Guide to creating Sinhala and Tamil Unicode fonts

Preface

ICTA has identified the enabling of ICT in local languages, Sinhala and Tamil, as a priority area in the e-Sri Lanka initiative. ICTA is promoting the use of ICT in Sinhala and Tamil, and is addressing issues relating to enabling ICT in local languages such as the relevant standards with regard to encoding, keyboard layouts, collation sequences and also the availability of standards based (Unicode compliant) fonts.

With the advent word processing and desktop publishing there was an increase in the development of English fonts. Now there is a wide choice with regard to standards based fonts in English. But this is not so for Unicode compliant Sinhala fonts. Therefore ICTA has been promoting the development of Unicode compliant local language fonts. Several Unicode fonts such as MalithWeb, Dinamina initially were developed. Subsequently ICTA ensured the development of the Unicode Sinhala font Bhashitha and the Unicode Tamil font Sri Tamil. The font rules of these two fonts are given free to font developers.

But one of the impediments encountered in getting the fonts developed was that there was a dearth of proficient font developers, and it was necessary to get developers to develop Unicode compliant fonts. There was a necessity to develop awareness among font developers on how to create good quality standards based fonts. The necessity to develop awareness encompassed both the technical area covering font rules, and also designing.

In view of this ICTA, in partnership with the University of Colombo School of Computing (UCSC) implemented a training program with the objective of providing font developers with the knowledge and expertise on how to develop standards based, aesthetically correct local language fonts of good quality, in order to upgrade the general level of skill and knowledge in this area.

The training was conducted at the UCSC and was in the form of interactive workshops. An eminent panel of resource persons conducted the training. Participants were required to give a commitment that they would develop a portfolio of at least two glyphs.

As a result of this training there are font developers able to develop Sinhala fonts, several stylized Sinhala fonts have been developed by the trainees, and the proceedings of the training have been compiled into this book. This book fills a lacuna which has been felt in the font industry. It is expected that local language font developers will find this useful.

Reshan Dewapura,
Chief Operating Officer
Information and Communication Technology Agency of Sri Lanka (ICTA)
November 2010.

Resource team

Prof Gihan Dias

Dr. Ruvan Weerasinghe

Mr. Harsha Wijayawardhana

Prof. Rohini Paranavithana

Ms. Dineesha Ediriweera

Mr. G. Balachandran

Mr. Winnie Hettigoda

Mr. Sinnathambi Shanmugarajah

Mr. Chamara Dissanayake

Table of Contents

Introduction

Sinhala was encoded in Unicode in the latter part of 1990s. The encoding process was initially spearheaded by the Council for Information Technology (CINTEC) and by Prof. V.K. Samaranayake, Chairman of CINTEC and the founder director of University of Colombo School of Computing (UCSC), Prof. J. B. Disanayaka, presently Sri Lanka's Ambassador to Thailand and Professor emeritus of Sinhala, University of Colombo, and by Mr. Nandadsara of UCSC. Although the encoding was carried out in the 1990s, it took another seven years for a proper implementation of Sinhala Unicode to take place

The Information and Communication Technology Agency of Sri Lanka (ICTA) operational from July 2003, realized the importance of the local language initiative and agreed that this area should be taken over, given priority and promoted.

The objective was to ensure that the benefits of ICT should be taken to the the population in Sri Lanka most of whom, if given a choice, would prefer to use ICT in Sinhala or Tamil. Hence the Local Language Working group the members of which had with expertise and experience in this area was established.

Thereafter, joint research was carried out into Sinhala Unicode by UCSC and the Department of Computer Science and Engineering of the University of Moratuwa, under the aegis of ICTA. The input method for Sinhala Unicode was standardized in 2004 by the Sri Lanka Standards Institution (SLSI) as the Sri Lanka Standard Sinhala Character Code for Information Interchange SLS 1134 : 2004. The main author of this standardization document was Prof. Gihan Dias and the other two main researchers during this phase of the Sinhala Unicode implementation were Dr. Ruvan Weerasinghe and Mr. Harsha Wijayawardhana, both from UCSC.

With the release of Uniscribe (Shaping engine) of Microsoft in 2003-5, Microsoft introduced Iskoola Potha. ICTA's Local Language Initiative, under the direction of Prof. Gihan Dias spearheaded the development of Unicode compliant Sinhala fonts. ICTA held several awareness sessions on Unicode for Font Developers. During this period, Mr. Anura Tissera formerly of the Associated Newspapers of Sri Lanka (Lake House) created the "Dinamina" font. The development of several other fonts followed: Prof. Gihan Dias, Mr. Harsha Wijayawardhana, and the Linux User Group released Malithi, Sarasavi (Unicode) and LKLUG fonts respectively. All these fonts were results of research carried out into Unicode font creation in Sinhala and Tamil.

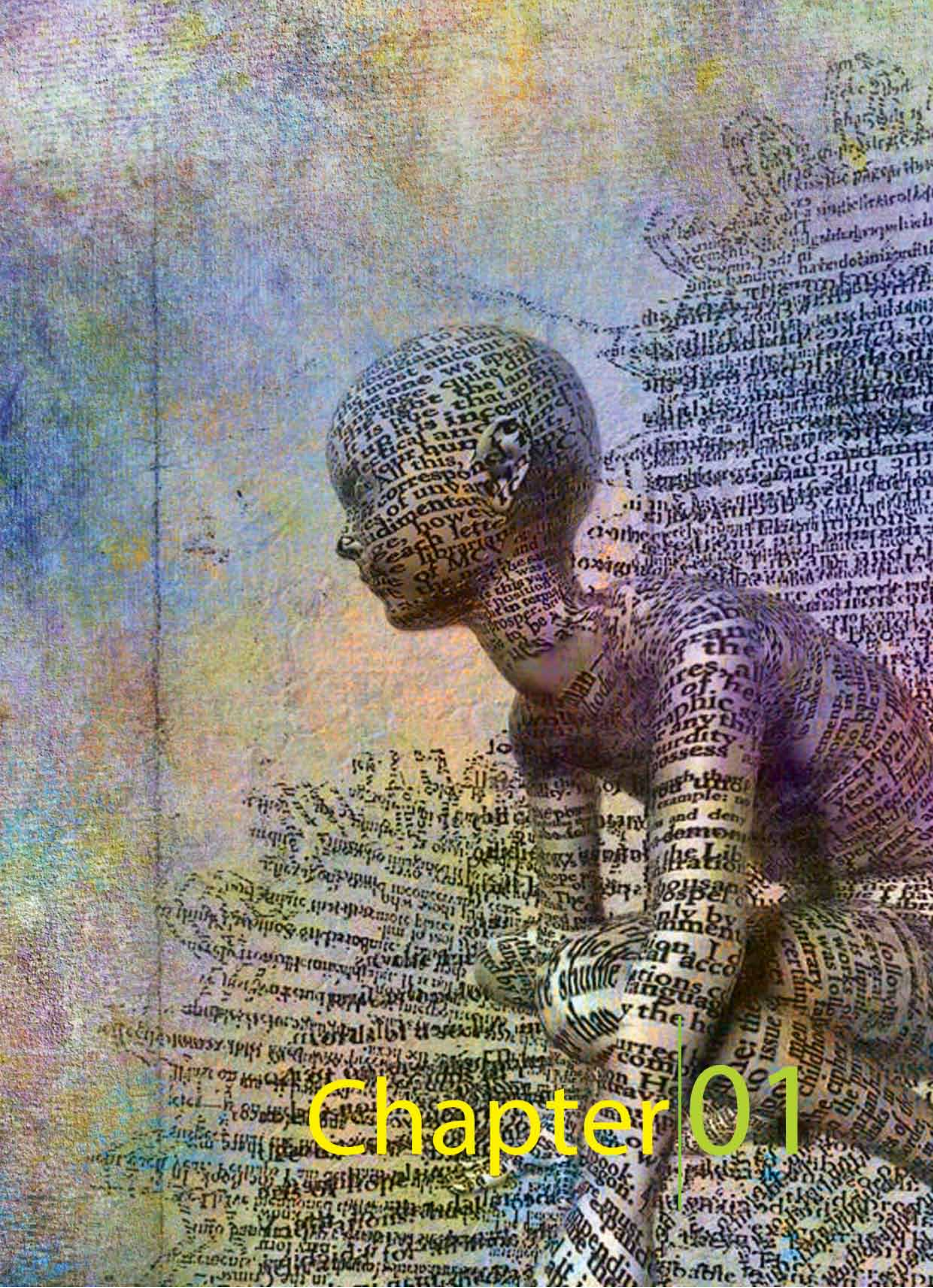
Mr. Sinnathambi Shanmugarajah thereafter developed a Sinhala "kit" incorporating standard Sinhala fonts, keyboard driver and installation instructions, which was made available free of charge on the site www.fonts.lk.

With the wider use of Sinhala and Tamil Unicode, the dearth of Sinhala and Tamil Unicode fonts had been identified as a major issue and as an obstacle for popularization of Sinhala and Tamil Unicode content. In order to remedy the above, ICTA initiated several projects.

Thereafter ICTA ensured the development of the Unicode Sinhala font Bhashitha. The objectives were to to make available freely 'Correct Sinhala SLS 1134 : 2004 compliant font rules, to generate accurate Sinhala documents with the use of full complement of characters facilitated in the standard SLS 1134 : 2004, to discourage the use of legacy Fonts that are not SLS 1134 : 2004 compliant and to enable Old Sinhala / Pali (written in Sinhala script) books and content to be published on the Internet for worldwide audiences.. Ms. Dineesha Ediriweera and Mr. E.D. Pemasiri developed the font rules and the font glyphs respectively for the Bhashitha font.

This training program was held to train font developers, on developing Unicode compliant, aesthetically correct fonts of good quality.

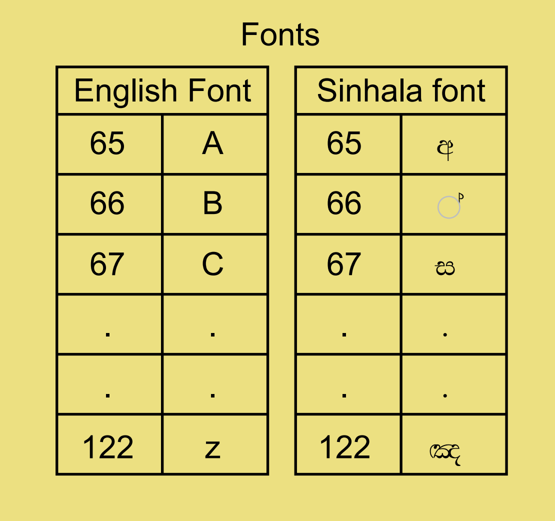
For this training program held at the UCSC, most of the pioneers in Sri Lanka's Unicode research in Sinhala and Tamil came together and it was coordinated at ICTA by Ms. Aruni Goonetilleke, program manager of ICTA.



Chapter 01

WHAT IS UNICODE

Dr. Ruvan Weerasinghe, University of Colombo School of Computing Computers store data as 1 and 0 and a decimal number expressed in 1 and 0 is known as a binary number. Characters are also stored as binary numbers in computers. The Latin script was the original script which was used in handling data in computers and subsequently the need for storing data in other languages became one of the major researches prior to the 1990s. Although the problem was solved in the pre-Unicode era by mapping numbers of English characters in a font to characters of another language (shapes of English characters were replaced by character shapes of another language), the need arose for storing data in different languages in a single document. When a document was prepared in two different languages, only one language could be viewed and the other appeared as garbage. Also in data transmission from one computer to another, it was essential to send the font which was used, with the document. Sorting of data was not possible since data was sorted according to English or any other Roman Script based language in the pre-Unicode era. By the early 1990s, Unicode was suggested as a possible solution to the above problems.

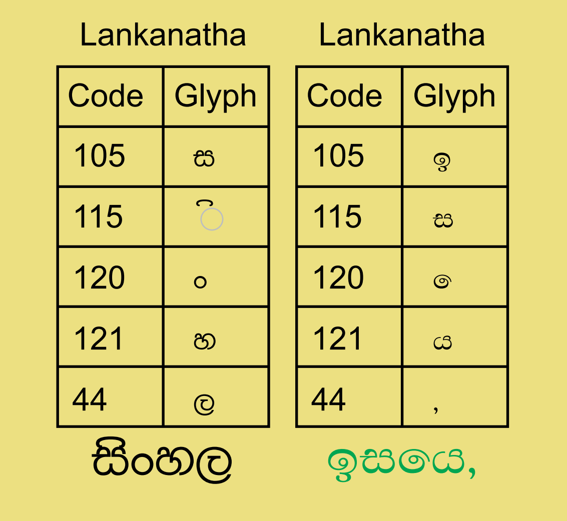


Fonts

English Font	
65	A
66	B
67	C
.	.
.	.
122	z

Sinhala font	
65	අ
66	◌ැ
67	ස
.	.
.	.
122	ඤ

Figure 1.1 Representing Sinhala and English in ASCII



Lankanatha

Code	Glyph
105	ස
115	◌ි
120	ං
121	හ
44	ළ

සිංහළ

Lankanatha

Code	Glyph
105	ඉ
115	ස
120	ෙ
121	ය
44	,

ඉසෙය,

Figure 1.2 Conflicts of using non Unicode Fonts

Unicode provides a unique number for every atomic character in different languages. This number is usually stored as two bytes inside the computer. The first byte of the two byte code of a character provides language identification and the second byte gives character identification in that language. Unicode is designed to work independent of platform and of the application program which uses Unicode font.

At present, more than fifty unique scripts have been encoded in Unicode.

i. The Unicode consortium

The Unicode consortium is a non-profit organization based in California, USA to develop, extend and promote the use of the Unicode Standard in the world. The membership of the Unicode consortium includes countries, large corporations and organizations in the computer and information processing industry. The consortium is financially supported solely by membership fees at present.

ii. Sinhala Unicode

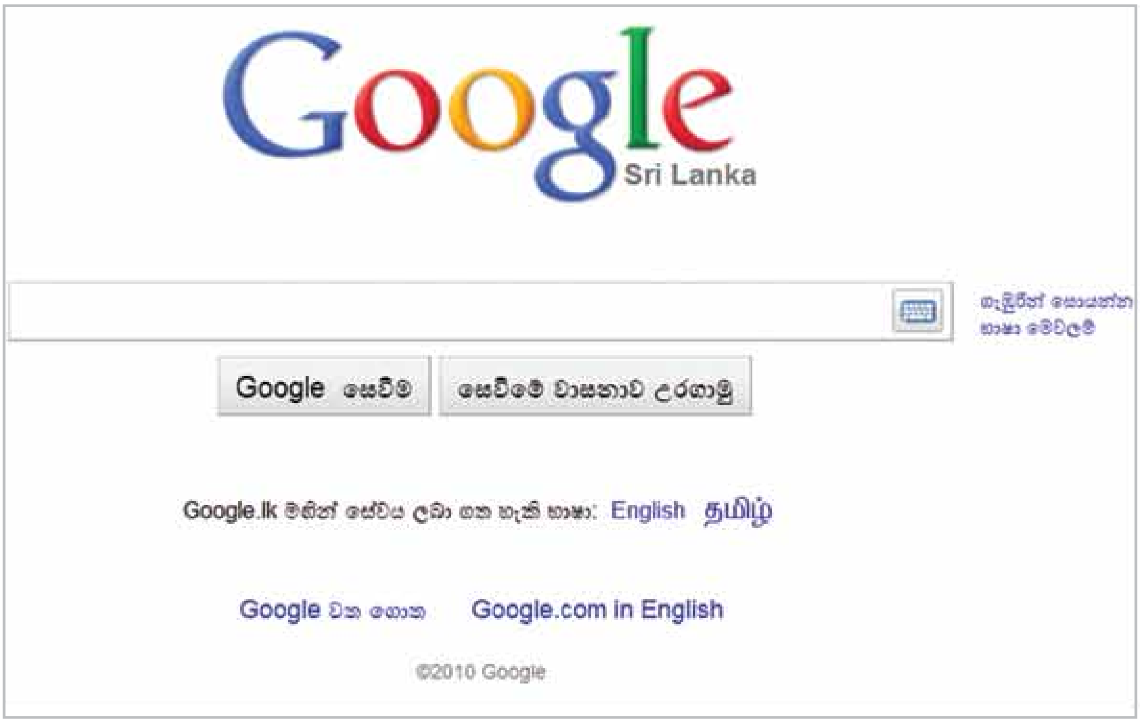
Computer literacy in Sri Lanka has overtaken English literacy and at present it is at 10%. It is now widely believed that the Computer literacy is at 20-25% in Sri Lanka. Localization had been a priority in Sri Lanka since the 1980s and Sinhala was encoded in Unicode in 1997 although the implementation of Sinhala Unicode was not initiated until 2003 in earnest. Sinhala, unlike English, is categorized under complex scripts. Sinhala belongs also to Phonetical languages or sound based languages and all Indic languages are categorized under this category. Sinhala, like other Indic languages, can make a single glyph with a consonant and vowel modifier. A glyph, always in the English language has a single code point and represents a single character. Unlike English, Sinhala glyphs can either have a single code point or a number or many code points or numbers.

In order for Sinhala Unicode to function properly, it requires the following:

1. Sinhala Unicode font
2. Shaping Engine such as Uniscribe for MS Windows, Pango Linux etc.
3. Sinhala Unicode Keyboard driver for input and this is not necessarily required for Unicode text rendering.

Tamil	2944-3071
Telugu	3072-3199
Kannada	3200-3327
Malayalam	3328-3455
Sinhala	3456-3583
Thai	3584-3711
Lao	3712-3839
Tibetan	3840-4095
Myanmar	4096-4255
Georgian	4256-4351
Hangul Jamo	4352-4607
Ethiopic	4608-4991
Ethiopic Supplement	4992-5023
Cherokee	5024-5119

Figure 2.3 Sinhala Unicode Range
Some Applications which use Sinhala Unicode fonts



Google Sri Lanka

ගැඹුරින් සොයන්න
භාෂා මෙවලම්

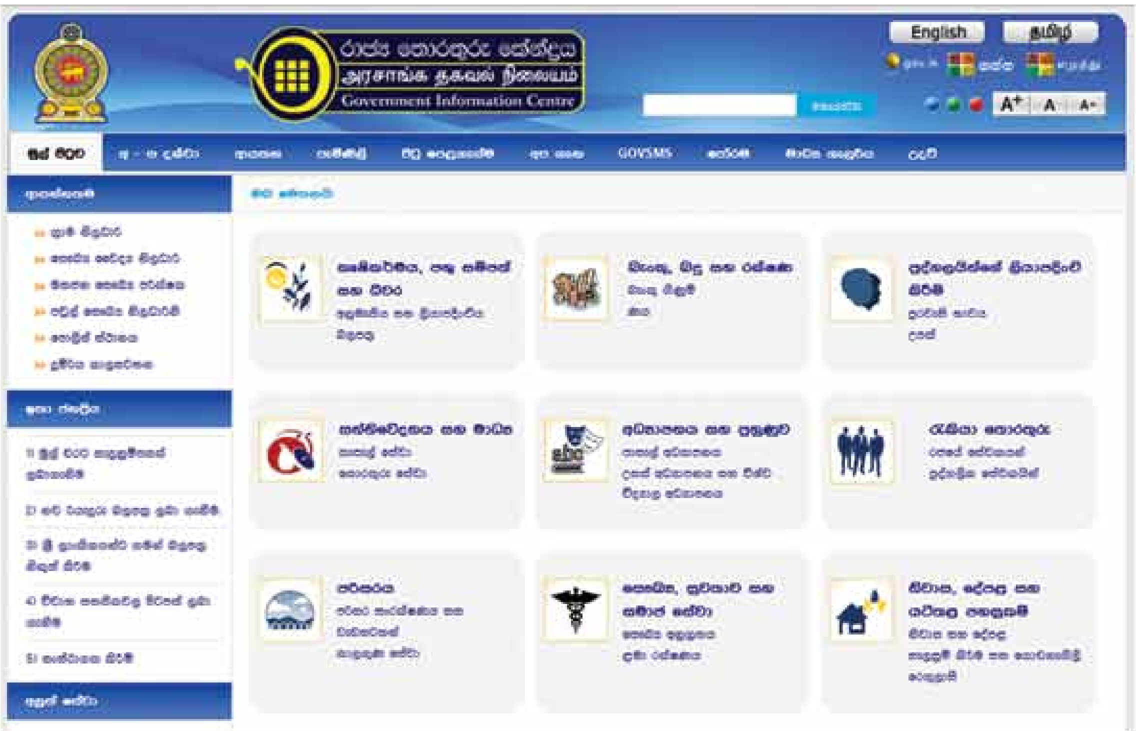
Google සෙවීම | සෙවීමේ වාසනාව උරගාමු

Google.lk මඟින් සේවය ලබා ගත හැකි භාෂා: English தமிழ்

Google වන ගොන   Google.com in English

©2010 Google

Figure 2.4 Google in Sinhala



රාජ්‍ය තොරතුරු කේන්ද්‍රය
அரசாங்க தகவல் நிலையம்
Government Information Centre

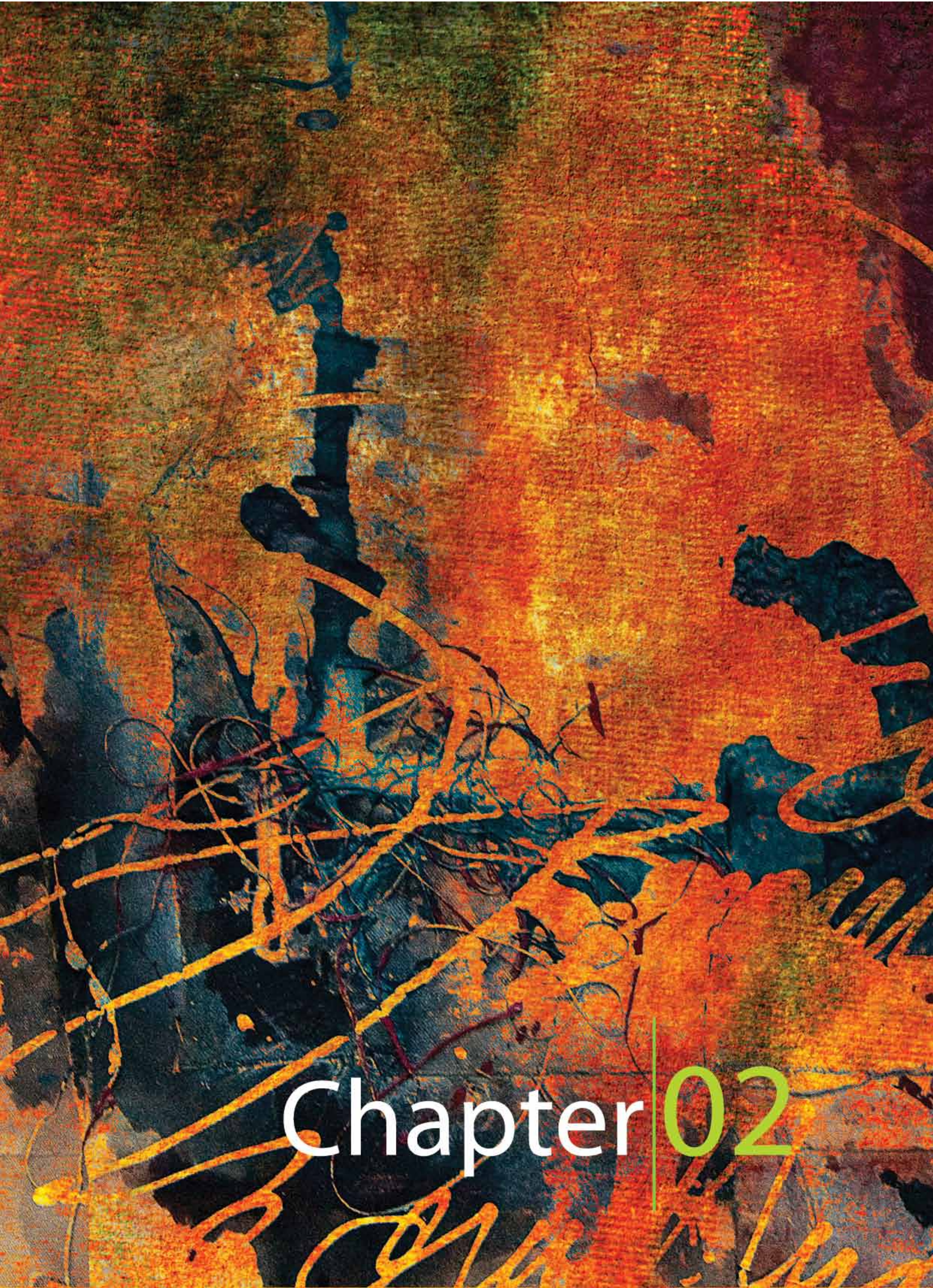
English தமிழ்

GOVSMS

A+ A A-

පරිසරය

Figure 2.5 A Sinhala website



Chapter 02

THE DEVELOPMENT OF THE SINHALA SCRIPT AND THE CHARACTERISTICS OF THE LETTERS

Professor Rohini Paranavitana
Department of Sinhala, University of Colombo.

Language is the basis of human communication and is an essential feature of human life and differentiates human beings from other living beings. Language thus interacts with every aspect of human life and enables one to exchange human thoughts with one another. As it is difficult to remember everything, which we have spoken, heard or thought, it is necessary to record them in writing, not only for the benefit of oneself but for the benefit of others in society and to transmit them further for future generations. Hence, writing paves way for communicatiion with people far in place and time. An alphabet is a set of standardized visual symbols, used to record/write the sounds of a language. The sounds and their symbols thus produce a perfect language. Therefore, the relationship between speech and written symbols is similar to the relationship between a human and his photograph.

Sinhala is the language spoken by the Sinhalese who constitute 75% of the island's population. The earliest extant Sinhala writings belong to the 3rd century B.C.E. These writings are to be found in inscriptions in rock caves of Mihintale, situated 8 kilometers from the ancient city of Anuradhapura. The earliest inscription there refers to a king by the name of Uttiya, the younger brother of King Devanampiya Tissa, who reigned in Sri Lanka in the 3rd century, B.C.E. These are called Brahmi inscriptions, because they are written in the script known as Brahmi. With the introduction of Buddhism to the island in the 2nd century B.C. by Arhant Mahinda, who was the son of Emperor Asoka in India, the inhabitants in the island became devotees of the Buddhism.

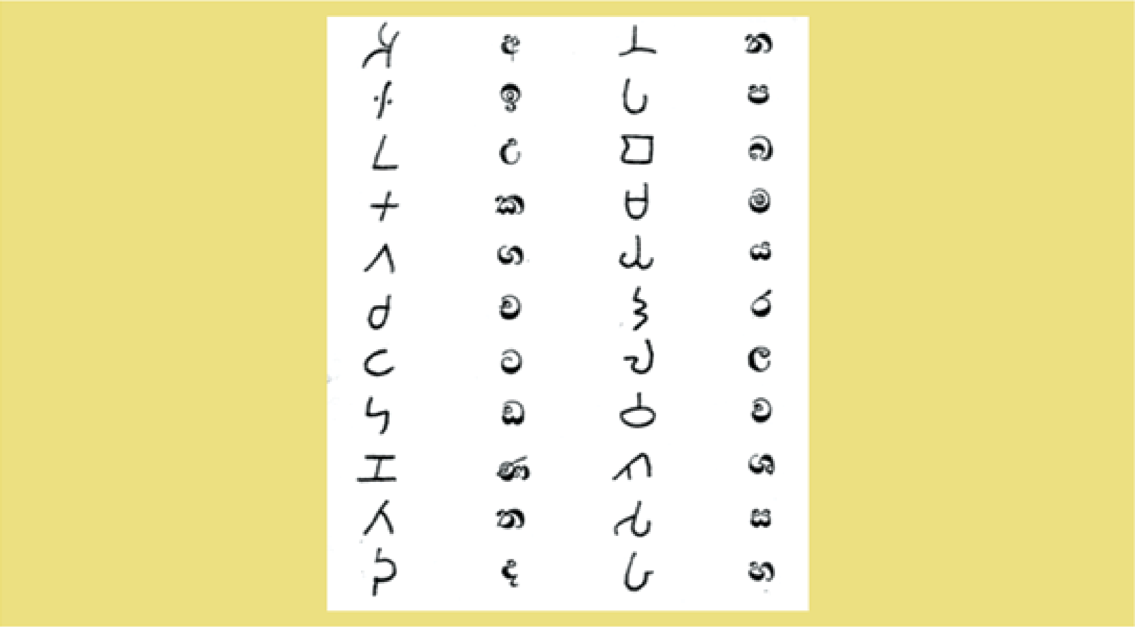


Fig 1□) Brahmi Script

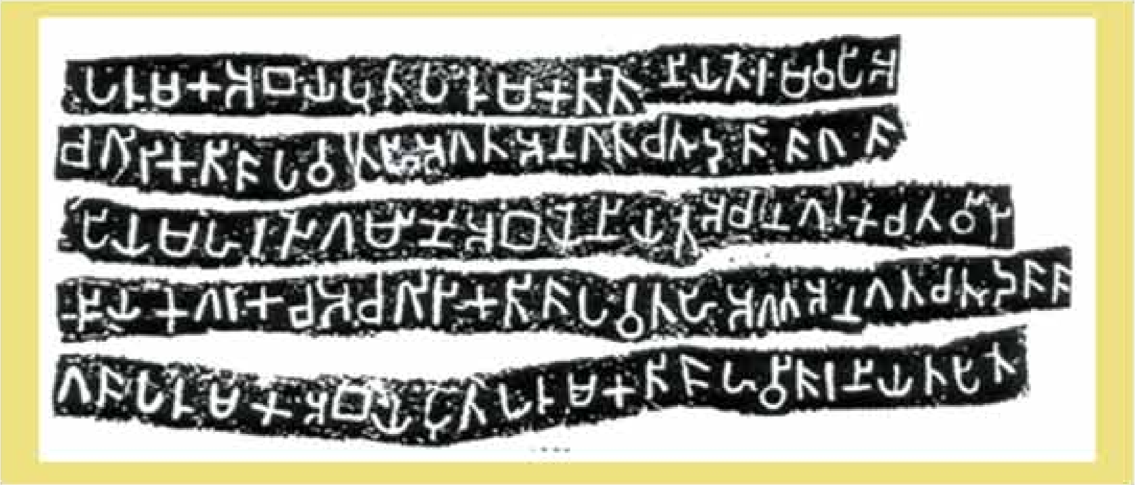


Fig2 :- Inscriptions (3rd century B.C.)2nd century A.D.

The primary task of laypersons was to provide food, lodging and medicine to the Buddhist monks. The content of the most of these inscriptions are records of the donations made to the Buddhist monks by their devotees. The names of the donors have been inscribed on the drip ledges of the caves, and were not only the names of the royal family but were also names of people from all walks of life. The Buddhist monks lead a simple life and they prefered to take their shelter in the caves, which were prepared by the devotees to suit their austere life. The earliest written records existing in the island today cannot be assigned to a period earlier than the 3rd century B.C.E. Archaeologists have argued with evidence that this script was brought from India even before the arrival of Arhant Mahinda. It has been found, that the earliest examples of Brahmi writings extant in the island contain some peculiar features, which are not noticed in the inscriptions of Emperor Asoka of India, who had extensively used Brahmi script to inscribe his discourses.

The earlier inscriptions written briefly, became longer and complicated with the development of the subject matter recorded in them such as regulations and administrative matters relating to the monasteries. As a result the Sri Lankan alphabet too began to invent its own features by the 4th century C.E. Over the ages, it accumulated new features and by the 8th century the long and short æ and nasal ng, nd, mb became evident in the Sinhala alphabet making them unique. The vowel sound æ had a short and a long vowel pronunciation, short æ being similar to the sound emanated when pronouncing English at and the long æ being similar to English ant.

The Aryan settlers, who migrated to Sri Lanka from time to time from India, brought with them the knowledge and methods of writing, which were in use in their motherland at that time. At the time of the arrival of Arhant Mahinda, the language of the island had attained a developed state for expressing a philosophy of highly complex thought to be given to people by way of free expression. Had the language not been developed, it would have been impossible to write commentaries for explaining the tenets of Buddhism during the lifetime of Arhant Mahinda. The

Pali Tripitaka and its commentaries were translated into Sinhala, the language of the inhabitants of the island. It is these commentaries, which were later translated into Pali by Buddhaghosha Thera, who arrived in Sri Lanka from South India in search of the Buddhist doctrine in the 5th century A.D.

There are hundreds of inscriptions representing the development of the Sinhala script which belong to every half a century of the Anuradhapura and Polonnaruwa periods. This unbroken series of epigraphic records indicates the evolution of the Sinhala Script. In the beginning, the subject matter and the number of letters used in the inscriptions were limited. In the evolution of the Sinhala Brahmi script in Sri Lanka several important stages have been mentioned. The first stage consists of the period from 3rd century B.C.E. to the end of the 4th century C.E. Subsequently at the second stage the letters began to take a slightly a circular shape. In the eighth century, it came under the influence of the South Indian Pallava Grantha script which influenced very much in taking the Sinhala script towards the modern form.

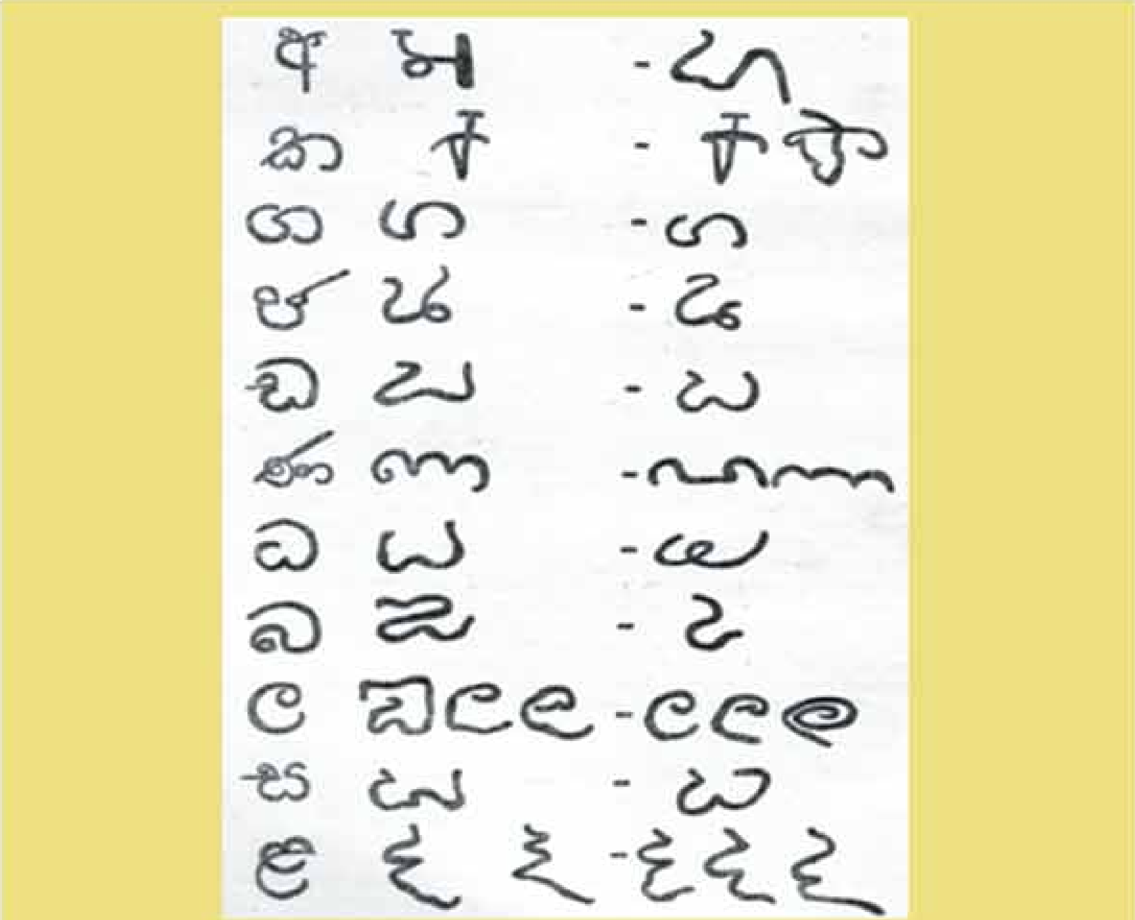
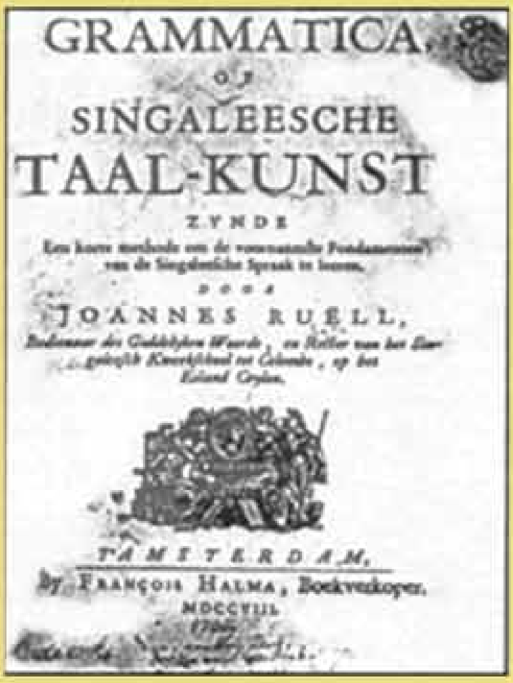


Fig 3:- Pallava - Sinhala akuru
Granthakshara(7-8 A.D.)   (8 century A.D.)

The types of material used for writing also had a tremendous impact on the evolution of the script. In the beginning stone would have been the most popular material on which the inscriptions were engraved. Stone was replaced by palm leaves by the eighth century. The texture of this material helped very much to move the stylus to write letters in a more circular manner in order to avoid the leaf being damaged by angular strokes. Sinhala was in use for writing various subjects with the influence from Pali and Sanskrit languages. In Sri Lanka, the compilation of books witnesses their origins from very early times. It has been suggested by S. Paranavitana that writing on palm leaves commenced in the 8th century C.E. but P. E. E. Fernando holds the view that the practice of using palm leaves as writing material was known in Sri Lanka even in the 5th century the latest.

When writing on palm leaves became popular, the books transmitted orally came to be written down on palm leaves . Henceforth, every composition except epigraphic records was written on palm leaves. The use of palm leaf as the material of writing was a great impetus for the improvement of literature as well as for the spread of learning. The palm leaf manuscripts succumbed to various forces of destruction and the oldest manuscripts available today are believed to have belonged to 13th century C.E. (Dambadeniya period). One of the manuscripts, the Cullavagga, is deposited in the Colombo Museum Library and the other manuscript, the Visuddhimagga tika, is available at the Peradeniya University Library. The majority of the manuscripts belong to the period of King Kirti Sri Rajasingha, who reigned in Kandy in the 18th century. As a result of the Buddhist and literary renaissance headed by venerable Weliwita Sri Saranankara thero, majority of the manuscripts, which were deposited in the temple libraries were copied and distributed among temples.

Printing of Sinhala books commenced in 1737 during the Dutch administration of the littoral of the island. A book on the grammar of the Sinhala language was published in Dutch by Johannes Ruell and was printed in Holland in 1707, even before the printing was introduced in the island. It contains examples from Sinhala language for various grammatical rules and those examples were printed in Sinhala characters. This can be presumed as the first time that Sinhala characters were printed in a book.



GRAMMATICA, OF SINGALEESCHE TAAL-KUNST

ZYNDE

Een korte methode om de voornaamste Fundamenten van de Singaleesche Spraak te leeren,

DOOR

JOANNES RUELL,

Bedienaar des Goddelyken Woords, en Rektor van het Eer- gelyk Kweekschool te Colombo, op het Eiland Ceylon.

T'AMSTERDAM,

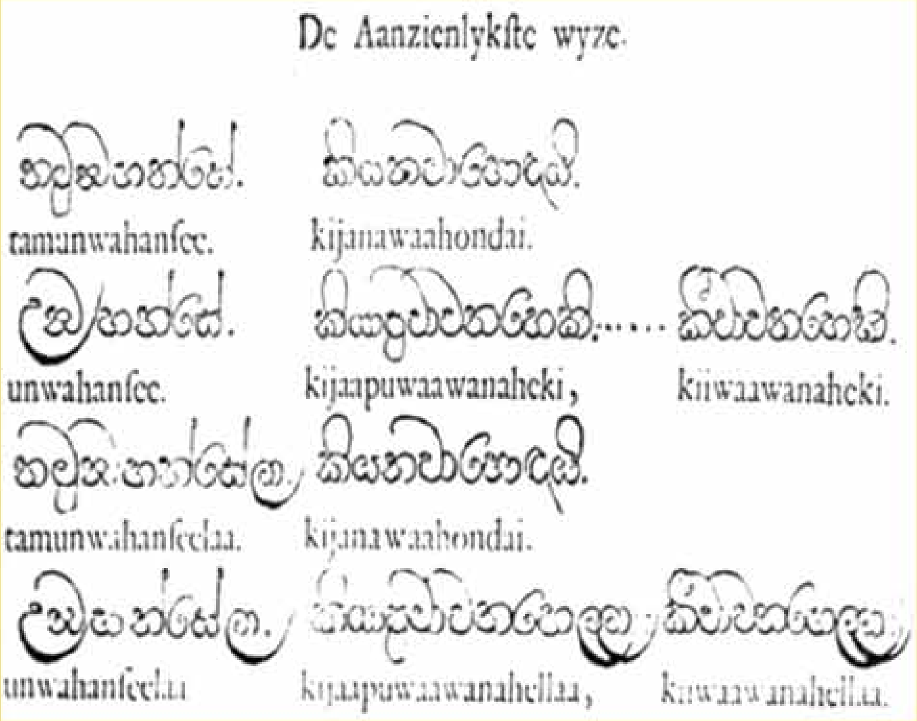
By FRANÇOIS HALMA, Boekverkoper.

MDCCVIII.

1708

1708 ඕලන්දයේ දි සිංහල අච්චු අකුරු තනා මුද්‍රිත යොහැන්නස් රුවල්ගේ
සිංහල ව්‍යාකරණ පොතේ ග්‍රන්ථනාම පිටුව

Fig 4:- Cover Page of the Grammatica of Singaleesche Taal)kunst (J. Ruell)



De Aanzienlykſte wyze.

තමුන්වහන්සේ. කියනවාහොඳයි.
tamunwahanſee. kijanawaahondai.

උන්වහන්සේ. කියාපුවාවනහෙකි. ..... කීවාවනහෙකි.
unwahanſee. kijaapuwaawanahcki, kiiwaawanahcki.

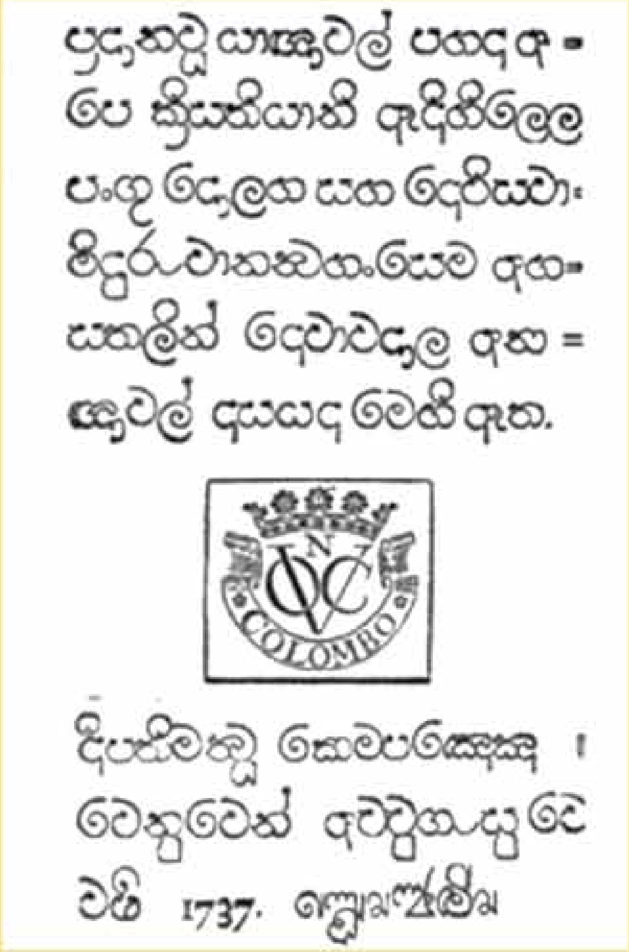
තමුන්වහන්සේලා. කියනවාහොඳයි.
tamunwahanſeelaa. kijanawaahondai.

උන්වහන්සේලා. කියාපුවාවනහෙළා. කීවාවනහෙළා.
unwahanſeelaa kijaapuwaawanahellaa, kiwaawanahellaa.

Fig 5:- Sinhala Characters 1707

It can be concluded that Sinhala printing was introduced in Sri Lanka in 1737 during the time of the Dutch Governor, Gustaaf Willem Baron van Imhoff. Sinhala letters, which were being written on palm leaves thus moved on to being printed on paper. The first book published in Sri Lanka is a prayer book in Sinhala. It was known as the Lord's Prayer. The proclamations of the Dutch Government called Plakaats were also printed in Sinhala, Tamil and Dutch, at the new press introduced by them. These Sinhala characters for printing were designed and cut on wood by a Dutch technician, Gabriel Schade . There, the printers have strictly followed the contemporary writing style extant on palm leaves without much of the standardized punctuations, separations of words, etc. The first printing press with Sinhala types had remained the only press for many years that followed. Several other printing presses were established by the Christian missionary societies during the early 19th century. Christian books, including the Sinhala. versions began to be published.

Buddhists started their own printing presses and published newspapers and journals in which they printed extracts from classical Sinhala texts among other things. These printers took the initiative of publishing the texts as part of their religious and indipendence movements. As the palm leaf writers were not in the habit of using many punctuation marks and separation of words, the modern text editors faced the difficult task of changing their system of writing. The script of the palm leaf manuscripts had peculiarities in writings particularly in the conjunct consonants, double consonants with repha and aspirates, half nasals, etc. With the introduction of the printing press the letters became still more circular.



පුද,කට්ටු යාඥාවල් පතද අ =
වෙ ක්‍රිස්තියානි ඇදිහිලෙල
පංඥ දෙ,ලහ සත දෙපීයාවෝ:
මිදුරුවොතනවතංයොම අය=
සතලිස් දෙවාවදුල අත =
ඥාවල් ඇයයද මෙහි ඇත.

N
VOC
COLOMBO

දිපශිමතුු කොමපකොටුව !
වෙතුවෙන් ඇවදුහලදු වෙ
වනි 1737. [illegible]

Fig 6 :- Page of the Prayer Book of the bible.
Printed in 1737

In the Sinhala alphabet, the letters take different shapes. A reader and the writer both have to differentiate one letter from the other by its particular shape. It is important to have a good knowledge of the shape of letters before starting to write. The letters are formed with the combination of several sub-components that are called subordinate parts of a letter, beginning with the main body followed by the strokes, the strokes that combine with the letter, the strokes that do not combine with the letter, height, straight lines, curves, space between letters, space between the letters, etc.

When writing a letter it has to start and end correctly. The letters are formed by combining several subordinate parts at the beginning. They are twofold. One is a dot attached to a line and the other is a line without a dot. These subordinate parts help to distinguish one letter from the other. The ending subordinate parts are also twofold viz. the endings of the letters ග, ත and ව curve down at the rear end. The ending subordinate parts of the letters ට, ඩ and ඩ curve to the left at the upper end. In letters එ, ඵ and ඨ the ending part bends to the left with a small hook like thing at the end. The nasalized consonants ඳ, ඟ and and the nasal consonant ඤ special kind of a subordinate part joins at the end.

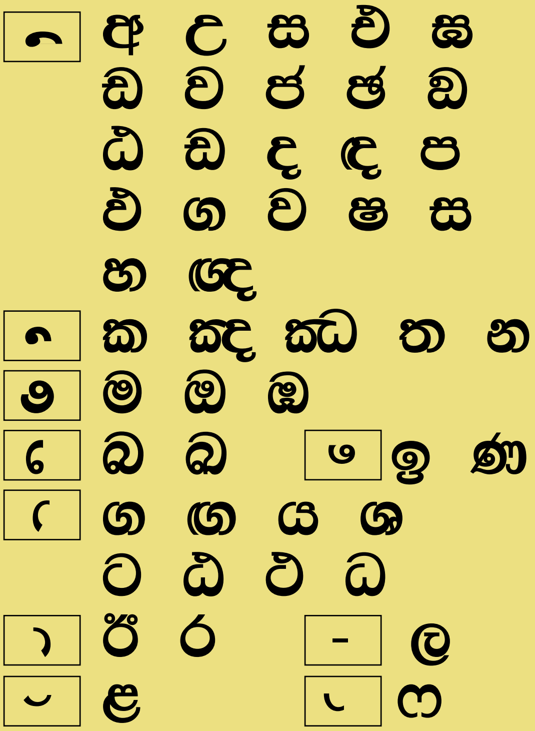


Fig 7:- Initial Components

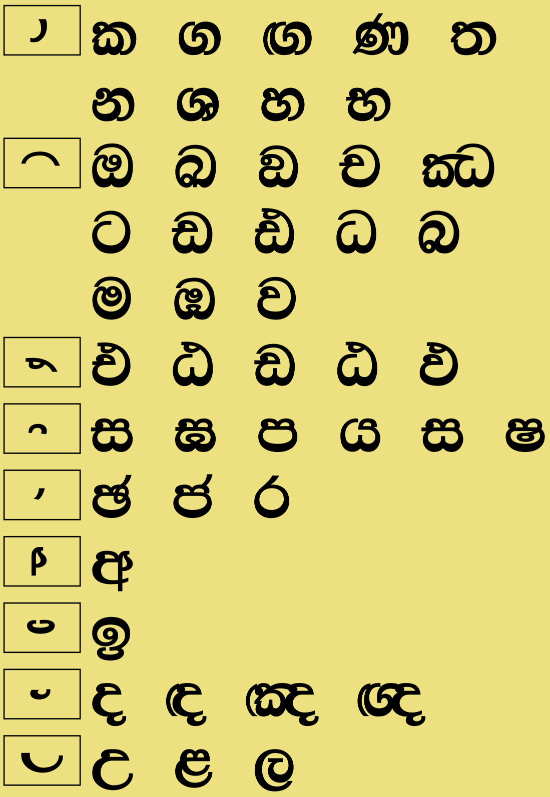


Fig 8:- Ending Components

Each and every letter has a specific height and according to the size of the letter they are classified into three, namely, medial letters, ascending letters and descending letters. The writing space should be divided into three phases horizontally. The letters written in the middle phase are medial letters. ග, හ, ස, ප are medial letters. The letters that spread up from the middle phase to the upper phase are called ascending letters. They are ට, ව, එ, ඩ, ඞ. The letters that spread down to the lower phase are called descending letters, අ, උ, ද.

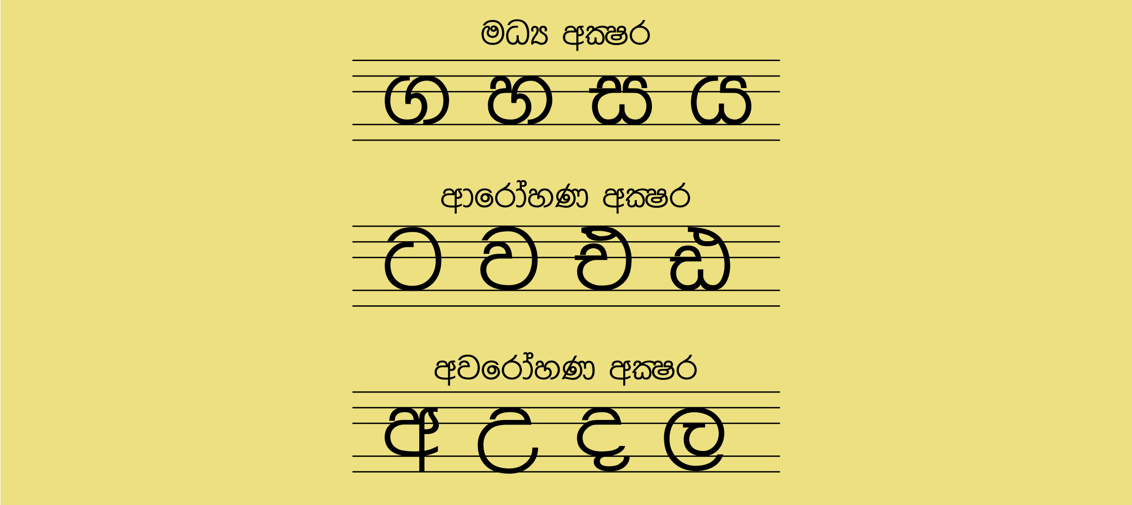
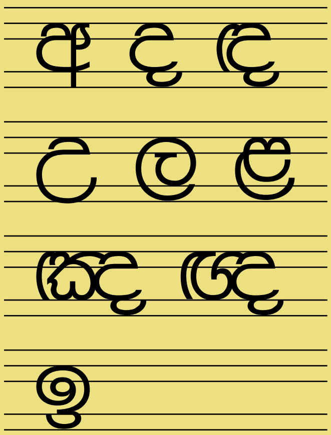


Fig 9:- Sizes of letters

The initial components of the medial letters are written in a small phase over the middle phase. The strokes that come after and before the consonants are written in the middle phase. The initial component of the letter ෆ begins at the lowest point of the middle phase. The stroke which denote උ vowel is written below the middle phase in ග, ඟ, ශ, හ and ක

There are two kinds of ascending letters, namely, letters without a stroke at the top and the letters with a stroke at the top. The letters of the first category are little shorter than the second category. ට ධ ව ම ඔ ඕ බ

The ඣ letter is formed by combining of two letters, one medial letter and an ascending letter. The ළු which represent cerebral ළ and vowel sound උ is also an ascending letter. The descending letters are written in the middle phase and spreads to the lower phase.



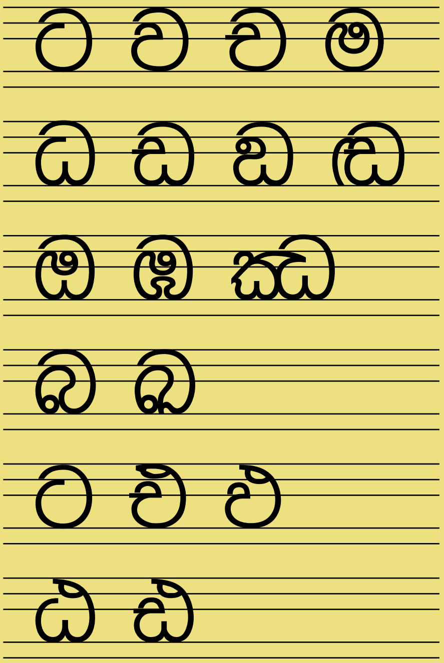


Fig 10:- Ascending letters *Fig 11:- Descending letters*

ඤ and ඥ are written with the combination of a medial letter and descending letter. All the medial strokes which combine with the descending letters are written in the same way as they are written with the medial letters.

The strokes that are written as a combination with consonants to represent a vowel sound can be divided in to three categories. The strokes are smaller than the letters. The medial strokes can be written in the same height as the medial letters දා, දැ, දෑ, දෘ, ද්‍ය, දෞ. The strokes are written usually after the letters except the (ෙ) kombuwa. It comes before the letter. The yansaya (්‍ය) that represents a consonant sound is also categorized as a medial stroke.

Traditionally the strokes that combine with the ද letter are written on the lower phase. It is more logical writing them on the middle phase. The ර letter with ඇ vowel sound is written with an ascending stroke.

The strokes ispilla, udupilla and uspilla represents vowel sounds. The cap stroke (raehaena) and flag stroke (kodiya) are written to denote that the vowel sound is not included in the consonant. The strokes which represents e vowel sound, udupilla and raehena are used with the ascending letters that ends at the left side of the upper phase. Those are ට ව ඩ. The strokes and kodiya are written with the other letters of that category to express the same objective.

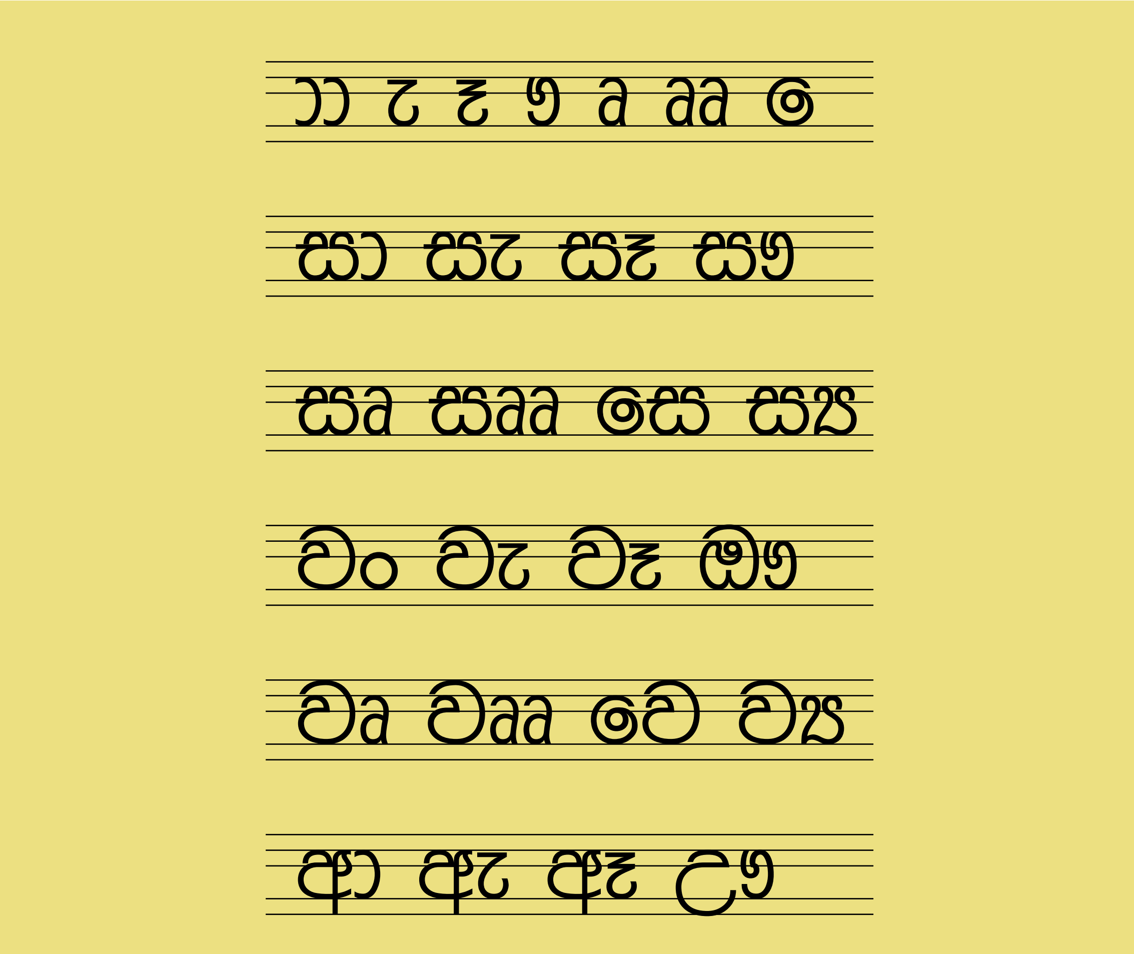


Fig12:- The Strokes

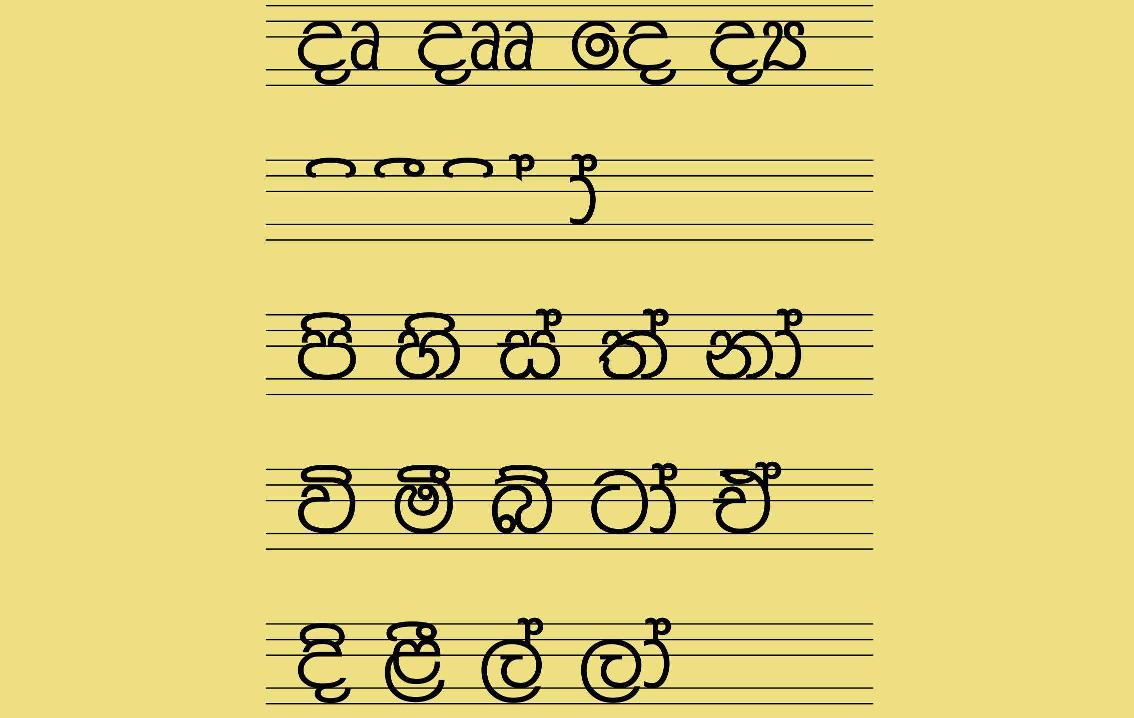


Fig:- 13 Ascending strokes

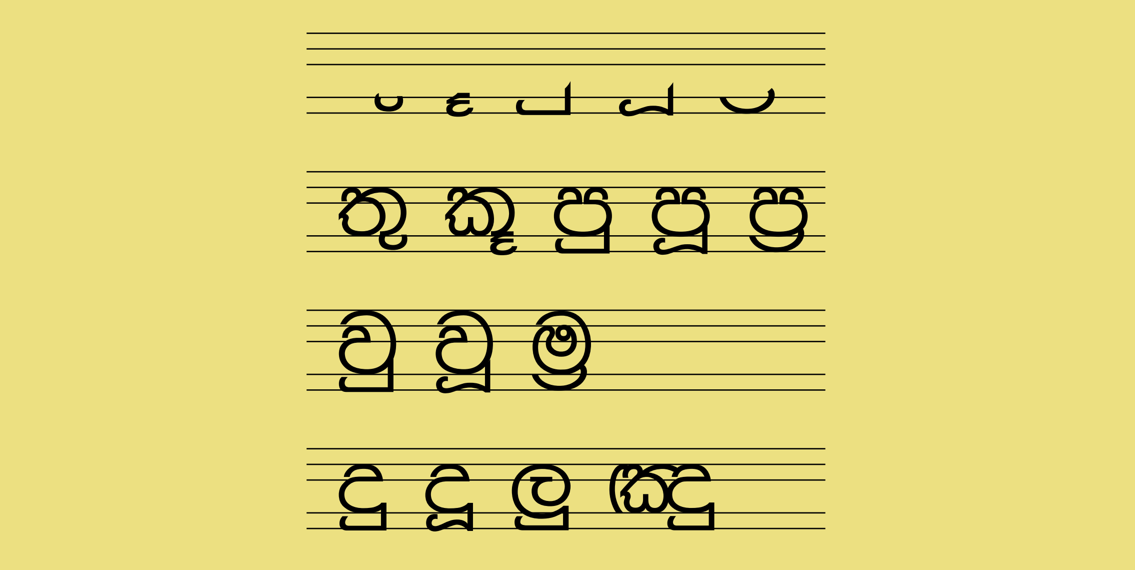


Fig:- 14 Dscending strokes

The uspilla that combines with the aelapilla denotes the ඕ vowel sound as in ගෝ and වෝ letters. Though there is a slightest difference between uspilla and udupilla it should be clearly understood and written. The aelapilla is combined with ර as an ascending stroke but it combines with other letters as an medial stroke රැ, කෑ.

The strokes that are written on the lower phase are descending strokes. Most of them denote the u vowel sound as in කු, තු, සු, වු. The rakaaraansaya which represents the consonant sound ර is also a descending stroke. There are two kinds of descending strokes. Kon paapilla and vak paapilla. Kon paapilla is the stroke that is written with the consonant ස, ද, ව; සු, දු, වු. Vak paapilla combines with the medial letters, ක, ශ, ග, ත, භ. These descending strokes are not used to denote the ර letter combine with the උ vowel sound. As a special case the aelapilla is written with ර to denote that sound. These strokes that denote උ sound are not used with the cerebral ළ either. The ascending letter ළු is written to denote that sound. A long aedapilla is used to denote the long ළූ sound. When the descending strokes are combine with ද (da) letter the ending component of that letter will be deleted as in ද, දු, ද්‍ය , ද්‍ර, දෘ (da, du,dya, dra, dru).

Chapter 03

TAMIL SCRIPT

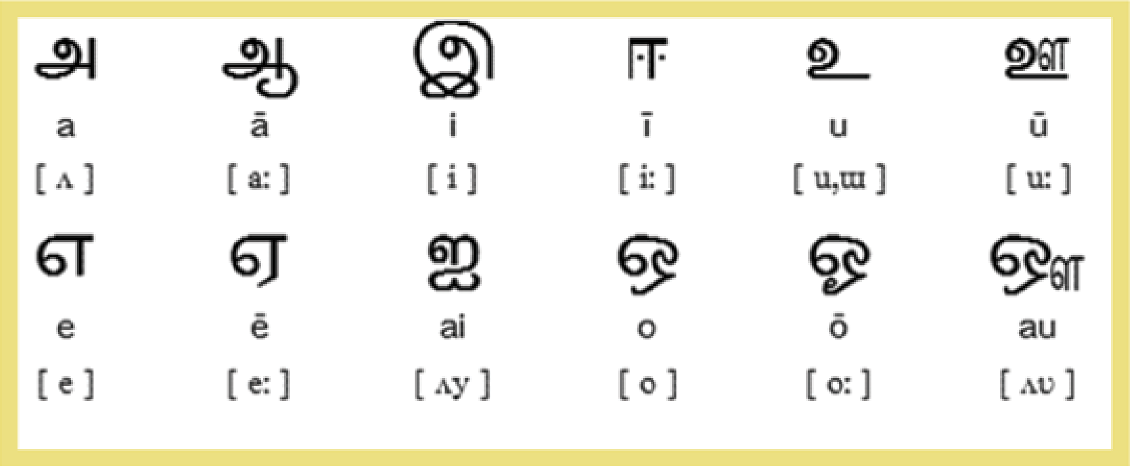
The Tamil language belongs to the Dravidian Family of languages. Tamil is one of the oldest languages in the Indian Subcontinent with a recorded history of two millennia. Tamil is the official language of the state of Tamil Nadu, India and is one of the twenty two scheduled languages of India. Tamil is spoken widely in India, Sri Lanka and is an official language in Sri Lanka and Singapore. The Tamil language was the first language to be declared a classical language by the government of India in 2004. Tamil is spoken by sizable minority populations in Mauritius, West Indies and Malaysia and it is commonly spoken by many emigrant communities worldwide. Tamil literature has a long history of two thousand years and the earliest period of Tamil literature which was known as “Sangam” period began in 300 BCE and ended in 300 CE. The earliest epigraphic evidence of Tamil goes back to 3rd century BCE and Tamil inscriptions from 1st BCE to 1 CE were found in Egypt and Thailand.

The closest relative to tamil language is Malayalam and until ninth century, Malayalam remained as a dialect of Tamil.The complete separation of Malayalam from Tamil to a separate language was not completed until 13th-14th century. Tamil descended from Proto Dravidian which was spoken in 3rd millennium BCE. Scholars believe that Proto Dravidian was spoken in the southern region of India around Godavari river basin. It is believed that in 3rd century BCE, Proto Dravidian would have given rise to Proto Tamil. The attested history of Tamil categorized into three periods : Old Tamil (300 BCE - 700 CE), Middle Tamil (700 - 1600), and Modern Tamil (1600 - Modern Present).

Tamil is written by using a script which known as Vetteluttu. Some of the earliest epigraphy of Tamil Scripts points that it had evolved from Southern Brahmi Script. Old inscriptions of Tamil found in the early period were short and were discovered in caves and on pottery. The longest text in old Tamil is known as Tollkappiyam, which was an early work of grammar and poetry , the oldest layer of this inscription may be going back to 1 BCE. A collection of literary work in Old Tamil had also survived which included two thousand poems and this collection is collectively known as Sangam Literature which could be dated back to 1st CE to 5th CE. Other literary works include Cilappatikaram and Manimekali. Especially Manimekalai is set in South India and in Jaffna Peninsula Sri Lanka. Subsequently the Tamil script was influenced by Pallava Grantha. Grantha, 'book' in Sanskrit was a script which was used for writing Sanskrit in the region where Dravidian was prevalent. Grantha had letters which represented sounds which were not in Tamil. The present look of the Tamil Script would have emerged in 8th Century CE and it has remained static since then.

The tamil Script has 12 vowels and 18 consonants and a special character which is known as Aytam. The combination of vowels and consonants make 216 combinations and this number is the smallest combinations, which is found in Indic Languages. Like other Indian Languages, all consonants carry an invisible Vowel 'a' and a dead consonant or pure sound is produced by placing a visible virama which is known as a 'Puli'.

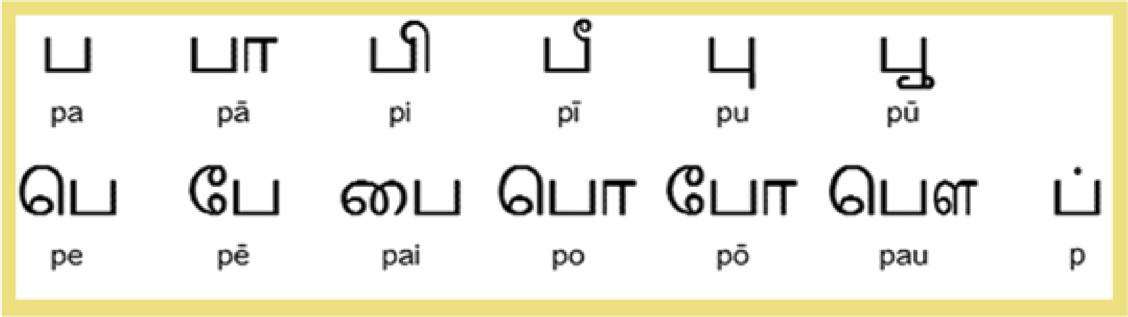
Tamil alphabet
Vowels
There are twelve vowels in Tamil script.



அ	ஆ	இ	ஈ	உ	ஊ
a	ā	i	ī	u	ū
[ʌ]	[a:]	[i]	[i:]	[u,ɯ]	[u:]
எ	ஏ	ஐ	ஒ	ஓ	ஔ
e	ē	ai	o	ō	au
[e]	[e:]	[ʌy]	[o]	[o:]	[ʌʊ]

(Source: http://www.omniglot.com)

Vowel diacritics

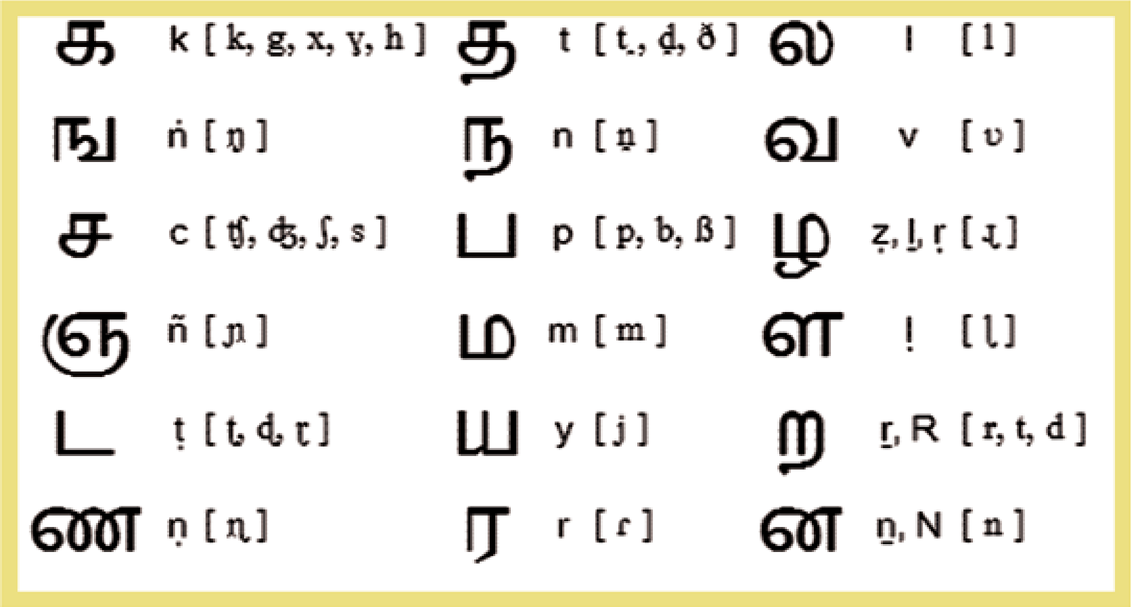


ப	பா	பி	பீ	பு	பூ	
pa	pā	pi	pī	pu	pū	
பெ	பே	பை	பொ	போ	பௌ	ப்
pe	pē	pai	po	pō	pau	p

(Source: http://www.omniglot.com)

Consonants

There are eighteen standard consonants in Tamil Alphabet

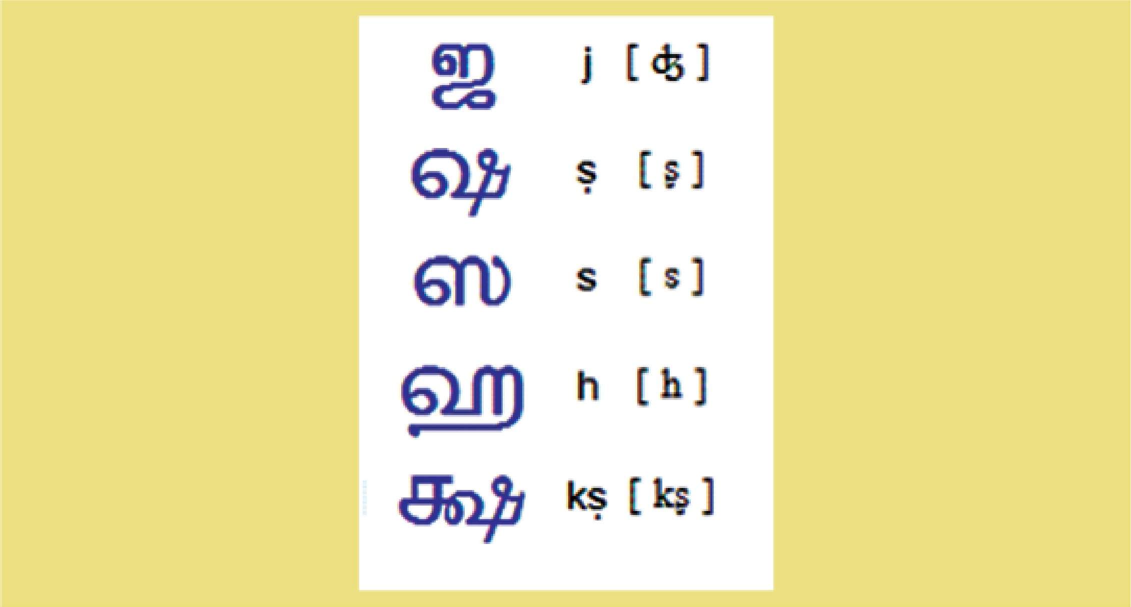


க	k [k, g, x, ɣ, ɦ]	த	t [t̪, d̪, ð]	ல	l [l]
ங	ṅ [ŋ]	ந	n [n̪]	வ	v [ʋ]
ச	c [ʧ, ʤ, ʃ, s]	ப	p [p, b, β]	ழ	ẓ, ḻ, ṛ [ɻ]
ஞ	ñ [ɲ]	ம	m [m]	ள	ḷ [ɭ]
ட	ṭ [ʈ, ɖ, ɽ]	ய	y [j]	ற	ṟ, R [r, t, d]
ண	ṇ [ɳ]	ர	r [ɾ]	ன	ṉ, N [n]

(Source: http://www.omniglot.com)

Others

There are some other consonants which are known as “grantha” letters and they are used to write consonants borrowed from Sanskrit, and also some words from English origin.

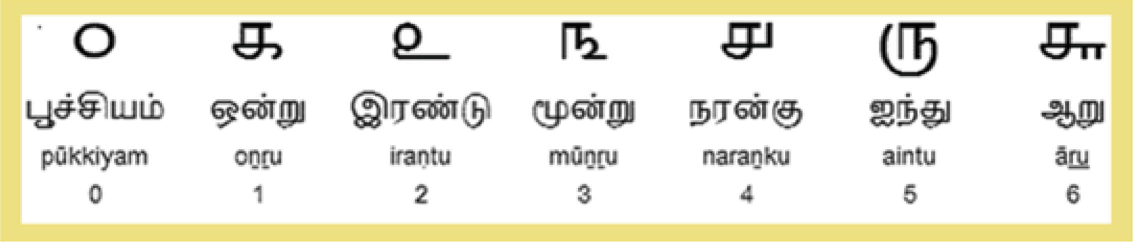


ஜ	j [ʤ]
ஷ	ṣ [ʂ]
ஸ	s [s]
ஹ	h [ɦ]
க்ஷ	kṣ [kʂ]

(Source: http://www.omniglot.com)

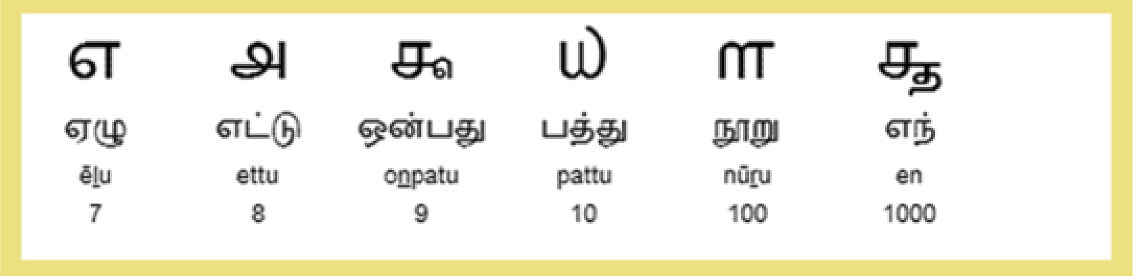
Numerals

Tamil numerals rarely appear in modern text. Instead Arabic numerals are used today.



௦	௧	௨	௩	௪	௫	௬
பூச்சியம்	ஒன்று	இரண்டு	மூன்று	நரன்கு	ஐந்து	ஆறு
pūkkiyam	oṉṟu	iraṇṭu	mūṉṟu	naraṉku	aintu	āṟu
0	1	2	3	4	5	6

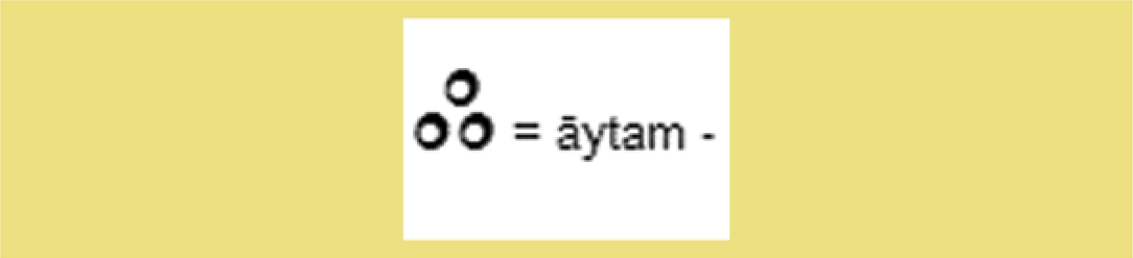
Āytam



௭	௮	௯	௰	௱	௲
ஏழு	எட்டு	ஒன்பது	பத்து	நூறு	எந்
ēḻu	ettu	oṉpatu	pattu	nūṟu	en
7	8	9	10	100	1000

(Source: http://www.omniglot.com)

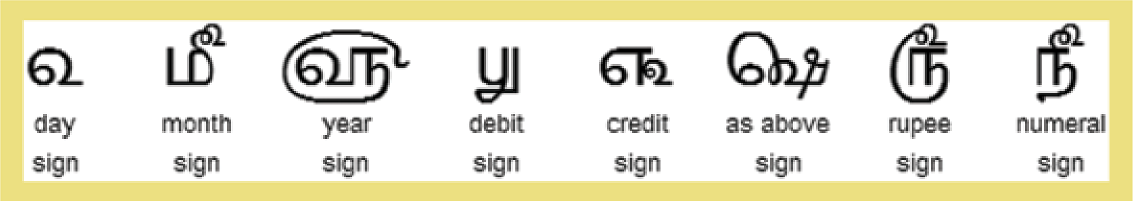
This is a special character in which is classified in Tamil grammar as being neither a consonant nor a vowel.



ஃ = āytam -

Symbols

There are some other symbols which are used for various purposes in Tamil.



௳	௴	௵	௶	௷	௸	௹	௺
day sign	month sign	year sign	debit sign	credit sign	as above sign	rupee sign	numeral sign

(Source: http://www.omniglot.com)

Tamil Unicode

Tamil Unicode was originally derived from ISCII standard. The Unicode range for Tamil is U+0B80–U+0BFF.

(Source: Wikipedia)

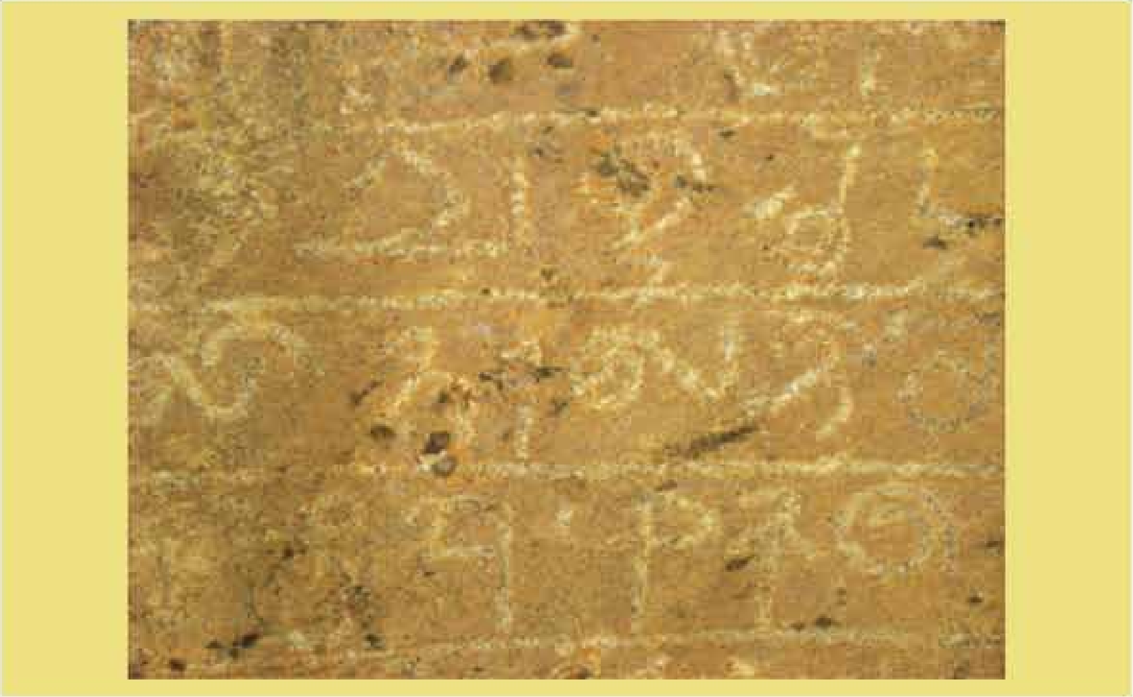


Figure 1 Tamil Inscription (Anuradhapura)

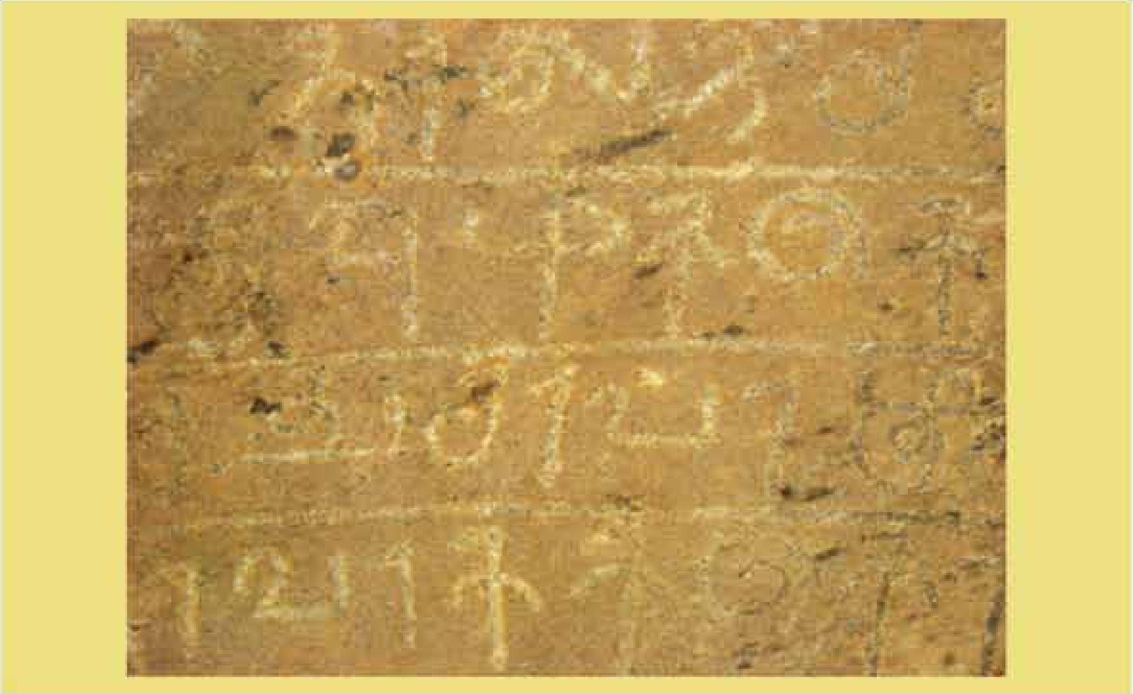
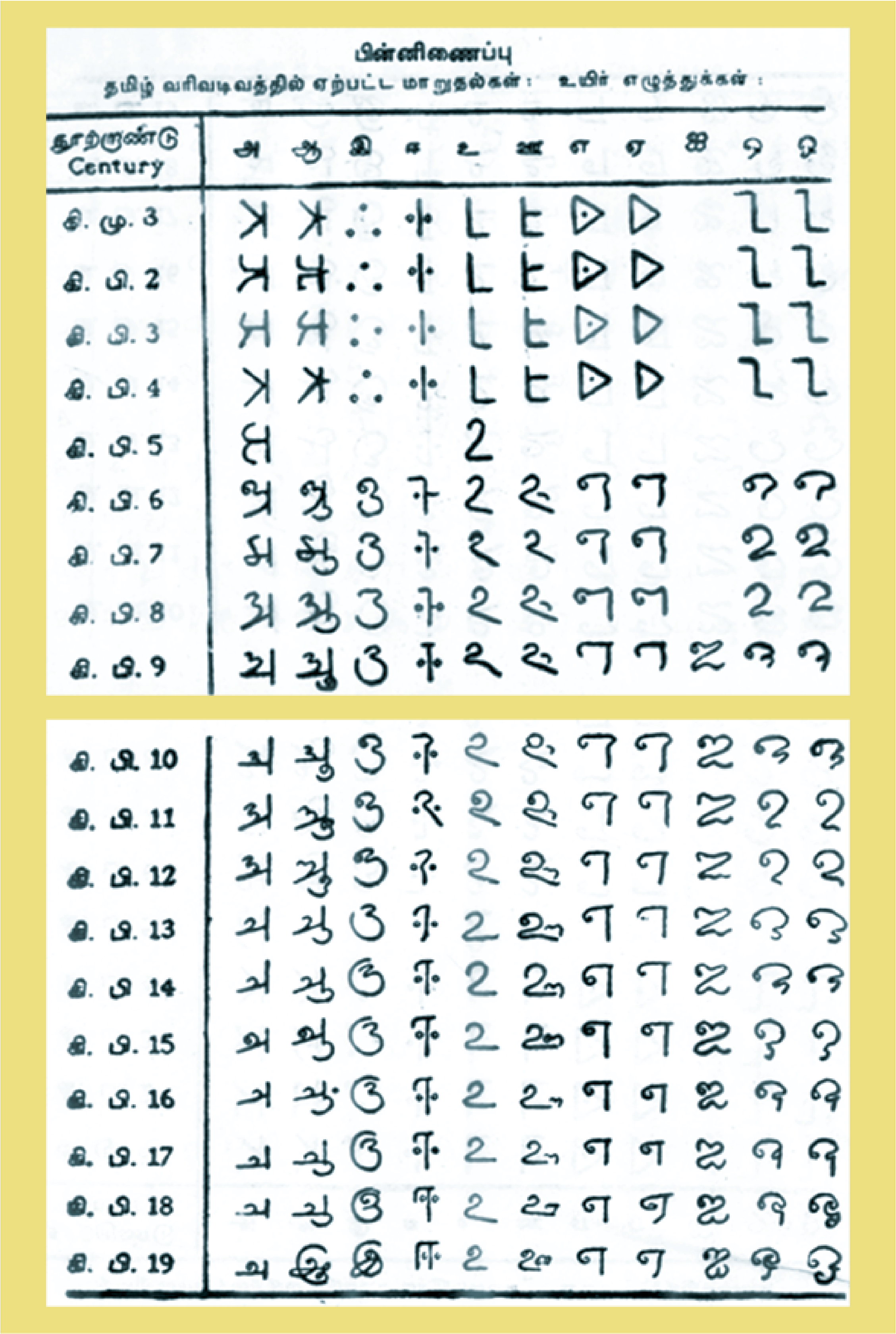
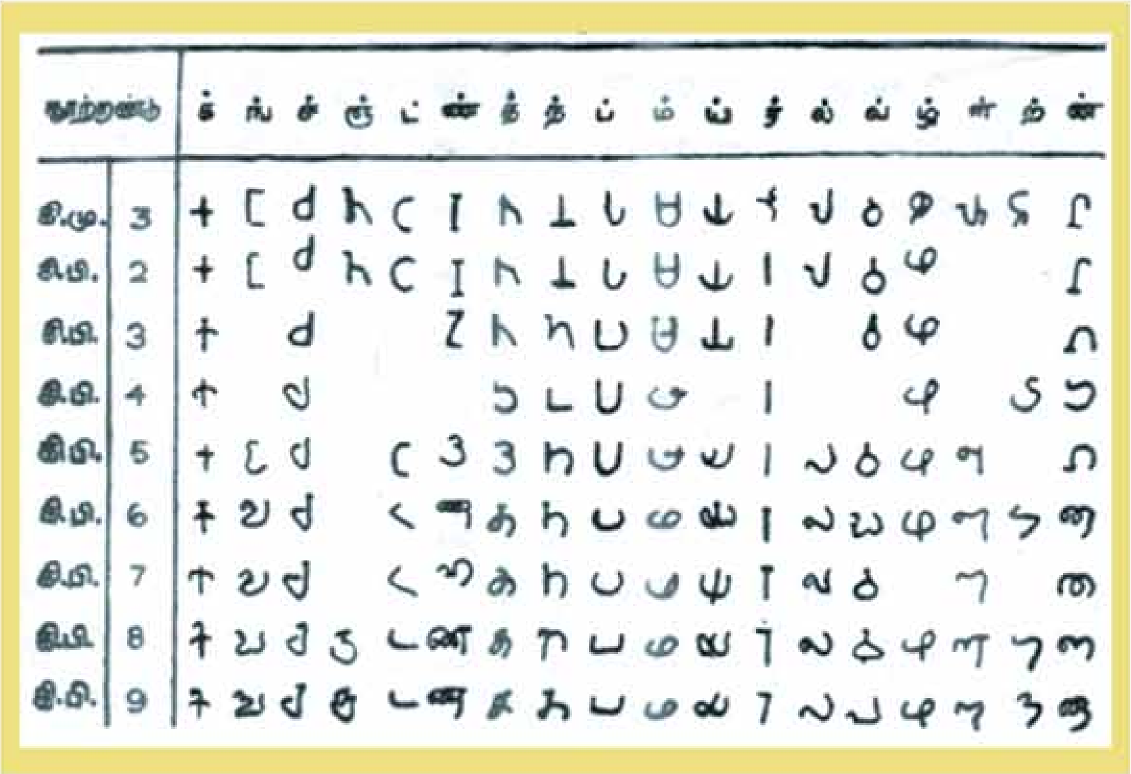


Figure 2 Tamil inscription


பின்னிணைப்பு
தமிழ் வரிவடிவத்தில் ஏற்பட்ட மாறுதல்கள் : உயிர் எழுத்துக்கள் :
நூற்றாண்டு
Century
அ
ஆ
இ
ஈ
உ
ஊ
எ
ஏ
ஐ
ஒ
ஓ
கி. மு. 3
கி. பி. 2
கி. பி. 3
கி. பி. 4
கி. பி. 5
கி. பி. 6
கி. பி. 7
கி. பி. 8
கி. பி. 9
கி. பி. 10
கி. பி. 11
கி. பி. 12
கி. பி. 13
கி. பி 14
கி. பி. 15
கி. பி. 16
கி. பி. 17
கி. பி. 18
கி. பி. 19


நூற்றாண்டு
க் ங் ச் ஞ் ட் ண் த் ந் ப் ம் ய் ர் ல் வ் ழ் ள் ற் ன்
கி.மு. 3
கி.பி. 2
கி.பி. 3
கி.பி. 4
கி.பி. 5
கி.பி. 6
கி.பி. 7
கி.பி. 8
கி.பி. 9

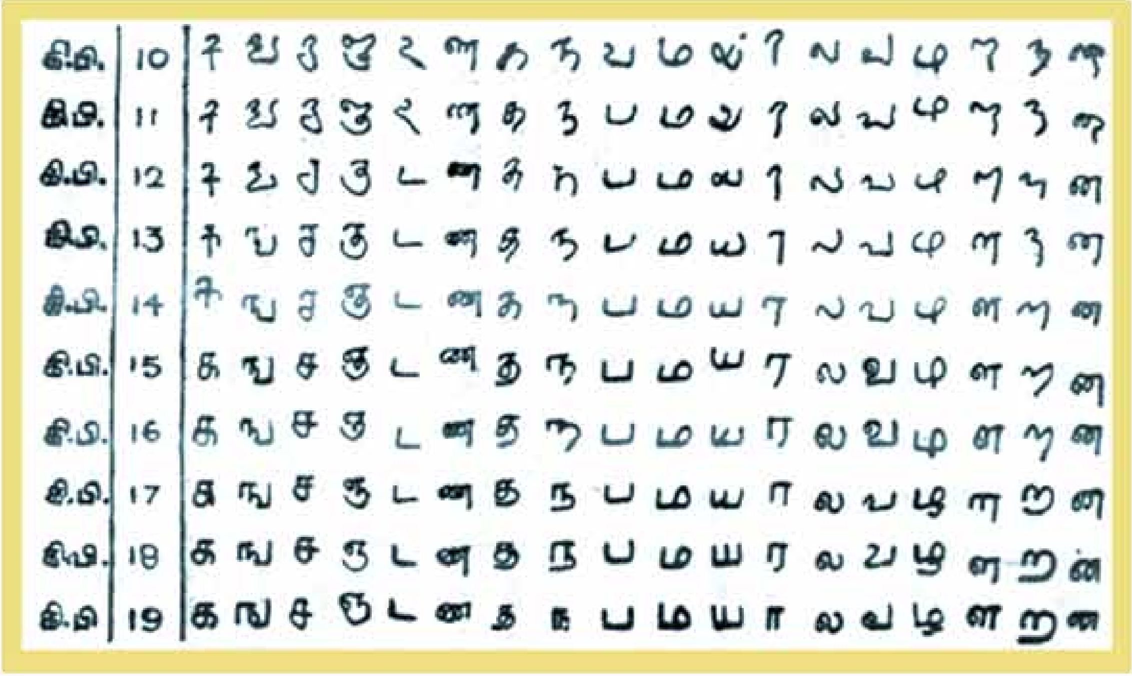
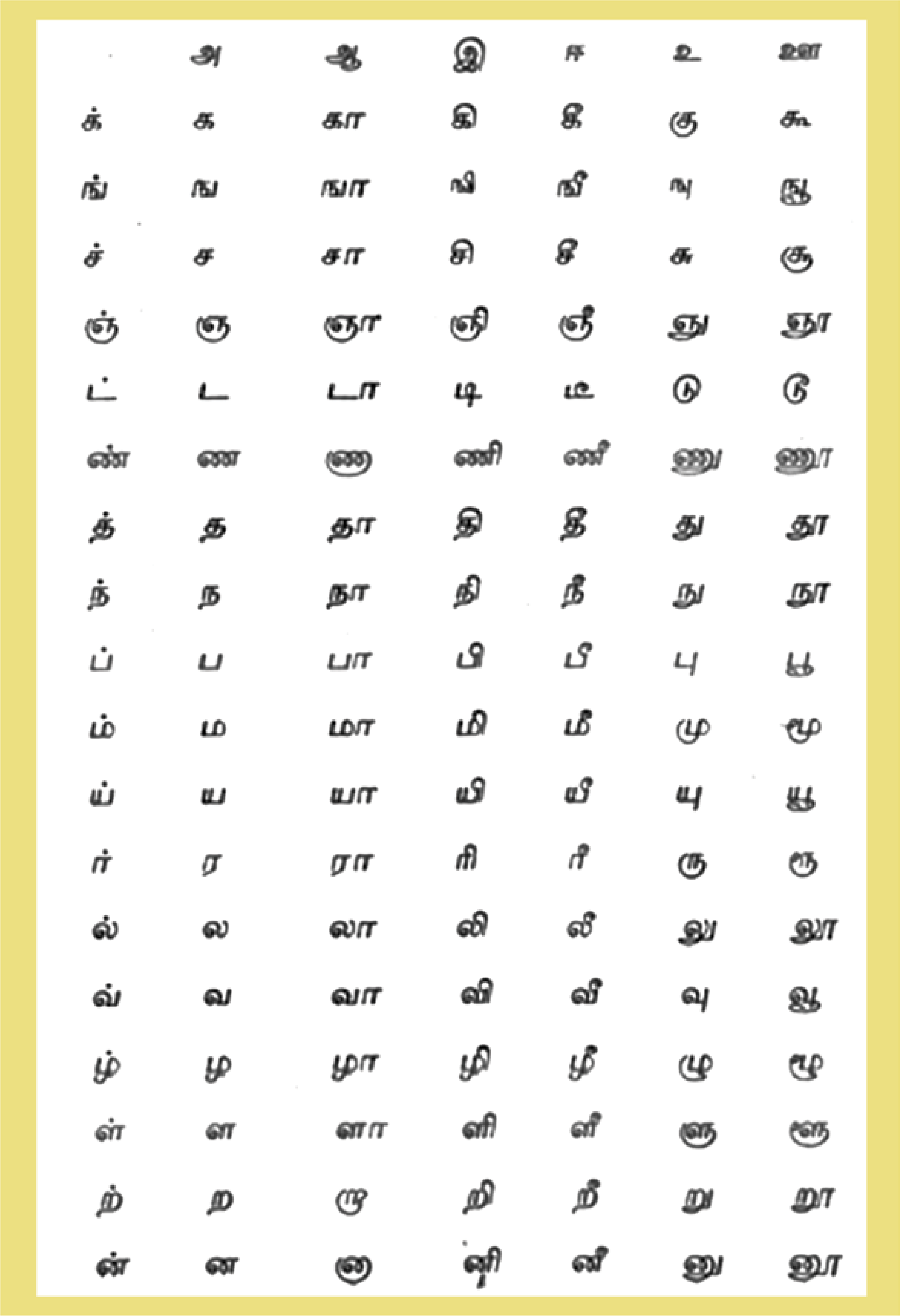
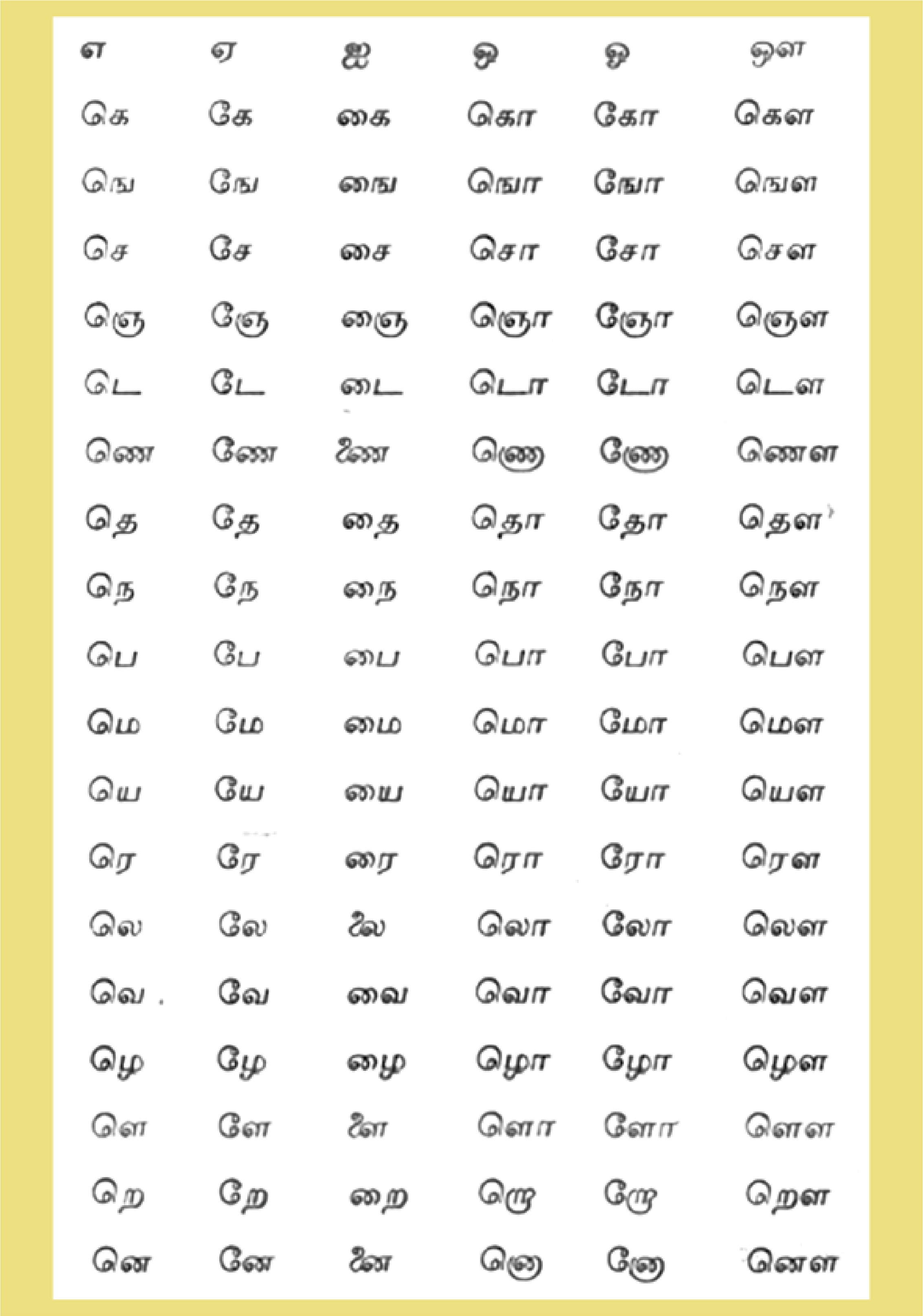

கி.பி. 10
கி.பி. 11
கி.பி. 12
கி.பி. 13
கி.பி. 14
கி.பி. 15
கி.பி. 16
கி.பி. 17
கி.பி. 18
கி.பி 19

Table of Vowels, Consonants

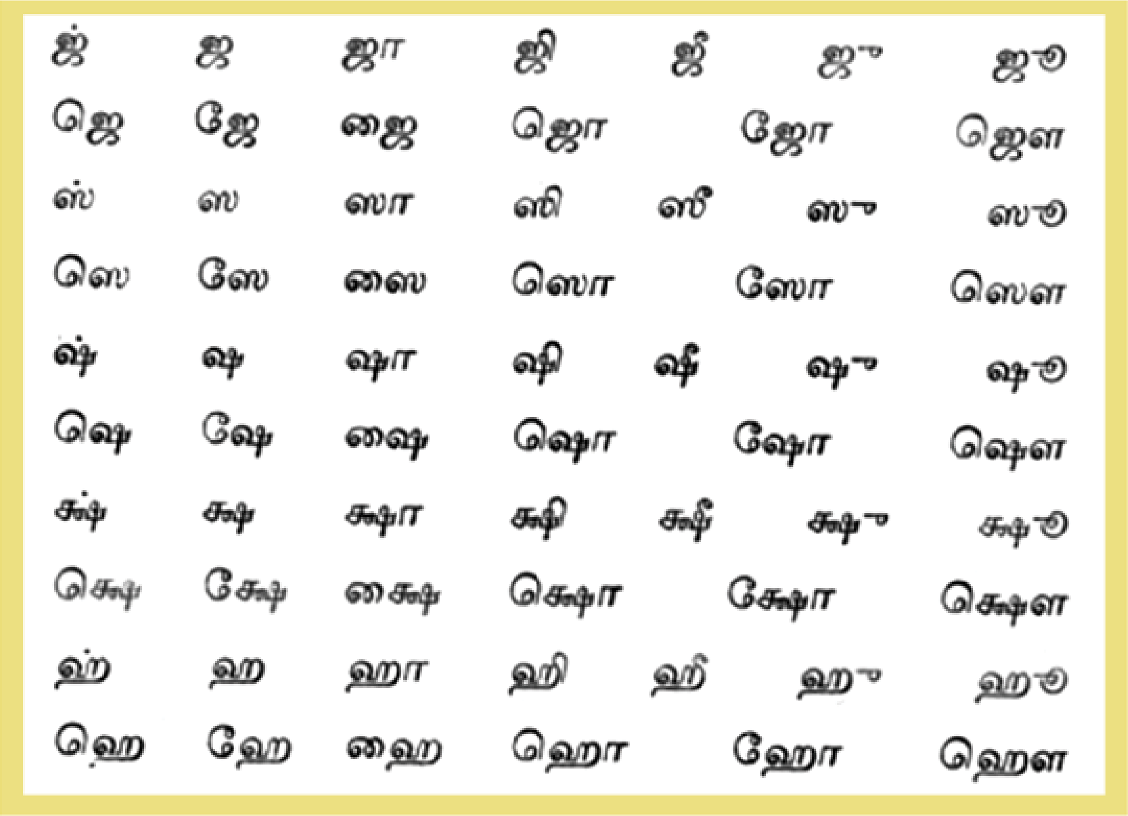


·	அ	ஆ	இ	ஈ	உ	ஊ
க்	க	கா	கி	கீ	கு	கூ
ங்	ங	ஙா	ஙி	ஙீ	ஙு	ஙூ
ச்	ச	சா	சி	சீ	சு	சூ
ஞ்	ஞ	ஞா	ஞி	ஞீ	ஞு	ஞூ
ட்	ட	டா	டி	டீ	டு	டூ
ண்	ண	ணா	ணி	ணீ	ணு	ணூ
த்	த	தா	தி	தீ	து	தூ
ந்	ந	நா	நி	நீ	நு	நூ
ப்	ப	பா	பி	பீ	பு	பூ
ம்	ம	மா	மி	மீ	மு	மூ
ய்	ய	யா	யி	யீ	யு	யூ
ர்	ர	ரா	ரி	ரீ	ரு	ரூ
ல்	ல	லா	லி	லீ	லு	லூ
வ்	வ	வா	வி	வீ	வு	வூ
ழ்	ழ	ழா	ழி	ழீ	ழு	ழூ
ள்	ள	ளா	ளி	ளீ	ளு	ளூ
ற்	ற	றா	றி	றீ	று	றூ
ன்	ன	னா	னி	னீ	னு	னூ

And Compound Letters



எ	ஏ	ஐ	ஒ	ஓ	ஔ
கெ	கே	கை	கொ	கோ	கௌ
ஙெ	ஙே	ஙை	ஙொ	ஙோ	ஙௌ
செ	சே	சை	சொ	சோ	சௌ
ஞெ	ஞே	ஞை	ஞொ	ஞோ	ஞௌ
டெ	டே	டை	டொ	டோ	டௌ
ணெ	ணே	ணை	ணொ	ணோ	ணௌ
தெ	தே	தை	தொ	தோ	தௌ
நெ	நே	நை	நொ	நோ	நௌ
பெ	பே	பை	பொ	போ	பௌ
மெ	மே	மை	மொ	மோ	மௌ
யெ	யே	யை	யொ	யோ	யௌ
ரெ	ரே	ரை	ரொ	ரோ	ரௌ
லெ	லே	லை	லொ	லோ	லௌ
வெ	வே	வை	வொ	வோ	வௌ
ழெ	ழே	ழை	ழொ	ழோ	ழௌ
ளெ	ளே	ளை	ளொ	ளோ	ளௌ
றெ	றே	றை	றொ	றோ	றௌ
னெ	னே	னை	னொ	னோ	னௌ



ஜ்	ஜ	ஜா	ஜி	ஜீ	ஜு	ஜூ
ஜெ	ஜே	ஜை	ஜொ		ஜோ	ஜௌ
ஸ்	ஸ	ஸா	ஸி	ஸீ	ஸு	ஸூ
ஸெ	ஸே	ஸை	ஸொ		ஸோ	ஸௌ
ஷ்	ஷ	ஷா	ஷி	ஷீ	ஷு	ஷூ
ஷெ	ஷே	ஷை	ஷொ		ஷோ	ஷௌ
க்ஷ்	க்ஷ	க்ஷா	க்ஷி	க்ஷீ	க்ஷு	க்ஷூ
க்ஷெ	க்ஷே	க்ஷை	க்ஷொ		க்ஷோ	க்ஷௌ
ஹ்	ஹ	ஹா	ஹி	ஹீ	ஹு	ஹூ
ஹெ	ஹே	ஹை	ஹொ		ஹோ	ஹௌ

Again & Again & Again & Again
10 Points Out of 10
You Might
PARK JAEBEOM
2PM2PM2PM

Chapter 04

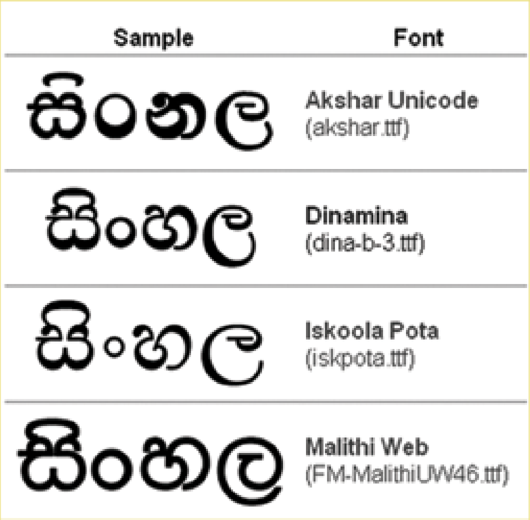
UNICODE ENCODING OF SINHALA

Prof. Gihan Dias
University of Moratuwa

Introduction and background

The need to support Sinhala on computers was identified some time back. In 1980's the main focus was on a hardware based solution. Later many Sinhala Fonts were developed.

Eg: DL-Anurada, DL-Lihini, DL-Manel, kaputa, kaputadotcom, Malithi Web, Thibus29STru



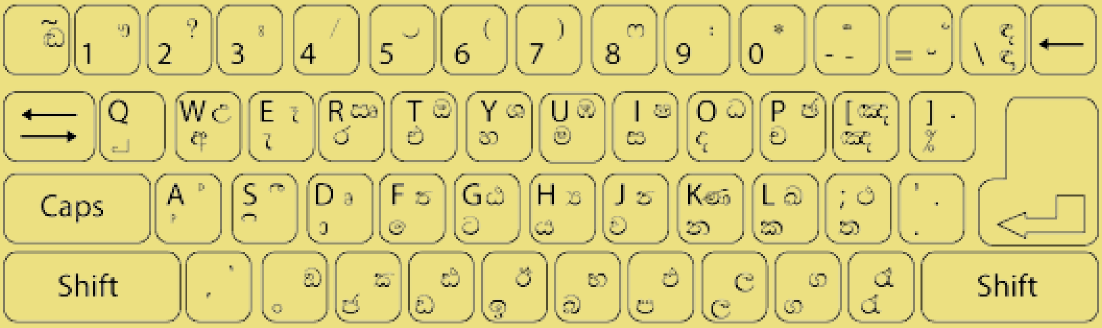
Sample	Font
සිංහල	Akshar Unicode (akshar.ttf)
සිංහල	Dinamina (dina-b-3.ttf)
සිංහල	Iskoola Pota (iskpota.ttf)
සිංහල	Malithi Web (FM-MalithiUW46.ttf)

When using a keyboard, different layouts were used:.

- Phonetic keyboard: In this layout the characters represent the sounds of spoken language.

E.g.: 'k' is used for 'ක'

- Wijesekara Keyboard: This keyboard is based on the Sinhala typewriter keyboard



When using the above keyboards several issues were identified. There was a problem of representing some characters such as කු, ළු, ඏ, රි, දු, න, ද, ඤ, ඹ.

To ensure that standards are followed regarding computing in Sinhala and Tamil, CINTEC established the Committee for Sinhala and Tamil in Computer technology in 1985.The Committee agreed on a unique Sinhala alphabet and an alphabetical order. At first the encoding was based on the ISCII (Indian Standard Code for Information Interchange) and later it was expanded to support Unicode. In 1997 it was adopted by Unicode with some minor modifications.

● SLS1134:2004: Sri Lanka Sinhala Character Code for Information Interchange in 2004 and this is the current standard which provides coding of the set of Sinhala characters for use in ICT

Sinhala Language and Alphabet
Sinhala Script

The Sinhala Language belongs to Indic family of languages. In the Sinhala script each letter may consist of a consonant and a vowel and so is called an “abugida”. Each syllable unit is an individual visual unit or glyph. Two glyphs can tie up and form conjunct consonants which are called “බැඳි අකුරු’’ Bandi akuru” in Sinhala. Sorting words which are written in different styles (Conjunct consonants as opposed to ordinary writing) should be done according to the sorting rules of the Sinhala language

Sinhala Alphabet

Sinhala Alphabet consists of 59 based letters. It includes

- 16 vowels
- 41 consonants
- 2 semi-consonants

Sinhala vowels

අ ආ ඇ ඈ ඉ ඊ උ ඌ ඍ එ ඒ ඓ ඔ ඕ ඖ

Consonants

ක ඛ ග ඝ ඟ

ච ඡ ජ ඣ ඤ

ට ඨ ඩ ඪ ණ ත ථ ද ධ න ඳ

ප ඵ බ භ ම ඹ

ය ර ල ව

ශ ෂ ස හ

ළ ෆ

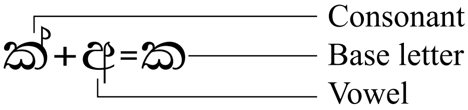
Semi-consonants

ං ඃ

Modifiers

ඍ ඏ

A base letter represents a consonant with a vowel.



When there is an Al-Lakuna(්) the default vowel is removed. You can change the vowel with other modifiers and these modifiers can be on any side of the letter (above, below, left or right).

ක + ඉ = කි— Base letter

Sinhala Unicode Chart - Range (0D80$0DFF)

	0D8	0D9	0DA	0DB	0DC	0DD	0DE	0DF
0		ඐ 0D90	ච 0DA0	ධ 0DB0	ව 0DC0	◌ැ 0DD0		
1		එ 0D91	ඡ 0DA1	න 0DB1	ශ 0DC1	◌ෑ 0DD1		
2	◌ං 0D82	ඒ 0D92	ජ 0DA2		ෂ 0DC2	◌ි 0DD2		◌ෲ 0DF2
3	◌ඃ 0D83	ඓ 0D93	ඣ 0DA3	ඳ 0DB3	ස 0DC3	◌ී 0DD3		◌ෳ 0DF3
4		ඔ 0D94	ඤ 0DA4	ප 0DB4	හ 0DC4	◌ු 0DD4		0DF4
5	අ 0D85	ඕ 0D95	ඥ 0DA5	ඵ 0DB5	ළ 0DC5			
6	ආ 0D86	ඖ 0D96	ඦ 0DA6	බ 0DB6	ෆ 0DC6	◌ූ 0DD6		
7	ඇ 0D87		ට 0DA7	භ 0DB7				
8	ඈ 0D88		ඨ 0DA8	ම 0DB8		◌ෘ 0DD8		
9	ඉ 0D89		ඩ 0DA9	ඹ 0DB9		ෙ◌ 0DD9		
A	ඊ 0D8A	ක 0D9A	ඪ 0DAA	ය 0DBA	◌් 0DCA	ෙ◌් 0DDA		
B	උ 0D8B	ඛ 0D9B	ණ 0DAB	ර 0DBB		ෛ◌ 0DDB		
C	ඌ 0D8C	ග 0D9C	ඬ 0DAC			ෙ◌ා 0DDC		
D	ඍ 0D8D	ඝ 0D9D	ත 0DAD	ල 0DBD		ෙ◌ා් 0DDD		
E	ඎ 0D8E	ඞ 0D9E	ථ 0DAE			ෙ◌ෟ 0DDE		
F	ඏ 0D8F	ඟ 0D9F	ද 0DAF		◌ා 0DCF	◌ෟ 0DDF		

Sinhala Unicode encoding

In Encoding each Sinhala letter, is represented by a sequence of one or more Unicode characters.

E.g.: ක, කා, ක

Letter type	Description	Encoding

Vowels	Each vowel is represented by one character in the range 0D85 - 0D96
	A vowel such as ආ should not be represented as a character sequence such as 0D85 - 0DCF
		අ = 0D85,
		ආ = 0D86,
		ඍ = 0D8D

Consonants	Each consonant is represented by one character in the range 0D9A - 0DC9	ක = 0D9Aල
		ඤ = 0DA4

Pure Consonants	A pure consonant (i.e.ල without an implicit vowel) is represented by a two character sequence ශ්‍රීcons 0DCA□ (cons + ්) where cons represents a consonant	ක් = 0D9A 0DCA ට = 0DA7 0DCA
Consonants with vowel signs	A consonant with a vowel sign is represented by a two character sequence "cons + vs" where vs represents a vowel sign.	කා = 0D9A 0DCF කෞ = 0D9A 0DDE කෛ = 0D9A 0DDB
රැ,රෑ,රු,රූ,ළූ,ළු	These are non standard letters and encoded accordingly	රැ = 0DBB 0DD0 රෑ= 0DBB 0DD1 රු = 0DBB 0DD4 රූ= 0DDB 0DD6 ළු = 0DC5 0DD4 ළූ= 0DC5 0DD6
Semi-consonant signs anusvaraya (ං) orvisargaya (ඃ)	These signs may follow a vowelල a consonant or a vowel sign. If there is a semi consonant it is always the last character in a combining character sequence	අං = 0D85 0D82 (අ + අඃ = 0D85 0D83 (අ = ඃ) කං = 0D9A 0D82 (ක + ං) කෝං = 0D9A 0DDD 0D82 (ක + ෙ ෝ+ ං) කුං = 0D9A 0DD4 0D83 (ක +ු + ං)
්‍යං (yansaya)	Yansaya is a consonant modifierwhich represent ශ්‍රීya□ which follows a pure consonant	් + zwj + ය : ්‍ය 0DCA 200D 0DBA(yansaya) ත්‍යA = 0DAD 0DCA 200D 0DBA

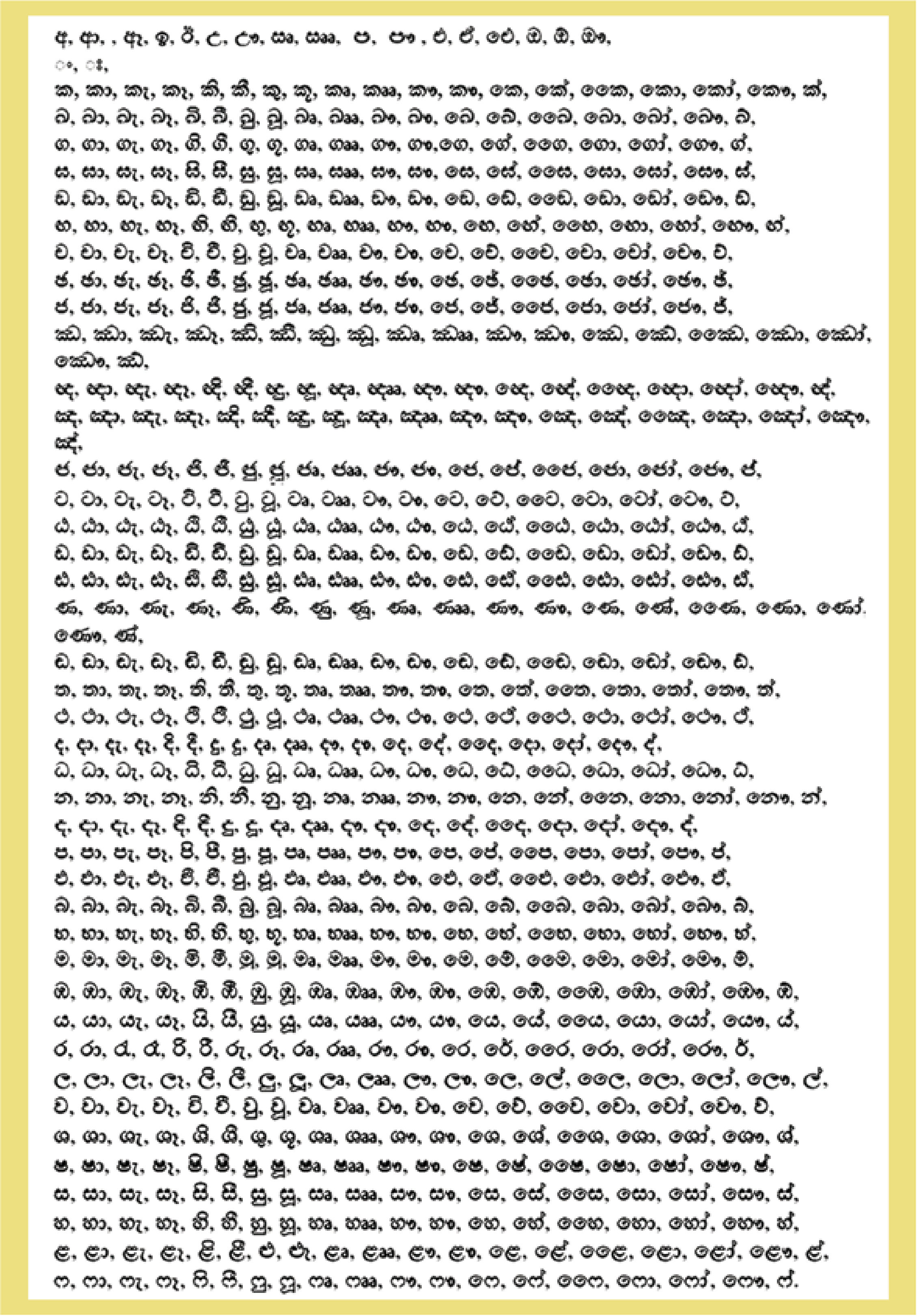
(Rakaransaya)	Rakaransaya is a consonant modifier which represent a "ra" which follows a pure consonant Rakaransaya can be represented by (cons + ් + zwj + ර) where cons represents a consonant	ක්‍ර = 0D9A 0DCA 200D 0DBB (ක + ් + zwj + ර) ක්ෙර් ර් 0D9A 0DCA 200D 0DBB 0DD9 (ක + ් + zwj + ර + ෙ)
Repaya	The repaya represents the letter ර් preceding a consonant	කර්ම + 0D9A 0DBB 0DCA 200D 0DB8(ක + ර + ් + zwj + ම)
Conjunct letters (බැඳි අකුරු)	Conjunct letters are represented as (cons + ් + zwj + cons) where second consonant may optionally be followed by a vowel sign Conjunct letters may be further joined by a rakaaraansaya or yansaya	න්‍ද = 0DB1 0DCA 200D 0DAF (න + ් + zwj + ද) ක්‍ෂේ = 0D9A 0DCA 200D 0DC2 0DDA (ක + ් + zwj + ෂ +ෙ ්) න්‍ද්‍රා = 0DB1 0DCA 200D 0DAF 0DCA 200D 0DBB 0DCF (න + ් + zwj + ද + ් + zwj + ර + ා)
Touching letters	A pure consonant written touching the following letter instead of using an al*lakuna used in classical and Buddhist	සස = 0DC3 200D 0DCA 0DC3 (ස + zwj + ් + ස)

Zwj(Zero - width Joiner) - ZWJ joins two or more consonants to form a single unit (conjunct consonants) and it can also alter shape of preceding consonants (cursive ness of the consonant).This has the value 200D.

ZWNJ (Zero - width non joiner) - This can be used to disjoin a single ligature into two or more units. This has the value 200C

The Sinhala Collation Sequence

This refers to the order in which Sinhala character strings should be placed when sorting them

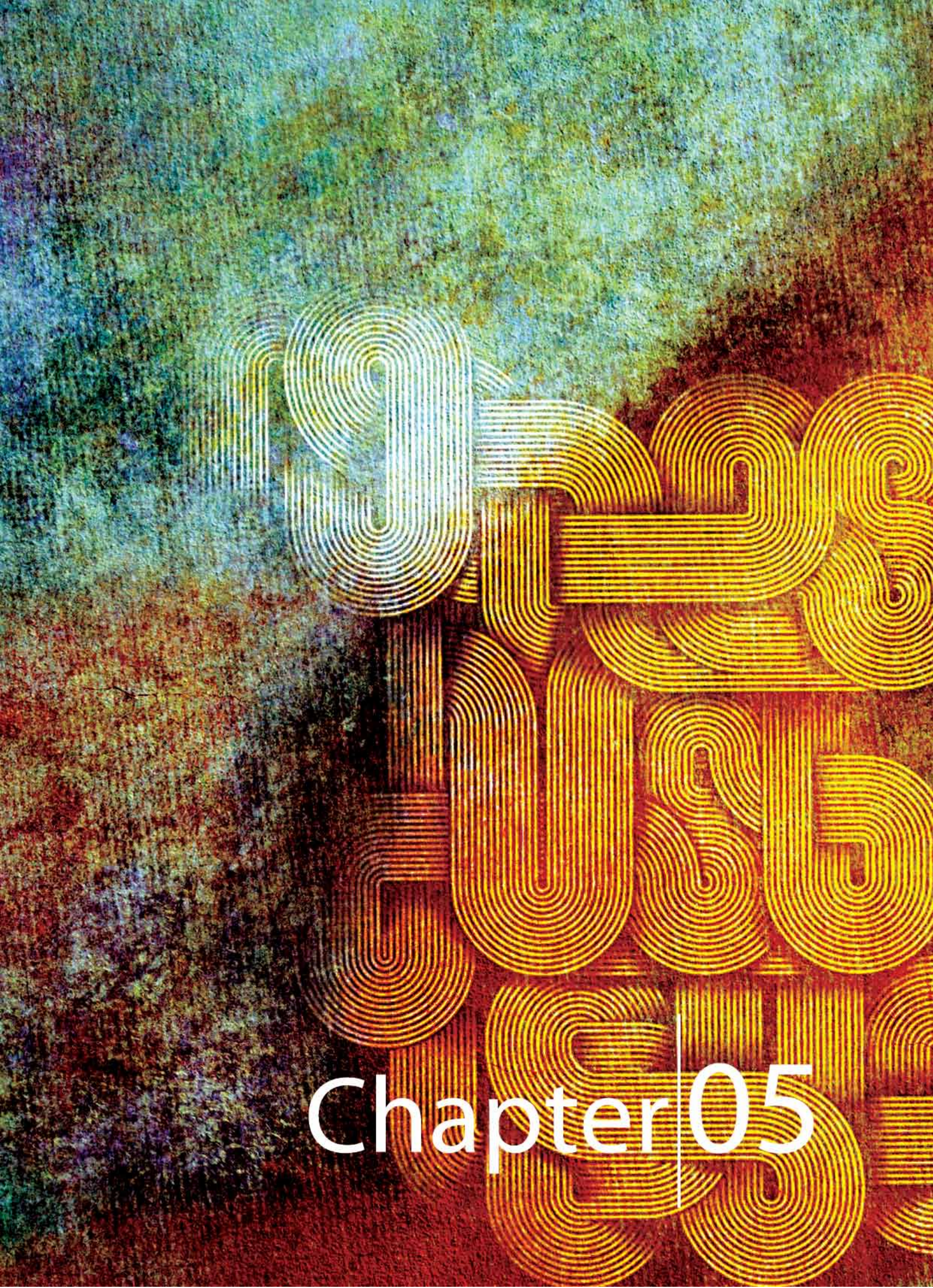


අ, ආ, , ඇ, ඉ, ඊ, උ, ඌ, ඍ, ඎ, ඏ, ඐ , එ, ඒ, ඓ, ඔ, ඕ, ඖ,
◌ං, ◌ඃ,
ක, කා, කැ, කෑ, කි, කී, කු, කූ, කෘ, කෲ, කෟ, කෳ, කෙ, කේ, කෛ, කො, කෝ, කෞ, ක්,
ඛ, ඛා, ඛැ, ඛෑ, ඛි, ඛී, ඛු, ඛූ, ඛෘ, ඛෲ, ඛෟ, ඛෳ, ඛෙ, ඛේ, ඛෛ, ඛො, ඛෝ, ඛෞ, ඛ්,
ග, ගා, ගැ, ගෑ, ගි, ගී, ගු, ගූ, ගෘ, ගෲ, ගෟ, ගෳ,ගෙ, ගේ, ගෛ, ගො, ගෝ, ගෞ, ග්,
ඝ, ඝා, ඝැ, ඝෑ, ඝි, ඝී, ඝු, ඝූ, ඝෘ, ඝෲ, ඝෟ, ඝෳ, ඝෙ, ඝේ, ඝෛ, ඝො, ඝෝ, ඝෞ, ඝ්,
ඞ, ඞා, ඞැ, ඞෑ, ඞි, ඞී, ඞු, ඞූ, ඞෘ, ඞෲ, ඞෟ, ඞෳ, ඞෙ, ඞේ, ඞෛ, ඞො, ඞෝ, ඞෞ, ඞ්,
ඟ, ඟා, ඟැ, ඟෑ, ඟි, ඟී, ඟු, ඟූ, ඟෘ, ඟෲ, ඟෟ, ඟෳ, ඟෙ, ඟේ, ඟෛ, ඟො, ඟෝ, ඟෞ, ඟ්,
ච, චා, චැ, චෑ, චි, චී, චු, චූ, චෘ, චෲ, චෟ, චෳ, චෙ, චේ, චෛ, චො, චෝ, චෞ, ච්,
ඡ, ඡා, ඡැ, ඡෑ, ඡි, ඡී, ඡු, ඡූ, ඡෘ, ඡෲ, ඡෟ, ඡෳ, ඡෙ, ඡේ, ඡෛ, ඡො, ඡෝ, ඡෞ, ඡ්,
ජ, ජා, ජැ, ජෑ, ජි, ජී, ජු, ජූ, ජෘ, ජෲ, ජෟ, ජෳ, ජෙ, ජේ, ජෛ, ජො, ජෝ, ජෞ, ජ්,
ඣ, ඣා, ඣැ, ඣෑ, ඣි, ඣී, ඣු, ඣූ, ඣෘ, ඣෲ, ඣෟ, ඣෳ, ඣෙ, ඣේ, ඣෛ, ඣො, ඣෝ,
ඣෞ, ඣ්,
ඤ, ඤා, ඤැ, ඤෑ, ඤි, ඤී, ඤු, ඤූ, ඤෘ, ඤෲ, ඤෟ, ඤෳ, ඤෙ, ඤේ, ඤෛ, ඤො, ඤෝ, ඤෞ, ඤ්,
ඥ, ඥා, ඥැ, ඥෑ, ඥි, ඥී, ඥු, ඥූ, ඥෘ, ඥෲ, ඥෟ, ඥෳ, ඥෙ, ඥේ, ඥෛ, ඥො, ඥෝ, ඥෞ,
ඥ්,
ඦ, ඦා, ඦැ, ඦෑ, ඦි, ඦී, ඦු, ඦූ, ඦෘ, ඦෲ, ඦෟ, ඦෳ, ඦෙ, ඦේ, ඦෛ, ඦො, ඦෝ, ඦෞ, ඦ්,
ට, ටා, ටැ, ටෑ, ටි, ටී, ටු, ටූ, ටෘ, ටෲ, ටෟ, ටෳ, ටෙ, ටේ, ටෛ, ටො, ටෝ, ටෞ, ට්,
ඨ, ඨා, ඨැ, ඨෑ, ඨි, ඨී, ඨු, ඨූ, ඨෘ, ඨෲ, ඨෟ, ඨෳ, ඨෙ, ඨේ, ඨෛ, ඨො, ඨෝ, ඨෞ, ඨ්,
ඩ, ඩා, ඩැ, ඩෑ, ඩි, ඩී, ඩු, ඩූ, ඩෘ, ඩෲ, ඩෟ, ඩෳ, ඩෙ, ඩේ, ඩෛ, ඩො, ඩෝ, ඩෞ, ඩ්,
ඪ, ඪා, ඪැ, ඪෑ, ඪි, ඪී, ඪු, ඪූ, ඪෘ, ඪෲ, ඪෟ, ඪෳ, ඪෙ, ඪේ, ඪෛ, ඪො, ඪෝ, ඪෞ, ඪ්,
ණ, ණා, ණැ, ණෑ, ණි, ණී, ණු, ණූ, ණෘ, ණෲ, ණෟ, ණෳ, ණෙ, ණේ, ණෛ, ණො, ණෝ.
ණෞ, ණ්,
ඬ, ඬා, ඬැ, ඬෑ, ඬි, ඬී, ඬු, ඬූ, ඬෘ, ඬෲ, ඬෟ, ඬෳ, ඬෙ, ඬේ, ඬෛ, ඬො, ඬෝ, ඬෞ, ඬ්,
ත, තා, තැ, තෑ, ති, තී, තු, තූ, තෘ, තෲ, තෟ, තෳ, තෙ, තේ, තෛ, තො, තෝ, තෞ, ත්,
ථ, ථා, ථැ, ථෑ, ථි, ථී, ථු, ථූ, ථෘ, ථෲ, ථෟ, ථෳ, ථෙ, ථේ, ථෛ, ථො, ථෝ, ථෞ, ථ්,
ද, දා, දැ, දෑ, දි, දී, දු, දූ, දෘ, දෲ, දෟ, දෳ, දෙ, දේ, දෛ, දො, දෝ, දෞ, ද්,
ධ, ධා, ධැ, ධෑ, ධි, ධී, ධු, ධූ, ධෘ, ධෲ, ධෟ, ධෳ, ධෙ, ධේ, ධෛ, ධො, ධෝ, ධෞ, ධ්,
න, නා, නැ, නෑ, නි, නී, නු, නූ, නෘ, නෲ, නෟ, නෳ, නෙ, නේ, නෛ, නො, නෝ, නෞ, න්,
ඳ, ඳා, ඳැ, ඳෑ, ඳි, ඳී, ඳු, ඳූ, ඳෘ, ඳෲ, ඳෟ, ඳෳ, ඳෙ, ඳේ, ඳෛ, ඳො, ඳෝ, ඳෞ, ඳ්,
ප, පා, පැ, පෑ, පි, පී, පු, පූ, පෘ, පෲ, පෟ, පෳ, පෙ, පේ, පෛ, පො, පෝ, පෞ, ප්,
ඵ, ඵා, ඵැ, ඵෑ, ඵි, ඵී, ඵු, ඵූ, ඵෘ, ඵෲ, ඵෟ, ඵෳ, ඵෙ, ඵේ, ඵෛ, ඵො, ඵෝ, ඵෞ, ඵ්,
බ, බා, බැ, බෑ, බි, බී, බු, බූ, බෘ, බෲ, බෟ, බෳ, බෙ, බේ, බෛ, බො, බෝ, බෞ, බ්,
භ, භා, භැ, භෑ, භි, භී, භු, භූ, භෘ, භෲ, භෟ, භෳ, භෙ, භේ, භෛ, භො, භෝ, භෞ, භ්,
ම, මා, මැ, මෑ, මි, මී, මු, මූ, මෘ, මෲ, මෟ, මෳ, මෙ, මේ, මෛ, මො, මෝ, මෞ, ම්,
ඹ, ඹා, ඹැ, ඹෑ, ඹි, ඹී, ඹු, ඹූ, ඹෘ, ඹෲ, ඹෟ, ඹෳ, ඹෙ, ඹේ, ඹෛ, ඹො, ඹෝ, ඹෞ, ඹ්,
ය, යා, යැ, යෑ, යි, යී, යු, යූ, යෘ, යෲ, යෟ, යෳ, යෙ, යේ, යෛ, යො, යෝ, යෞ, ය්,
ර, රා, රැ, රෑ, රි, රී, රු, රූ, රෘ, රෲ, රෟ, රෳ, රෙ, රේ, රෛ, රො, රෝ, රෞ, ර්,
ල, ලා, ලැ, ලෑ, ලි, ලී, ලු, ලූ, ලෘ, ලෲ, ලෟ, ලෳ, ලෙ, ලේ, ලෛ, ලො, ලෝ, ලෞ, ල්,
ව, වා, වැ, වෑ, වි, වී, වු, වූ, වෘ, වෲ, වෟ, වෳ, වෙ, වේ, වෛ, වො, වෝ, වෞ, ව්,
ශ, ශා, ශැ, ශෑ, ශි, ශී, ශු, ශූ, ශෘ, ශෲ, ශෟ, ශෳ, ශෙ, ශේ, ශෛ, ශො, ශෝ, ශෞ, ශ්,
ෂ, ෂා, ෂැ, ෂෑ, ෂි, ෂී, ෂු, ෂූ, ෂෘ, ෂෲ, ෂෟ, ෂෳ, ෂෙ, ෂේ, ෂෛ, ෂො, ෂෝ, ෂෞ, ෂ්,
ස, සා, සැ, සෑ, සි, සී, සු, සූ, සෘ, සෲ, සෟ, සෳ, සෙ, සේ, සෛ, සො, සෝ, සෞ, ස්,
හ, හා, හැ, හෑ, හි, හී, හු, හූ, හෘ, හෲ, හෟ, හෳ, හෙ, හේ, හෛ, හො, හෝ, හෞ, හ්,
ළ, ළා, ළැ, ළෑ, ළි, ළී, ළු, ළූ, ළෘ, ළෲ, ළෟ, ළෳ, ළෙ, ළේ, ළෛ, ළො, ළෝ, ළෞ, ළ්,
ෆ, ෆා, ෆැ, ෆෑ, ෆි, ෆී, ෆු, ෆූ, ෆෘ, ෆෲ, ෆෟ, ෆෳ, ෆෙ, ෆේ, ෆෛ, ෆො, ෆෝ, ෆෞ, ෆ්.

Font levels

Sinhala fonts are standardized in to three different levels. Level 1 font provides basic Sinhala support whereas higher level fonts have more advanced features.

Level 1 fonts	Level 2 fonts	Level 3 fonts
These fonts commonly used vowels, consonants and consonants with modifiers and they are intended to be used in mobile devices. This level supports special characters such as yansaya,rakaransaya and repaya and “ක්ෂe”	These fonts shall have all the features of level one font and additionally support existing combination of Sinhala consonants with repaya. These fonts are intended for general applications such as documents and books.	These fonts support special characters such as ඍ and ඎ and also support all combinations of strokes with conjuncts including “repaya+ispilla” combinations. Touching letters are also supported. These are intended for advanced publications and printing applications which supports pali and Sanskrit scripts and historic documents.



Chapter 05

INTRODUCTION TO FONTS

by Harsha Wijayawardhana,

Theekshana

1. Introduction

In typography, a font is regarded as a set of characters (it is usually a Alphabet) of a single size and style of a particular typeface. A typeface can include one or many fonts with a single style or stylistic unity, but can vary in sizes with a coordinated set of glyphs which represent characters, symbols punctuation, and numerals. The word 'glyph' plays a major role in a font and a typeface where a glyph can be defined as an element of writing and is a vague term in typography especially in multi lingual fonts. In Indic scripts, a set of conjunct characters or letters (conjunction of atomic characters) can make a single glyph. Prior to the advent of desktop publishing, a font and a typeface had very distinct meanings. In those days, a font was a member of a family of fonts whereas a typeface was a family of fonts with a similar visual appearance with specific sizes. An example, a member of 6 point Roman Sans Serif bold or italic is a font and a Roman Sans Serif is a typeface. The size of a font is measured in points and in the pre computer age, a particular point size was a different font, the distinction of which, in the era of computer desktop publishing where size is easily scalable in outline-fonts, has become obsolete. In other words, a single font is scalable to any point size.

2. Font Sizes

A point size of a font was mentioned in the previous section and a point had been defined differently in different periods. In the present day of desktop publishing the most popular way of defining a point of a font or typeface is as a Desktop Publishing unit, which is 1/72 of an inch or 0.0139 in/0.35mm. Specified size of the Desktop Publishing Unit is set to scale size of an em-square which is an invisible box, the height of which is typically bit larger than distance from the tallest ascender to lowest descender. Measurements of non typographic units (meter, feet, and inches) are given as Caps height or height of Capital Letters.

3. Digital fonts

Digital fonts can be divided into two main categories: Bitmap fonts and outline fonts. Bitmap fonts store an image of character or a letter as a Bitmap whereas, in an outline font, an image of a character or a letter is stored as mathematical descriptions of lines or curves. Outline fonts are also known as Vector Fonts. When outlines fonts are used, Rasterizing routines are carried out either by application software, Operating System or printers and outlines of characters in a font are rendered according to vector interpretations which are stored in the font. Digital fonts also contain information on font metrics for composition as well as information on kerning pairs. Common fonts formats include METAFONT, PostScript Type1, True Type Fonts and Open Type fonts. Digital fonts are created using font editors such as Fontlab and Fontforge.

Typefaces are divided into two main categories: Serif and Sans serif. Serifs are a decorative type of typeface where a serif font comprises small features at the end of strokes within a letter. Fonts which do not have these small features are known as Sans Serif. 'Sans' means 'Without' in French. In other words Sans Serif means 'without serif'. Usually, printed works such as newspapers and books are printed with Serif fonts. Web sites do not require specifying a particular type of font and are normally displayed with a font which is specified by the browser settings. But those web sites that are normally specified with a specific font are displayed with Sans Serif fonts as a general rule.

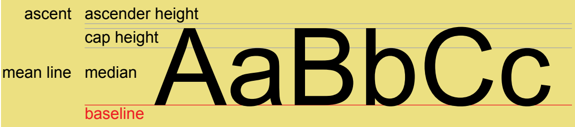


Figure 3 Sans Serif Typeface

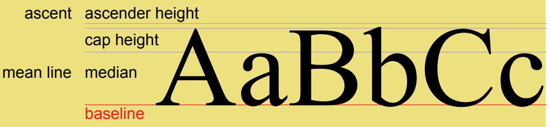


Figure 4 Serif (courtesy of Wikipedia)

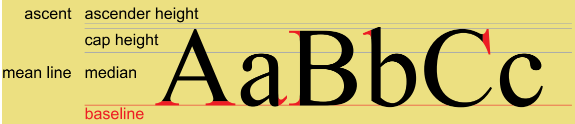


Figure 5 Red Color provides for marking Serif

Some Sinhala Typefaces

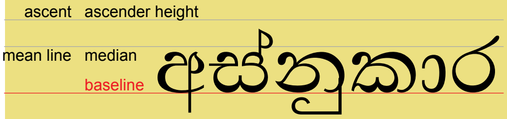


Figure 4 Bhashitha

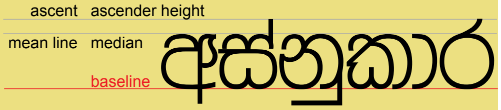


Figure 5

4. Font Metrics

Understanding of font metrics is of great importance when designing fonts. Characters of most scripts of the world rest on an imaginery horizontal line. This horizontal line is called a Baseline. Some glyphs (such as English letter 'y') of scripts have certain parts below baseline, and distance between edge of lowest descending part of the character of a script and the baseline is known as descent. And the part which descends below the baseline is called Descender. As opposed to the descent, certain characters of scripts have parts spanning upwards vertically such as letter 'h' and the distance between the character or the glyph of a script which spans farthest upwards and the baseline is known as ascent. Since languages such as English and German have capital and simple letters, heights that are taken by capital letters are known as Cap Height. When glyphs are aligned horizontally, an imaginary median line can be drawn above most of upper edge of English simple letters and this line is known as median. The distance between median line and baseline is called x-height. Sinhala and Tamil belong to Indic languages and they do not have Capital and Simple letters.

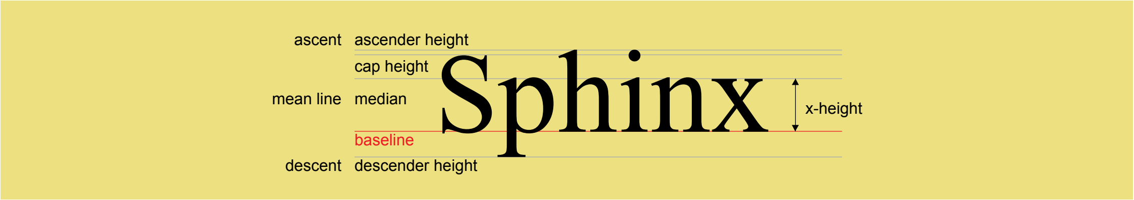


Figure 6 Courtesy of Wikipedia

5. Bimap fonts

Bitmap fonts store their glyphs as an array of pixels or bitmaps. These fonts are also known less commonly as Raster fonts since bitmap fonts are simply a collection of Raster images of Glyphs. For each character of a font with single size, an image is created. For different sizes of a font, different sets of images to suit those sizes are created. As an example, if three different sizes of font are created with any combination of Bold and Italic sets, 12 different completely sets of fonts have to be created. Bitmap fonts are easier to create, and faster to render, but they have certain drawbacks. If Bitmap fonts are arbitrarily scaled up or magnified to a larger size, the visual appearance of the font may not be appealing due to jagged edges of bitmap. Some of the advantages of Bitmap fonts are that they are extremely fast and easier to render; and they are easier to create than the other kinds. In addition, bitmaps give the same output when unscaled.

Due to speed and ease of rendering, Bitmap fonts are used for displaying on Linux and Windows recovery consoles. Bitmaps are also used in systems which have fewer resources such as embedded systems.

6. Outline fonts

Outline fonts are also known as Vector fonts as mentioned previously since borders and curves of characters of an outline font are given mathematically and are rendered by either Operating System, Software Application or printer. Outline fonts use Bezier curves in creating Vector fonts. Since curves and lines are given in mathematical formula, outline fonts can be scaled bigger without causing pixilation. Although outline fonts are aesthetically appealing, they require more processing and could produce undesirable rendering depending upon the font, and the rendering software, and output size. One of the major problems of outline fonts is the inability of accurately rendering of Bezier curves onto raster display. This has been overcome by methods, such as Font Hinting, which require more sophisticated and expensive software.

Historically, Adobe developed outline fonts and Type I and Type 3 were some of the original fonts which were developed. Using Postscript, the glyphs of Type I are outline fonts with Bezier curves. Unlike Type 1 of Adobe, Type 3 used more widely of PostScript language. But Type 3 had not used Font hinting and had issues when scaling up and down especially rendering on low resolution devices.

7. True Type fonts (TTF)

True type fonts (TTF) were developed by Apple to replace adobe's Type 1 font in the 1980s and belong to the outline font category. Unlike type 1 and type 3 of Adobe, True type fonts have become more popular and are widely used today. In those days, True type fonts gave font developers a greater degree of freedom and control, allowing them to manipulate to pixel level of a True Type Font which was not possible earlier and which may not be allowed any more in the modern day TTF fonts.

Apple licensed True Type fonts to Microsoft Microsoft developed a Postscript like language to be used with Laser printers which was used by Apple on their Apple laser printers. Microsoft partnered with Monotype Corporation. This was a major company tracing its history to the latter part of the 1800s. Monotype Corporationwas in manufacturing and production of printing machinery, for introduction of introduce several new technologies in Font hinting such anti aliasing which was later developed into full blown True Type hinting. This enabled Microsoft and Monotype engineers to use TTF fonts in low resolution devices such as Dot Matrix printers, a major breakthrough indeed. Microsoft began using TTF fonts on Windows 3.1 and Microsoft and Monotype Corporation developed several typefaces such as Courier new compatible with Courier, and Time New Roman compatible with Roman. These fonts have survived on some of the latest Microsoft Operating Systems such as Vista, Microsoft Windows 2008 and Windows 7.

In 1994, Microsoft introduced smart font technology which was known as True Type Open which was later developed into Open Type Fonts with the merger of Adobe's Type 1 font outlines in 1996. Today almost every Operating System in the world uses True Type fonts.

8. Open Type fonts

As mentioned previously, Open Type fonts are an extension of True Type fonts with support for Postscript font data. Open Type fonts are also known as Smart Fonts. By 1996, Microsoft and Adobe had already formulated required standards for Open Type fonts. In western writing, usually each character is mapped to a shape hence a one to one map exits. Open Type fonts are important for non Western writing which may have glyphs that are created by the conjunction of two or more atomic characters or syllables.

In order for Open Type fonts to function, operating systems, or applications must support rendering of open type fonts and an operating system accomplishes this by having a rendering engine. Microsoft operating systems have a rendering engine which is known as Uniscribe. Linux has several rendering engines or shaping engines: Pango, QT etc.

Data on shapes of Glyphs and also rules on how they will be rendered are stored in Open Type Fonts as Open Type Layout tables. These rules are parsed by a rendering engine and according to rules, the shaper will either create a single glyph joining two or more glyphs which represent atomic syllables or will place a glyph, which represents several atomic syllables.

When creating an Open Type font, initially a TTF is created using tools such as FontLab, Fontforge or any other vector graphic based font creating tool (TTF creation is explained in the previous chapters). Once the font is created according a desired shape, it has to be programmed as to how the font will be rendered. For placing rules, different tools are used depending upon the programmer. Some use Microsoft VOLT (Visual Open type Layout Tool) or others FontForge; and some create rules by using Text editor which are then later compiled by using MS Volt, Fontforge or any other compilers, some of which are command line based.

9. How to place rules in Open Type fonts

Open Type fonts retain rendering information by having Open Type tables. Some of main Open Type tables used in Non Western Writing are substitution (GSUB) , Positioning (GPOS) and Glyph Definition (GDEF) tables which are categorized as advanced Typographic tables. Several other tables which are used in Open Type Fonts are listed below:

i. CMAP: character to Glyph mapping
ii. Head: Font Header
iii. Name: Naming table etc
iv. Post: PostScript information etc.

The above mentioned advanced Typographic tables are created based upon Open Type layout common format. This format allows storing of Open Type Layout data as Script, Language System, Features, and Lookup. The following gives a basic description of each of the above:

Script: Scripts are defined at the top level and may contain a collection of glyphs which belong to several languages. For instance Latin Script supports English, French, and German etc. In Open Type Fonts, multiple scripts can be supported by a single font. Arabic Script is used to write Farsi, and Urdu. Sinhala is defined under Sinhala Script and Language.

Language System: the Language system modifies the appearance and function of a particular language. For instance Urdu and Farsi have additional glyphs more than what Arabic requires. This is because Urdu and Farsi belong to Indo-European languages although written in Arabic which belongs to the Semitic language group. By defining a language system as Urdu, an Open Type font is able to render Urdu with those which are additional and specific for Urdu language.

Features: A language defines basic features of its typographic characteristics. Certain features are language specific and others apply to a Language group. In the absence of Language specific features, default features are applied.

Lookup tables: Features are implemented or stored by using Lookup tables. Two specific lookup tables are used: GPOS (Positioning) and GSUB (Substitution). Features of Open Type fonts can be implemented using either GPOS or GSUB lookup tables or the both. GPOS and GSUB can be applied to a group or a class of glyphs.

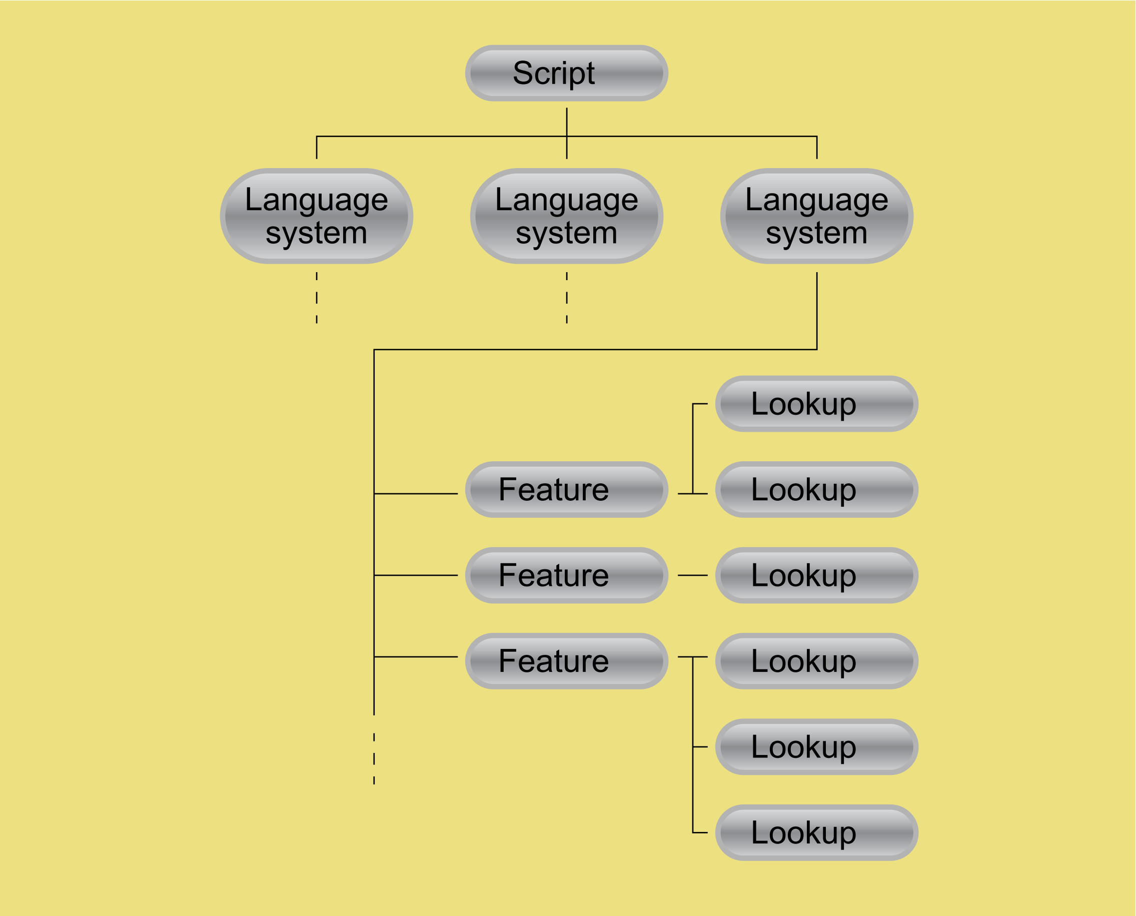


Figure 7 Open Type Layout Common format.(Ref: http://www.microsoft.com/Typography/OTSpec/chapter2.htm)

.

References

1. http://www.microsoft.com/Typography/OTSpec/chapter2.htm
2. http://en.wikipedia.org/wiki/Typeface
3. http://www.w3.org/TR/CSS2/fonts.html
4. Foley, Van Dam, Feiner , Hughs, Computer Graphics principles and Practice, Second Edition Addison Wesley Longman, 2001

Appendix I

How to determine and calculate height and width of a font

1. EM Unit measure of a Font is 16 points: 1EM = 16 points
2. Fontlab measures height and width UPM or Unit per EM
3. A glyph must be within a squire of following UPM unit for different fonts (FontLab calculations)
   i. PS fonts 1000
   ii. TTF 1024
   iii. TTF on a Microsoft platform must be 2048
4. Assuming 1EM square is 260 mm:

- Caps = 200 mm (not needed for Sinhala and Tamil)
- Ascender = 200 mm
- x-height = 140 mm
- Descender = -60

For PS font: using 1000 UPM as measure,

- Caps 200/260 * 1000 = 769
- Ascender 200/260 * 1000 = 538
- x-height 140/260 * 1000 = 769
- Descender *60/260 * 1000 = *231

Appendix II

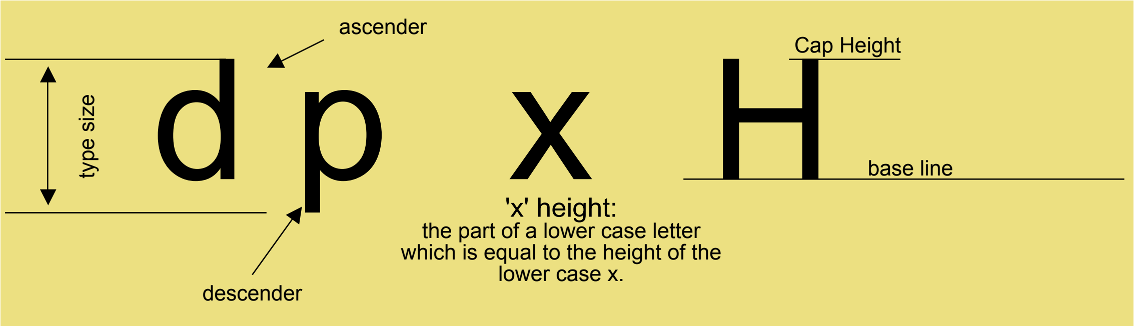


Figure 8 Source: Wikipedia

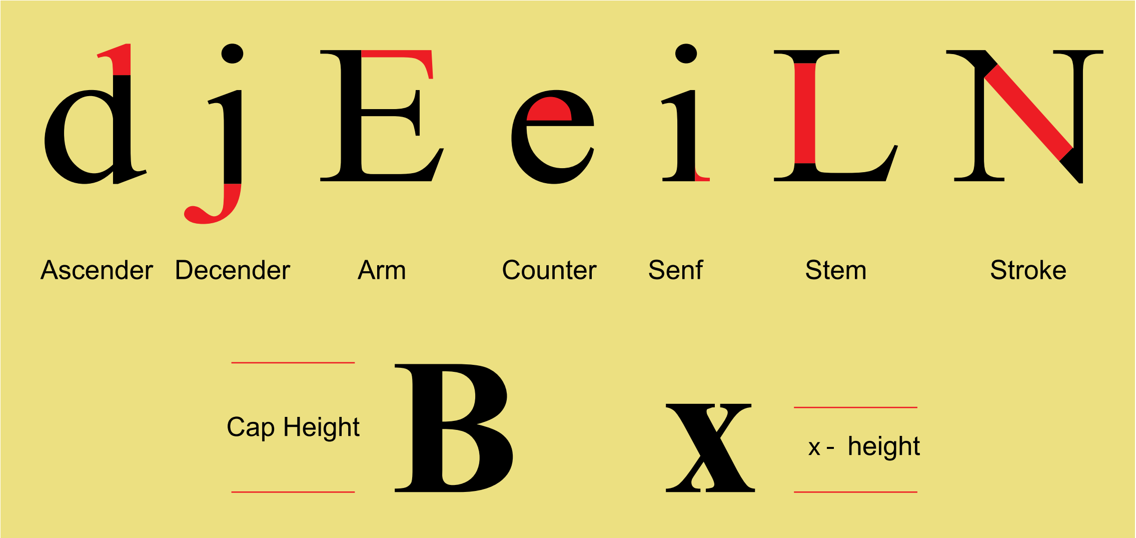
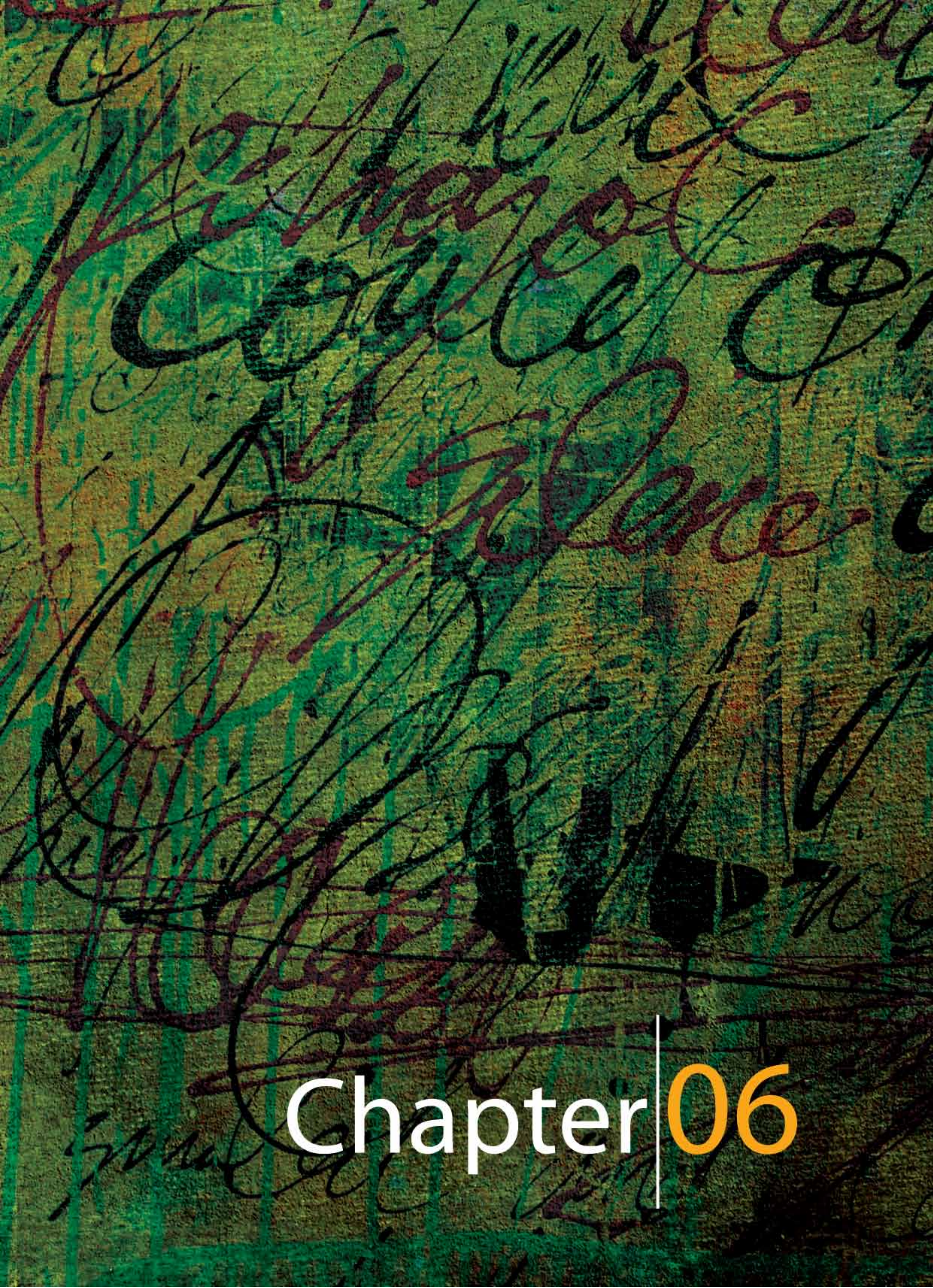


Figure 9 Source: Wikipedia



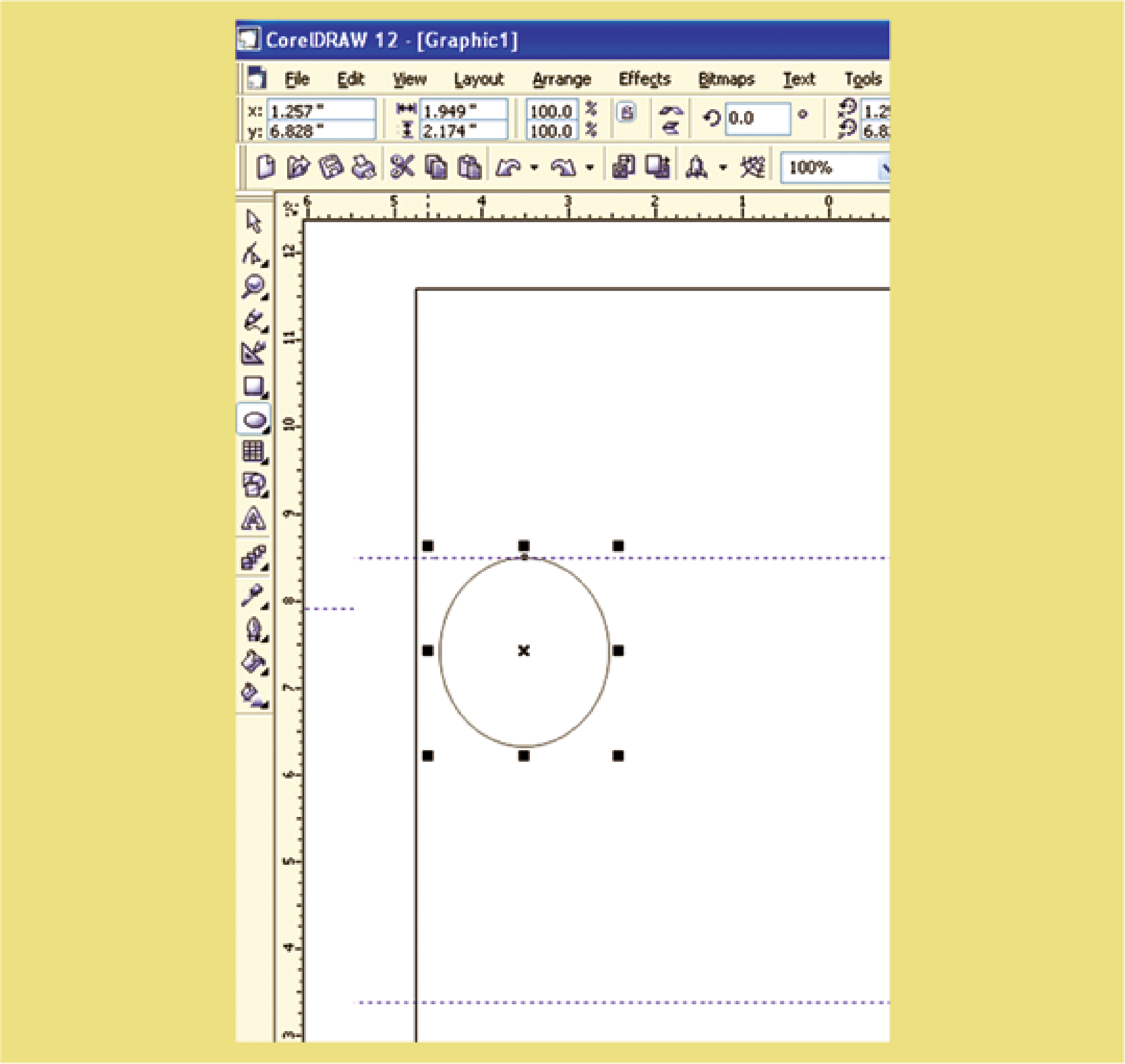
Chapter 06

FONT DESIGN

Winnie Vitharana

University of Visual and PerformingArts

මෙම වැඩසටහන සඳහා Core Draw මෘදුකාංගය උපයෝගි කරගෙන ඇත. එහි Tool box එකේ ඇති වෘත්තය අඳින මෙවලම භාවිතයට ගෙන අකුරු නිර්මාණය කරමි. මෙහි දැක්වෙන්නේ එම මෙවලම භාවිතයට ගනිමින් බින්දුවක් නිර්මාණය කර ආකාරය ය.

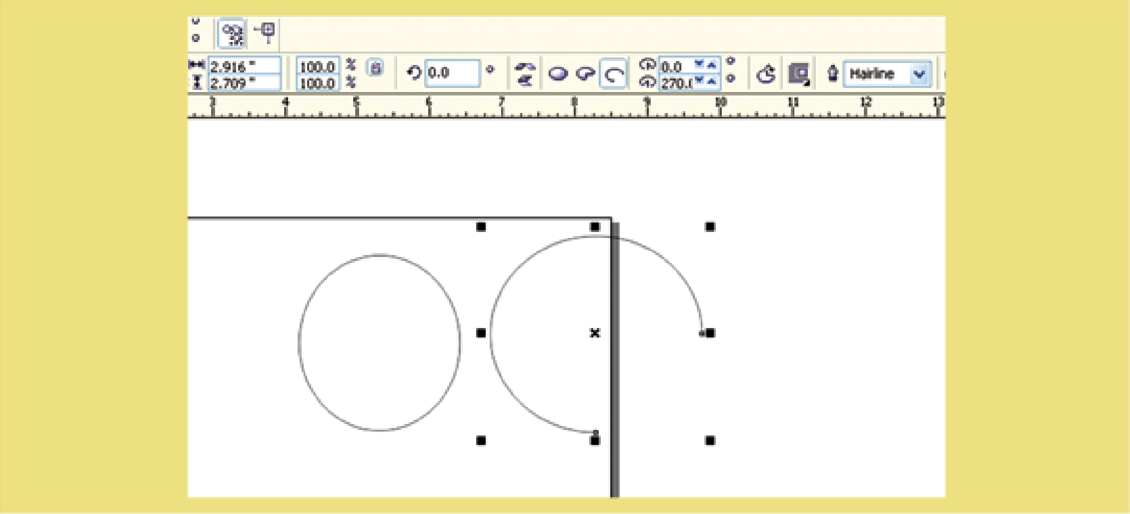


CorelDRAW 12 - [Graphic1]

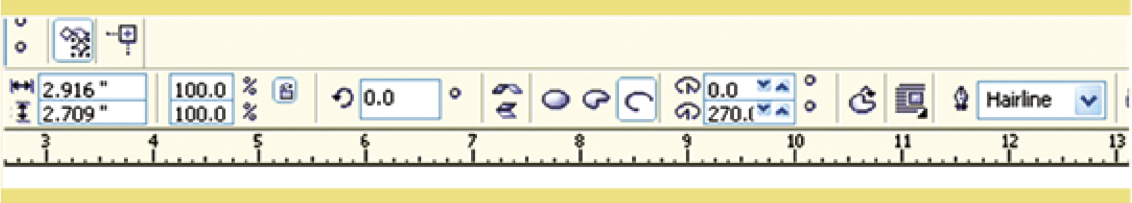
x: 1.257"
y: 6.828"
1.949"
2.174"
100.0 %
100.0 %
0.0 °
100%

(1 රූප සටහන)

අකුරක කොටස් වෙන වෙන ම නිර්මාණ කොට ඒවා, නිසි ලෙස එක් තැන් කිරීමෙන් අකුරක් නිර්මාණය වේ. මෙතැන් සිට පෙන්වීමට අදහස් කරන්නේ සිංහල හෝඩියේ එන “න” අක්ෂරය නිර්මාණය කරන ආකාරයයි.

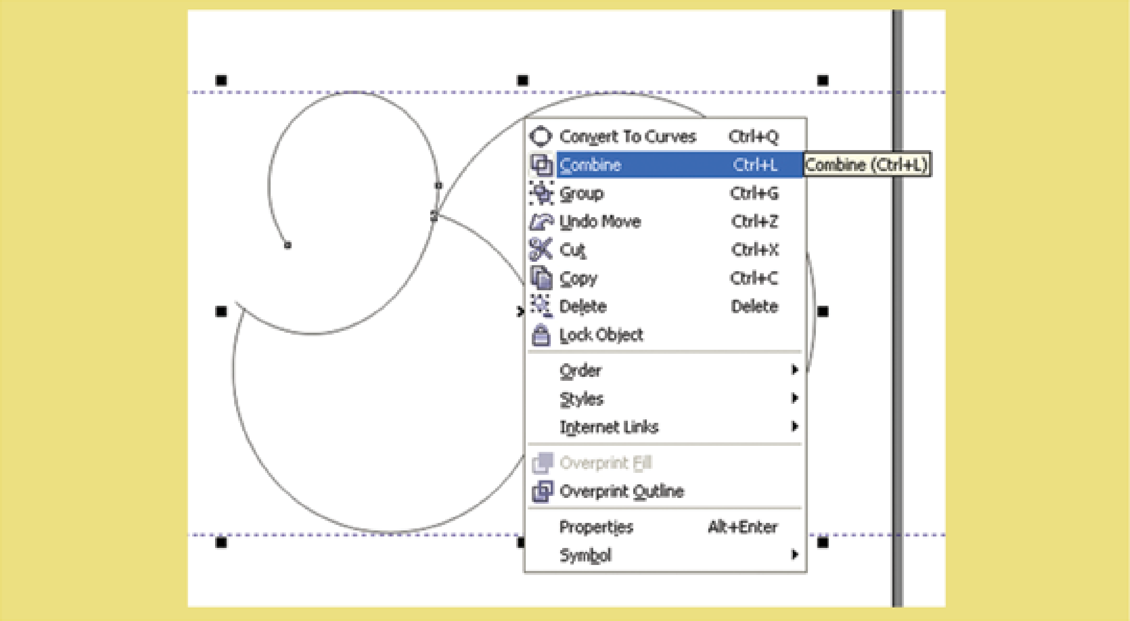


(2 රූප සටහන)



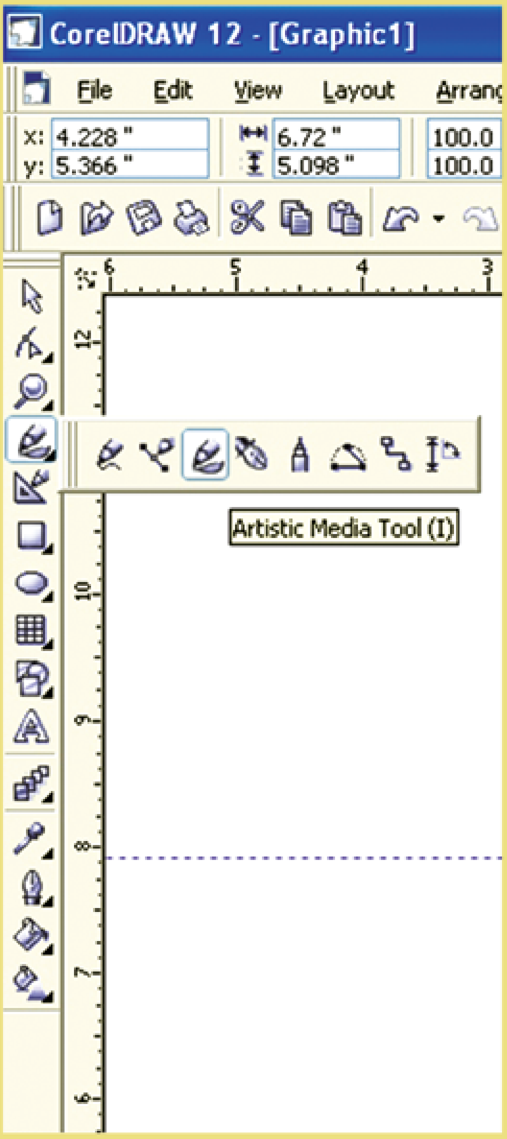
(3 වන රූප සටහන)

වෘත්තය අඳින මෙවලම මගින් වෘත්තය ඇඳ පසුව අපට අවශ්‍ය ආකාරයට එහි රේඛාව කෙටි කළ හැකි ය. ඒ සඳහා කළ යුත්තේ ඉහළින් දැක්වෙන වෘත්තය සහ ඒ ළඟ ඇඳි රේඛාව කෙටි කරන ලද වෘත්තය දැක්වෙන සංඛේතවලින් පෙන්නුම් කරන මෙවලම භාවිත කිරීමයි. 2 සහ 3 රූප සටහන්වලින් වෘත්තය මෙවලම පමණක් භාවිත කරමින් නිර්මාණය කරන ලද “න” අක්ෂරය ඔබට දැකිය හැකිය.



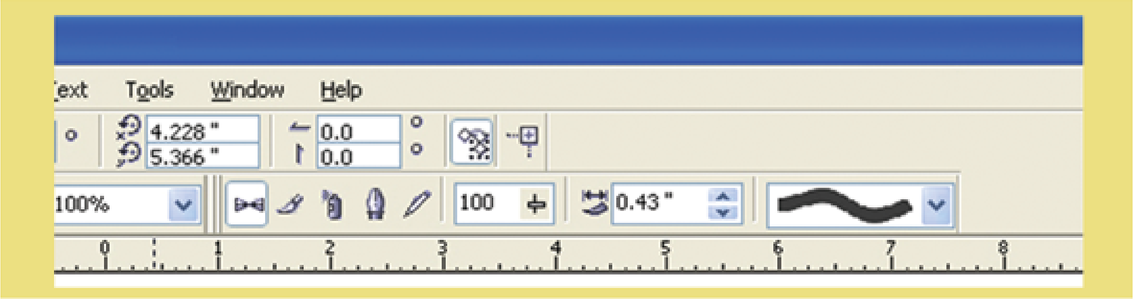
(4 වන රූප සටහන)

මෙහි නිර්මාණය කරන ලද “න” අක්ෂරයේ කොටස් වෙන වෙන ම select වේ. දැන් අප කළ යුත්තේ අක්ෂරය සම්පූර්ණ තෝරා අක්ෂරයේ රේඛාව මත කර්සරය තබා මූසිකය දකුණු ක්ලික් කිරීමය. එවිට විවෘත වන මෙනුවේ ඇති විධාන අතරින් combine යන වචනය ක්ලික් කිරීමෙන් අක්ෂරය නිර්මාණය කර ඇති සියලු කොටස් ඒකාබද්ධ වී තනි අනුරුවක් සෑදේ.



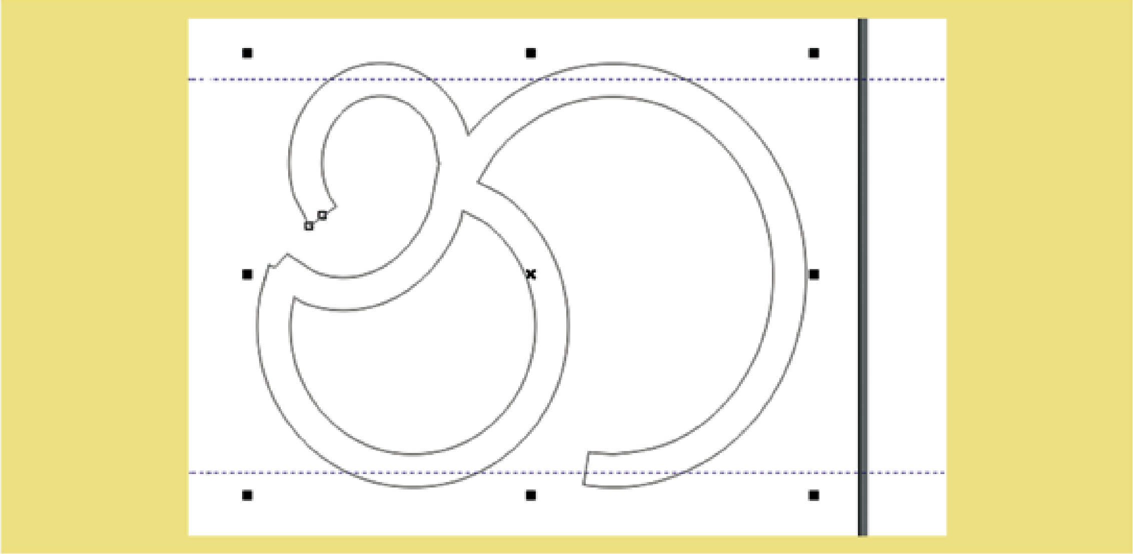
(5 වන රූප සටහන)

ඊළඟට ඔබ කවුළුවේ වම් පැත්තේ ඇති tool box එක වෙත යායුතුය. එහි Artist Media Tool (1) තෝරා ගත් විට 6 වන රූප සටහනේ දැක්වෙන brush tool එක ක්ලික් කළ විට ඒ තුළ ඇති විවිධ brush වර්ග විවෘත වෙයි.



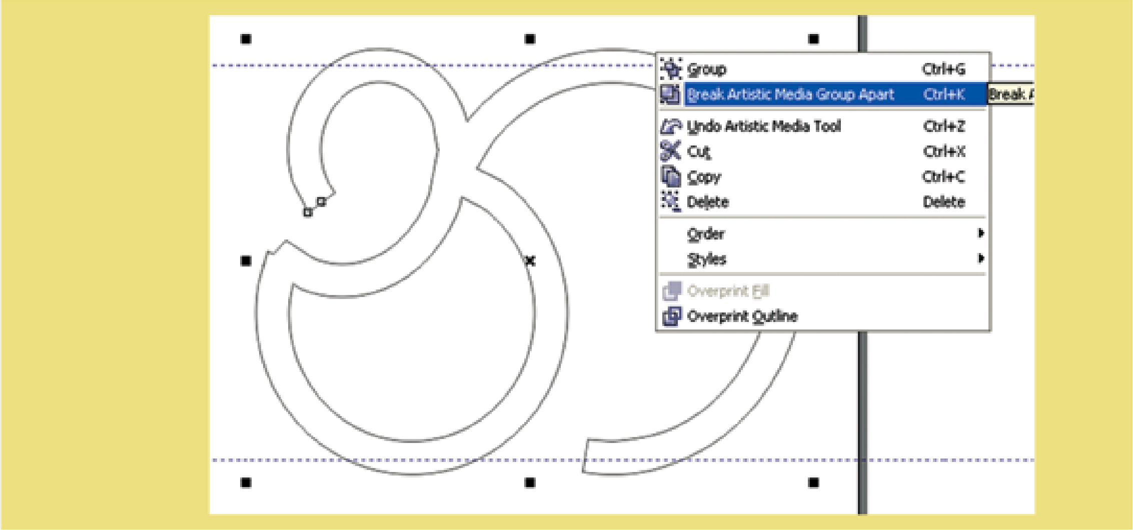
(6 වන රූප සටහන)

මේ අතරින් මා තෝරා ගත් *brush* එක *7* වන රූප සටහනේ දැක්වෙයි.



(7 වන රූප සටහන)

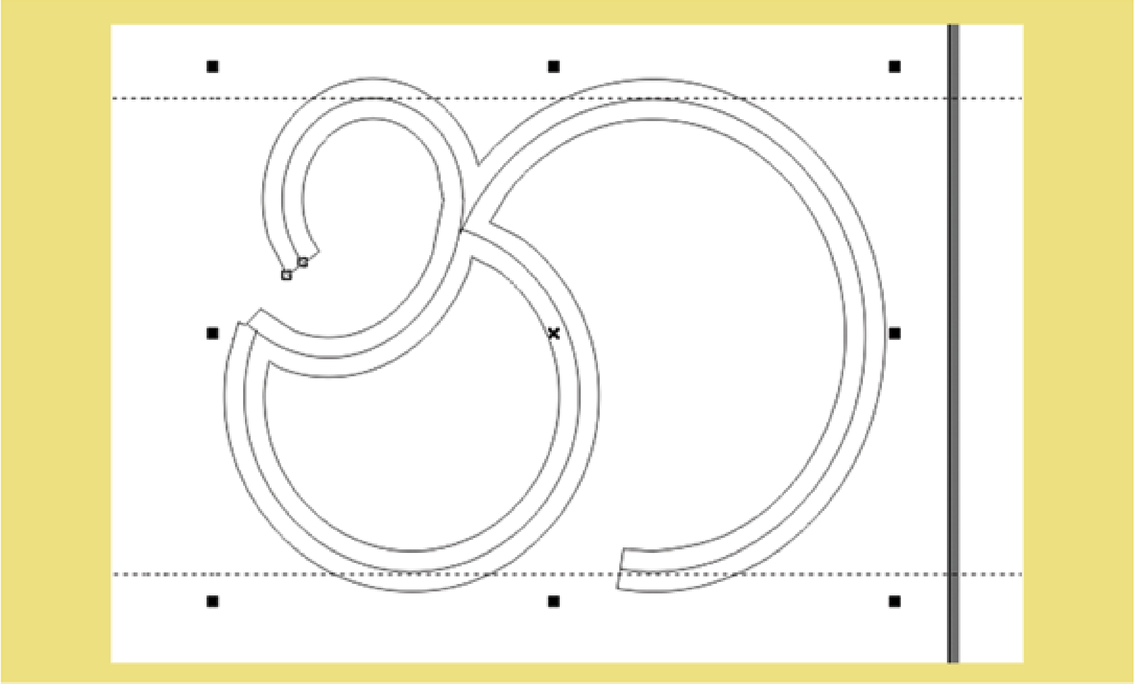
එසේ තෝරා ගත් කළ මෙතෙක් hairline රේඛාවෙන් තිබූ අක්ෂරය 8 වන රූප සටහනේ “න” අක්ෂරය මෙන් bold වේ. ඔබට අවශ්‍ය නම් brush size වෙනස් කිරීමෙන් අක්ෂරය නිර්මාණය කරන ලද රේඛාව සිහින් කිරීම හෝ මහත් කිරීම කළ හැකිය.



(8 වන රූප සටහන)

(8) ඔබේ curserයේ තවමත් ඇත්තේ brush tool එකයි. දැන් ඔබ එම tool එක අක්ෂරයේ රේඛාවක් මත තබා දකුණු ක්ලික් කළ යුතුය. එවිට මෙහි දැක්වෙන මෙනුව විවෘත වේ. එහි දෙවැනියට ඇති Break Artistic Media Group apart නම් විධානය දිස්වෙයි. ඇතැම් විට මෙනුවේ මෙය දිස් නොවිය හැකියි. එසේ වන්නේ ඔබ කර්සරය රේඛාව මත තබා ක්ලික් කිරීමේදී සිදුව ඇති කුඩා වෙනස්කමක් නිසාය. එය නැවත කිරීමෙන් ඔබට නිවැරදි ලෙස මෙනුව තෝරා ගත හැකිය.

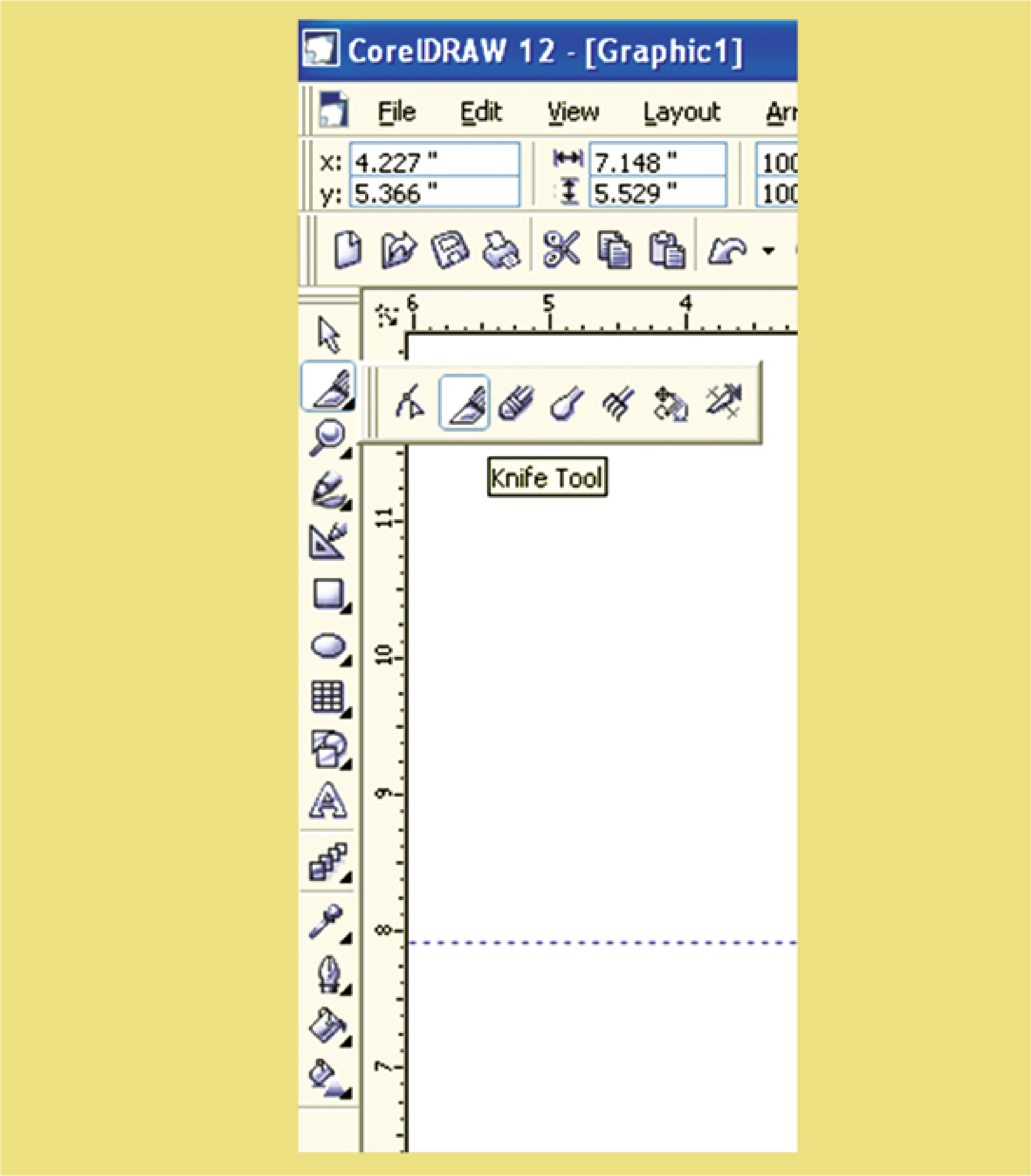
එම විධානය දීමෙන් පසු 9 වන රූප සටහනේ දැක්වෙන අක්ෂරය මධ්‍යයේ ඇති hairline එක මතු වේ.



(9 වන රූප සටහන)

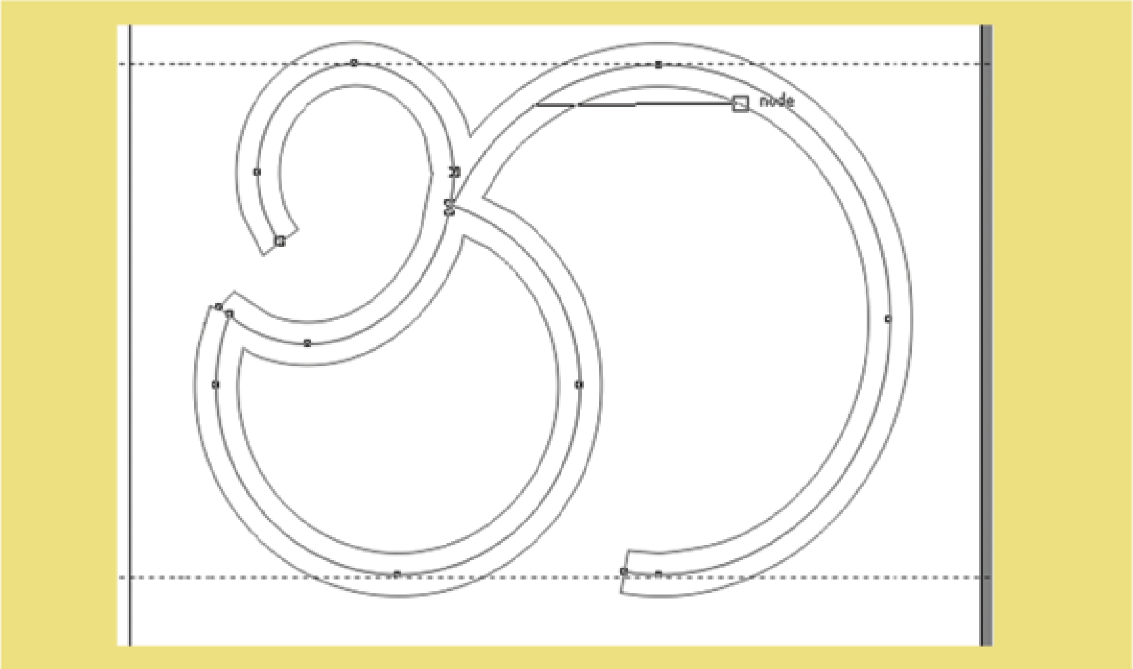
අප එය ඉවත් කළ යුතුය. එය ඉවත් කරන්නේ මෙසේ ය.

(10) රූප සටහනේ දැක්වෙන කවුළුවේ වම්පස ඇති tool box එකේ ඇති දෙවන නිරූපකය (icon එක) ක්ලික් කළ විට රූප සටහනේ දැක්වෙන icon හත දිස්වේ. එයින් knife tool එක තෝරා ගන්න.



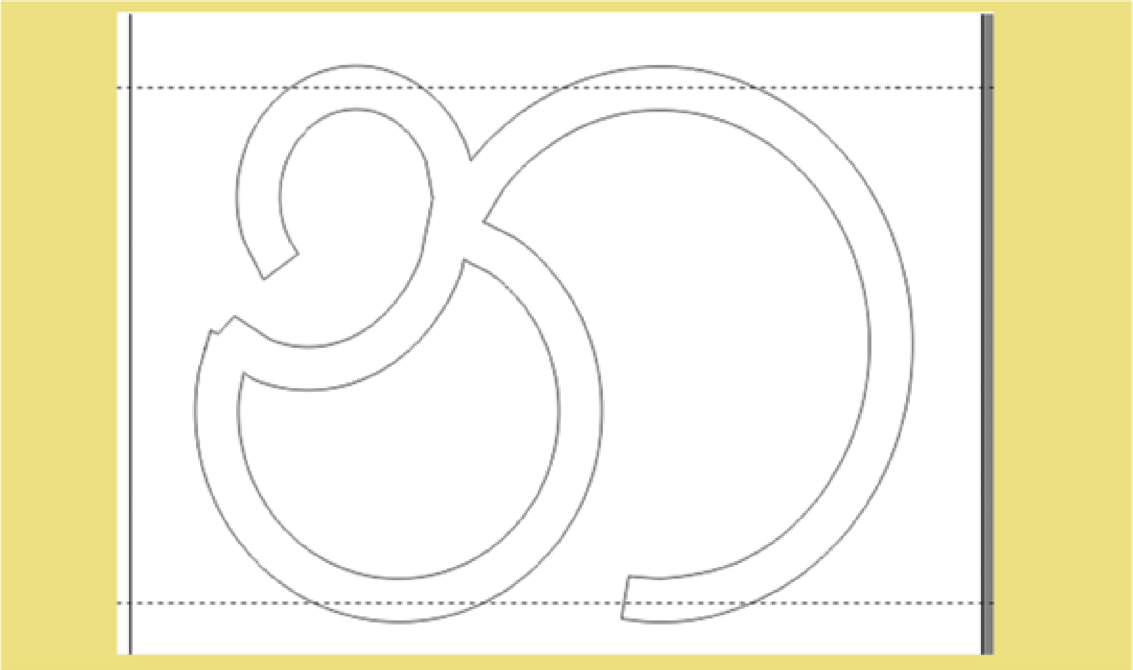
(10 වන රූප සටහන)

දැන් (11) වන රූප සටහනේ දැක්වෙන ආකාටෙ knife tool එක ගෙන අක්ෂරයේ ඇති hairline එක මත තබා ක්ලික් කොට නැවත රේඛාවේ වෙනස් තැනක තබා ක්ලික් කරන්න. එවිට node දෙක අතර රේඛාව සෘජු රේඛාවක් ලෙස දිස් වෙයි. අනතුරුව මකන්න. එවිට අක්ෂරයේ තිබූ hairline එක මුළුමනින් ම මැකෙයි.



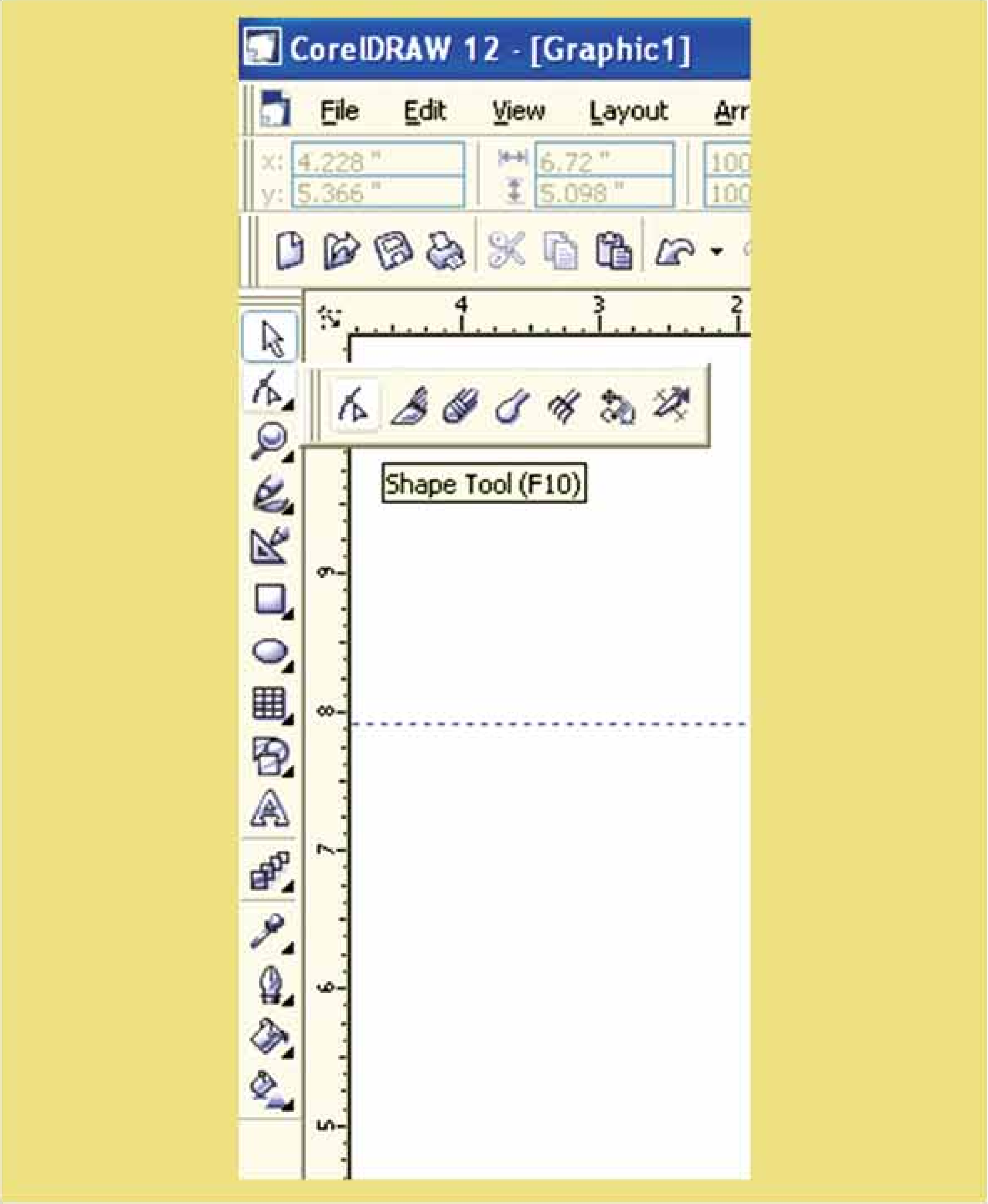
(11 වන රූප සටහන)

දැන් අක්ෂරය 12 වැනි රූප සටහනේ ඇති පරිදි දිස්වේ.



(12 වන රූප සටහන)

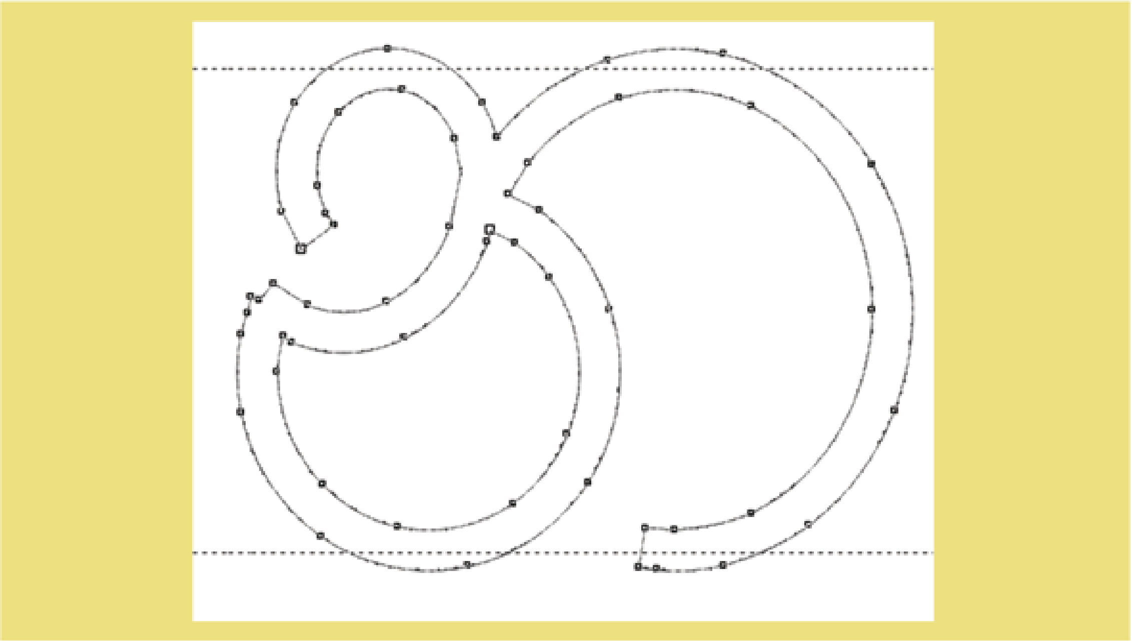
මීළඟට කලින් knife tool එක ගත් තැනම ඇති (13 රූප සටහන) shape tool (F10) තෝරා ගන්න.



x: 4.228 "
y: 5.366 "
6.72 "
5.098 "
100
100

(13 වන රූප සටහන)

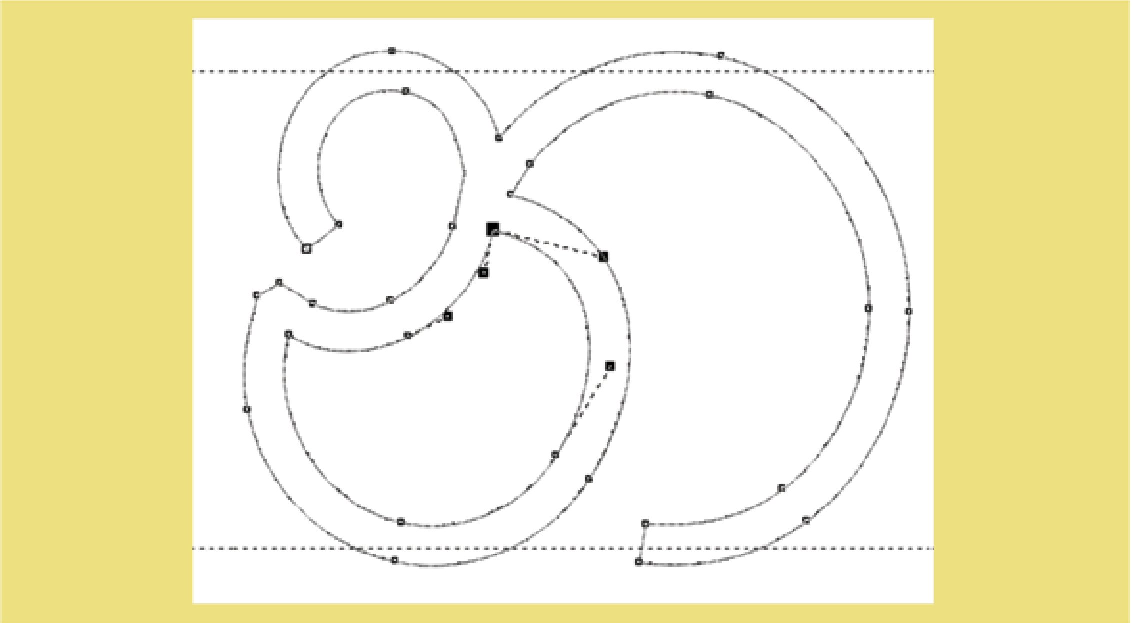
එය අකුර මත කොතැනක හෝ තබා ක්ලික් කරන්න. එවිට ඔබට 14 වන රූප සටහනේ මෙන් අක්ෂරය පුරා තිබෙන නෝඩු දැක ගත හැකි වනු ඇත.



(14 වන රූප සටහන)

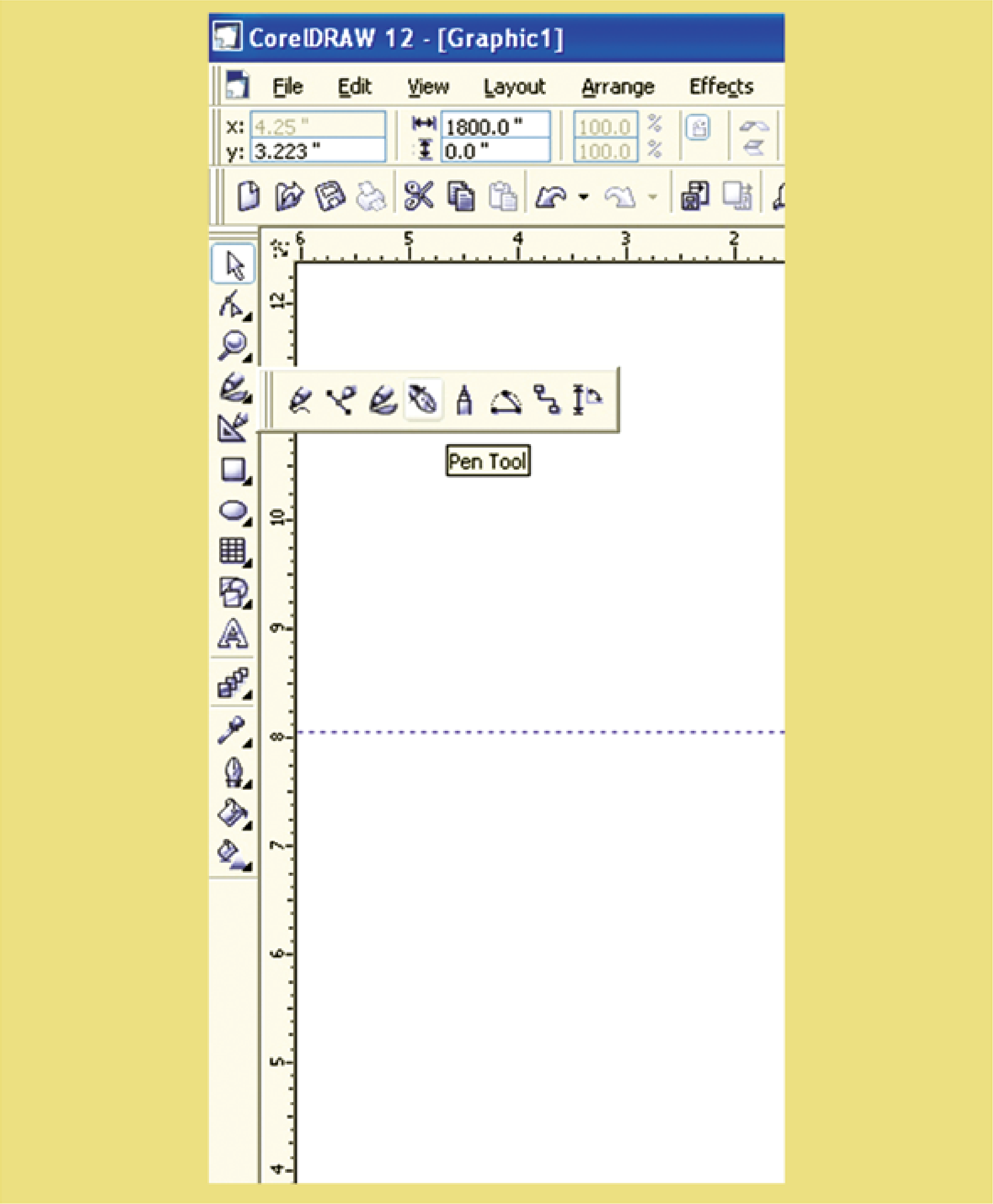
Shape tool එක nod එකක් මත තබා දෙවරක් click කළ විට හෝ එක් වරක් ක්ලික් කොට මැකූ විට හෝ node එක මැකේ. හැම node එකකම රේඛාව එහාට මෙහාට හැරවිය හැකි පහසුකම් දැකිය හැකිය.

(15 වන රූප සටහන) දැන් ඔබ කළ යුත්තේ අනවශ්‍ය nodes මකා අක්ෂරය වඩාත් ඔප මට්ටම් වන ලෙස රේඛාව හසුරුවා අවසාන නිර්මාණය සකස් කිරීමය.



(15 වන රූප සටහන)

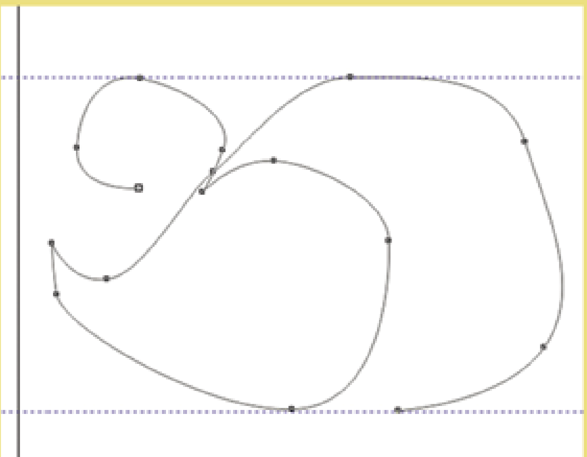
ජ්‍යාමිතික මෙවලම් උපයෝගී කරගෙන අක්ෂර නිර්මාණය කිරීම හැර free hand අක්ෂර නිර්මාණය කිරීම ද කළ හැකිය.



x: 4.25"
y: 3.223"
1800.0"
0.0"
100.0 %
100.0 %

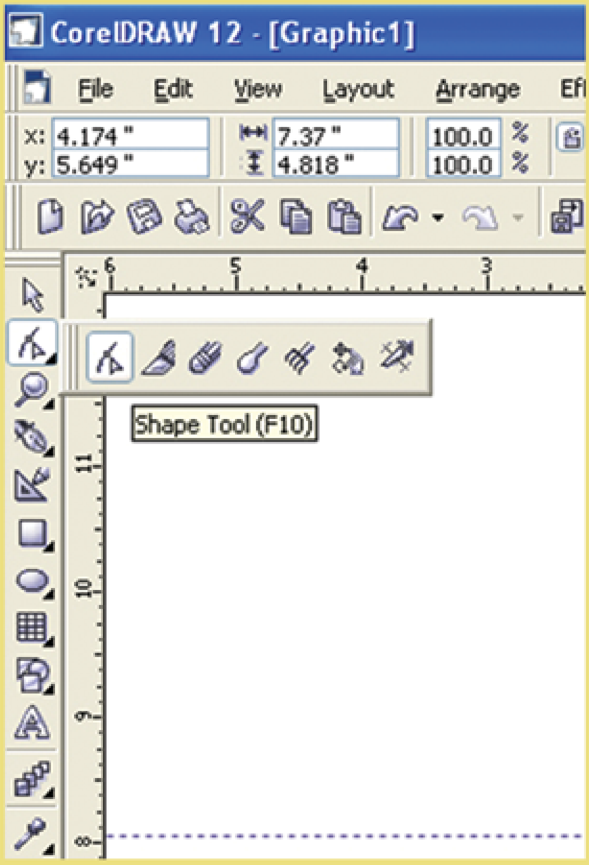
(16 වන රූප සටහන)

ඒ සඳහා ඔබ කවුළුවේ වම් පස ඇති tool box එකෙන් pen tool එක තෝරා ගත යුතුය. අනතුරුව pen tool එක භාවිත කරමින් අකුරක් නිර්මාණය කළ යුතුය. එවිට 17 වන රූප සටහනේ මෙන් අකුරක් ඔබට දැක ගත හැකිය.



(17 වන රූප සටහන)

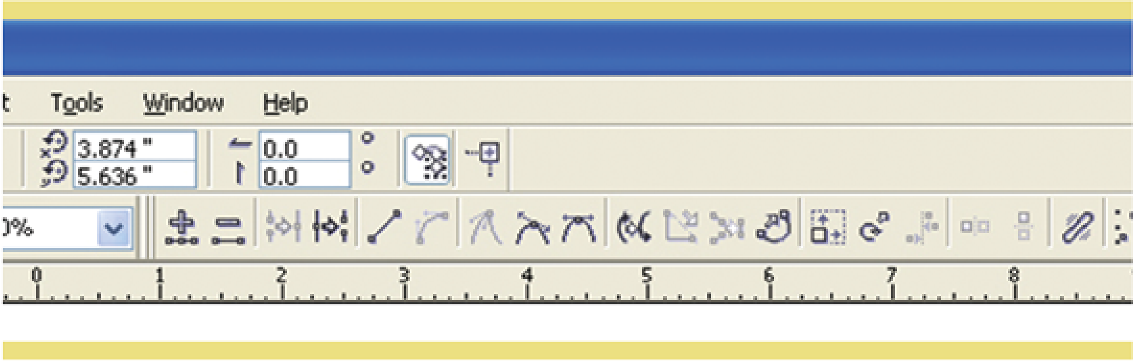
අනතුරුව ඔබ කලින් මෙන් shape tool එක තෝරා ගෙන පෙර නිර්මාණය කළ අක්ෂරයේ හැඩය නිවැරදි කළ යුතුය. (18 වන රූප සටහන)



x: 4.174 "
y: 5.649 "
7.37 "
4.818 "
100.0 %
100.0 %

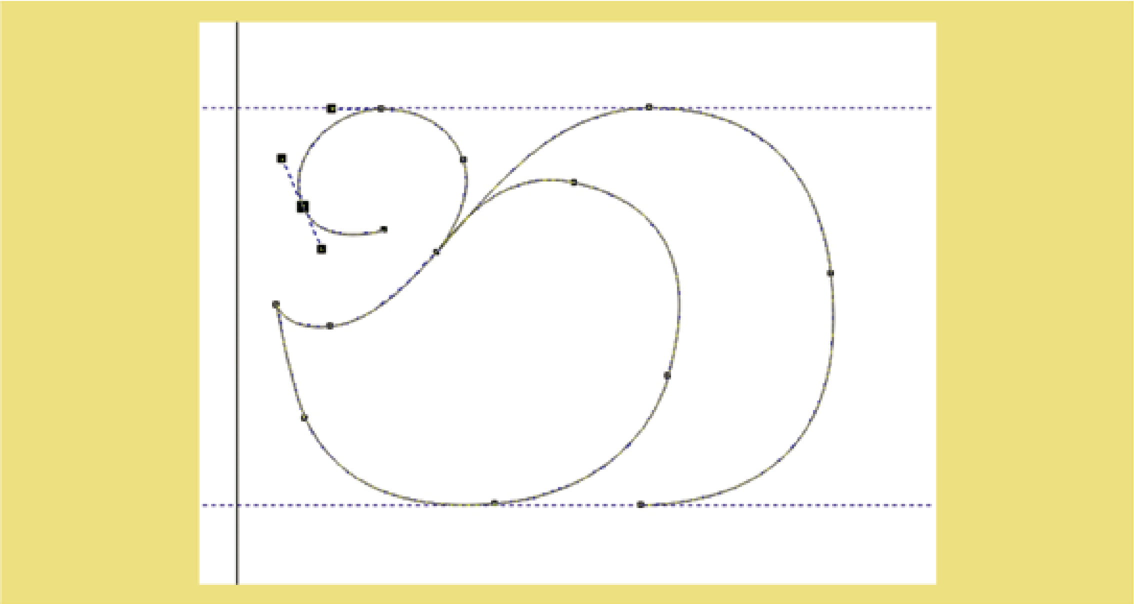
(18 වන රූප සටහන)

ඒ සඳහා අවශ්‍ය කරන පහසුකම් 19 වන රූප සටහනේ දක්වා ඇති මෙවලම් වලින් ලබා ගත හැකිය.



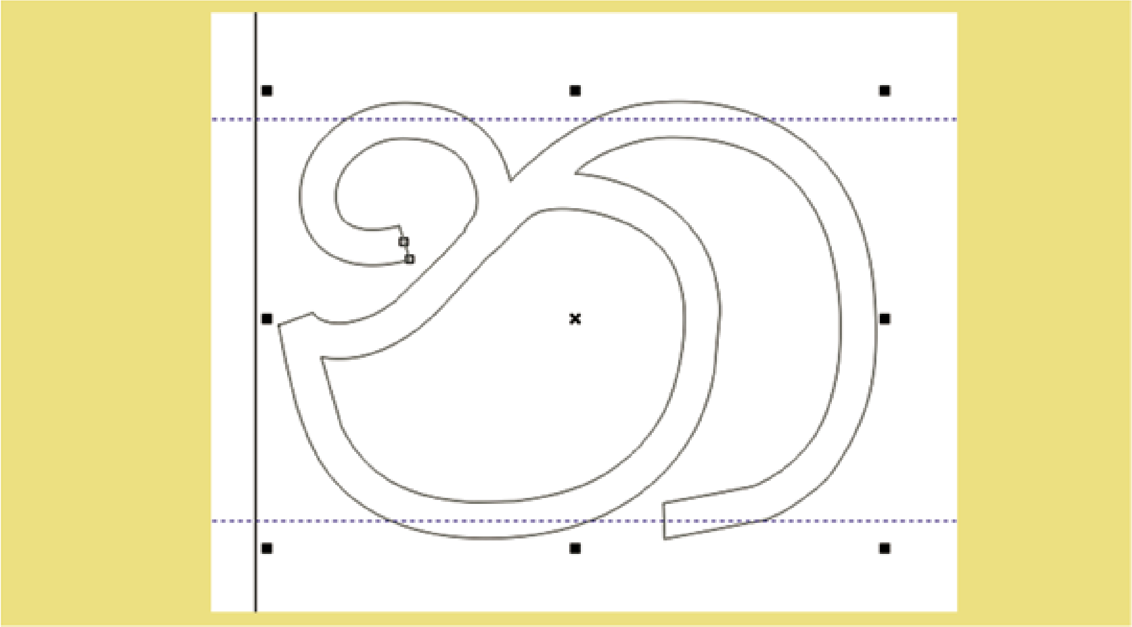
(19 වන රූප සටහන)

එසේ නිර්මාණය කළ අක්ෂරය 20 රූප සටහනින් දැක්වේ.

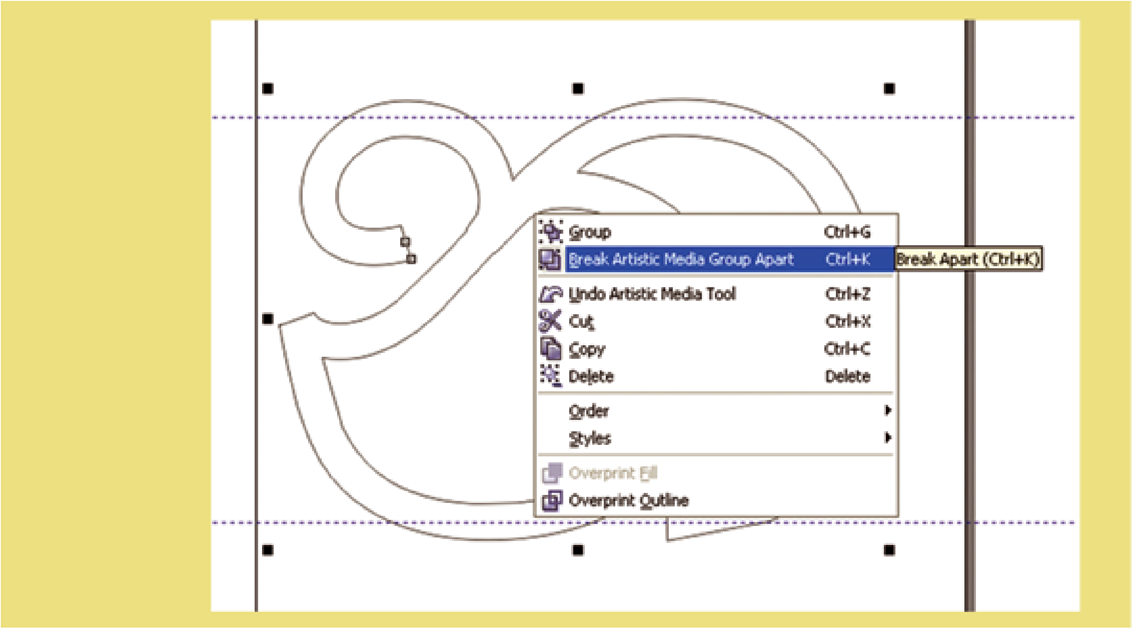


(20 වන රූප සටහන)

මෙතැන් සිට කලින් අක්ෂරය නිර්මාණය කිරීමේ දී කළ දෑ අනුගමනය කළ යුතුය. 21, 22 රූප සටහන් දෙකෙන් පෙන්නුම් කරන්නේ එසේ කරන ලද නිර්මාණයේ අවස්ථා දෙකකි.

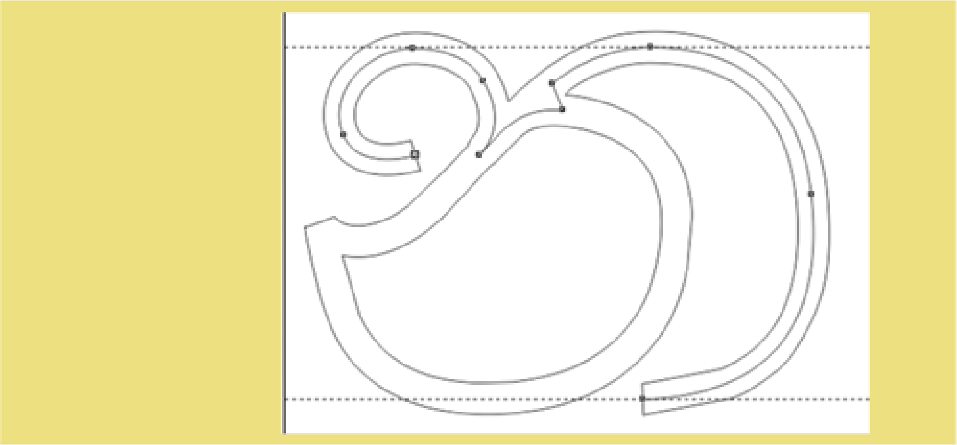


(21 වන රූප සටහන)

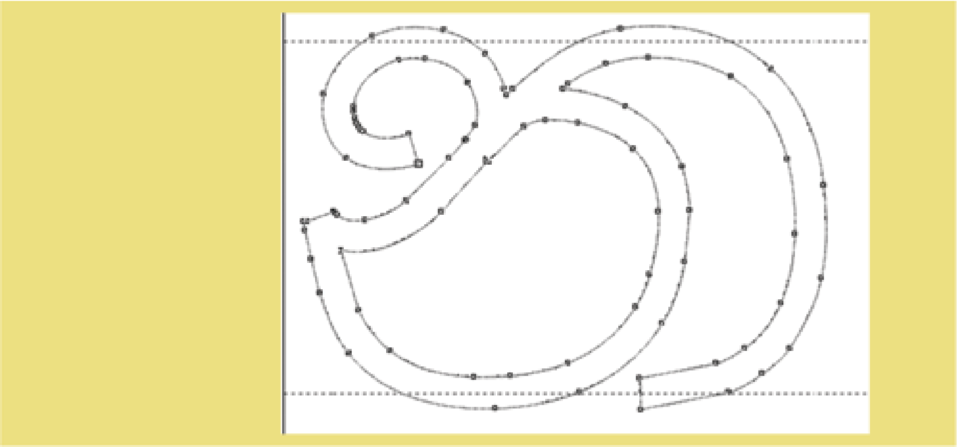


(22 වන රූප සටහන)

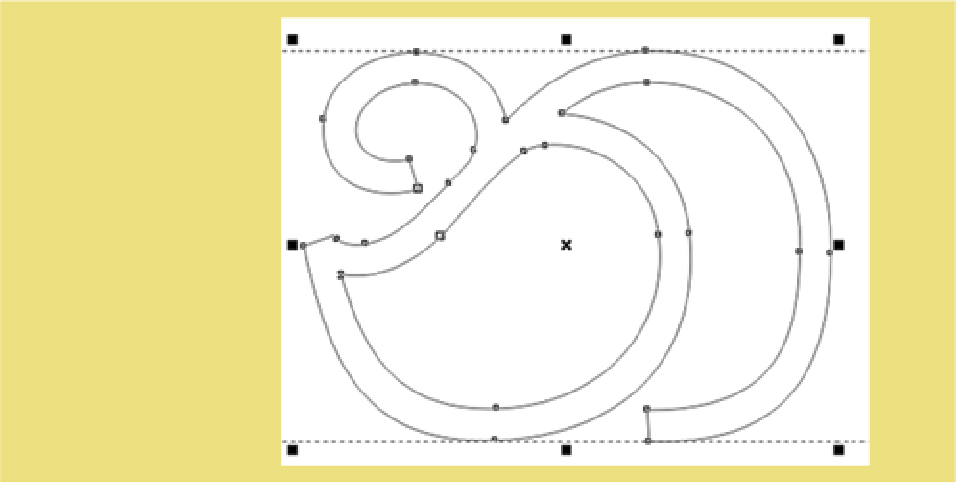
23, 24 සහ 25 යන රූප සටහන් ද එහිම තවත් අවස්ථා තුනකි.



(23 වන රූප සටහන)

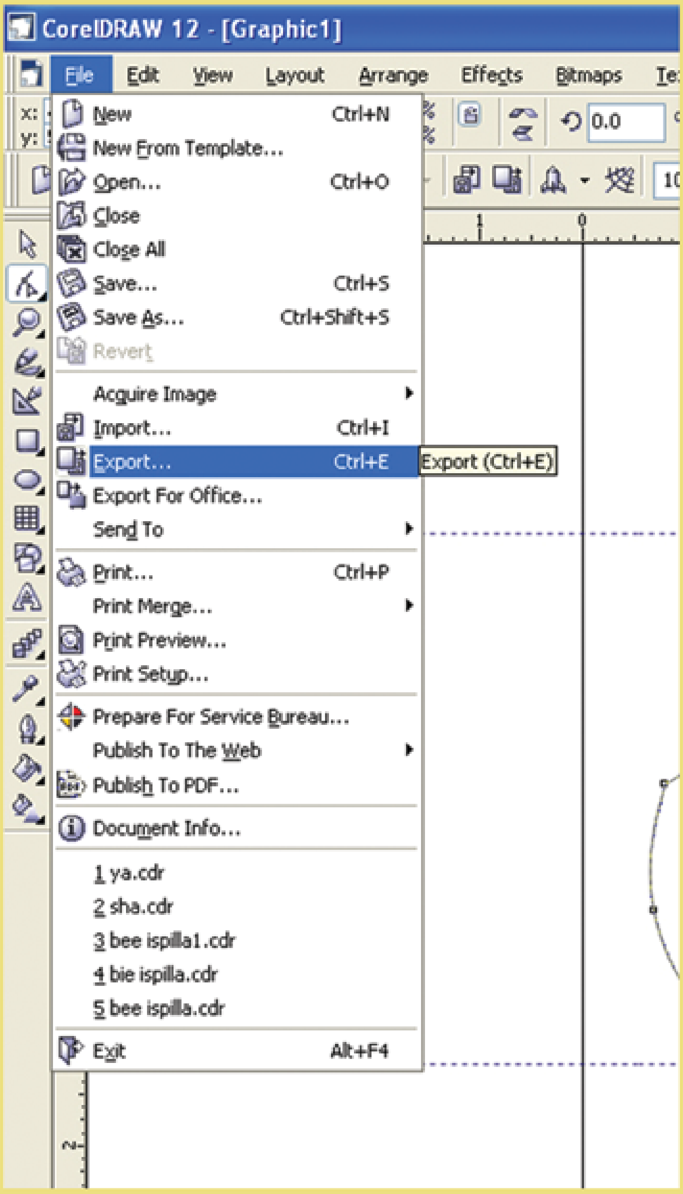


(24 වන රූප සටහන)



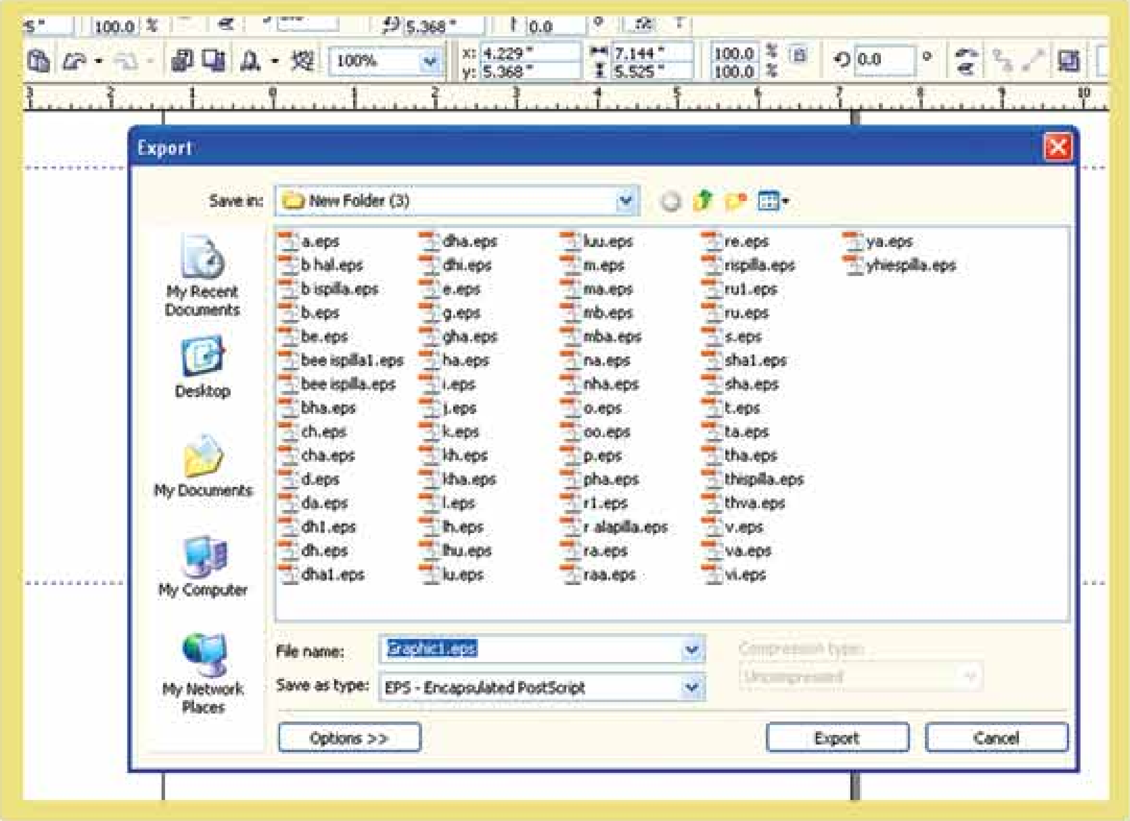
(25 වන රූප සටහන)

දැන් ඔබ නිර්මාණය කරන ලද අක්ෂරය සුරැකිය කළ යුතුය. එය මෙසේ කළ යුතුය. File>export. 26 වන රූප සටහන බලන්න.



(26 වන රූප සටහන)

Export විධානය දුන් විට 27 වන රූප සටහනේ දැක්වෙන සුරැකෙන ස්ථානය විවෘතවේ. එය සුරැකිය යුත්තේ EPS - (Encapsulted Post.Script) ගොනුවක් ලෙසය. දැන් export කරන්න.

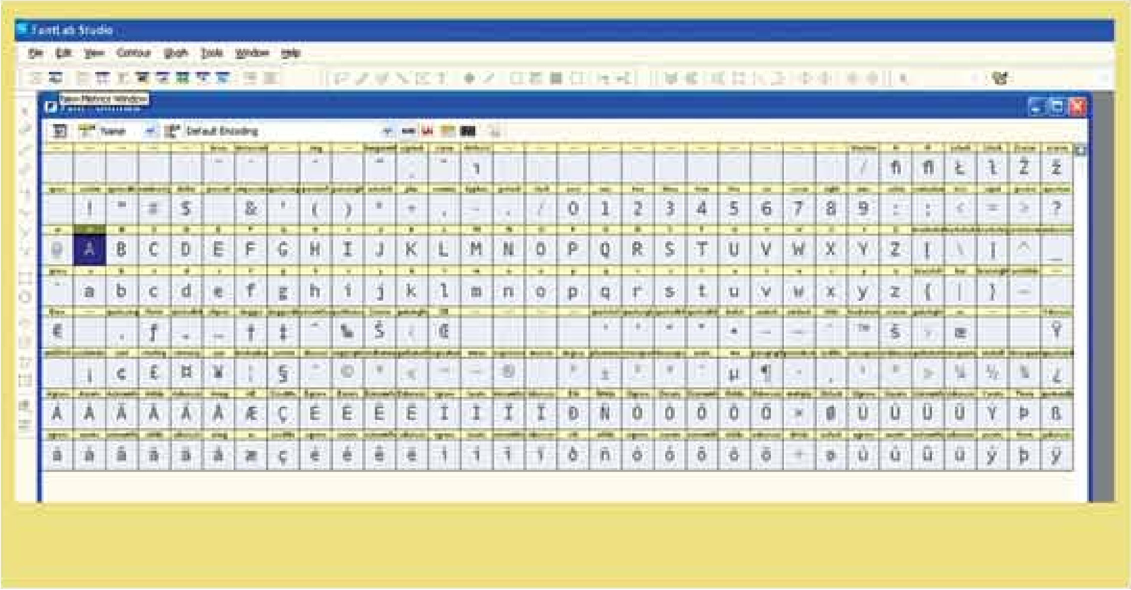


File name: Graphic1.eps

Save as type: EPS - Encapsulated PostScript

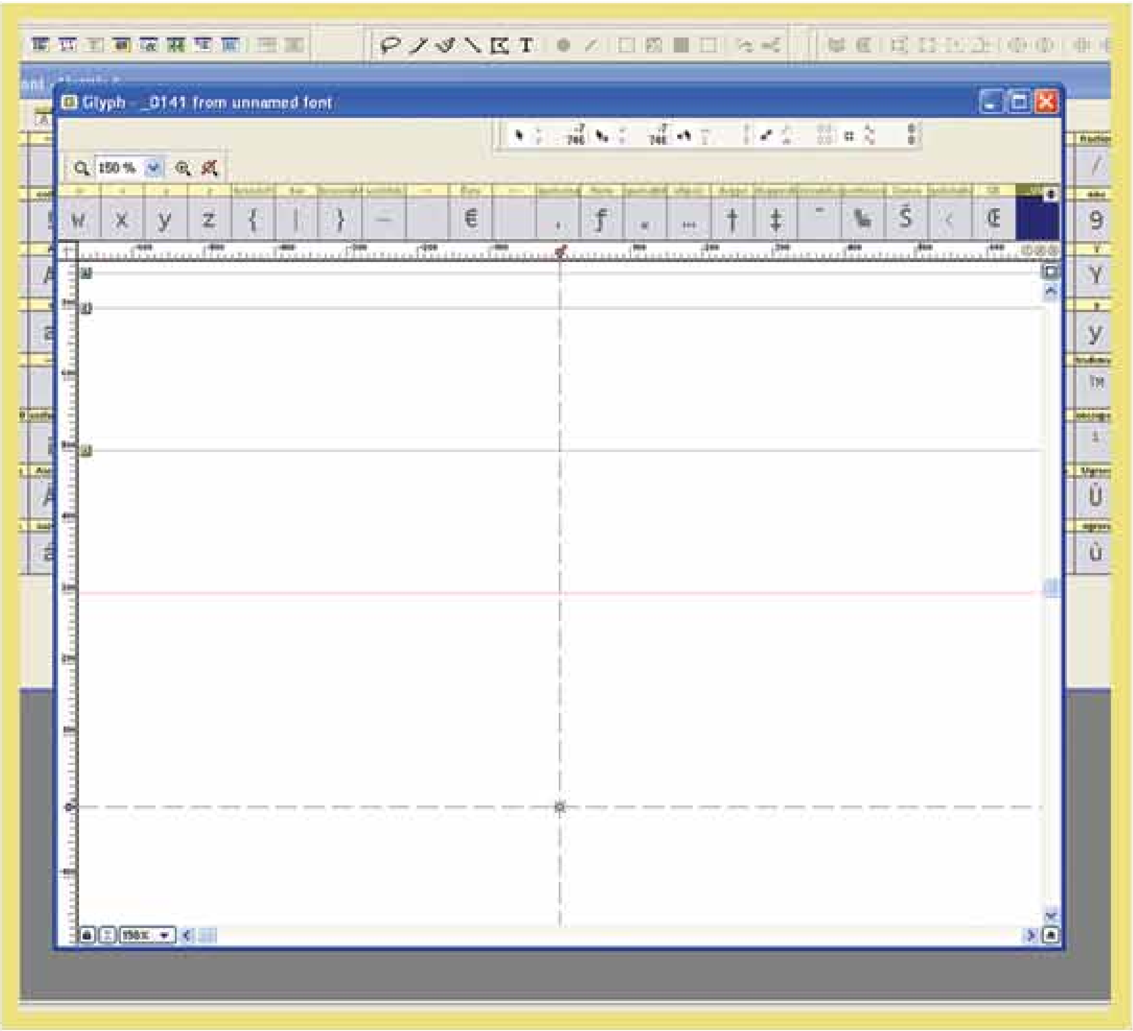
(27 වන රූප සටහන)

CorelDraw මෘදුකාංගය මගින් කරනු ලැබූ වැඩ කොටස එතෙකින් නිමවෙයි. දැන් ඔබ පිවිසිය යුත්තේ Font lab මෘදුකාංගයටය. Font lab මෘදුකාංගය විවෘත කොට එහි ගොනුව විවෘතකොට glyph එකක් විවෘත කරන්න. එවිට 28 වැනි රූප සටහනේ දැක්වෙන රූප සටහනට සමාන තත්ත්වයක් ඔබේ කවුළුවේ දැකිය හැකිය.


FontLab Studio
File Edit View Contour Glyph Tools Window Help
Name
Default Encoding
! " # $ & ' ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9 : ; < = > ?
@ A B C D E F G H I J K L M N O P Q R S T U V W X Y Z [\ ] ^ _
` a b c d e f g h i j k l m n o p q r s t u v w x y z { | } ~

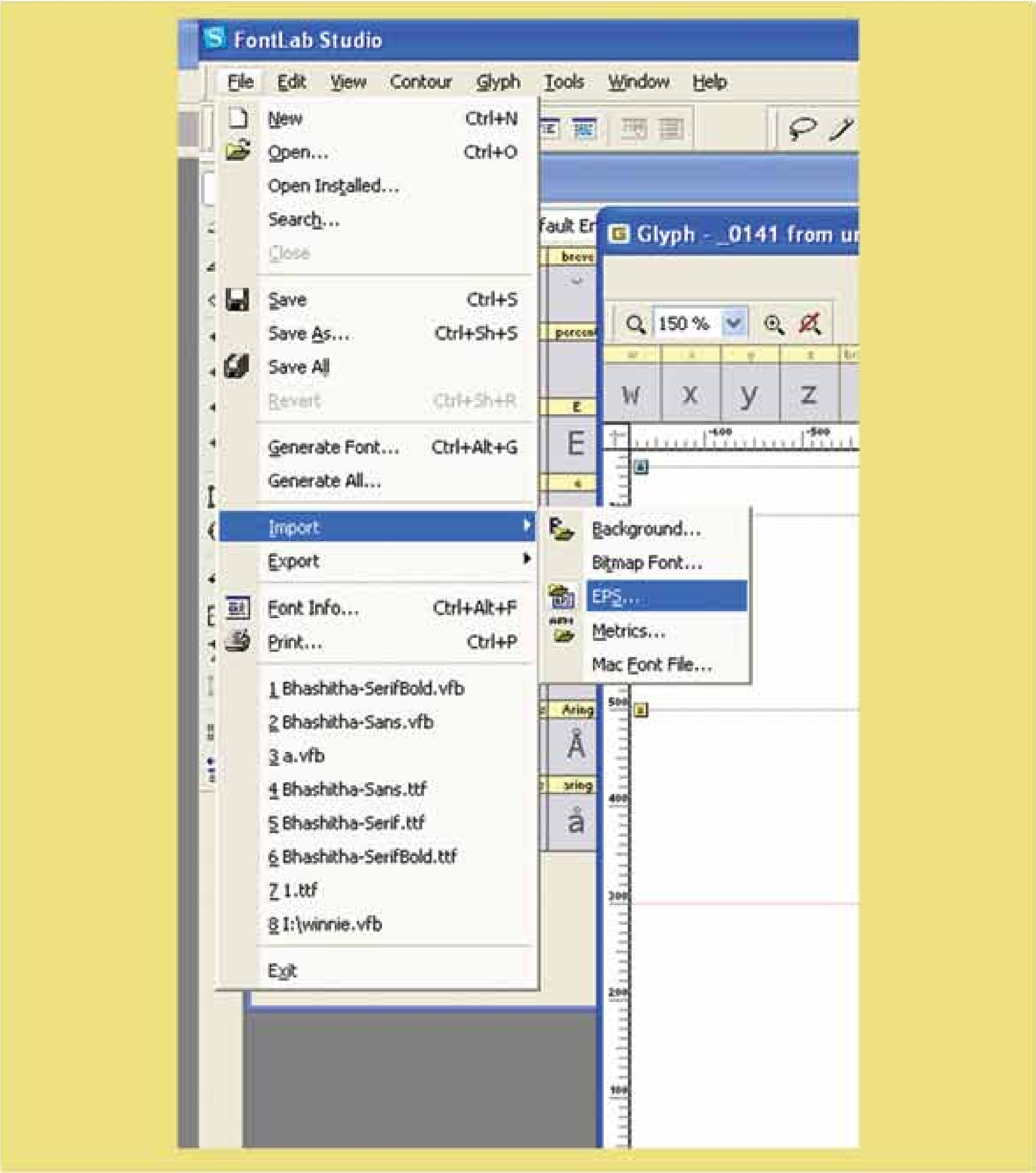
(28 වන රූප සටහන)

මෙහි එක කොටුවකට එක අක්ෂරය බැගින් ඇති අතර ඉන් එකක් මත කර්සරය තබා ද්වි ක්ලික් කළ විට 29 වැනි පිටුවේ දැක්වෙන ආකාරය පරිගණක තිරයේ දැකිය හැකිය.


Glyph - _0141 from unnamed font
150 %
w x y z { | } — € , ƒ „ … † ‡ ˆ ‰ Š ‹ Œ
/ 9 Y y ™ ¹ Ù ù
150%

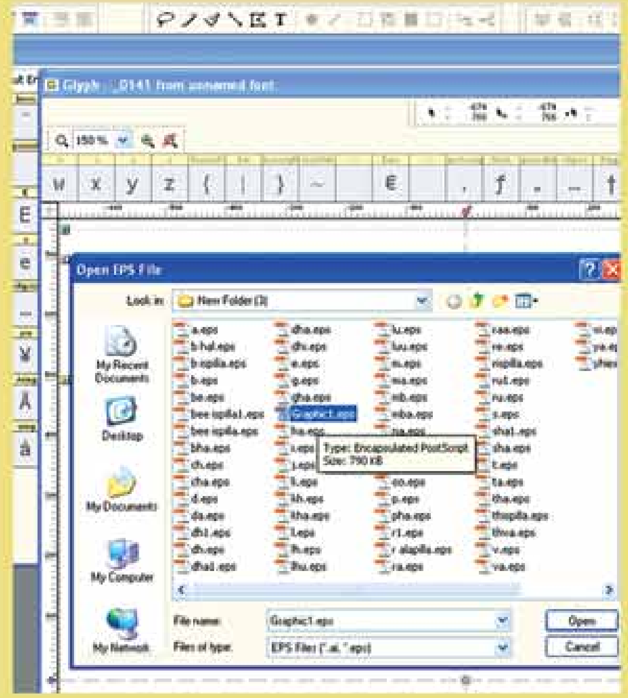
(29 වන රූප සටහන)

දැන් ඔබ කළ යුත්තේ 30 වැනි රූප සටහනේ දැක්වෙන පරිදි font lab හි file විවෘත කොට මෙනුව දිගේ පහළට විත් import හි නතර වීමය. එවිට එහි අකුරු මෙනුවේ විවෘත වේ. ඒ අකුරු මෙනු එකේ ඇති EPS යන්න ක්ලික් කරන්න.



(30 වන රූප සටහන)

එවිට 31 රූප සටහනේ දැක්වෙන ආකාරයට ඔබ විසින් නිර්මාණය කොට සුරකින ලද අක්ෂරය ඇති ගොනුව එක විවෘත වනු ඇත.

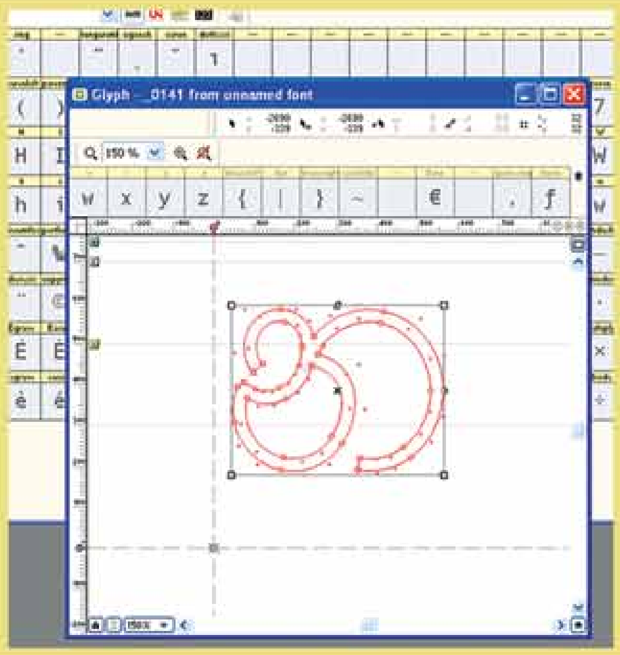


File name: Graphic1.eps

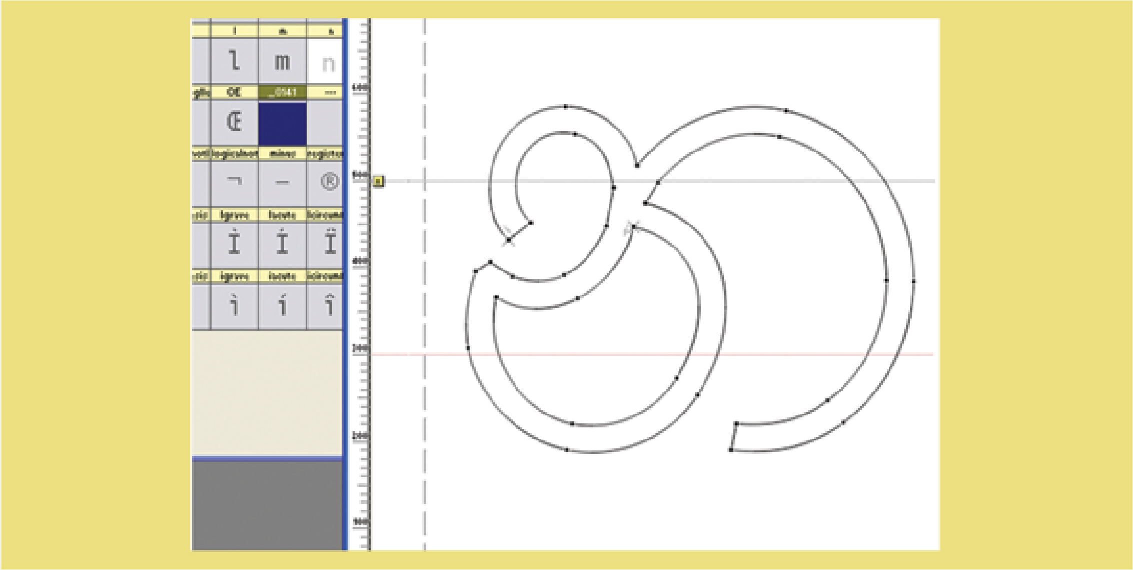
Files of type: EPS Files (*.ai, *.eps)

(31 වන රූප සටහන)

ඒ තුළ ඇති ඔබ නිර්මාණය කළ අක්ෂරය විවෘත කරන්න. එවිට 32 වැනි රූප සටහනේ දැක්වෙන පරිදි අක්ෂරය glyph එක තුළ විවෘතවේ. අවශ්‍ය නම් මෙතැනදීත් ඔබට අක්ෂරයේ වැරදි නිවැරදි කළ හැකිය.

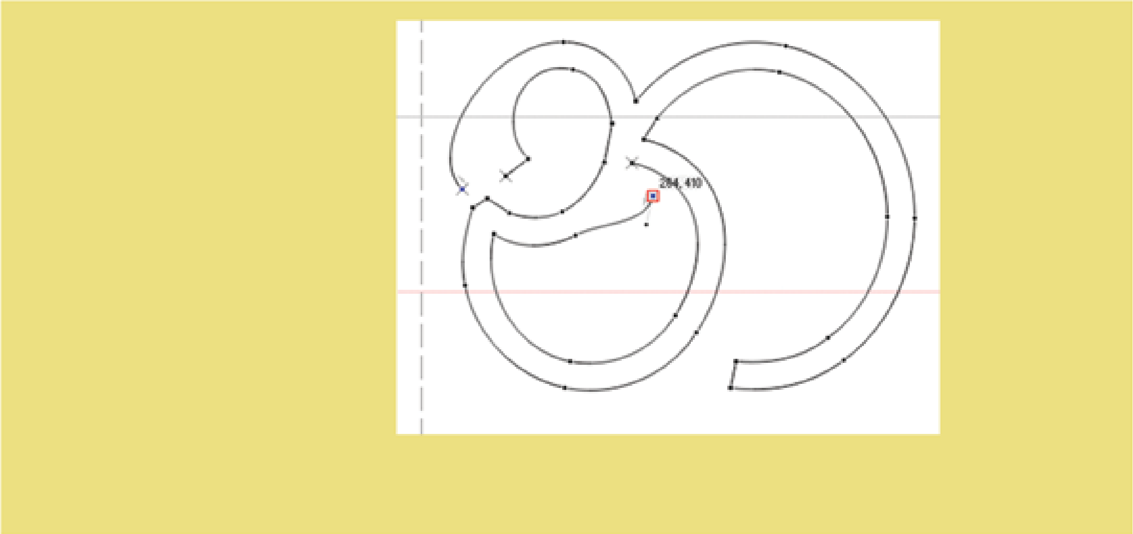


(32 වන රූප සටහන)



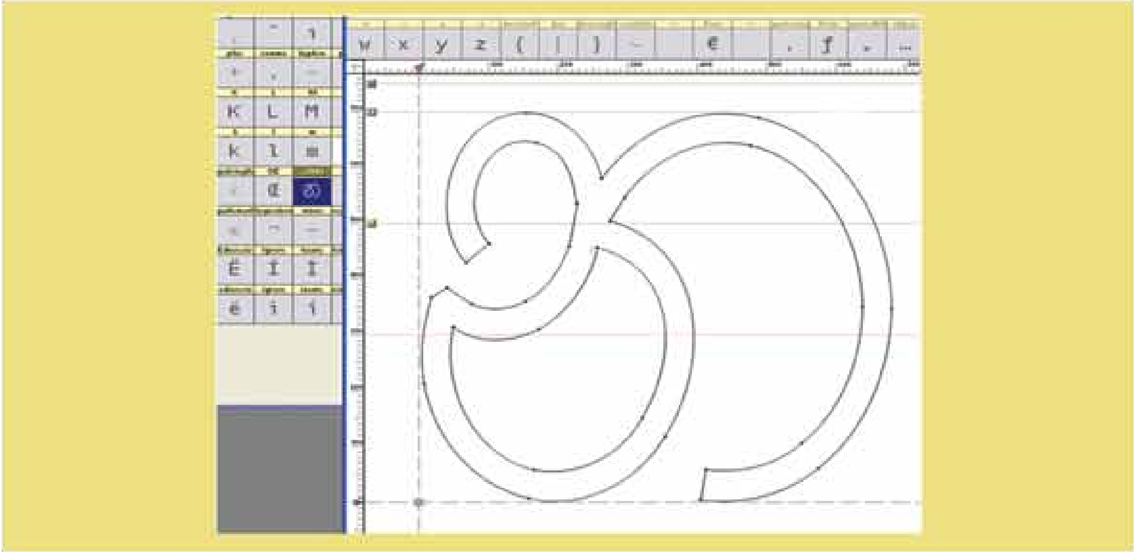
(33 වන රූප සටහන)

එහෙත් මෙහි විශාල රූපයේ පෙනෙන අක්ෂරය glyph හි අදාළ කොටුවේ නොපෙනෙනු ඇත. එයට හේතුව අක්ෂරය නිර්මාණය වී ඇති රේඛාව කොතැනකින් හෝ විසන්ධි වී තිබීමයි. එලෙස විසන්ධි වී ඇති තැන් රේඛාව මත ඇති කුඩා කතිර ලෙසින් දිස්වේ. ඒ තැන්වලට Arrow එක තබා ඇද්දොත් 34 වන රූප සටහනේ මෙන් රේඛාව විසන්ධි වේ.



(34 වන රූප සටහන)

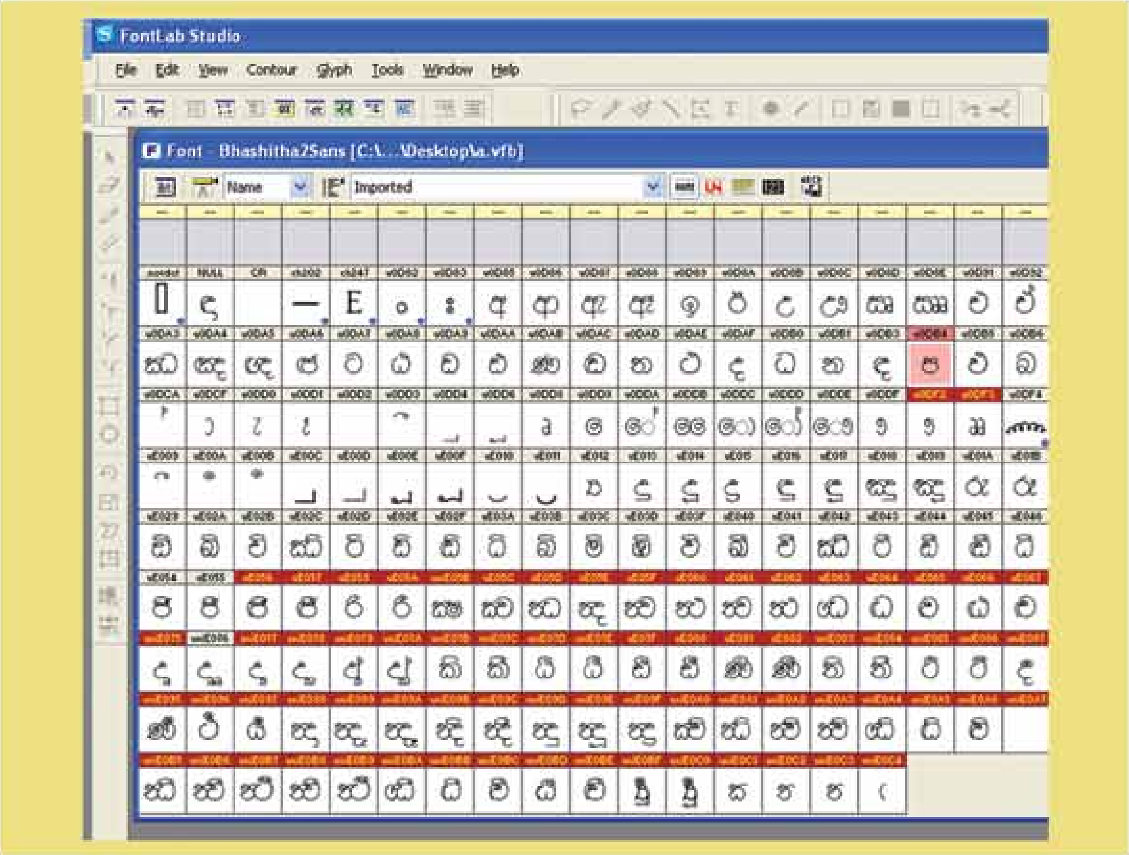
මෙවැනි තැන් සොයා ඒ මත Arrow එක තබා මදක් එළියට ඇද යළි තිබුණු තැනටම ගෙන යාමෙන් විසන්ධිවී තිබූ රේඛාව ස්ථිර ලෙසට සන්ධි වේ. ස්ථිර ලෙස සන්ධි වූ පසු glyph එක තුළ අක්ෂරය දැකිය හැකිය. (35 රූපය).



(35 වන රූප සටහන)

තවද අපට නිර්මාණය කරන ලද අක්ෂරය උස හෝ පළල අතට ඇදීමෙන් මුලින් නිර්මාණය කළ අක්ෂරය මදක් වෙනස් කළ හැකිය.

ඒ ආකාරයට මුළු හෝඩියම නව රටාවකින් නිර්මාණය කර glyph වලට යොදා ගැනීමෙන් අලුත් ෆොන්ටයක් සෑදිය හැකිය. (36 රූපය)



(36 වන රූප සටහන)