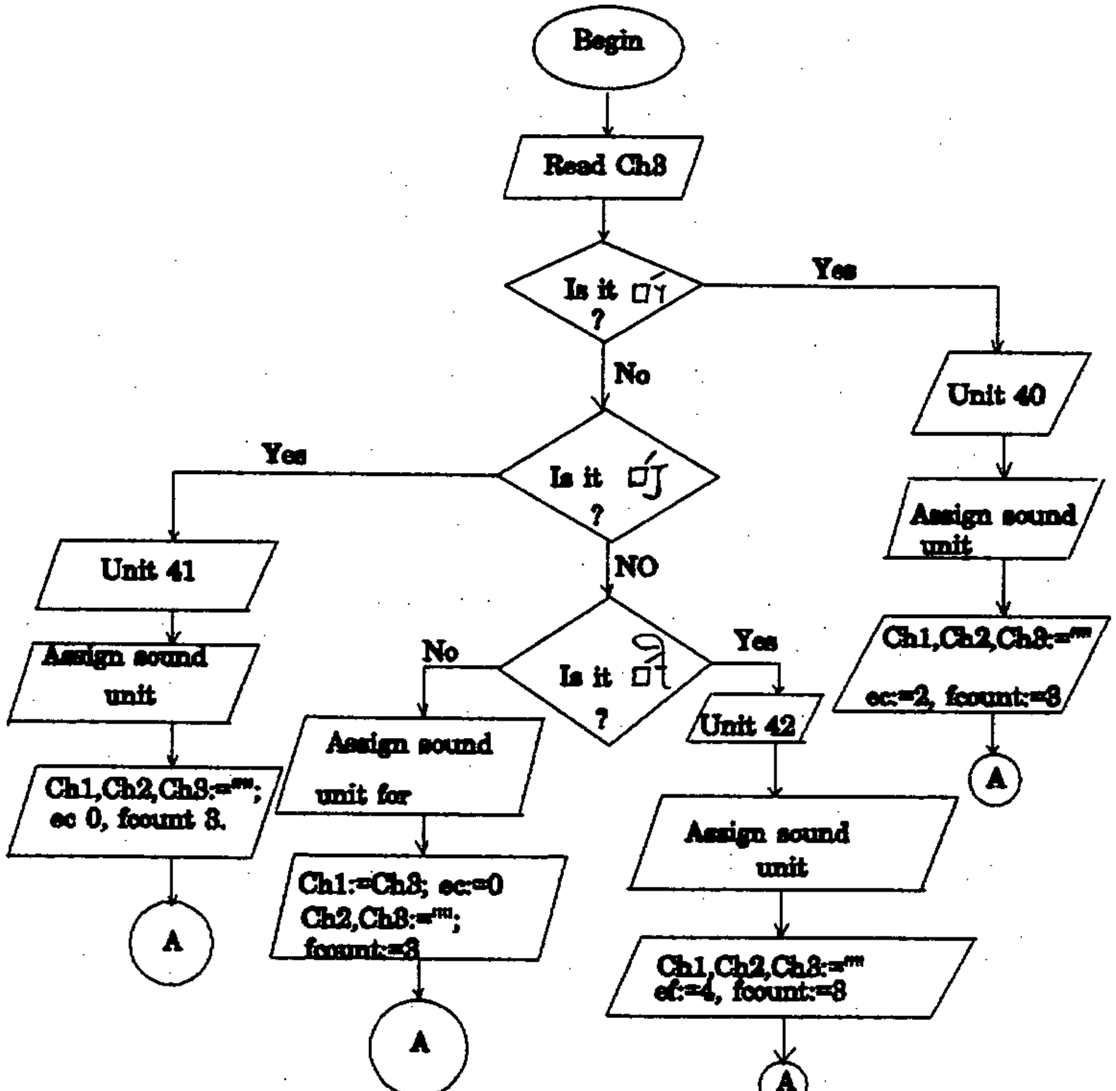


Fig4.6: Flow diagram for unit 4

The blank box □ represents a consonant listed in appendix A (except ৎ,ঃ and ং)

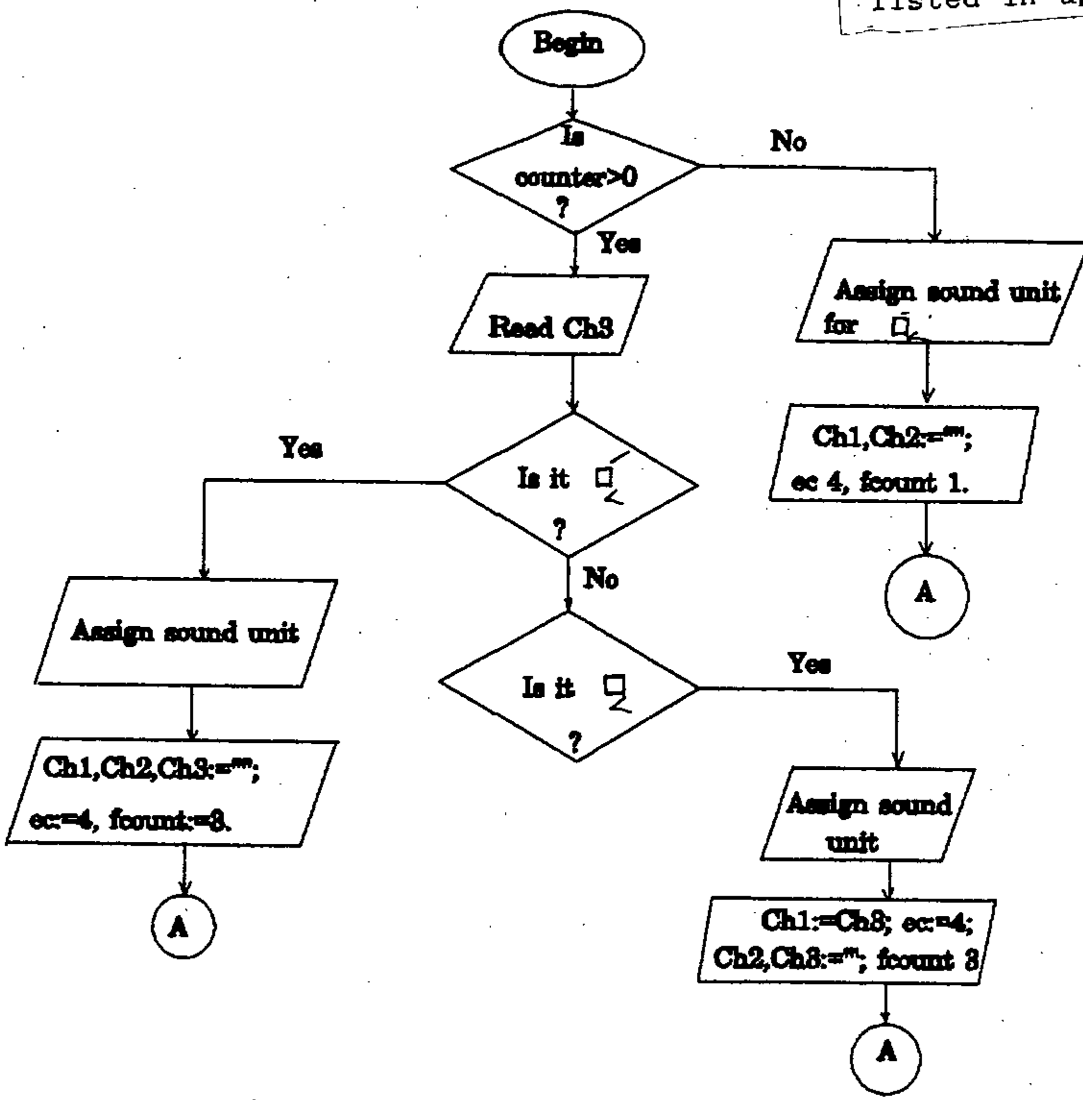


Fig4.7: Flow diagram for unit 5

The blank box ◻ represents a consonant listed in appendix A (except ...and ...)

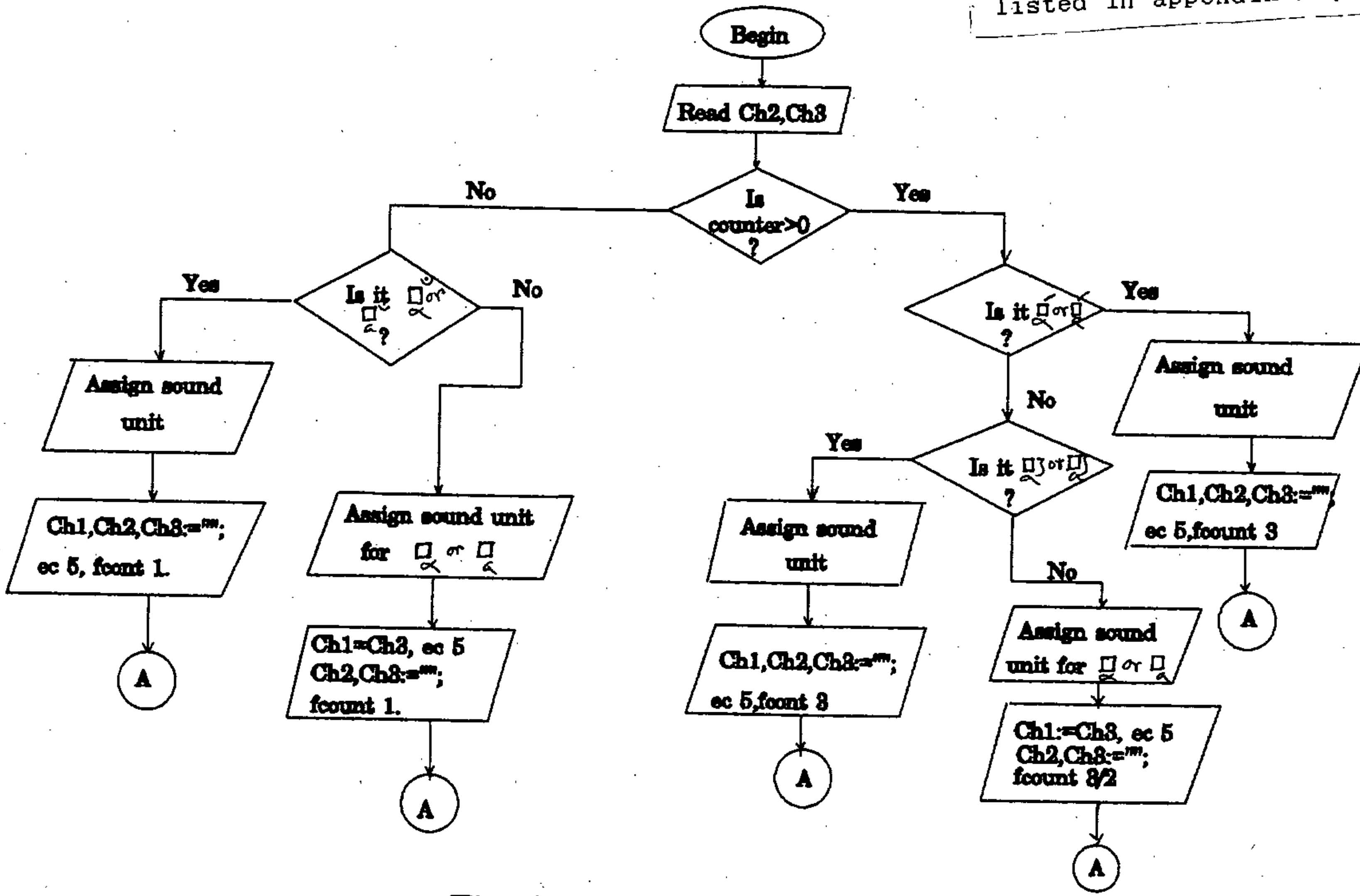


Fig4.8: Flow diagram for unit 6

The blank box □ represents a consonant listed in appendix A (except ꞉,: and ◌)

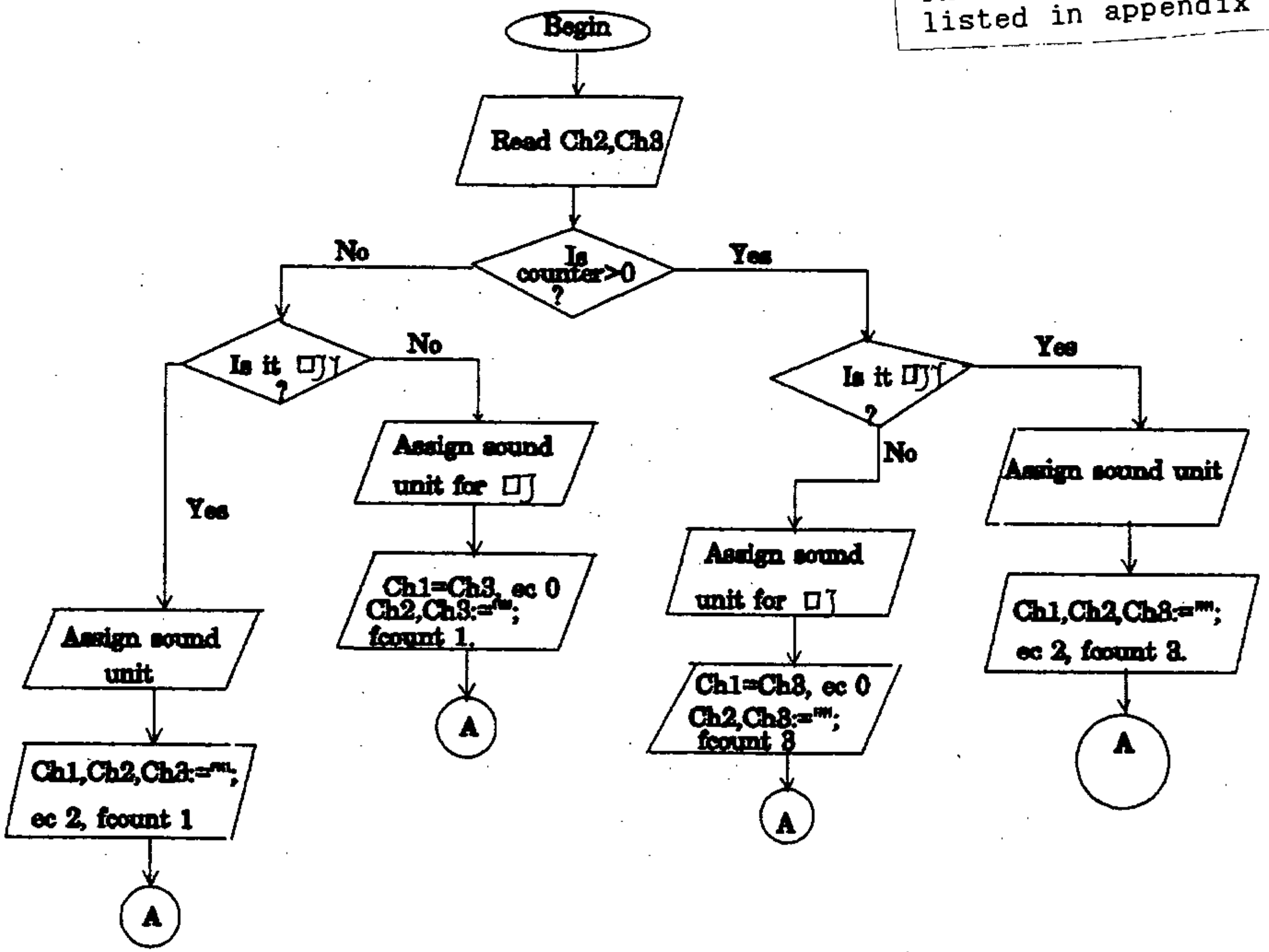


Fig 4.9: Flow diagram for unit 7

The blank box □ represents a consonant listed in appendix A (except ◌,◌ and ◌)

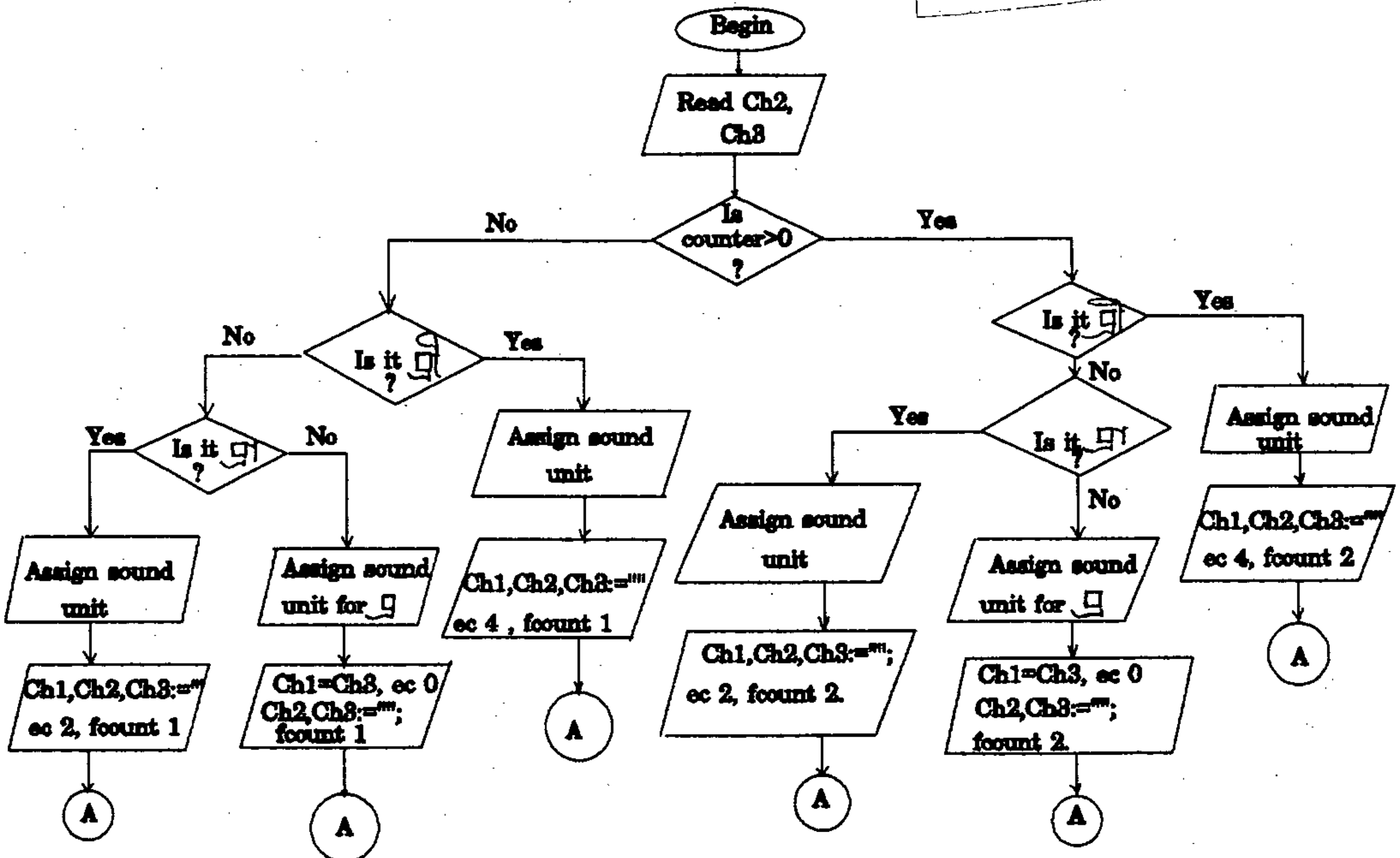


Fig 4.10: Flow diagram for unit 8

The blank box □ represents a consonant listed in appendix A (except , and)

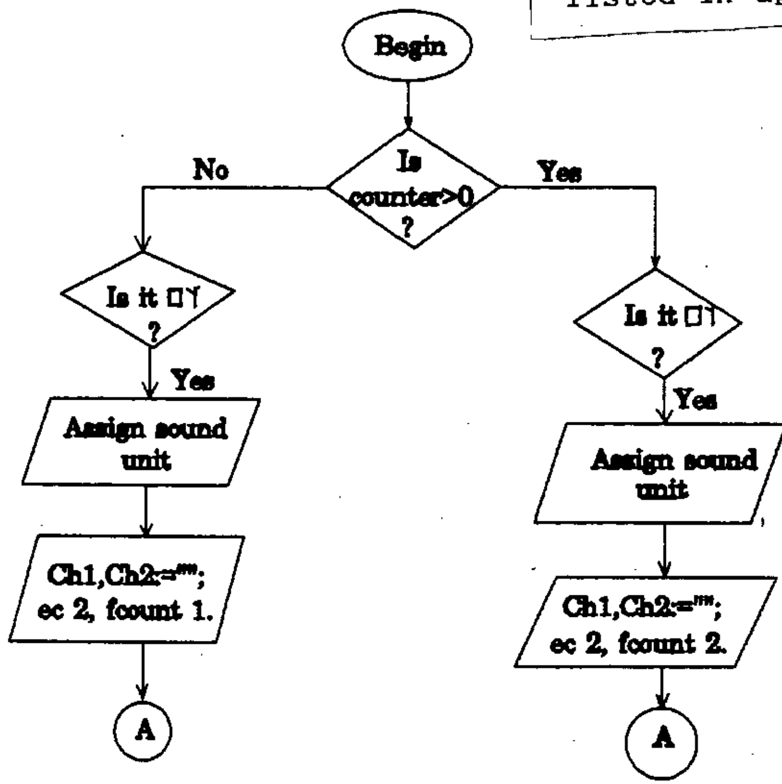


Fig 4.11: Flow diagram for unit 9

The blank box □ represents a consonant listed in appendix A (except [illegible] and [illegible])

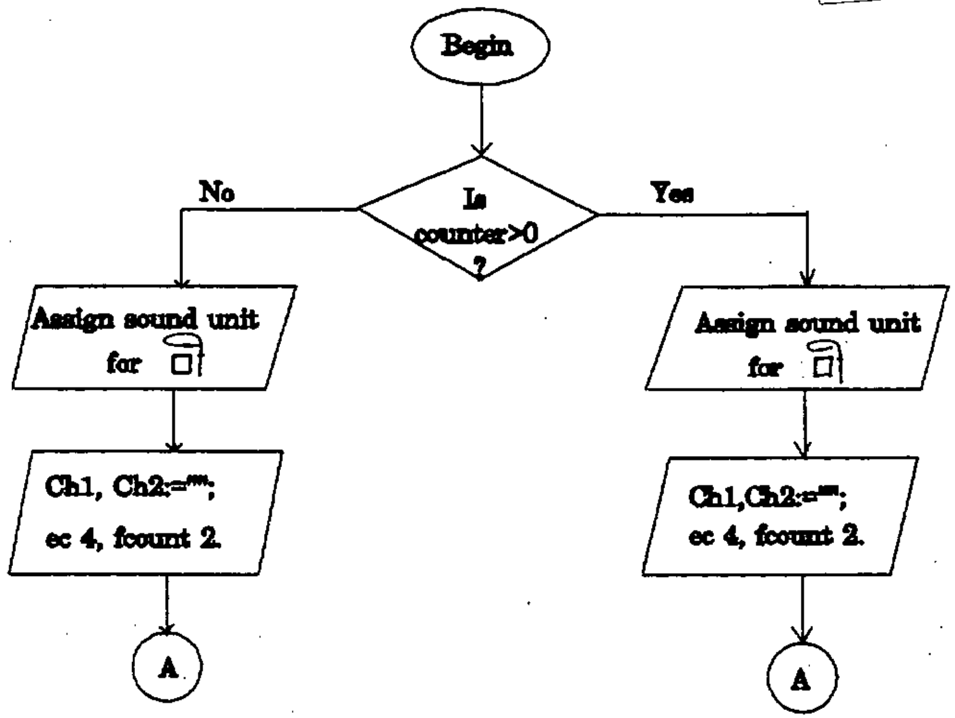


Fig 4.12: Flow diagram for unit 10

The blank box □ represents a letter listed in appendix A (except ی)

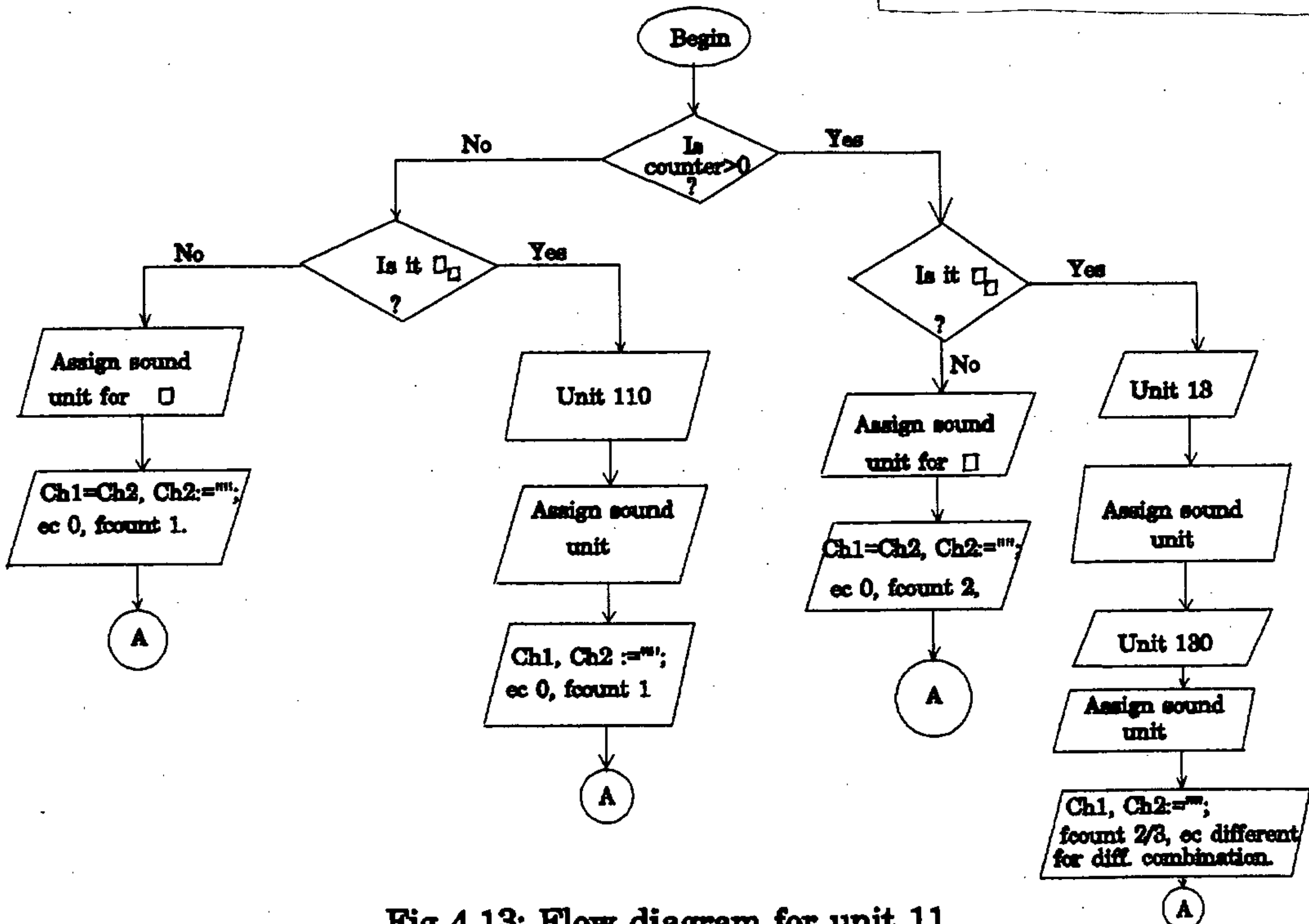


Fig 4.13: Flow diagram for unit 11

The blank box □ represents a consonant listed in appendix A (except ະ,ະ and ◌ັ)

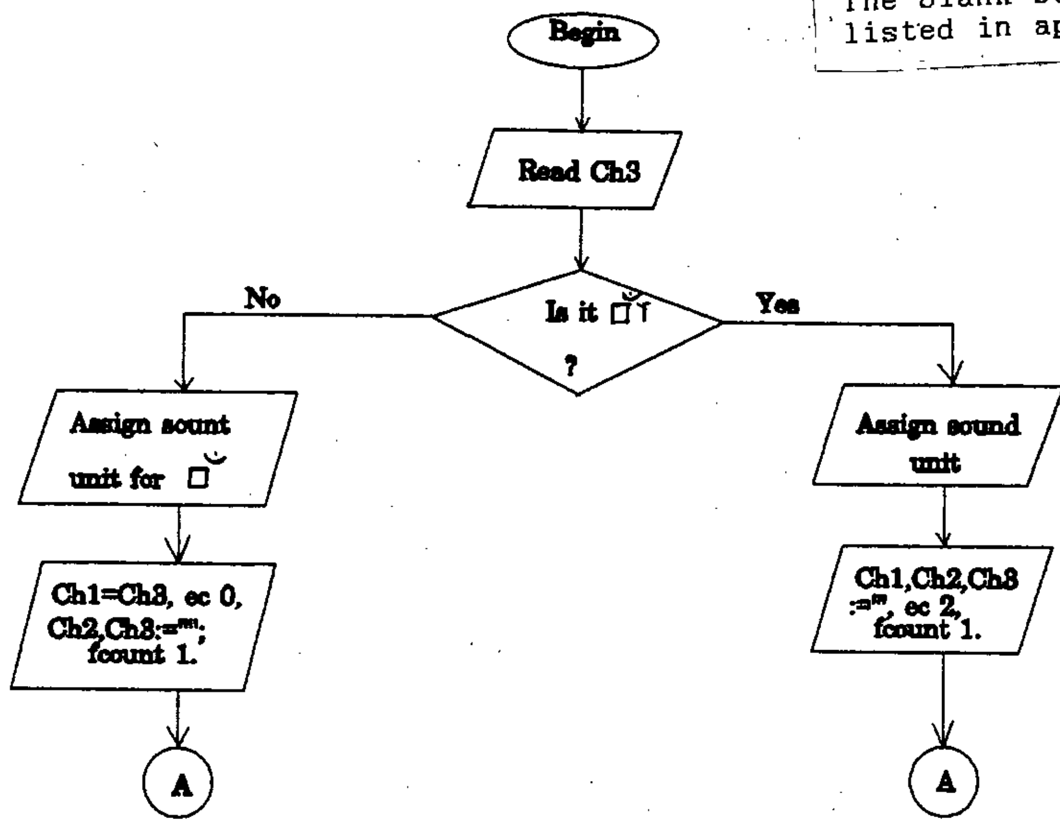


Fig 4.14: Flow diagram for unit 12

particular example in section 3.1. The units 2, 3, 4, 8, and 11 checks the "end code" of the previous utterance and goes to assign the actual sound unit/units required to form the correct pronunciation of a word.

The Bangla text-to-speech converter algorithm has been briefly described in section 4.2. Two other aspects, which are, the file merging technique and the play back procedure need to be described here. They are briefly described in sections 4.3 and 4.4, respectively.

4.3 The File Merging Technique

A few Bangla words have single sound units in them. However, most of the Bangla words require multiple sound units to form the corresponding speech sounds.

After assigning the sound units for each speech segment of a word, it is necessary to merge the sound units together to reconstruct the speech sound corresponding to the word. Depending on the sound units present in a word, the file merging technique differs. An individual file corresponding to a sound unit has its own header, which contains useful information about the type and length of the sound file. This information is vital for producing correct speech sound when the sound file is played back. When two such sound files are merged together to form a new sound file, the headers of both the files are omitted, and a new header is inserted followed by the speech information of the two sound units. When three sound units are merged together, another new header is generated omitting the headers of each sound unit and merging all speech information sequentially.

In our Bangla 'text-to-speech converter' algorithm, we have kept provision for merging a maximum of ten sound units together. Because, so far as we have come across, the longest Bangla word (অঘটনঘটনপটীয়সী) requires eleven sound units অ+য+ট+অঁন+য+টে+অঁন্+স্যা+টি+ষল to be merged together to form the corresponding speech sound.

4.5 The Play Back Procedure

The play back facility of the Sound Galaxy software, provided with the Sound Galaxy Nx Pro 16 sound card, has been utilized to play back the play files of the Bangla 'text-to-speech converter' algorithm. After merging the sound units for each word of a sentence, it is essential to play back the words of a sentence one after another. In our Bangla 'text-to-speech' converter we have kept a provision to play back a maximum of eleven words for a sentence.

This chapter has described the main algorithm of a Bangla text-to-speech converter and has given the flow diagrams of all the units used by the main program. The next chapter explores some of the possible applications of the Bangla 'text-to-speech converter' algorithm, which has been developed and implemented in PASCAL programming language.

CHAPTER 5

APPLICATIONS OF THE BANGLA TEXT TO SPEECH CONVERTER

This chapter describes some of the possible applications of the Bangla text to speech converter. There is a special feature of the Bangla text-to-speech converter, and this text-to-speech converter has its own drawbacks as well. These two aspects are also described in this chapter.

5.1 Possible applications of the Bangla text-to-speech converter

The Bangla text-to-speech converter, which has been developed in PASCAL programming language, can be used in a variety of applications. They are enumerated below.

(a) As a language tool for the dumb

The Bangla text-to-speech converter can be used as a language tool for the dumb who can write, but cannot speak. If the desired texts are input from the keyboard of a PC, the text-to-speech converter will convert these texts into audible speech sounds. Thus, a dumb can communicate with others using the text-to-speech converter. A disabled person, who after an accident has got his/her speech organ damaged, can also make use of this text-to-speech converter to communicate with others.

(b) Text-to-speech converter to read books or newspapers

The Bangla text-to-speech converter can read books or newspapers coupled with another converter that can convert standard Bangla printed texts to word processor text files (This converter has yet to be developed!). First, the 'printed text to word processor file' converter will convert the printed Bangla texts to corresponding word processor file, and then the Bangla text-to-speech converter will operate and convert the texts to audible speech sounds. The combination may be named as a 'book reader'. This so called book or newspaper reader can be of great help to those who want to hear, but do not want to read, or do not know how to read, or cannot read because he/she is a blind.

(c) Generation of speech of a particular person from arbitrary text input

In our Bangla text-to-speech converter, the sound units of a particular person are recorded from different sentences pronounced by that person. Later on, the text-to-speech converter generates speech sounds from arbitrary text input, by merging the appropriate sound units together. This may have manyfold applications, some of which are mentioned below.

(i) Voice generation of a person in an audio transmission studio

Once the basic sound units of a particular person are recorded, the Bangla text-to-speech converter simply requires to input the corresponding texts from the keyboard of a PC, to produce any speech sound in that person's voice. This voioe generation in the absence of a person may be used in an audio transmission studio. It may totally replace the persons who read news in Bangla, make announcements, and to some extent, those who participate in a drama, in story telling etc.

(ii) To produce actor's/actress' voice in an unfinished Bangla film

If the sound units of a person are once recorded, his/her voice may be generated by using the Bangla text-to-speech converter even after his/her death. Also, the voice of a person at young age can be generated at a later stage. This facility may be used along with the video processing systems to finish the dialogues of an unfinished Bangla film, which could not be finished otherwise owing to sudden death or illness of an actor/actress.

(iii) As a talking computer in the classroom

In educational institutions, if a teacher becomes so busy that he cannot be present in the classroom, he may input the texts of his lecture beforehand, and the text-to-speech converter will operate and take the class on his absence. Thus, it will be a talking computer in the class room.

(iv) For advertisement and announcement purposes in public places

If. the sound units are recorded from the voice of a person with desired inflection and personality, the Bangla text-to-speech converter can be used for advertisement and announcement purposes in public places (like at a bus stop, air-port, railway station, hospital, fair etc.). It can also act as a receptionist at a party or function.

(v) In Telecommunication and electronic mail systems

Bangla text-to-speech converter can be used in Telecommunication and electronic mail systems to produce audio messages from the corresponding text inputs.

(vi) For generating speech sounds to develop various video games and education programs on a PC.

The Bangla text-to-speech converter can be used to develop useful education programs (like teaching Bangla alphabets using audio sounds from the converter), and also to generate Bangla sounds for the video games on a PC.

5.2 Special feature of the Bangla text-to-speech converter

A 'text-to-speech synthesizer' produces synthetic or artificial speech from arbitrary text input. Our algorithm to convert text input to corresponding sound output has been named as a 'text-to-speech converter'. We have preferred the word 'converter' instead of the word 'synthesizer' to choose the name of the text-to-speech conversion algorithm, as it was initially designed to convert arbitrary text input to corresponding speech sound of a particular person.

Conventional text-to-speech synthesizers of other languages can produce either artificial speech, or the speech of a particular person. However, our text-to-speech converter can not only convert text input to speech sound of a particular person, it can also convert the text input to synthetic speech, if the sound units are generated from appropriate mathematical models. Here lies the speciality of this text-to-speech converter. This is possible greatly because of the unique property that the Bangla language has a fixed pronunciation rule for a specific arrangement of Bangla letters in a word (This has also been discussed in section 2.4). This reduces the number of required sound units markedly (approximately 1500). On the contrary, there is no such definite rule for the English language, and the morpho dictionary method of text-to-speech conversion by Allen [2] requires 10000 speech segments.

5.3 Drawbacks of the Bangla text-to-speech converter

Although, the Bangla text-to-speech converter could be developed with less numbers of basic sound units compared to its English language counterpart (the morpho dictionary method [2]), it has some drawbacks as well.

The major drawback is that the resulting speech, while certainly intelligible on a basic sound unit basis, tends to be perceived, at some instances, as mechanical and non human. There still remains the problem of intonation, because words formed from different sound units has different intonations. There are also discontinuities from word to word, which may detract the naturalness of a sentence. Slapping the sound units together from different words recorded at different times also result in a choppy 'sing song' effect, since the pitch and duration of the words do not necessarily match.

The sound units generated from this text-to-speech converter does not indicate where the speaker wishes to place emphasis on. In Bangla language, there are imperative, interrogative and exclamatory sentences, which express the feeling of the speaker differently at different situations. Different words are also pronounced differently depending on the environment and situation in which the speaker is speaking. This text-to-speech converter, however, can not provide us with these variations.

In the next chapter conclusions of the current research are drawn, and some suggestions are put forward for future improvements and applications of the Bangla text-to-speech converter.

CHAPTER 6

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

In this chapter we have concluded the results of the current research and have proposed a few topics for future research in this field.

6.1 Conclusions

The text-to-speech converter which has been described in this dissertation is the first one developed for the Bangla language. Various methods are available to develop a text-to-speech converter. They are: (a) the phoneme method, and (b) the sentence, word, and part of a word storage methods. A summary of the development of the Bangla text-to-speech converter is described as follows.

To develop the Bangla text-to-speech converter we did not make use the phoneme method, as it requires complex analysis and difficult software development. Moreover, text-to-speech conversion in this method would be a bit slow. The Bangla text-to-speech converter was developed using the method that stores part of the word (syllable or speech segment). The advantage of this method is that it is easy to implement using a commercially available sound card. Our text-to-speech converter requires less number of sound units to be stored, and it produces less 'sing song' effect compared to its English language counterpart (the morpho dictionary method [2]).

The Bangla text-to-speech converter can generate speech sounds of a particular person from arbitrary text input. If the basic sound units are generated from the voice of another person, it would generate speech sounds of that person from arbitrary text input.

English text-to-speech converters, using the phoneme method, can produce synthetic speech only. However, in our text-to-speech converter, if the basic sound units are generated from mathematical models, it will generate synthetic speech as well.

The Bangla text-to-speech converter can be widely used as a language tool for the dumb (who can write, but can not speak), a speech generator of a particular person from arbitrary text input, and a talking computer in a classroom. This may also be used for advertisement and announcement purposes in public places and in telecommunication.

Although, the Bangla text-to-speech converter could be developed with less numbers of basic sound units compared to its English language counterpart, it has some drawbacks as well. As it merges different sound units together to form a word, there remains the problem of intonation. Slapping the sound units together from different words, recorded at different times, may also result in a choppy 'sing song' effect. This text-to-speech converter can not generate variations of speech sounds of an imperative, or interrogative, or exclamatory sentence.

In Bangla language, when a joined-letter is involved in a word, the rule for converting it to speech sound is a bit complex. There are also some exceptions in pronouncing some words involving particular Bangla letters. Such exceptions have been taken care of in unit 11 of the text-to-speech converter algorithm. However, as this text-to-speech converter is first of its kind, and as this is the first version of it, all the problems could not be solved in a single piece of work. A few other problems remain to be solved. For example, the converter produces the correct speech sound of the Bangla word 'বার' when it represents a day. However, when the same word 'বার' is used to represent a number (which represents twelve in Bangla), the converter will still pronounce the word as to mean 'day' of a week. Many other problems might have remained unnoticed. These problems will, hopefully, be eliminated in the subsequent versions of the Bangla-text-to-speech converter.

6.2 Suggestions for further research in this field

The Bangla text-to-speech converter, which has been developed and described in this dissertation, first converts the words of a sentence to speech files, and when the conversion of a sentence is finished it plays back the speech files sequentially. This takes time to generate the speech files before another sentence could be played back. There may be modifications of this converter to suit specific needs. These and the other topics of research are described below.

(a) A provision can be kept in the Bangla-text-to speech converter to generate play files corresponding to text inputs of a paragraph or of a page, and then the speech files of the paragraph or page should be played back sequentially at a time. This will produce speech sounds of the texts of a paragraph or of a page continuously. However, this method would increase the hard disk surface requirement to store many speech files, and would take more time to convert texts to speech files before the actual playing of the files can begin.

(b) However, a sort of parallel processing might also be incorporated, which would, perhaps, be a better alternative to it, and would reduce the hard disk surface requirement. The interrupt facility of the PC could be used to play the speech files corresponding to a sentence. While the sound card is playing a sound file back, the processor will carry on the identification, sound file assignment, and file merging tasks of a new sentence. Thus, playing the current speech files and generating new speech files for the next sentence together will facilitate faster conversion with reduced hard disk surface requirement.

(c) Further research may be carried out to produce naturalness of the speeoh sounds produced by the Bangla text-to-speech converter by reducing the 'sing-song' effect and the problem of intonation. In this regard, extensive analysis of the Bangla sound units may be carried out.

(d) Research may be carried out to develop proper mathematical models of the Bangla sound units, which would enable the current Bangla text-to-speech converter to produce synthetic Bangla speech.

(e) Research may also be carried out to develop Bangla speech recognition systems using linear predictive models, and neural network systems.

(f) Research may be carried out to develop application softwares (the applications of the Bangla text-to-speech converter are described in section 5.1) and education tools for PC, using the speech sounds produced by the Bangla text-to-speech converter.

(g) The sound units generated from the voice of a particular person may be used to produce speech sounds of other languages by using the Bangla text-to-speech converter. A dictionary of that language has to be prepared in a tabular form so that the word of that language and its pronunciation has to be written side by side in Bangla letters. Then a special software has to be developed, which will translate the texts of that language into the pronunciation sequence in Bangla (may be termed words, which may not carry any meaning in Bangla) using the 'look-up-table' type dictionary. Then, the Bangla text-to-speech converter will be invoked to produce the speech sounds corresponding to the pronunciation sequence. The success of this will however, depend on whether the speech sounds corresponding to the Bangla sound units are sufficient to pronounce all the words in that language.

REFERENCES

[1] Bristow, G.,"**Electronic Speech Synthesis**", McGraw Hill Book Company, pp. 99-113, 1984.

[2] Owens, F.J, "**Signal Processing of Speech**",The Macmillan Press Ltd., London, UK, 1993.

[3] Ainsworth, W.A,"**Performance of a Speech Synthesis System**", Int.J. Man-machine Studies, 6, pp.493-511.

[4] Allen, T., Hunnicut, M.S. and Klatt, D.,"**From text to Speech: The MI talk System**", Cambridge University Press, Cambridge, UK, 1987.

[5] Flanagan, J.L,"**Speech Analysis, Synthesis and Perception**", Springer-Verlag, New York, USA, 1972.

[6] Jamal Uddin, M., and Sobhan, M.A., " **Studies on the effects of number of periods on the formant frequencies of Bangla voice using Hamming window**", 38th Annual Convention of IEB, 1994.

[7] Gottfried, B.S.,"**Schaum's outline of theory and problems of programming with PASCAL**", TMH Edition, India, 1991.

[8] Vincent, A.T., "**Computer assisted support for blind students-the use of microcomputer linked voice Synthesizer**", Computer and Education 6, pp. 55-60, 1982.

[9] Cohen, R.J., "**Segmenting speech using dynamic programming**", J. Acoust. Soc. Am. 69, pp. 1430-1440, 1981.

[10] Huges, G.W., "**Identification of speech sounds by means of a digital computer**", J. Acoust. Soc. Am., pp. 31, 113-121, 1959.

[11] Jassem, W. and Domagala, P., "**Phonetic segmentation in a bottom-up speech analysis**", Intern. Conf. Speech I/O-Techniques and Applications, London, pp. 59-62, 1986.

[12] Roach, P., **"English Phonetics and Phonology"**, Cambridge: Cambridge University Press, 1983.

[13] Jassem, W.(ed), Warsawa: Panstw. Wyd. Naukowe, **"In speech analysis and synthesis"**, Vol.3., pp. 53-74.

[14] Board, G., and Scully, J., **"Communicating with computers in human terms"**, Proceeding of Graphics Interface ´82, pp. 329-335, 1983.

[15] Fry, D.B., **"The Physics of Speech"**, Cambridge: Cambridge University Press, 1979.

[16] Martin, T.B, **"Practical applications of voice input to machines"**, Proc. IEEE, pp. 64, 487-501, 1976.

[17] Simon, J.C. (ed), **"Spoken Language Generation and Understanding"**, D. Reidel Publishing Co., Dordrecht, Holland, 1980.

[18] Lee, F.F., **"Reading machine: From text to speech"**, IEEE Trans. Audio Electroacoustics, Vol. AU-17, No-6, pp. 26-34, December 1969.

[19] Ainsworth, W.A, **"A system for converting English text into speech"**, IEEE Trans Audio Electroacoustics, Vol. AU-21, NO.3, pp. 288-290, June 1974.

[20] Hertz, S.R., **"From text to speech with SRS"**, J. Acoust. Soc. Am., Vol. 72, No. 5, pp. 1155-1170, October 1982.

[21] Sinha, D.K. (ed), **"Special issue on computer applications in processing of indian languages and scripts"**, JIETE, Vol. 30, No.6, November 1984.

[22] **"Sound Galaxy NX PRO 16 Sound Card User Manual"**, Version 1.1, 1992.

APPENDIX A

VOWELS OF BANGLA WORD:

অ, আ, ই, ঈ, উ, ঊ, ঋ, এ, ঐ, ও, ঔ

CONSONANTS OF BANGLA WORD:

ক, খ, গ, ঘ, ঙ, চ, ছ, জ, ঝ, ঞ, ট, ঠ , ড, ঢ, ণ, ত, থ
দ, ধ, ন, প, ফ, ব, ভ, ম, য, র, ল, শ, ষ, স, হ, ড়, ঢ়,
য়, ৎ , ং, ঃ, ঁ

VOWEL AUXILIARIES OF BANGLA WORD:

া, ি, ী, ে, ু, ূ, ৈ, ৗ

OTHERS SYMBOLS:

্, ´, ্য, ্র, ্ব

APPENDIX B

BANGLA SOUND UNITS AND THEIR FILE NAMES

Sound Unit	File Name
অ	V1.WAV
আ	V2.WAV
ই	V3.WAV
ঈ	V4.WAV
উ	V5.WAV
ঊ	V6.WAV
ঋ	V7.WAV
এ	V8.WAV
ঐ	V9.WAV
ও	V10.WAV
ঔ	V11.WAV
অই্	V12.WAV
আই্	V13.WAV
অঈ্	V14.WAV
আঈ্	V15.WAV
অউ্	V16.WAV
আউ্	V17.WAV
অঊ্	V18.WAV
আঊ্	V19.WAV
অও্	V20.WAV
আও্	V21.WAV
এই্	V22.WAV
এঈ্	V23.WAV

Sound Unit	File Name
অঁ	V1C.WAV
আঁ	V2C.WAV
ইঁ	V3C.WAV
ঈঁ	V4C.WAV
উঁ	V5C.WAV
ঊঁ	V6C.WAV
ঋঁ	V7C.WAV
এঁ	V8C.WAV
ঐঁ	V9C.WAV
ওঁ	V10C.WAV
ঔঁ	V11C.WAV

Sound Unit	File Name
ক	C1.WAV
কে	C2.WAV
কা	C3.WAV
কো	C4.WAV
কি	C5.WAV
কী	C6.WAV
কু	C7.WAV
কূ	C8.WAV
কৈ	C9.WAV
কৌ	C10.WAV
কৃ	C11.WAV
ক্র	C12.WAV
ক্রে	C13.WAV
ক্রা	C14.WAV
ক্রো	C15.WAV
ক্রি	C16.WAV
অক্	C17.WAV
আক্	C18.WAV
এক্	C19.WAV
ইক্	C20.WAV
উক্	C21.WAV
ঐক্	C22.WAV
খ	C23.WAV
খা	C24.WAV
খা	C25.WAV
খো	C26.WAV
খি	C27.WAV
খী	C28.WAV
খু	C29.WAV

Sound Unit	File Name
খূ	C30.WAV
খৈ	C31.WAV
খৌ	C32.WAV
খৃ	C33.WAV
খ্র	C34.WAV
খ্রে	C35.WAV
খ্রা	C36.WAV
খ্রো	C37.WAV
খ্রি	C38.WAV
অখ্	C39.WAV
আখ্	C40.WAV
এখ্	C41.WAV
ইখ্	C42.WAV
উখ্	C43.WAV
ঐখ্	C44.WAV
গ	C45.WAV
গে	C46.WAV
গা	C47.WAV
গো	C48.WAV
গি	C49.WAV
গী	C50.WAV
গু	C51.WAV
গূ	C52.WAV
গৈ	C53.WAV
গৌ	C54.WAV
গৃ	C55.WAV
গ্র	C56.WAV
গ্রে	C57.WAV
গ্রা	C58.WAV

Sound Unit	File Name
গ্রো	C59.WAV
গ্রি	C60.WAV
·অগ্	C61.WAV
·এগ্	C62.WAV
·আগ্	C63.WAV
·ইগ্	C64.WAV
·উগ্	C65.WAV
·ঐগ্	C66.WAV
ঘ	C67.WAV
ঘে	C68.WAV
ঘা	C89.WAV
ঘো	C70.WAV
ঘি	C71.WAV
ঘী	C72.WAV
ঘু	C73.WAV
ঘূ	C74.WAV
ঘৈ	C75.WAV
ঘৌ	C76.WAV
ঘৃ	C77.WAV
ঘ্র	C78.WAV
ঘ্রে	C79.WAV
ঘ্রা	C80.WAV
ঘ্রো	C81.WAV
ঘ্রি	C82.WAV
·অঘ্	C83.WAV
·এঘ্	C84.WAV
·আঘ্	C85.WAV
·ইঘ্	C86.WAV
·উঘ্	C87.WAV

Sound Unit	File Name
·ঐঘ্	C88.WAV
·অঙ	C89.WAV
·আঙ	C90.WAV
·এঙ	C91.WAV
·ইঙ	C92.WAV
·উঙ	C93.WAV
·ঐঙ	C94.WAV
চ	C95.WAV
চে	C96.WAV
চা	C97.WAV
চো	C98.WAV
চি	C99.WAV
চী	C100.WAV
	C101.WAV
	C102.WAV
	C103.WAV
	C104.WAV
	C105.WAV
	C106.WAV
	C107.WAV
	C108.WAV
	C109.WAV
	C110.WAV
চু	C111.WAV
চূ	C112.WAV
চৈ	C113.WAV
চৌ	C114.WAV
চৃ	C115.WAV
চ্র	C116.WAV

Sound Unit	File Name
চ্চে	C117.WAV
চ্চা	C118.WAV
চ্চো	C119.WAV
চ্চি	C120.WAV
অঁচ্	C121.WAV
এঁচ্	C122.WAV
আঁচ্	C123.WAV
ইঁচ্	C124.WAV
উঁচ্	C125.WAV
ছ	C126.WAV
ছে	C127.WAV
ছা	C128.WAV
ছো	C129.WAV
ছি	C130.WAV
ছী	C131.WAV
ছু	C132.WAV
ছূ	C133.WAV
ছৈ	C134.WAV
ছৌ	C135.WAV
ছৃ	C136.WAV
চ্ছ	C137.WAV
চ্ছে	C138.WAV
চ্ছা	C139.WAV
চ্ছো	C140.WAV
চ্ছি	C141.WAV
অঁছ্	C142.WAV
এঁছ্	C143.WAV
আঁছ্	C144.WAV
ইঁছ্	C145.WAV

Sound Unit	File Name
উঁছ্	C146.WAV
ঐঁছ্	C147.WAV
জ	C148.WAV
জে	C149.WAV
জা	C150.WAV
জো	C151.WAV
জি	C152.WAV
জী	C153.WAV
জু	C154.WAV
জূ	C155.WAV
জৈ	C156.WAV
জৌ	C157.WAV
জৃ	C158.WAV
জ্জ	C159.WAV
জ্জে	C160.WAV
জ্জা	C161.WAV
জ্জো	C162.WAV
জ্জি	C163.WAV
অঁজ্	C164.WAV
এঁজ্	C165.WAV
আঁজ্	C166.WAV
ইঁজ্	C167.WAV
উঁজ্	C168.WAV
ঐঁজ্	C169.WAV
ঝ	C170.WAV
ঝে	C171.WAV
ঝা	C172.WAV
ঝো	C173.WAV
ঝি	C174.WAV

Sound Unit	File Name
ঝী	C175.WAV
ঝু	C176.WAV
ঝূ	C177.WAV
ঝৈ	C178.WAV
ঝৌ	C179.WAV
ঝৃ	C180.WAV
ঝ্র	C181.WAV
ঝে	C182.WAV
ঝা	C183.WAV
ঝো	C184.WAV
ঝি	C185.WAV
অঝ্	C186.WAV
এঝ্	C187.WAV
আঝ্	C188.WAV
ইঝ্	C189.WAV
উঝ্	C190.WAV
ইঞা	C191.WAV
ইঞো	C192.WAV
ট	C193.WAV
টে	C194.WAV
টো	C195.WAV
টা	C196.WAV
টি	C197.WAV
টী	C198.WAV
টু	C199.WAV
টূ	C200.WAV
টৈ	C201.WAV
টৌ	C202.WAV
টৃ	C203.WAV

Sound Unit	File Name
ট্র	C204.WAV
ট্রে	C205.WAV
ট্রা	C206.WAV
ট্রো	C207.WAV
ট্রি	C208.WAV
অট্	C209.WAV
এট্	C210.WAV
আট্	C211.WAV
ইট্	C212.WAV
উট্	C213.WAV
ঐট্	C214.WAV
ঠ	C215.WAV
ঠে	C216.WAV
ঠা	C217.WAV
ঠৌ	C218.WAV
ঠি	C219.WAV
ঠী	C220.WAV
ঠু	C221.WAV
ঠূ	C222.WAV
ঠৈ	C223.WAV
ঠৌ	C224.WAV
ঠৃ	C225.WAV
ঠ	C226.WAV
ঠে	C227.WAV
ঠা	C228.WAV
ঠো	C229.WAV
ঠি	C230.WAV
অঠ্	C231.WAV
এঠ্	C232.WAV

Sound Unit	File Name
'আঠ্	C233.WAV
'ইঠ্	C234.WAV
'উঠ্	C235.WAV
ড	C236.WAV
ডে	C237.WAV
ডা	C238.WAV
ডো	C239.WAV
ডি	C240.WAV
ডী	C241.WAV
ডু	C242.WAV
ডূ	C243.WAV
ডৈ	C244.WAV
ডৌ	C245.WAV
ডৃ	C246.WAV
ড্র	C247.WAV
ড্রে	C248.WAV
ড্রা	C249.WAV
ড্রো	C250.WAV
ড্রি	C251.WAV
'অড্	C252.WAV
'এড্	C253.WAV
'আড্	C254.WAV
'ইড্	C255.WAV
'উড্	C256.WAV
'ঐড্	C257.WAV
ঢ	C258.WAV
ঢে	C259.WAV
ঢা	C260.WAV
ঢো	C261.WAV
ঢি	C262.WAV
ঢী	C263.WAV
ঢু	C264.WAV
ঢূ	C265.WAV
ঢৈ	C266.WAV
ঢৌ	C267.WAV
ঢৃ	C268.WAV
ঢ্র	C269.WAV
ঢ্রে	C270.WAV
ঢ্রা	C271.WAV
ঢ্রো	C272.WAV
ঢ্রি	C273.WAV
'অঢ্	C274.WAV
'এঢ্	C275.WAV
'আঢ্	C276.WAV
'ইঢ্	C277.WAV
'উঢ্	C278.WAV
'ঐঢ্	C279.WAV
ণ	C280.WAV
ণে	C281.WAV
ণা	C282.WAV
ণো	C283.WAV
ণি	C284.WAV
ণী	C285.WAV
ণু	C286.WAV
ণূ	C287.WAV
ণৈ	C288.WAV
ণৌ	C289.WAV
ণৃ	C290.WAV

Sound Unit	File Name
ণ	C291.WAV
ণে	C292.WAV
ণা	C293.WAV
ণো	C294.WAV
ণি	C295.WAV
অণ্	C296.WAV
এণ্	C297.WAV
আণ্	C298.WAV
ইণ্	C299.WAV
উণ্	C300.WAV
ঐণ্	C301.WAV
ত	C302.WAV
তে	C303.WAV
তা	C304.WAV
তো	C305.WAV
তি	C306.WAV
তী	C307.WAV
তু	C308.WAV
তূ	C309.WAV
তৈ	C310.WAV
তৌ	C311.WAV
তৃ	C312.WAV
ত্র	C313.WAV
ত্রে	C314.WAV
ত্রা	C315.WAV
ত্রো	C316.WAV
ত্রি	C317.WAV
অত্	C318.WAV
এত্	C319.WAV

Sound Unit	File Name
আত্	C320.WAV
ইত্	C321.WAV
উত্	C322.WAV
ঐত্	C323.WAV
থ	C324.WAV
থে	C325.WAV
থা	C326.WAV
থো	C327.WAV
থি	C328.WAV
থী	C329.WAV
থু	C330.WAV
থূ	C331.WAV
থৈ	C332.WAV
থৌ	C333.WAV
থৃ	C334.WAV
থ্র	C335.WAV
থ্রে	C336.WAV
থ্রা	C337.WAV
থ্রো	C338.WAV
থ্রি	C339.WAV
অথ্	C340.WAV
এথ্	C341.WAV
আথ্	C342.WAV
ইথ্	C343.WAV
উথ্	C344.WAV
ঐথ্	C345.WAV
দ	C346.WAV
দে	C347.WAV
দা	C348.WAV

Sound Unit	File Name
দো	C349.WAV
দি	C350.WAV
দী	C351.WAV
দু	C352.WAV
দূ	C353.WAV
দৈ	C354.WAV
দৌ	C355.WAV
দৃ	C356.WAV
দ্র	C357.WAV
দ্রে	C358.WAV
দ্রা	C359.WAV
দ্রো	C360.WAV
দ্রি	C361.WAV
'অদ্	C362.WAV
'এদ্	C363.WAV
'আদ্	C364.WAV
'ইদ্	C365.WAV
'উদ্	C366.WAV
'ঐদ্	C367.WAV
ধ	C368.WAV
ধে	C369.WAV
ধা	C370.WAV
ধি	C371.WAV
ধী	C372.WAV
ধু	C373.WAV
ধূ	C374.WAV
ধৈ	C375.WAV
ধৌ	C376.WAV
ধৃ	C377.WAV

Sound Unit	File Name
ধ্র	C378.WAV
ধ্রে	C379.WAV
ধ্রা	C380.WAV
ধ্রো	C381.WAV
ধ্রি	C382.WAV
'অধ্	C383.WAV
'এধ্	C384.WAV
'আধ্	C385.WAV
'ইধ্	C386.WAV
'উধ্	C387.WAV
'ঐধ্	C388.WAV
ন	C389.WAV
নে	C390.WAV
না	C391.WAV
নো	C392.WAV
নি	C393.WAV
নী	C394.WAV
নু	C395.WAV
নূ	C396.WAV
নৈ	C397.WAV
নৌ	C398.WAV
নৃ	C399.WAV
ন্র	C400.WAV
ন্রে	C401.WAV
ন্রা	C402.WAV
ন্রো	C403.WAV
ন্রি	C404.WAV
'অন্	C405.WAV
'এন্	C406.WAV

Sound Unit	File Name
আন্	C407.WAV
ইন্	C408.WAV
উন্	C409.WAV
ওন্	C410.WAV
প	C411.WAV
পে	C412.WAV
পা	C413.WAV
পো	C414.WAV
পি	C415.WAV
পী	C416.WAV
পু	C417.WAV
পূ	C418.WAV
পৈ	C419.WAV
পৌ	C420.WAV
পৃ	C421.WAV
প্র	C422.WAV
প্রে	C423.WAV
প্রা	C424.WAV
প্রো	C425.WAV
প্রি	C426.WAV
অপ্	C427.WAV
এপ্	C428.WAV
আপ্	C429.WAV
ইপ্	C430.WAV
উপ্	C431.WAV
ওপ্	C432.WAV
ফ	C433.WAV
ফে	C434.WAV
ফা	C435.WAV

Sound Unit	File Name
ফো	C436.WAV
ফি	C437.WAV
ফী	C438.WAV
ফু	C439.WAV
ফূ	C440.WAV
ফৈ	C441.WAV
ফৌ	C442.WAV
ফৃ	C443.WAV
ফ্র	C444.WAV
ফ্রে	C445.WAV
ফ্রা	C446.WAV
ফ্রো	C447.WAV
ফ্রি	C448.WAV
অফ্	C449.WAV
এফ্	C450.WAV
আফ্	C451.WAV
ইফ্	C452.WAV
উফ্	C453.WAV
ওফ্	C454.WAV
ব	C455.WAV
বে	C456.WAV
বা	C457.WAV
বো	C458.WAV
বি	C459.WAV
বী	C460.WAV
বু	C461.WAV
বূ	C462.WAV
বৈ	C463.WAV
বৌ	C464.WAV

Sound Unit	File Name
বৃ	C465.WAV
ব্র	C466.WAV
ব্রে	C467.WAV
ব্রা	C468.WAV
ব্রো	C469.WAV
ব্রি	C470.WAV
•অব্	C471.WAV
•এব্	C472.WAV
•আব্	C473.WAV
•ইক্	C474.WAV
•উব্	C475.WAV
•ঐব্	C476.WAV
ভ	C477.WAV
ভে	C478.WAV
ভা	C479.WAV
ভো	C480.WAV
ভি	C481.WAV
ভী	C482.WAV
ভু	C483.WAV
ভূ	C484.WAV
ভৈ	C485.WAV
ভৌ	C486.WAV
ভৃ	C487.WAV
ভ্র	C488.WAV
ভ্রে	C489.WAV
ভ্রা	C490.WAV
ভ্রো	C491.WAV
ভ্রি	C492.WAV
•অভ্	C493.WAV

Sound Unit	File Name
•এভ্	C494.WAV
´আভ্	C495.WAV
ˊ•ইভ্	C496.WAV
•উভ্	C497.WAV
•ঐভ্	C498.WAV
ম	C499.WAV
মে	C500.WAV
মা	C501.WAV
মো	C502.WAV
মি	C503.WAV
মী	C504.WAV
মু	C505.WAV
মূ	C506.WAV
মৈ	C507.WAV
মৌ	C508.WAV
মৃ	C509.WAV
ম্র	C510.WAV
ম্রে	C511.WAV
ম্রো	C512.WAV
ম্রা	C513.WAV
ম্রি	C514.WAV
•অম্	C515.WAV
•এম্	C516.WAV
•আম্	C517.WAV
•ইম্	C518.WAV
•উম্	C519.WAV
•ঐম্	C520.WAV
য়	C521.WAV
যে	C522.WAV

Sound Unit	File Name
যা	C523.WAV
যো	C524.WAV
যি	C525.WAV
যী	C526.WAV
যু	C527.WAV
যূ	C528.WAV
যৈ	C529.WAV
যৌ	C530.WAV
যৃ	C531.WAV
য়	C532.WAV
য়ে	C533.WAV
য়ো	C534.WAV
য়া	C535.WAV
য়ি	C536.WAV
অয্	C537.WAV
এয্	C538.WAV
আয্	C539.WAV
ইয্	C540.WAV
উয্	C541.WAV
ঐয্	C542.WAV
র	C543.WAV
রে	C544.WAV
রা	C545.WAV
রো	C546.WAV
রি	C547.WAV
রী	C548.WAV
রু	C549.WAV
রৃ	C550.WAV
রৈ	C551.WAV

Sound Unit	File Name
রৌ	C552.WAV
অর্	C553.WAV
এর্	C554.WAV
আর্	C555.WAV
ইর্	C556.WAV
উর্	C557.WAV
ঐর্	C558.WAV
ল	C559.WAV
লে	C560.WAV
লা	C561.WAV
লো	C562.WAV
লি	C563.WAV
লী	C564.WAV
লু	C565.WAV
লূ	C566.WAV
লৈ	C567.WAV
লৌ	C568.WAV
লৃ	C569.WAV
ল্র	C570.WAV
ল্রে	C571.WAV
ল্রো	C572.WAV
ল্রা	C573.WAV
ল্রি	C574.WAV
অল্	C575.WAV
এল্	C576.WAV
আল্	C577.WAV
ইল্	C578.WAV
উল্	C579.WAV
ঐল্	C580.WAV

Sound Unit	File Name
শ	C581.WAV
শে	C582.WAV
শা	C583.WAV
শো	C584.WAV
শি	C585.WAV
শী	C586.WAV
শু	C587.WAV
শূ	C588.WAV
শৈ	C589.WAV
শৌ	C590.WAV
শৃ	C591.WAV
শ্র	C592.WAV
শ্রে	C593.WAV
শ্রা	C594.WAV
শ্রো	C595.WAV
শ্রি	C596.WAV
অঁশ্	C597.WAV
এঁশ্	C598.WAV
আঁশ্	C599.WAV
ইঁশ্	C600.WAV
উঁশ্	C601.WAV
ঐঁশ্	C602.WAV
ষ	C603.WAV
ষে	C604.WAV
ষা	C605.WAV
ষো	C606.WAV
ষি	C607.WAV
ষী	C608.WAV
ষু	C609.WAV

Sound Unit	File Name
ষূ	C610.WAV
ষৈ	C611.WAV
ষৌ	C612.WAV
ষৃ	C613.WAV
ষ্র	C614.WAV
ষ্রে	C615.WAV
ষ্রা	C616.WAV
ষ্রো	C617.WAV
ষ্রি	C618.WAV
অঁষ্	C619.WAV
এঁষ্	C620.WAV
আঁষ্	C621.WAV
ইঁষ্	C622.WAV
উঁষ্	C623.WAV
ঐঁষ্	C624.WAV
স	C625.WAV
সে	C626.WAV
সা	C627.WAV
সো	C628.WAV
সি	C629.WAV
সী	C630.WAV
সু	C631.WAV
সূ	C632.WAV
সৈ	C633.WAV
সৌ	C634.WAV
সৃ	C635.WAV
স্র	C636.WAV
স্রে	C637.WAV
স্রা	C638.WAV

Sound Unit	File Name
স্রো	C639.WAV
স্রি	C640.WAV
অঁস্	C641.WAV
এঁস্	C642.WAV
আঁস্	C643.WAV
ইঁস্	C644.WAV
উঁস্	C645.WAV
ঐঁস্	C646.WAV
হ	C647.WAV
হে	C648.WAV
হা	C649.WAV
হো	C650.WAV
হি	C651.WAV
হী	C652.WAV
হু	C653.WAV
হূ	C654.WAV
হৈ	C655.WAV
হৌ	C656.WAV
হৃ	C657.WAV
হ্র	C658.WAV
হ্রে	C659.WAV
হ্রা	C660.WAV
হ্রো	C661.WAV
হ্রি	C662.WAV
অঁহ্	C663.WAV
এঁহ্	C664.WAV
আঁহ্	C665.WAV
ইঁহ্	C666.WAV
উঁহ্	C667.WAV

Sound Unit	File Name
ক	C668.WAV
কে	C669.WAV
কা	C670.WAV
কো	C671.WAV
কি	C672.WAV
কী	C673.WAV
কু	C674.WAV
কূ	C675.WAV
কৈ	C676.WAV
কৌ	C67.WAV7
কৃ	C678.WAV
ক্র	C679.WAV
ক্রে	C680.WAV
ক্রা	C681.WAV
ক্রো	C682.WAV
ক্রি	C683.WAV
অঁক্	C684.WAV
এঁক্	C685.WAV
আঁক্	C686.WAV
ইঁক্	C687.WAV
উঁক্	C688.WAV
ঐঁক্	C689.WAV
ড়	C690.WAV
ড়ে	C691.WAV
ড়া	C692.WAV
ড়ো	C693.WAV
ড়ি	C694.WAV
ড়ী	C695.WAV
ড়ু	C696.WAV

Sound Unit	File Name
ড়ু	C697.WAV
ড়ৈ	C698.WAV
ড়ৌ	C699.WAV
অঁড়্	C700.WAV
এঁড়্	C701.WAV
আঁড়্	C702.WAV
ইঁড়্	C703.WAV
উঁড়্	C704.WAV
ঐঁড়্	C705.WAV
ঢ়	C706.WAV
ঢ়ে	C707.WAV
ঢ়া	C708.WAV
ঢ়ো	C709.WAV
ঢ়ি	C710.WAV
ঢ়ী	C711.WAV
ঢ়ু	C712.WAV
ঢ়ূ	C713.WAV
ঢ়ৈ	C714.WAV
ঢ়ৌ	C715.WAV
অঁঢ়্	C716.WAV
এঁঢ়্	C717.WAV
আঁঢ়্	C718.WAV
ইঁঢ়্	C719.WAV
উঁঢ়্	C720.WAV
ঐঁঢ়্	C721.WAV
য়	C722.WAV
য়ে	C723.WAV
য়া	C724.WAV
য়ো	C725.WAV
অং	C726.WAV
এং	C727.WAV
ইং	C728.WAV
আং	C729.WAV
উং	C730.WAV
ঐং	C731.WAV
এ ‹	C732.WAV
অ ‹	C733.WAV
ই ‹	C734.WAV
আ ‹	C735.WAV
উ ‹	C736.WAV
ঐ ‹	C737.WAV
অঃ	C738.WAV
এঃ	C739.WAV
আঃ	C740.WAV
ইঃ	C741.WAV
উঃ	C742.WAV
ঐঃ	C743.WAV
ত্র	C744.WAV
ত্রা	C745.WAV
ত্রে	C746.WAV
ত্রো	C747.WAV
ত্রৈ	C748.WAV
ত্রু	C749.WAV
ত্রূ	C750.WAV
ত্রি	C751.WAV
ত্রী	C752.WAV
ত্রৌ	C753.WAV
শৃ	C754.WAV

Sound Unit	File Name
স্মৃ	C755.WAV
স্পু	C756.WAV
স্ফু	C757.WAV
স্কু	C758.WAV
স্পী	C759.WAV
স্ফী	C760.WAV
স্বী	C761.WAV
স্রা	C762.WAV
স্মা	C763.WAV
স্ফা	C764.WAV
স্হা	C765.WAV
স্টা	C766.WAV
স্ট্র	C767.WAV
স্ব	C768.WAV
স্ম	C769.WAV
স্ত	C770.WAV
স্প	C771.WAV
স্ফ	C772.WAV
স্ক	C773.WAV
স্হ	C774.WAV
শ্বা	C775.WAV
শ্ব	C776.WAV
শ্ম	C777.WAV
জ্বা	C778.WAV
জ্ব	C779.WAV
দ্বী	C780.WAV
জ্ঞা	C781.WAV
ষ্থা	C782.WAV
স্টি	C783.WAV

Sound Unit	File Name
অয়্	C784.WAV
এয়্	C785.WAV
আয়্	C786.WAV
ইয়্	C787.WAV
উয়্	C788.WAV
স্প্রি	C789.WAV
স্প্রে	C790.WAV
স্পে	C791.WAV
স্কো	C792.WAV
স্কে	C793.WAV
স্ক্রে	C794.WAV
স্ট্রো	C795.WAV
স্ট্রে	C796.WAV
স্টে	C797.WAV
প্লে	C798.WAV
ক্য	C799.WAV
ক্যা	C800.WAV
খ্য	C801.WAV
খ্যা	C802.WAV
গ্য	C803.WAV
গ্যা	C804.WAV
ঘ্য	C805.WAV
ঘ্যা	C806.WAV
চ্য	C807.WAV
চ্যা	C808.WAV
ছ্য	C809.WAV
ছ্যা	C810.WAV
জ্য	C811.WAV
জ্যা	C812.WAV

Sound Unit	File Name
ঝ্য	C813.WAV
ঝ্যা	C814.WAV
ট্য	C815.WAV
ট্যা	C816.WAV
ঠ্য	C817.WAV
ঠ্যা	C818.WAV
ড্য	C819.WAV
ড্যা	C820.WAV
ঢ্য	C821.WAV
ঢ্যা	C822.WAV
ণ্য	C823.WAV
ণ্যা	C824.WAV
ত্য	C825.WAV
ত্যা	C826.WAV
থ্য	C827.WAV
থ্যা	C828.WAV
দ্য	C829.WAV
দ্যা	C830.WAV
ধ্য	C831.WAV
ধ্যা	C832.WAV
ন্য	C833.WAV
ন্যা	C834.WAV
প্য	C835.WAV
প্যা	C836.WAV
ফ্য	C837.WAV
ফ্যা	C838.WAV
ব্য	C839.WAV
ব্যা	C840.WAV
ভ্য	C841.WAV

Sound Unit	File Name
ভ্যা	C842.WAV
ম্য	C843.WAV
ম্যা	C844.WAV
য্য	C845.WAV
য্যা	C846.WAV
র্য	C847.WAV
র্যা	C848.WAV
ল্য	C849.WAV
ল্যা	C850.WAV
শ্য	C851.WAV
শ্যা	C852.WAV
ষ্য	C853.WAV
ষ্যা	C854.WAV
স্য	C855.WAV
স্যা	C856.WAV
হ্য	C857.WAV
হ্যা	C858.WAV
ক্ষ্য	C859.WAV
ক্ষ্যা	C860.WAV
ড়্য	C861.WAV
ড়্যা	C862.WAV
ঢ়্য	C863.WAV
ঢ়্যা	C864.WAV
	C865.WAV
	C866.WAV
ঢো	C867.WAV
	C868.WAV
	C869.WAV
	C870.WAV

Sound Unit	File Name
কঁ	C1C.WAV
কেঁ	C2C.WAV
কাঁ	C3C.WAV
কোঁ	C4C.WAV
কিঁ	C5C.WAV
কীঁ	C6C.WAV
কুঁ	C7C.WAV
কূঁ	C8C.WAV
কৃঁ	C9C.WAV
কৈঁ	C10C.WAV
কৌঁ	C11C.WAV
খঁ	C23C.WAV
খেঁ	C24C.WAV
খাঁ	C25C.WAV
খোঁ	C26C.WAV
খিঁ	C27C.WAV
খীঁ	C28C.WAV
খুঁ	C29C.WAV
খূঁ	C30C.WAV
খৃঁ	C31C.WAV
খৈঁ	C32C.WAV
খৌঁ	C33C.WAV
গঁ	C45C.WAV
গেঁ	C46C.WAV
গাঁ	C47C.WAV
গোঁ	C48C.WAV
গিঁ	C49C.WAV
গীঁ	C50C.WAV
গুঁ	C51C.WAV

Sound Unit	File Name
গূঁ	C52C.WAV
গৃঁ	C53C.WAV
গৈঁ	C54C.WAV
গৌঁ	C55C.WAV
ঘঁ	C67C.WAV
ঘেঁ	C68C.WAV
ঘাঁ	C69C.WAV
ঘোঁ	C70C.WAV
ঘিঁ	C71C.WAV
ঘীঁ	C72C.WAV
ঘুঁ	C73C.WAV
ঘূঁ	C74C.WAV
ঘৃঁ	C75C.WAV
ঘৈঁ	C76C.WAV
ঘৌঁ	C77C.WAV
চঁ	C95C.WAV
চেঁ	C96C.WAV
চাঁ	C97C.WAV
চোঁ	C98C.WAV
চিঁ	C99C.WAV
চীঁ	C100C.WAV
চুঁ	C111C.WAV
চূঁ	C112C.WAV
চৈঁ	C113C.WAV
চৌঁ	C114C.WAV
চৃঁ	C115C.WAV
ছঁ	C126C.WAV
ছেঁ	C127C.WAV
ছাঁ	C128C.WAV

Sound Unit	File Name
ছোঁ	C129C.WAV
ছিঁ	C130C.WAV
ছীঁ	C131C.WAV
ছুঁ	C132C.WAV
ছূঁ	C133C.WAV
ছৈঁ	C134C.WAV
ছৌঁ	C135C.WAV
ছৃঁ	C136C.WAV
জঁ	C148C.WAV
জেঁ	C149C.WAV
জাঁ	C150C.WAV
জোঁ	C151C.WAV
জিঁ	C152C.WAV
জীঁ	C153C.WAV
জুঁ	C154C.WAV
জূঁ	C155C.WAV
জৈঁ	C156C.WAV
জৌঁ	C157C.WAV
জৃঁ	C158C.WAV
ঝঁ	C170C.WAV
ঝেঁ	C171C.WAV
ঝাঁ	C172C.WAV
ঝোঁ	C173C.WAV
ঝিঁ	C174C.WAV
ঝীঁ	C175C.WAV
ঝুঁ	C176C.WAV
ঝূঁ	C177C.WAV
ঝৈঁ	C178C.WAV
ঝৌঁ	C179C.WAV

Sound Unit	File Name
ঝৃঁ	C180C.WAV
টঁ	C193C.WAV
টেঁ	C194C.WAV
টোঁ	C195C.WAV
টাঁ	C196C.WAV
টিঁ	C197C.WAV
টীঁ	C198C.WAV
টুঁ	C199C.WAV
টূঁ	C200C.WAV
টৈঁ	C201C.WAV
টৌঁ	C202C.WAV
টৃঁ	C203C.WAV
ঠঁ	C215C.WAV
ঠেঁ	C216C.WAV
ঠাঁ	C217C.WAV
ঠোঁ	C218C.WAV
ঠিঁ	C219C.WAV
ঠীঁ	C220C.WAV
ঠুঁ	C221C.WAV
ঠূঁ	C222C.WAV
ঠৈঁ	C223C.WAV
ঠৌঁ	C224C.WAV
ঠৃঁ	C225C.WAV
ডঁ	C236C.WAV
ডেঁ	C237C.WAV
ডাঁ	C238C.WAV
ডোঁ	C239C.WAV
ডিঁ	C240C.WAV
ডীঁ	C241C.WAV

Sound Unit	File Name
ডুঁ	C242C.WAV
ডূঁ	C243C.WAV
ডৈঁ	C244C.WAV
ডৌঁ	C245C.WAV
ডৃঁ	C246C.WAV
টঁ	C258C.WAV
টেঁ	C259C.WAV
টাঁ	C260C.WAV
টোঁ	C261C.WAV
টিঁ	C262C.WAV
টীঁ	C263C.WAV
টুঁ	C264C.WAV
টূঁ	C265C.WAV
টৈঁ	C266C.WAV
টৌঁ	C267C.WAV
টৃঁ	C268C.WAV
ণঁ	C280C.WAV
ণেঁ	C281C.WAV
ণাঁ	C282C.WAV
ণোঁ	C283C.WAV
ণিঁ	C284C.WAV
ণীঁ	C285C.WAV
ণুঁ	C286C.WAV
ণূঁ	C287C.WAV
ণৈঁ	C288C.WAV
ণৌঁ	C289C.WAV
ণৃঁ	C290C.WAV
তঁ	C302C.WAV
তেঁ	C303C.WAV

Sound Unit	File Name
তাঁ	C304C.WAV
তোঁ	C305C.WAV
তিঁ	C306C.WAV
তীঁ	C307C.WAV
তুঁ	C308C.WAV
তূঁ	C309C.WAV
তৈঁ	C310C.WAV
তৌঁ	C311C.WAV
তৃঁ	C312C.WAV
থঁ	C324C.WAV
থেঁ	C325C.WAV
থাঁ	C326C.WAV
থোঁ	C327C.WAV
থিঁ	C328C.WAV
থীঁ	C329C.WAV
থুঁ	C330C.WAV
থূঁ	C331C.WAV
থৈঁ	C332C.WAV
থৌঁ	C333C.WAV
থৃঁ	C334C.WAV
দঁ	C346C.WAV
দেঁ	C347C.WAV
দাঁ	C348C.WAV
দোঁ	C349C.WAV
দিঁ	C350C.WAV
দীঁ	C351C.WAV
দুঁ	C352C.WAV
দূঁ	C353C.WAV
দৈঁ	C354C.WAV

Sound Unit	File Name
দোঁ	C355C.WAV
দৃঁ	C356C.WAV
ধঁ	C368C.WAV
ধেঁ	C369C.WAV
ধাঁ	C370C.WAV
ধিঁ	C371C.WAV
ধীঁ	C372C.WAV
ধুঁ	C373C.WAV
ধূঁ	C374C.WAV
ধৈঁ	C375C.WAV
ধোঁ	C376C.WAV
ধৃঁ	C377C.WAV
ধৌঁ	C867C.WAV
নঁ	C389C.WAV
নেঁ	C390C.WAV
নাঁ	C391C.WAV
নোঁ	C392C.WAV
নিঁ	C393C.WAV
নীঁ	C394C.WAV
নুঁ	C395C.WAV
নূঁ	C396C.WAV
নৈঁ	C397C.WAV
নৌঁ	C398C.WAV
নৃঁ	C399C.WAV
পঁ	C411C.WAV
পেঁ	C412C.WAV
পাঁ	C413C.WAV
পোঁ	C414C.WAV
পিঁ	C415C.WAV
পীঁ	C416C.WAV
পুঁ	C417C.WAV
পূঁ	C418C.WAV
পৈঁ	C419C.WAV
পৌঁ	C420C.WAV
পৃঁ	C421C.WAV
ফঁ	C433C.WAV
ফেঁ	C434C.WAV
ফাঁ	C435C.WAV
ফোঁ	C436C.WAV
ফিঁ	C437C.WAV
ফীঁ	C438C.WAV
ফুঁ	C439C.WAV
ফূঁ	C440C.WAV
ফৈঁ	C441C.WAV
ফৌঁ	C442C.WAV
ফৃঁ	C443C.WAV
বঁ	C455C.WAV
বেঁ	C456C.WAV
বাঁ	C457C.WAV
বোঁ	C458C.WAV
বিঁ	C459C.WAV
বীঁ	C460C.WAV
বুঁ	C461C.WAV
বূঁ	C462C.WAV
বৈঁ	C463C.WAV
বৌঁ	C464C.WAV
বৃঁ	C465C.WAV
ভঁ	C477C.WAV

Sound Unit	File Name
ভেঁ	C478C.WAV
ভাঁ	C479C.WAV
ভোঁ	C480C.WAV
ভিঁ	C481C.WAV
ভীঁ	C482C.WAV
ভুঁ	C483C.WAV
ভূঁ	C484C.WAV
ভৈঁ	C485C.WAV
ভৌঁ	C486C.WAV
ভৃঁ	C487C.WAV
মঁ	C499C.WAV
মেঁ	C500C.WAV
মাঁ	C501C.WAV
মোঁ	C502C.WAV
মিঁ	C503C.WAV
মীঁ	C504C.WAV
মুঁ	C505C.WAV
মূঁ	C506C.WAV
মৈঁ	C507C.WAV
মৌঁ	C508C.WAV
মৃঁ	C509C.WAV
যঁ	C521C.WAV
যেঁ	C522C.WAV
যাঁ	C523C.WAV
যোঁ	C524C.WAV
যিঁ	C525C.WAV
যীঁ	C526C.WAV
যুঁ	C527C.WAV
যূঁ	C528C.WAV

Sound Unit	File Name
যৈঁ	C529C.WAV
যৌঁ	C530C.WAV
যৃঁ	C531C.WAV
রঁ	C543C.WAV
রেঁ	C544C.WAV
রাঁ	C545C.WAV
রোঁ	C546C.WAV
রিঁ	C547C.WAV
রীঁ	C548C.WAV
রুঁ	C549C.WAV
রূঁ	C550C.WAV
রৈঁ	C551C.WAV
রৌঁ	C552C.WAV
লঁ	C559C.WAV
লেঁ	C560C.WAV
লাঁ	C561C.WAV
লোঁ	C562C.WAV
লিঁ	C563C.WAV
লীঁ	C564C.WAV
লুঁ	C565C.WAV
লূঁ	C566C.WAV
লৈঁ	C567C.WAV
লৌঁ	C568C.WAV
লৃঁ	C569C.WAV
শঁ	C581C.WAV
শেঁ	C582C.WAV
শাঁ	C583C.WAV
শোঁ	C584C.WAV
শিঁ	C585C.WAV

Sound Unit	File Name
শীঁ	C586C.WAV
শুঁ	C587C.WAV
শূঁ	C588C.WAV
শৈঁ	C589C.WAV
শৌঁ	C590C.WAV
শৃঁ	C591C.WAV
ষঁ	C603C.WAV
ষেঁ	C604C.WAV
ষাঁ	C605C.WAV
ষোঁ	C606C.WAV
ষিঁ	C607C.WAV
ষীঁ	C608C.WAV
ষুঁ	C609C.WAV
ষূঁ	C610C.WAV
ষৈঁ	C611C.WAV
ষৌঁ	C612C.WAV
ষৃঁ	C613C.WAV
সঁ	C625C.WAV
সেঁ	C626C.WAV
সাঁ	C627C.WAV
সোঁ	C628C.WAV
সিঁ	C629C.WAV
সীঁ	C630C.WAV
সুঁ	C631C.WAV
সূঁ	C632C.WAV
সৈঁ	C633C.WAV
সৌঁ	C634C.WAV
সৃঁ	C635C.WAV
হঁ	C647C.WAV
হেঁ	C648C.WAV
হাঁ	C649C.WAV
হোঁ	C650C.WAV
হিঁ	C651C.WAV
হীঁ	C652C.WAV
হুঁ	C653C.WAV
হূঁ	C654C.WAV
হৈঁ	C655C.WAV
হৌঁ	C656C.WAV
হৃঁ	C657C.WAV
ক্ষঁ	C668C.WAV
ক্ষেঁ	C669C.WAV
ক্ষাঁ	C670C.WAV
ক্ষোঁ	C671C.WAV
ক্ষিঁ	C672C.WAV
ক্ষীঁ	C673C.WAV
ক্ষুঁ	C674C.WAV
ক্ষূঁ	C675C.WAV
ক্ষৈঁ	C676C.WAV
ক্ষৌঁ	C677C.WAV
ক্ষৃঁ	C678C.WAV
ড়ঁ	C690C.WAV
ড়েঁ	C691C.WAV
ড়াঁ	C692C.WAV
ড়োঁ	C693C.WAV
ড়িঁ	C694C.WAV
ড়ীঁ	C695C.WAV
ড়ুঁ	C696C.WAV
ড়ূঁ	C697C.WAV

Sound Unit	File Name
ড়ৈঁ	C698C.WAV
ড়ৌঁ	C699C.WAV
ঢ়ঁ	C706C.WAV
ঢ়েঁ	C707C.WAV
ঢ়াঁ	C708C.WAV
ঢ়োঁ	C709C.WAV
ঢ়িঁ	C710C.WAV
ঢ়ীঁ	C711C.WAV
ঢ়ুঁ	C712C.WAV
ঢ়ূঁ	C713C.WAV
ঢ়ৈঁ	C714C.WAV
ঢ়ৌঁ	C715C.WAV

Sound Unit	File Name

APPENDIX C

Bangla symbols and their corresponding ASCII code number

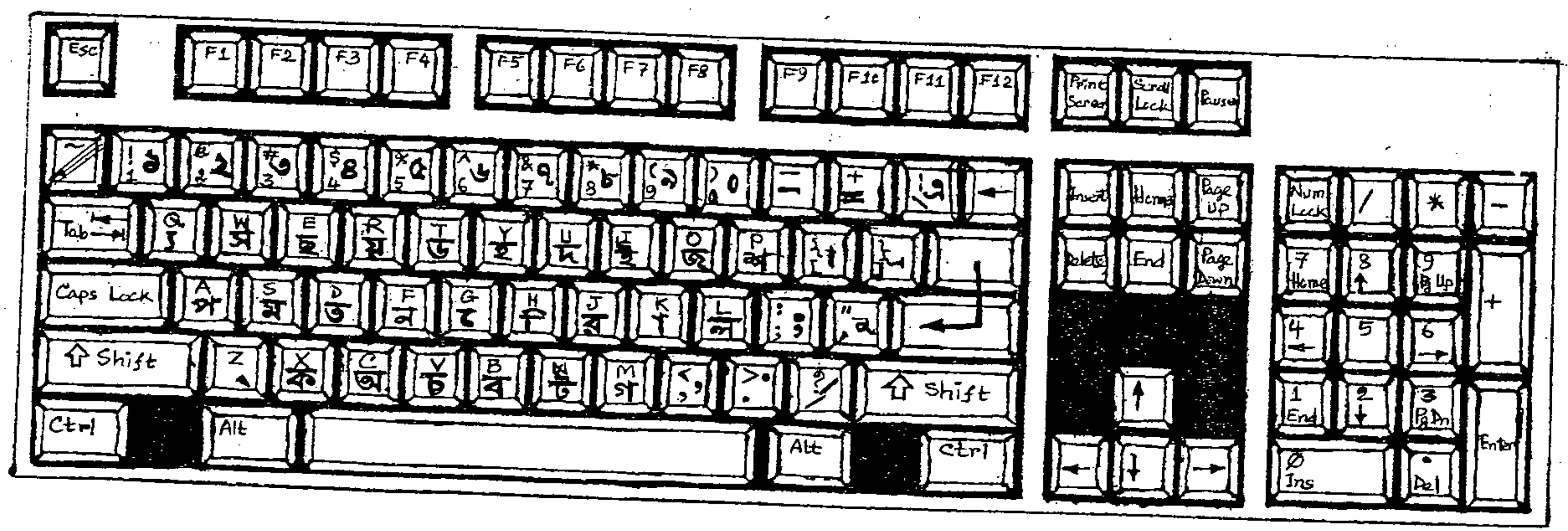
Bangla Symbol	ASCII Code
অ	99
আ	208
ই	105
ঈ	73
উ	84
ঊ	219
ঋ	205
এ	92
ঐ	213
ও	67
ঔ	214
ক	120
খ	35
গ	109
ঘ	77
ঙ	211
চ	118
ছ	101
জ	111
ঝ	209
ঞ	212
ট	110
ঠ	78
ড	116
ঢ	71
ণ	125
ত	100
থ	74
দ	117

Bangla Symbol	ASCII Code
ধ	72
ন	102
প	97
ফ	90
ব	98
ভ	75
ম	115
য	82
র	106
ল	108
শ	112
ষ	222
স	119
হ	121
য়	114
ং	69
ৎ	89
ঃ	58
ঁ	218
[illegible]	94
ে	103
া	107
ি	104
ী	81
-	39
[illegible]	220
ৈ	223
ৗ	221
্	38

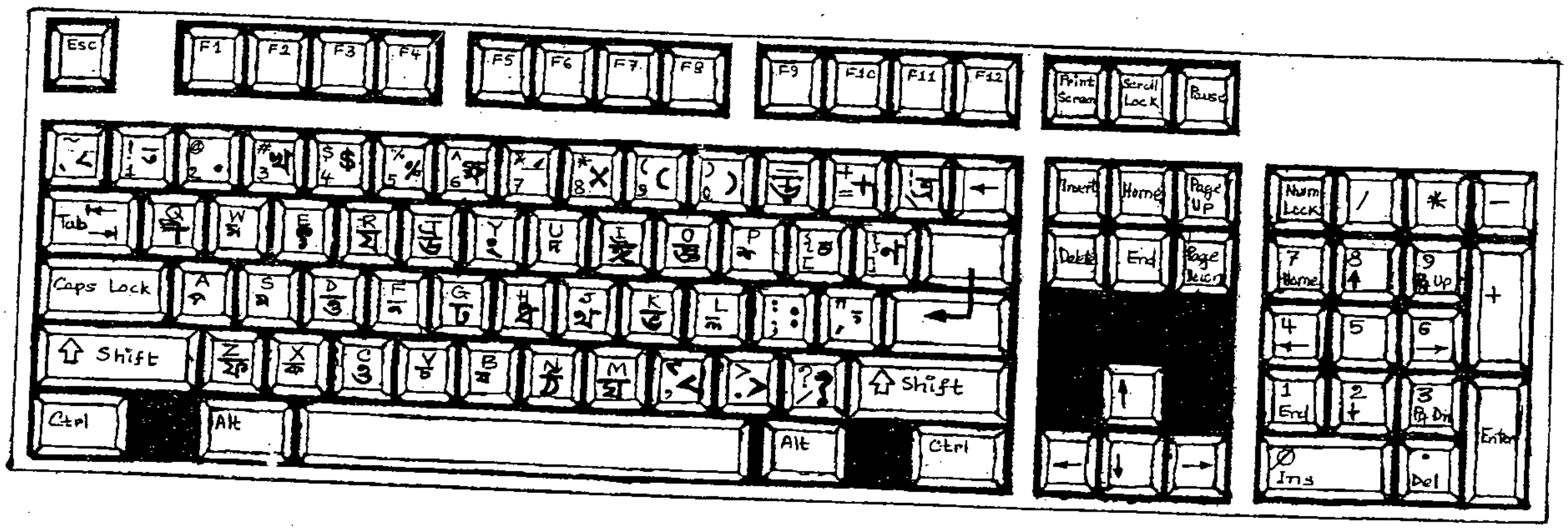
Bangla Symbol	ASCII Code
`	122
[illegible]	113
[illegible]	126
[illegible]	93
[illegible]	87
[illegible]	80
[illegible]	123
প	65
[illegible]	83
[illegible]	70
ব	216
ক	88
[illegible]	86
[illegible]	66
[illegible]	85
[illegible]	76
[illegible]	230
[illegible]	193
[illegible]	197
[illegible]	199
[illegible]	227
[illegible]	34
[illegible]	33
[illegible]	195
ত	68
থ	200
[illegible]	79
ত্র	124
[illegible]	206

Bangla Symbol	ASCII Code
[illegible]	196
[illegible]	202
[illegible]	192
[illegible]	201
[illegible]	224
[illegible]	207
ড়	95
[illegible]	203
য়	114
।	91
,	44
;	59
:	58
?	63

Keyboard (NORMAL KEY)



Keyboard (SHIFT KEY)



Keyboard (COMPOSE KEY)
/(RIGHT CONTROL KEY)