

Table of Contents

1	Introduction	3
2	Running the Application.....	3
3	Using Application	3
3.1	File Browse Button.....	4
3.2	Syllabify Text Button	5
3.3	Text Input Area	5
3.4	Clear Button	6
3.5	Output Area	6
3.6	Phoneme Input Area	6
3.7	Syllabify Phonemes Button	7
3.8	Sinhala Phonetic Guide	7
4	<i>Appendix A: Phoneme Representation Scheme</i>	8

1 Introduction

This document explains how to use the Sinhala Syllabification Tool (Demo version)¹. The software demonstrates the syllabification of Sinhala Unicode text, words or phonemes. Sinhala Grapheme-to-Phoneme conversion is also demonstrated by this tool.

2 Running the Application

After successful installation of software and all its dependencies, follow the steps below to start the application.

Go to, Windows **Start** > **All programs** > **Sinhala Syllabification Tool** and click on **Sinhala Syllabification Tool**.

3 Using Application

The start up screen of Sinhala Syllabification Tool is appeared as Figure 3.1. (The controls are described in the latter part of this manual by the labels referred in the figure.)

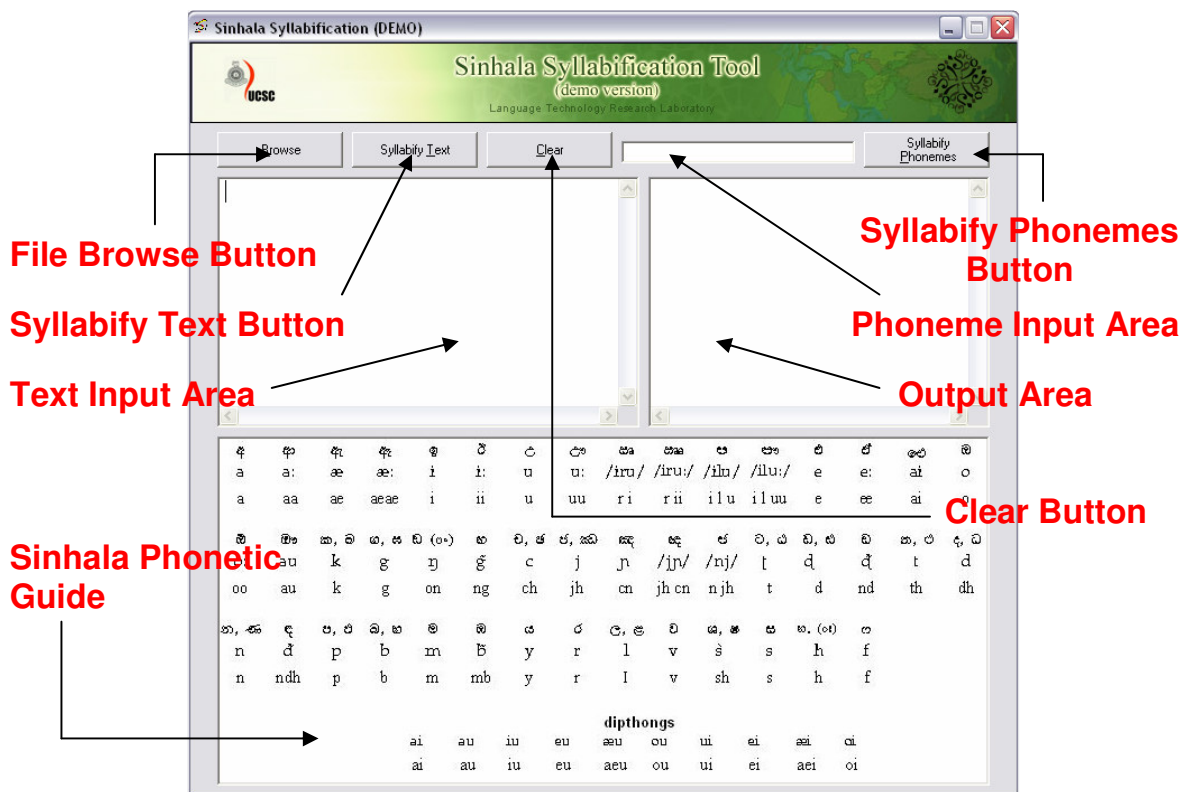


Figure 3.1: Sinhala Syllabification Tool

¹ The demo version illustrates the syllabification algorithm and letter-to-sound conversion algorithm used in the Sinhala Text to Speech system that is being developed by the Language Technology Research Laboratory of University of Colombo School of Computing.

3.1 File Browse Button

1. Use this button to open a file containing **Sinhala Unicode text** in to **Text Input Area**. (Software will only process Unicode Sinhala words in the given text. The characters, symbols and words typed in any other language are ignored). When clicked, the standard Windows File Open dialogue will appear as shown in Figure 3.2.

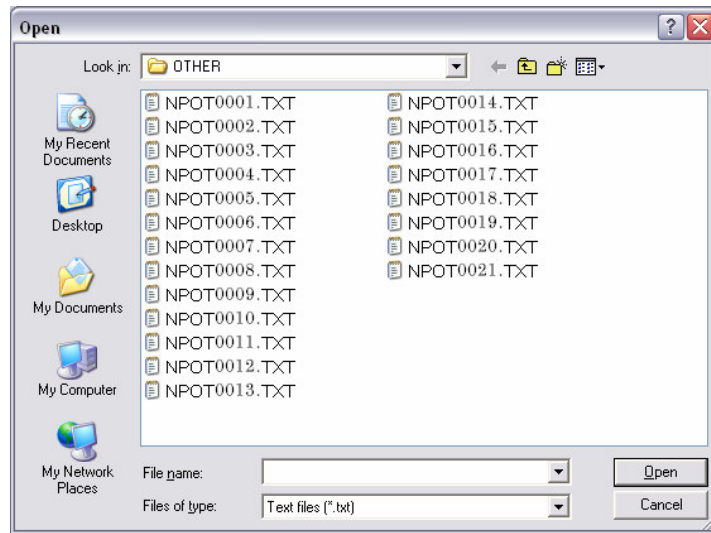


Figure 3.2: File Open Dialogue

2. Browse and select the file. Then click **open**. The content of the file will appear in the **Text Input Area** as shown in Figure 3.3.

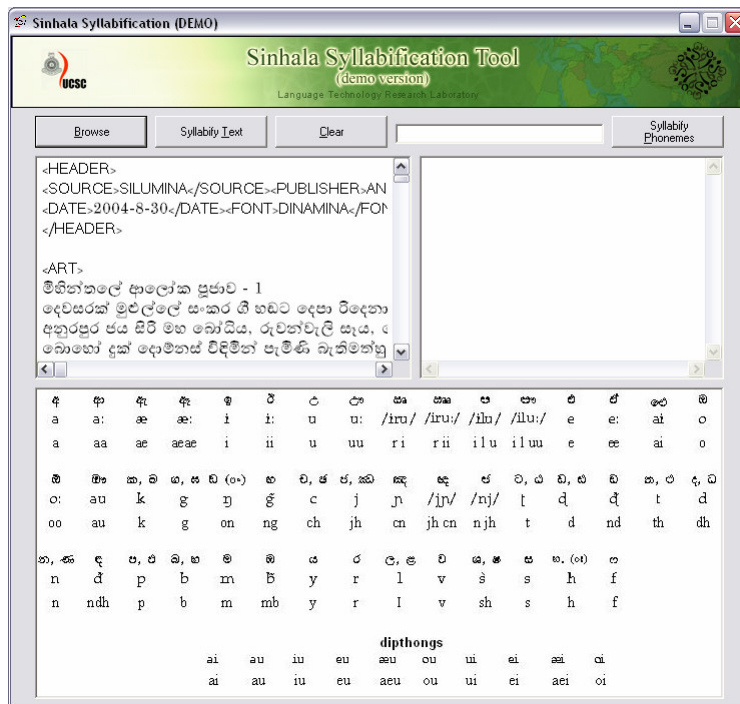


Figure 3.3: File loaded in to Text Input Area

3.2 Syllabify Text Button

By clicking this button, Sinhala Unicode words in the **Text Input Area** are examined by the phonetic analyzer built in to the software and produce the pronunciation form of each word along with the syllable boundaries marked by brackets in the **Output Area**. Please refer Figure 3.4.

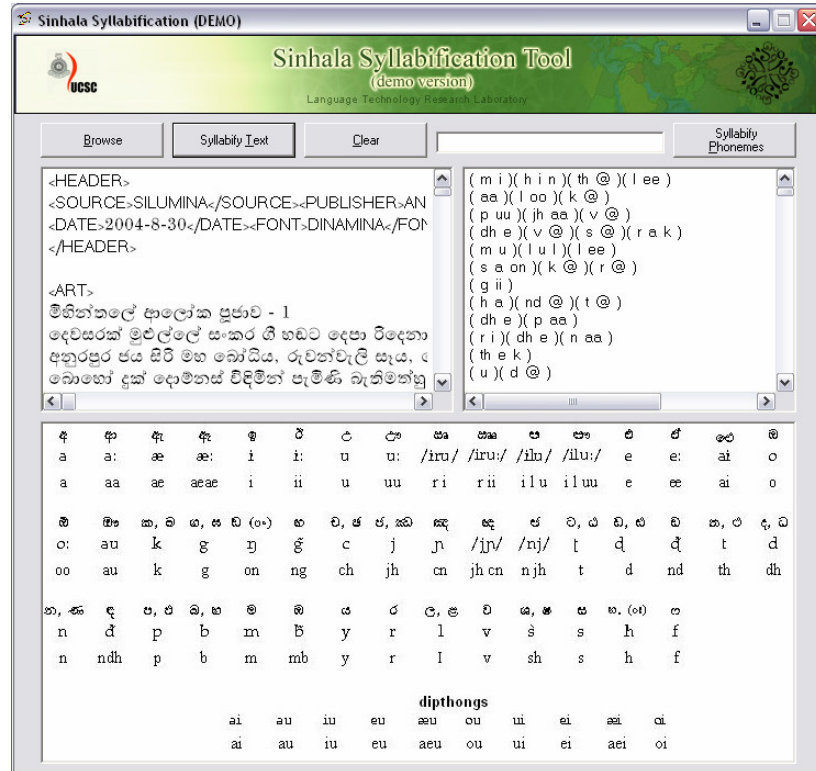


Figure 3.4: Syllabified words displayed in the Output Area

NOTE:

The phonetic notation used in the output is a modified version of ARPABET scheme which can be found on URL: <http://www.billnet.org/phon/arpabet.html>. The **Sinhala Phonetic Guide** in the bottom of the Sinhala Syllabification Tool describes how Sinhala alphabet is pronounced and the phoneme representation scheme used by this tool. Please refer section 3.8 Sinhala phonemic guide, and Appendix A: Phoneme Representation Scheme.

3.3 Text Input Area

In this area, either you can load Sinhala Unicode text in a file, as described in the section 3.1 or type Sinhala Unicode text by keyboard. After typing, click **Syllabify Text Button** to syllabify words and display the syllabified words in **Output Area**. Refer Figure 3.5.

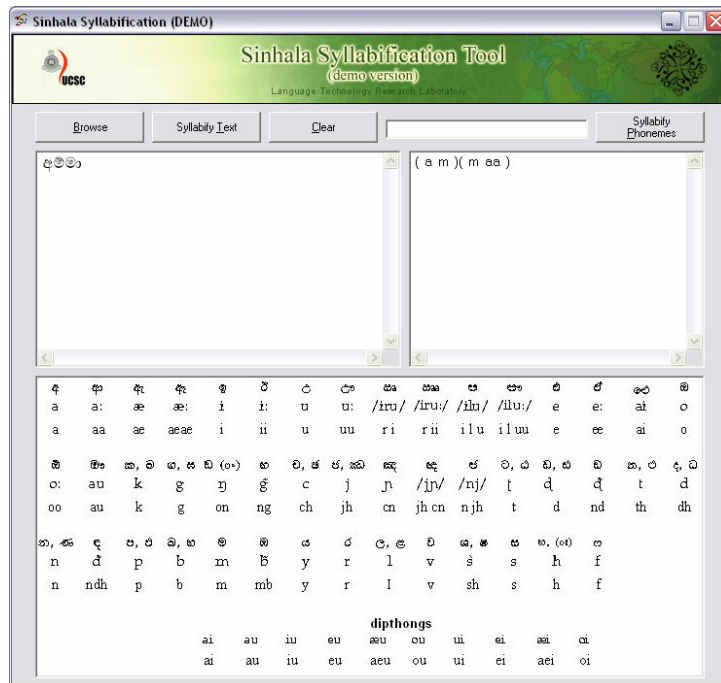


Figure 3.5: Use keyboard to enter text in the **Text Input Area**

3.4 Clear Button

Use this button to clear contents in both **Text Input Area** and **Output Area**.

3.5 Output Area

Output Area displays the syllabified words or phonemes. Here, the syllables are separated by brackets and phonemes² are separated by spaces. Each line displays a single syllabified word.
eg. (a m) (m aa).

NOTE:

The output can be copied into memory by right click on the **Output Area** and selecting **Copy**. The content can be then pasted in to any **Text Editor**.

3.6 Phoneme Input Area

Accepts a list of phonemes² (in a given word) separated by spaces.
eg. a m m aa

² The notation used to describe phonemes is presented in Appendix A.

NOTE:

Make sure no space appears at the end of the last phoneme.

3.7 Syllabify Phonemes Button

By clicking this button, the phonemes presented in the **Phoneme Input Area** are syllabified and displayed in the **Output Area**.

3.8 Sinhala Phonetic Guide

This guide describes how the Sinhala alphabet is pronounced. The pronunciation forms are given in IPA and notation described in Appendix A.

අ	ආ	ඇ	←	Sinhala letters
a	a:	æ	←	IPA
ා	aa	ae	←	Modified ARPABET scheme presented in appendix A.

4 Appendix A: Phoneme Representation Scheme

Sinhala Letter	Corresponding Phoneme(s)	
	IPA	Scheme used in this tool
අ	a	a
ආ	a:	aa
ඇ	æ	ae
ඈ	æ:	aeae
ඉ	i	i
ඊ	i:	ii
උ	u	u
ඌ	u:	uu
සා	/iru/	r i
සාා	/iru:/	r ii
ඊ	/ilu/	i l u
ඊා	/ilu:/	l l uu
එ	e	e
ඒ	e:	ee
ඒඒ	ai	ai
ඔ	o	o
ඔා	o:	oo
ඔා	au	au
ක, ඛ	k	k
ග, ඝ	g	g
ඛ (ච්ච)	ŋ	q
ඟ	g̃	ng
ච, ඡ	c	ch
ජ, ඣ	j	jh
ඤ	ɲ	cn
ඤ	/jɲ/	jh cn
ඤ	/nj/	cn jh
ට, ඨ	t	t
ඩ, ඩ	d	d
ඩ	ḍ	nd
ණ, ට	t	th
ඳ, ධ	d	dh
න, ඹ	n	n
ඳ	ḍ	ndh
ප, ඵ	p	p
බ, භ	b	b

Appendix A: Continued..

Sinhala Letter	Corresponding Phoneme(s)	
	IPA	Scheme used in this tool
ම	m	m
ඹ	ɓ	mb
ය	y	y
ර	r	r
ල, ළ	l	l
ව	v	v
ශ, ෂ	ʃ	sh
ස	s	s
හ, (ඔ)	h	h
ඞ	f	f