

---

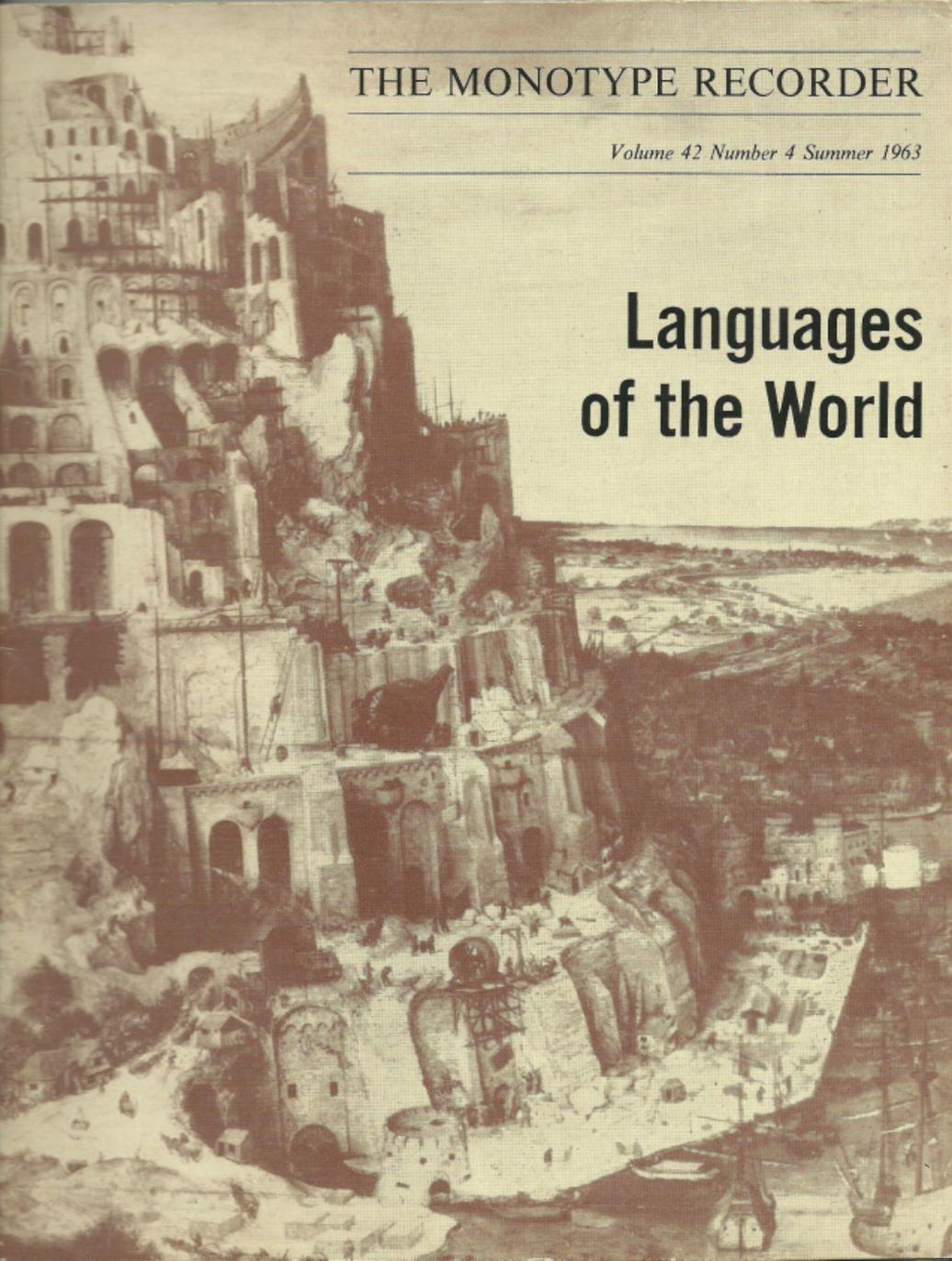
THE MONOTYPE RECORDER

---

*Volume 42 Number 4 Summer 1963*

---

# Languages of the World





**LANGUAGES OF THE WORLD**

This number of *The Monotype Recorder* is set in 'Monotype' Times New Roman Series 327 (which typeface comes of age in October 1963) and was printed in England by Stephen Austin & Sons Limited of Hertford.

The publication of this number of *The Monotype Recorder* coincides with the 11th International Printing Machinery and Allied Trades Exhibition (IPEX) held at Olympia and Earls Court, London, in July 1963.

A copy of *The Monotype Recorder* is sent gratis to every printing office with 'Monotype' machines. Copies of this number are on sale at five shillings each.

May we remind our friends and the trade generally that the words 'Monotype' and 'Monophoto' are our Registered Trade Marks and indicate that the goods to which they are applied are of our manufacture or merchandise.

**THE MONOTYPE RECORDER**

*Volume 42 Number 4 Summer 1963*

# **LANGUAGES OF THE WORLD**

**that can be set on 'Monotype' machines**

*compiled by R. A. Downie*

**THE MONOTYPE CORPORATION LIMITED**

## **Contents**

### **INTRODUCTION**

### **PART ONE : SCRIPTS**

Amharic:	page 12
Arabic:	page 9
Armenian:	page 12
Bengali:	page 22
Burmese:	page 26
Cyrillic:	page 14
Devanagari:	page 21
Georgian:	page 16
Glagolitic:	page 16
Greek:	page 13
Gujerathi:	page 22
Gurmukhi:	page 23
Hebrew:	page 7
Kannada:	page 24
Latin:	page 17
Malayalam:	page 25
Sinhalese:	page 23
Syriac Estrangelo:	page 8
Tamil:	page 25
Telugu:	page 24
Thai:	page 27

### **PART TWO : LANGUAGES**

Arranged alphabetically: pages 29-55

## Introduction

The whole earth today is not "of one language and of one speech". In spite of the fact that we have reproduced Brueghel's well-known painting of the Tower of Babel on the front cover of this issue of *The Monotype Recorder*, it is not our intention to emulate the generations of the sons of Shem, Ham and Japheth. We do not entirely despair of man's achieving a supra-national language sometime in the future but, although we have included one or two artificial languages in this survey, we are interested in them only in so far as they are present-day realities.

The structure we are building does not aspire to the heights; it sticks closely to the ground, concerning itself with what is, not what will be. It is constantly being repaired and renovated, and we cannot be sure that it is fully up to date. Language is so very much a living thing that it is almost impossible to call a halt and say that these are the definitive facts in A.D. 1963. Some written languages have roots stretching deep down into the soil, with branches spreading out in all directions. Some are of recent growth, with insecure roots that can easily be dug up and transplanted. Others are no longer spoken but are still employed for literary and sacred purposes, about which scholars may disagree.

Of one thing we can be certain. With the rapid growth of literacy throughout the world and the remarkable progress being made in education, there will be an ever increasing demand for vernacular literature of all kinds, particularly in those newly independent states where language is a symbol of nationhood and its typographical appearance a matter of pride. It is therefore important that all those concerned with the printing of foreign literature should have some knowledge of the basic principles of language, even if they have no understanding of the languages themselves.

The mathematician and philosopher, Leibniz, who is believed by some to have anticipated Newton in the invention of differential calculus, was undoubtedly responsible for an early anticipation of the present work. He collected specimens of languages with the idea of devising an international language, and to this end corresponded with Jesuits in the Far East. After his death the Empress Catherine II of Russia commissioned the German traveller, P. S. Pallas, to compile a collection of languages. His *Linguarum totius orbis vocabularia comparativa*, when it appeared in 1786, contained 285 words in two hundred languages, 149 Asian and 51 European. A later edition included words from a further eighty African and American Indian languages. The Spaniard Hervas, in his *Catálogo de las Lenguas* (1800-05), and the German

J. C. Adelung in his *Mithridates* (1806-17), threw their nets much wider, and the four volumes of Adelung's work contain the Lord's Prayer in close on five hundred languages and dialects. These works were inspired by philosophical curiosity and the contemporary fashion for collecting specimens, whether fossils or statues or languages.

The object here attempted is strictly practical. It is to give people some basic information about the scripts and languages of the world; to show how they have developed and where they are spoken; to demonstrate their alphabets and to explain how they are used; to provide some notion of the equivalent value of the various symbols by indicating how they can be transliterated; and to illustrate the various alphabets in action with specimens set in 'Monotype' or 'Monophoto' faces. The work is divided into two parts - the one covering written scripts, the other dealing with the many languages within the script groups. The former has been grouped roughly according to families, and the latter has been arranged alphabetically.

This is a field in which complete and exhaustive treatment is not to be contemplated. There is no agreement as to how many languages there are in the world. In the continent of Africa alone, no one can say to the nearest hundred how many languages and dialects are spoken. The general policy here followed, particularly with regard to Africa, has been to give examples of written languages that are used in printed form as means of communication and instruction rather than those that have merely been described by outsiders, and to illustrate the languages of the different regions, the various speech-families, and the diverse systems of reproducing the speech sounds in print.

This is a field of knowledge in which it is temptingly easy to be dogmatic, and just as easy to be wrong. One might, for example, be inclined to follow the grammarians and say that the letters *j* and *x* do not exist in Welsh, only to find that in the latest translation of the Bible both are admitted. If such traps await us on our doorstep, as it were, what chance is there of avoiding all the pitfalls in dealing with the languages of more remote parts? In attempting a task that is virtually impossible, one is perhaps entitled to expect a measure of forbearance from critical readers.

Among general books that have been consulted are Ballhorn's *Alphabete orientalischer und okzidentalischer Sprachen*, first published about a century ago, von Ostermann's *Manual of Foreign Languages* (Washington, 1935), Gleichen and Reynold's *Alphabets of Foreign Languages*

(R. G. S., 1933), Diringer's *The Alphabet* (London, 1953), and Gelyarevsky and Grevene's *Определитель Языков Мира по Письменностям* (Moscow, 1961).

Valuable assistance has been obtained from *Notices sur les Caractères Étrangers anciens et modernes*, edited by Charles Fossey and published by the Imprimerie Nationale de France in 1948. Even more comprehensive is Manuel Aguirre's *La Escritura en el Mundo* (Madrid, 1961), which arrived too late to be used in the preparation of this work but which has been extensively and profitably consulted in checking details.

In dealing with the Cyrillic alphabet and Slav tongues, considerable use has been made of R. G. A. de Bray's *Guide to the Slavonic languages* (London, 1951). The many massive volumes of Grierson's *Linguistic Survey of India* have been frequently consulted, never in vain.

It would be a long and fruitless task to enumerate the grammars, handbooks, manuals, dictionaries and anthologies of the individual languages that have been used for reference. Various libraries have given willing help, in particular the London Library, the India Office Library, the special language library of the City of Westminster in South Audley Street, and London University Library.

The British and Foreign Bible Society and the Scripture Gift Mission have courteously answered inquiries; Longmans Green & Co. Ltd. have permitted inspection of the many books they have published in African and Asian languages; and the printing firms of Stephen Austin & Sons Ltd. of Hertford, A. Brown & Sons Ltd. of Hull, and the Morija Printing Works of Basutoland, were kind

enough to supply specimens of their printing in African languages.

Brueghel's painting of the Tower of Babel was reproduced by courtesy of the Kunsthistorisches Museum, Vienna. The pages of Walton's Polyglot Bible (1657) and the verses from the Great King James Bible (1611) were reproduced by permission of the St. Bride Printing Library.

We owe a great debt to the many Continental and Overseas managers of The Monotype Corporation, particularly the Eastern Area manager, Mr. L. A. Collier, who assisted in the correction of the proofs. Professor A. N. Tucker of the School of Oriental and African Studies gave generously of his time when consulted on the use of Africa script. Mr. J. S. Beckett of the Admiralty was most helpful on Semitic languages and Sanskrit, in spite of his firm belief that true scholarship should not be of the slightest practical use to anyone. Mr. Colm O Lochlainn provided us with the material on Irish. In addition to these people, we must not fail to mention several foreign visitors to Monotype House, who were practically press-ganged into the position of having to check proofs on the spot. We also acknowledge our indebtedness to Mr. Tibor Szántó for originally suggesting the basic idea of this project.

Last but not least, we could not have contemplated the production of this work without the knowledgeable and efficient service provided by the printers, Stephen Austin & Sons Ltd. of Hertford, whose co-operation made our task so very much easier.

R. A. D.  
A. D. B. J.

# PART ONE: SCRIPTS

## Hebrew

The modern Hebrew alphabet is the direct descendant of an alphabet invented more than 3,500 years ago, possibly in Palestine itself, which according to some scholars is the ancestor of all alphabets. The order of the letters has remained the same from earliest times, and the 22 characters have always indicated consonants, although four of them are now used to indicate long vowels.

There were various ways of writing Hebrew in early times, but just before the Christian era a distinctive Palestinian Jewish type of script became the standardized form, and from this the modern Hebrew alphabet is derived. It was called, from its shape, Square Hebrew Script or *ketabh meruba*. Detailed rules laid down by the Talmud as to calligraphy and consonantal orthography made deviation from the standard impossible, and the modern printed versions of the Old Testament closely resemble in script the Torah scrolls of two thousand years ago.

Hebrew is always written from right to left. (So that it is immediately ready for proofing, the lines of type are stacked on the left-hand side of the Composition Caster by means of a Reverse Delivery Attachment.)

Hebrew has no capital letters or italics. There are five letters which have two forms, one when initial or medial and another when final. Some letters can be printed with or without an internal point, which affects their pronunciation. (See alphabet on right.)

The following notes on distinguishing certain characters which are similar may be useful:

כ (*kaph*) is round; ב (*beth*) has a slight projection at the lower right-hand corner;

נ (*nun*) is straight at the foot; ג (*gimel*) is broken at the 'foot';

ד (*daleth*) overlaps at the top; ר (*resh*) curves; ך (final *kaph*) is like ד (*daleth*) but projects below the line;

ה (*he*) is open at the top left; ח (*cheth*) is closed; ט (*tau*) has a foot at the left;

ׁ (*yod*) does not touch the line; ׃ (*vau*) does touch the line;

ז (*zayin*) has a cross stroke at the top; ן (final *nun*) projects below the line;

ס (*samech*) is round; ם (final *mem*) is square;

ת (*teth*) is open at the top; ׀ (*mem*) is open at the foot;

ׂ (*ayin*) has its tail turned towards the left; ׃ (*tzadde*) curves first towards the right; ׅ (final *tzadde*) drops its tail straight down.

The earlier texts were without vowels or punctuation down to about 500 A.D. A system of punctuation and vowels was added during the succeeding centuries by scholars called *Massoretes* (who derived their name from the Hebrew word for "tradition" or "writing") and was completed by the 10th century. The system of vocalization is called the *Tiberian* or *Tiberiadic* and consists of dots and little dashes both above and below the line.

The Massoretic vowel signs are as follows:

- = a as in בַּדְּחִי (*badhi*)      . = i as in מִינִי (*mini*)      ׁ = u as in גִּמְלוֹ (*gimlo*)  
 ׃ = e as in בְּנֵי (*beni*)      . = a or o as in הוֹדְשָׁם (*hodsham*)

There are also two vertical dots below a letter, which may either be used on their own, as in ׃, or combined with other vowel signs, thus ׃.

Usual form	Final form	Name	Transliteration	Numerical value
א		'aleph	'	1
ב	בּ	beth	bh(v), b	2
ג	גּ	gimel	gh, g	3
ד	דּ	daleth	dh, d	4
ה		he	h	5
ו	וּ	vau	v	6
ז		zayin	z	7
ח		cheth	h̄	8
ט		teth	t	9
י		yod	y	10
כ	כּ	kaph	kh, k	20
ל		lamedh	l	30
מ	מּ	mem	m	40
נ	נּ	nun	n	50
ס		samech	ṣ	60
ע		'ayin	'	70
פ	פּ	pe	ph, p	80
צ	צּ	tzadde	ṣ	90
ק		koph	q	100
ר		resh	r	200
ש	שׁ	sin, shin	s, sh	300
ת	תּ	tau	th, t	400

### 'Monotype' Hebrew Faces

Ashurith (219) *unpointed*  
 Hebrew (222) *unpointed*  
 Levenim (221) *unpointed*  
 Peninim (220) *unpointed*  
 Peninim Pointed (217)  
 Rabbinic (228) *unpointed*  
 Sonzino (218) *pointed*

A final *kaph* takes these two dots inside rather than beneath  $\aleph$ . Another vowel sign is a dot to the top left of a letter, as in *ph* (*hoq*).

Other signs were introduced to give guidance on pronunciation and to act as punctuation. The end of a verse was marked by two points like a colon, and the last word of each verse might have a small vertical line below it. Sometimes two words are joined by a small horizontal stroke at the top of the line. These signs are illustrated in the specimen from the *Book of Genesis* (see *Hebrew in Part Two*).

In modern Hebrew the same characters and vowel signs are used, but the full-stop and other punctuation marks have been introduced.

**NUMERALS** Hebrew letters were also used to indicate numbers. This was not a Biblical usage and may have been introduced in imitation of Greek. The letters from *aleph* to *teth* stand for the numbers 1 to 9, *yod* to *tzadde* for the tens, and *koph* to *tau* for the hundreds up to 400. 500 to 900 were generally expressed by the combination of *tau* with the other signs for hundreds, or occasionally by the final letters *kaph* (500), *mem* (600), *nun* (700), *pe* (800), *tzadde* (900). Two dots over a letter indicated a thousand. Compounds of numbers were expressed by a combination of letters, those denoting the higher numbers being placed towards the right. In modern Hebrew the Arabic numerals are commonly used.

Hebrew has been used to write other Semitic languages, including Aramaic (see *Part Two*), and also Yiddish (see *Part Two*).

## Syriac Estrangelo

The Syrian or Syriac alphabet, like the language, is Semitic in origin and shares many features with Hebrew and Arabic. All are written from right to left, necessitating the use of a reverse delivery attachment (see *Hebrew: Part One*), all lack capitals, and all consisted originally of consonants only. In Syriac three consonant characters  $\aleph$ ,  $\beth$ , and  $\aleph$ , later came to be used as well to indicate vowels. The order of the letters in Syriac is the same as in the Hebrew alphabet and many of the letter forms closely resemble Hebrew.

As in the much later Arabic alphabet which, if it was not entirely derived from the Syriac was at least influenced by it, there are different forms of the letter for different places within a word. In Syriac these differences are very slight.

The full point is used to indicate the end of a sentence. The signs  $\cdot$  and  $\cdot$  are used to mark breaks in a sentence and  $\diamond$  indicates the end of a paragraph. Two dots over a letter and one below indicate that the word in which this letter occurs is in the plural, e.g.  $\aleph$ . As in Hebrew, letters are also used as numerals.

The most important variety of Syriac is Estrangelo, which derives its name from the Arabic words meaning "the writing of the Scriptures". It was employed exclusively for the best part of five hundred years. One 'Monotype' face is available, Syriac Estrangelo (564).

independent form	final form	initial form	medial form	transliteration	numerical value
$\aleph$	$\aleph$	—	—	'	1
$\beth$	$\beth$	—	—	b	2
$\gamma$	$\gamma$	—	—	g	3
$\delta$	$\delta$	—	—	d	4
$\epsilon$	$\epsilon$	—	—	h	5
$\zeta$	$\zeta$	—	—	w	6
$\eta$	$\eta$	—	—	z	7
$\theta$	$\theta$	—	—	h	8
$\iota$	$\iota$	—	—	t	9
$\kappa$	$\kappa$	—	—	i	10
$\lambda$	$\lambda$	—	—	k	20
$\mu$	$\mu$	—	—	l	30
$\nu$	$\nu$	—	—	m	40
$\xi$	$\xi$	—	—	n	50
$\omicron$	$\omicron$	—	—	s	60
$\pi$	$\pi$	—	—	'	70
$\rho$	$\rho$	—	—	p	80
$\sigma$	$\sigma$	—	—	q	90
$\tau$	$\tau$	—	—	r	100
$\upsilon$	$\upsilon$	—	—	sh	200
$\phi$	$\phi$	—	—	t	300
$\chi$	$\chi$	—	—		400

## Arabic

The sudden emergence of the Arabs upon the stage of world affairs is one of the most surprising events in history. Under the impulse of Mohammed and his religion the scattered tribes of the Arabian Peninsula became first a unified nation and then the conquerors of continents. Using the desert as the English were later to use the sea, with cavalry and camels in place of ships, they extended their conquests eastwards through the mighty empires of Persia and Byzantium, westwards through the whole of North Africa and into Spain, southwards to Nubia and the Sudan and northwards to the Caucasus. At the height of their power the Islam Empire extended from the Atlantic to the boundary of India over a territory that was twice as large as the Roman Empire at its largest. Their trading influence went even wider, for they dealt with India, Ceylon, the East Indies and China. Their caravan routes criss-crossed Africa, and their trade with the far north has been attested by the discovery of thousands of their coins in Scandinavia and along the Volga, and the presence of several Arabic words in old Icelandic literature.

Wherever they went they took their language, their religion and their script, and in fact all three are still widely spread throughout the region of their original conquests. There are today some seventy million people who speak Arabic, and some 350 million Mohammedans, whose Bible is the Arabic Koran.

The Koran was composed by Mohammed in a language which, previously unrecorded, suddenly displays itself in all its perfection, with its flexibility and its infinite richness so completely developed that from that time to this it has undergone no important change. The Koran was committed to the prodigious memory of Mohammed's followers. Part of it was written down during his lifetime and the rest within two years of his death in 632 A.D. It represents the earliest recorded use of Arabic for literary purposes.

The earliest manuscripts of the Koran were written in Kufic script, which takes its name from the town of Kufa, the seat of a famous Moslem school. Kufic was a monumental script, and its straight lines make it suitable for carving on stone or on metal. It is often to be seen in inscriptions on the walls of mosques, for example, and on coins.

A variety of Kufic known as Maghribi (or "Western") was developed by the Moslem peoples of West and North Africa, including Tangiers and Morocco, in the 10th century.

Kufic had several defects. The same symbols were used for two or more different letters and, as in other Semitic alphabets, all vowels were omitted. During the early centuries of Islam, however, the urgent need to preserve the sacred text of the Koran from corruption and error forced their scribes to devise conventions to overcome these defects. To distinguish letters represented by the same symbol one or more dots were placed above or below it. Vowels were also indicated by signs placed above or below the letter affected. These vowel signs are still used in the Koran and in books of instruction, but in ordinary writing the vowels are omitted, since the practised Arabic reader can recognize a word from its consonantal framework.

When the language came to be written on papyrus, or on the paper which the Arabs first introduced into Europe from China, the straight lines of Kufic gave way to the rounded cursive script known as Naskh, which is the parent of the modern Arabic script.

A variety of Naskh known as Solloss is composed of the large letters

## ARABIC

of Naskh as they were used for the titles of books and the headings of chapters. It is widely used for ornamental and ceremonial purposes.

Arabic script spread even more extensively than the language and it is, after Latin, the most widely used in the world today. It was adopted in Persia and throughout the Ottoman Empire and it is still used, with a few modifications, to write modern Persian today. It was the official script in modern Turkey until the Latin alphabet was officially adopted by Kemal Atatürk. It was also used for many forms of Turkic languages outside Turkey itself, for example those spoken in Azerbaijan in the U.S.S.R., in Turkestan, Tibet, Daghestan, the Crimea and Mongolia. Most if not all of these have now given up Arabic and use Cyrillic script. Eastwards, in India and Pakistan, Arabic characters are used to write Urdu, Sindhi, and Pashto. It has also been employed for other Indian languages, some of which have their own script as well, such as Punjabi and Kashmiri.

Further east, Arabic script was used by the scattered inhabitants of the East Indies, but here it has lost ground to the Latin alphabet, for the many millions of Indonesia are committed to the use of the Latin script and those of Malaya are in a state of transition from one to the other.

The same may be said of Africa, for while the Arabic language and script are still prevalent in North Africa, further south languages like Swahili, Hausa and Fulani which were once written in Arabic script now use Latin.

Arabic is written from right to left (except the numerals), and the letters of each word are joined together. There are no capitals. The forms of the letters differ according to their position in a word, and these are shown in the table on the right. Usually there are four different forms: (a) when standing alone; (b) when final (joined to the preceding letter); (c) when medial (connected to both preceding and following letters); (d) when initial (joined to succeeding letter). Six letters can be joined only to the preceding, not to the following letters, and in these cases the initial and medial forms are the same as the independent and final forms respectively.

The letter *alif* has no sound of its own but is commonly used to support other signs, thus:  $\text{ا}$   $\text{آ}$ . Combined with *lam* it forms the article *al* or *el*, and is written thus:  $\text{أ}$  or  $\text{إ}$ . This is sometimes regarded as the 29th letter of the alphabet.

If short vowels are to be shown, they are indicated by the addition of signs above or below the consonants which they follow.  $\text{َ}$  signifies *a*, thus  $\text{دَا}$ ;  $\text{ِ}$  signifies *i*, thus  $\text{دِي}$ ;  $\text{ُ}$  signifies *u* or *o*, thus  $\text{دُو}$ . Other signs are:  $\text{ْ}$  indicating the absence of a vowel;  $\text{ّ}$  indicating that a consonant is to be doubled;  $\text{ّّ}$  indicating that a consonant is repeated without a vowel between.

## Persian

The same letters and vowel signs are used as for Arabic, with an additional four characters:  $\text{پ}$ ,  $\text{چ}$ ,  $\text{ژ}$ ,  $\text{گ}$ .

## North-West Africa

The great expanse of territory that extends from Egypt to the Atlantic, comprising the whole of Libya, Tunisia, Algeria and Morocco, is known in Arabic as Maghrib or "west". This is the home of some 26 million people of whom three-quarters speak some form of Arabic, the other quarter being Berber-speaking.

In this area the writing and printing of Arabic developed certain peculiar characteristics. Many of these have disappeared, but there are still four letters in common use which are lacking in standard Arabic.

Name	Independent form	Final form	Medial form	Initial form	Transliteration					
Alif	ا	ا			a					
Be	ب	با	ب	ب	b					
Te	ت	تا	ت	ت	t					
The	ث	ثا	ث	ث	th					
Jim	ج	جا	ج	ج	j					
He	ح	حا	ح	ح	h					
Khe	خ	خا	خ	خ	kh					
Dal	د	دا	د	د	d					
Dhal	ذ	ذا	ذ	ذ	dh					
Re	ر	را	ر	ر	r					
Ze	ز	زا	ز	ز	z					
Sin	س	سا	س	س	s					
Shin	ش	شا	ش	ش	sh					
Sad	ص	صا	ص	ص	s					
Zwad	ض	ضا	ض	ض	zw					
Toe	ط	طا	ط	ط	t					
Zhoe	ظ	ظا	ظ	ظ	dh, z					
'Ain	ع	عا	ع	ع	'					
Ghain	غ	غا	غ	غ	gh					
Fe	ف	فا	ف	ف	f					
Qaf	ق	قا	ق	ق	q					
Kaf	ك	كا	ك	ك	k					
Lam	ل	لا	ل	ل	l					
Mim	م	ما	م	م	m					
Nun	ن	نا	ن	ن	n					
Wau	و	وا	و	و	w					
Ha	ه	ها	ه	ه	h					
Ye	ي	يا	ي	ي	y					
Numerals										
	١	٢	٣	٤	٥	٦	٧	٨	٩	٠
	1	2	3	4	5	6	7	8	9	0



## Amharic

Amharic script is derived from the early script of South Arabia, and an early form of it has been found that dates back to the 4th century A.D. The practice of writing it from left to right (unlike Arabic and Hebrew and other Semitic scripts) may be due to Greek influence. It has no capital letters, and the order of letters in the alphabet is unlike any other script. (For language specimen see *Part Two*.)

There are 33 basic characters, shown in the alphabet on the right, in which the vowel *a* is implied. Each of these letters has seven forms indicating different vowels thus *h ka, h k̄ū, h k̄l, h k̄ā, h k̄ē, h ke, h kō*. In addition, there are 45 characters representing certain consonants followed by the sound *w*, thus *h kwā, h kwā, h kwē, h kwe*.

The punctuation marks used are: *ˆ* to separate words, *ˆ* and *ˆ* to mark breaks in a sentence, and *ˆ* for a full point.

Two 'Monotype' faces are available: Amharic (599) and Amharic (624).

ሀ	ha	ቸ	cha	የ	ya
ለ	la	ኀ	ha	ደ	da
ሐ	ha	ነ	na	ጅ	ja
መ	ma	ኘ	ña	ጎ	ga
ሠ	sa	ሐ	'a	ጠ	ta
ረ	ra	ከ	ka	ጪ	cha
ረ	sa	ኸ	kha	አ	pa
ሸ	sha	ወ	wa	ኧ	tsa
ቀ	qa	ዐ	'a	ፀ	ta
በ	ba	ዘ	za	ፈ	fa
ተ	ta	ዠ	zha	ፐ	pa

ሸ	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ
1	2	3	4	5	6	7	8	9	10

## Armenian

Armenian is a member of the Indo-European family of languages. Its alphabet is derived from the Greek, which it follows closely in having consonants and vowels. Traditionally it was devised about 400 A.D. by St. Mesrop in collaboration with St. Sahak and a Greek from Samosata called Rupanos. It was bound up with the spread of Christianity in Armenia and the growth of an independent church there.

As elsewhere an established alphabet contributed to the development of a literary tradition. The 5th century became the Golden Age of Armenian literature, and a famous school of translators, founded by St. Sahak, produced versions of the Bible (from Syriac and Greek) and of some of the great Greek and Latin masterpieces. The earliest literature is religious, but there has been a lively modern literature since the middle of the 19th century.

The Armenian language is now spoken by some 2½ million people in the Armenian Republic of the U.S.S.R., in Turkey and in other Middle East countries. It is estimated that some seventy newspapers are printed in it in different parts of the world.

The ancient alphabet consisted of 36 letters, of which 22 corresponded to Greek characters and 14 were introduced to denote sounds unknown in Greek. In medieval times two further letters were added, namely *ֆ* to denote *f*, which occurs only in foreign words, and *օ* which takes the place of an old diphthong.

Armenian is available in one 'Monotype' face - Series 638 - and the alphabet is shown on the right. The sloping letters are used for normal text composition, generally with upright capitals although sloping capitals are becoming increasingly common. The so-called italic, which is upright, is derived from cursive and called *notrigir*.

inclined	upright	transl.	inclined	upright	transl.
Ա	ա	a	Ե	ե	h
Բ	բ	b	Զ	զ	n
Գ	գ	g	Ը	ւ	sh
Դ	դ	d	Պ	ո	v
Ե	ե	e	Ղ	չ	ch
Զ	զ	z	Պ	պ	b
Է	է	ē	Ղ	ղ	j
Ը	ը	e	Ռ	ր	r
Թ	թ	t	Ս	ս	s
Ճ	ճ	zh	Վ	վ	v
Ի	ի	i	Տ	տ	d
Լ	լ	l	Ր	ր	r
Խ	խ	kh	Յ	յ	ts
Մ	մ	dz	Ի	ի	v
Կ	կ	k	Փ	փ	p
Հ	չ	h	Ք	ք	k
Ձ	զ	ds	Օ	օ	o
Ղ	ղ	gh	Ֆ	ֆ	f
Ճ	ճ	j	Ligatures		
Մ	մ	m	և	ա	և
			և	ա	և



## Cyrillic

The Cyrillic alphabet takes its name from the Bulgarian apostle, St. Cyril, whose missionary activities in the ninth century are recounted below in the section dealing with the other Old Church Slavonic alphabet—Glagolitic. In spite of its name it is reasonably certain that it was not in fact invented by St. Cyril but was a later development. The Greek uncial alphabet of the 10th century served as the basis for most of its letters though certain of the sounds in Slavonic speech could not be rendered by Greek letters and special signs were adopted for these. The Cyrillic form of the Old Church Slavonic alphabet is shown side-by-side with the Glagolitic form in the table which accompanies the Glagolitic section following.

During the centuries that followed its invention, the Cyrillic alphabet underwent a number of changes, the most important occurring at the beginning of the 18th century with the advent of Peter the Great. This was the time of great reforms and, in keeping with his general policy of modernizing Russia, Peter did not forget his alphabet. He had it "romanized", modifying the appearance of some of the characters and dropping others altogether. Today, the original Cyrillic alphabet retains its former glory in the shape of ecclesiastic characters in liturgical works only, as used throughout the ages by the Orthodox Church—Russian, Bulgarian and Serbian. The 1917 Revolution brought in its wake further simplifications with the abolition of a number of characters.

The Cyrillic alphabet for Russian is shown on the right, in Times New Roman (327), together with the extra characters required for Ukrainian and Byelorussian. The Cyrillic alphabet for Serbo-Croatian, which omits nine of the Russian characters, is also shown, together with its Latin counterpart. The Macedonian alphabet is the same as for Serbo-Croatian, except that it replaces *h* and *h* by *f* and *k*, and adds the letter *s*.

There are no accents in Russian or the other Slavonic languages written in Cyrillic script, but it is common in textbooks and dictionaries for the stress in pronunciation to be marked by an acute accent for the benefit of students.

### Languages of the U.S.S.R.

Russian, Ukrainian and Byelorussian are Slavonic languages, but among the numerous languages in the U.S.S.R. there are many that are not in any way related to Russian. There are, in fact, at least 120 different languages spoken by the 216 million inhabitants of the U.S.S.R., some of them by only a few hundred speakers and some of them as yet never written down. Instruction and education are given in over a hundred different languages, and newspapers are published in 67 of them.

Among the language families strongly represented is Turkic, the name given to languages akin to Turkish. These languages are much more closely akin, for example, than the Romance languages, and with a knowledge of Turkish the traveller could make himself understood over a territory extending from Asia Minor to Mongolia. In earlier days several of these languages, like Turkish itself, were written with Arabic characters. After the 1917 Revolution there was a movement away from Arabic towards Latin characters. The lead was taken in the Soviet Republic of Azerbaijan, where the new Latin script was officially adopted in 1924. This was the inspiration for Kemal's introduction of Latin script in Turkey in 1928. It was also the beginning of a movement that was taken up by many other peoples in the U.S.S.R. including some, like the Kalmyks and Buryats of Mongolia and even the Jews and

## RUSSIAN

upright	inclined	translit.	upright	inclined	translit.
А а	<i>A a</i>	a	Р р	<i>P p</i>	r
Б б	<i>B b</i>	b	С с	<i>C c</i>	s
В в	<i>V v</i>	v	Т т	<i>T t</i>	t
Г г	<i>G g</i>	g	У у	<i>U u</i>	u
Д д	<i>D d</i>	d	Ф ф	<i>F f</i>	f
Е е	<i>E e</i>	e	Х х	<i>X x</i>	kh
Ё ё	<i>Ë ë</i>	e	Ц ц	<i>U u</i>	ts
Ж ж	<i>Ж жэ</i>	zh	Ч ч	<i>Ch ch</i>	ch
З з	<i>Z z</i>	z	Ш ш	<i>Sh sh</i>	sh
И и	<i>I i</i>	i	Щ щ	<i>Shch shch</i>	shch
Й й	<i>Ï ï</i>	y	Ъ ъ	<i>Ь ь</i>	hard sign
К к	<i>K k</i>	k	Ы ы	<i>I i</i>	i
Л л	<i>L l</i>	l	Ь ь	<i>Ь ь</i>	soft sign
М м	<i>M m</i>	m	Э э	<i>E e</i>	e
Н н	<i>N n</i>	n	Ю ю	<i>Iu iu</i>	yu
О о	<i>O o</i>	o	Я я	<i>Ya ya</i>	ya
П п	<i>P p</i>	p			

Extras for Ukrainian: *i, e, l*.

Extras for Byelorussian: *ÿ, i*.

## SERBO-CROATIAN

Cyrillic		Latin	Cyrillic		Latin
upright	inclined		upright	inclined	
А а	<i>A a</i>	A a	Н н	<i>HH</i>	N n
Б б	<i>B b</i>	B b	Њ њ	<i>Њ њ</i>	Nj nj
В в	<i>V v</i>	V v	О о	<i>O o</i>	O o
Г г	<i>G g</i>	G g	П п	<i>P p</i>	P p
Д д	<i>D d</i>	D d	Р р	<i>R r</i>	R r
Ђ ђ	<i>Ђ ђ</i>	Đ đ	С с	<i>C c</i>	S s
Е е	<i>E e</i>	E e	Т т	<i>T t</i>	T t
Ж ж	<i>Ж жэ</i>	Ž ž	У у	<i>U u</i>	U u
З з	<i>Z z</i>	Z z	Ф ф	<i>F f</i>	F f
И и	<i>I i</i>	I i	Х х	<i>X x</i>	H h
Ј ј	<i>J j</i>	J j	Ц ц	<i>U u</i>	C c
К к	<i>K k</i>	K k	Ч ч	<i>Ch ch</i>	Č č
Л л	<i>L l</i>	L l	Ш ш	<i>Sh sh</i>	Š š
Љ љ	<i>Lj lj</i>	Lj lj	У у	<i>U u</i>	Dž dž
М м	<i>M m</i>	M m	Ш ш	<i>Sh sh</i>	Š š

Chinese of the U.S.S.R., who already had a form of writing of their own. Lenin himself was enthusiastic for "latinization" as a weapon against illiteracy and he spoke of it as "the great revolution of the East".

The alphabet used was much like that still employed for Turkish and was known as the Unified Turkic Latin Alphabet (UTLA).

From 1939 onwards this alphabet fell out of official favour, and during the years that followed it was everywhere replaced by the Russian alphabet with many additional symbols. Cyrillic is thus in widespread use throughout the U.S.S.R. for nearly all languages found there.

Of the Turkic languages the most important are the following. **UZBEK**, spoken by about six million people not only in the Soviet Republic of Uzbekistan but also in Chinese Turkestan and in Afghanistan, goes back to the 15th century, and there has been a rapid growth in its recent use for educational purposes. **TATAR** (or Tartar) is a name applied to a group of kindred Turkic languages spoken by some 4½ million people in Western Siberia, on the Volga, and in the Crimea. **KAZAKH**, the language of the Soviet Republic of Kazakhstan, is numerically strong, being the native tongue of some 3½ million people. It has had a flourishing literature of its own since feudal Islamic times. **AZERBAIJANI**, the language of nearly three million people in the Soviet Republic of Azerbaijan, has a considerable literature that goes back to the 16th century. It is of growing importance as the *lingua franca* of Transcaucasia. **CHUVASH** is the native tongue of 1½ million people in the Autonomous Republic of the same name. **TURKMEN** is the language of the Soviet Republic of Turkmenistan, where it is spoken by about a million people. **KIRGIZ** is spoken by nearly a million people in the Soviet Republic of Kirgiz. **BASHKIR** is spoken by 590,000 people in the Autonomous Republic of the same name. **YAKUT** is the language of the Yakut Autonomous Republic, where it is spoken by 230,000 people. **KARAKALPAK** is spoken by 160,000 people in Uzbekistan. **KUMYK** is the language of 110,000 people in Daghestan. **KARACHAY** and **BALKAR**, two closely allied languages, are spoken by 110,000 people in the Caucasus.

Another group of languages has affinities with Finnish on the one hand and with Hungarian on the other, and is known as Finno-Ugric. In this group are: **MORDVIN**, spoken by some two million people in the Mordva Autonomous Republic; **CHEREMISS**, also called Mari, spoken by half a million people of the Mari Autonomous Republic; and, further east in Siberia, **VOTYAK**, **KOMI**, **KOMI-PERMYAK** and **OSTYAK**.

Another group of related languages is known from its geographical location as Caucasian. The most widely spoken is **CHECHEN**, the language of 420,000 people, followed by **AWAR**, **KABARDIN**, **LEZGIN**, **DARGWA**, **INGUSH**, **ADYGHE**, **ABKHASIAN**, **LAK** and **TABASSARAN**.

In Asiatic Russia, **BURYAT**, a Mongolian language, is spoken by 240,000 people, and **KALMYK**, another Mongolian language, by 95,000. Of Manchurian languages, **CHUKCHA** is spoken by 11,000 people, **KORYAK** by 6,000, and **GILYAK** by only 3,000.

Coming nearer home, there are several languages belonging to the Indo-European family but more closely related to the Persian than to the Slavonic branch. The chief of these are: **TAJIK**, spoken by 1½ million people in the Tajik Soviet Republic, **OSSETIC**, spoken by 370,000 people in a territory which extends both north and south of the central Caucasus range, and **KURDISH**, the language of the Kurds.

These then are the chief languages in the U.S.S.R. now printed with Cyrillic script. As shown elsewhere, Lithuanian, Latvian and Estonian are written with Latin script, and Armenian and Georgian have their own distinctive scripts.

#### 'Monotype' Cyrillic faces

Albion (42)  
 Albion (63)  
 Albion (70)  
 Baskerville (169)  
 Bembo Tiding (370)  
 Bodoni Bold (260)  
 Bodoni Bold Condensed (529)  
 Bodoni Ultra Bold (120)  
 Cushing (17)  
 French Old Style (68)  
 French Old Style Bold (69)  
 French Round Face (44)  
 Garamond Bold (201)  
 Gill Sans (262)  
 Gill Sans Bold (275)  
 Gill Sans Condensed Titling (373)  
 Gill Sans Cameo (231)  
 Gill Sans Condensed (485)  
 Gill Sans Light (362)  
 Grotesque (51)  
 Grotesque (96)  
 Grotesque Bold (97)  
 Grotesque Bold Condensed (166)  
 Grotesque Bold Condensed (111)  
 Grotesque Condensed (33)  
 Grotesque Condensed (383)  
 Modern Condensed (39)  
 Modern Condensed (79)  
 Modern Extended (77)  
 Modern Wide (75)  
 Neo Didot (27)  
 Placard Condensed (568)  
 Placard Condensed Titling (522)  
 Plantin Bold (194)  
 Plantin Light (113)  
 Times New Roman (327)  
 Times Bold (334)



## Latin

The Latin alphabet as we know it is derived from the Greek alphabet, as it was used by Greek colonists in Italy. The successful adaptation is due in part to the Etruscans, the mysterious people who occupied Italy before the Romans. The Etruscans applied the Greek alphabet to their own use (which included some nine thousand inscriptions as yet imperfectly deciphered) and passed it on to the Romans. The political supremacy of the Romans ensured the expansion of their alphabet as the official script of the Roman world.

An inscription carved on a fibula about the 7th century B.C. is the earliest known specimen of the Latin alphabet. The letters are Greek in form and read from right to left, as do many of the early examples down to the middle of the 4th century B.C. The early history of Roman usage is obscure, but from the 1st century B.C. there are abundant Latin inscriptions to allow a full study of its development.

The original Latin alphabet, like the Etruscan, had 21 letters. The additional letters were acquired gradually. The letter C was at first used to denote the sounds of both *c* and *g*, and later it was given a bar to distinguish the sound *g*; in this way the letter G was acquired. X was added at the end of the alphabet to denote the sound *ks*. In the 1st century B.C. the letters Y and Z were introduced to represent Greek sounds and were placed after X in the alphabet.

In the Christian era further changes were made. The letters U and V, hitherto interchangeable, were distinguished and given distinctive sounds. J, originally a form of initial I, was made to represent the letter when it was a consonant while I represented it as a vowel. Finally the letter W was introduced to represent the sound of *u* as a consonant. And so the total of 26 letters was achieved.

The earliest forms of the Latin alphabet were capitals, carved on stone with chisels. The best-known example, and the inspiration of many modern types, is the Trajan column in Rome (114 A.D.). With the use of other means of writing, by stylus on wax tablets, and by brush, quill and pen on papyrus, parchment and paper, more rounded cursive scripts were developed. Fragments of Latin writing on papyri show that, in the 2nd and 3rd centuries A.D., attempts were being made to evolve a new style based on cursive. And in the 4th century a fine round calligraphic script known as *uncial* emerged, and this flourished as a book script until the 8th century.

As the unifying power of the Roman Empire began to wane in the 5th century, so writing in both *majuscule* (capitals) and *minuscule* (small letters) took on different characteristics in different places at different times. One interesting example of this can still be appreciated in the Irish script of today (see alphabet on right). The uncial and half-uncial scripts, brought to Ireland by St. Patrick and his Roman missionaries in the 4th century, were perfected by the Celtic scribes and, because of Ireland's isolated position, the Irish script remained unchanged for generations.

The insular scripts of Ireland and England exerted considerable influence upon the Continent through the Irish and English missionaries who travelled all over Europe. At about 780 in France, Charlemagne called in the English scholar, Alcuin, to reorganize the educational system of his empire. Through his efforts a standardized style of writing was adopted for copying the Vulgate Bible, and this came to be known as *Carolingian minuscule*. This text letter was a new development of old cursive forms, particularly the half-uncial, and it is the forerunner of our own lower-case alphabet.

ABCDEFGHIJKLMNOPRSTU  
abcdefghilmnopprtu

Towards the end of the 9th century the Gothic script began to develop out of Carolingian minuscule; its slanting characters could be written faster, and the compressed nature of the letters saved space. In the following centuries this script became increasingly condensed and took on a blacker appearance. It flourished most in Germany and was the model for the first type faces of Johann Gutenberg and other early German printers. Black letter types, known as *Fraktur* (see alphabet on right), are still employed in Germany today although their use is somewhat on the wane.

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

At the time of the Italian Renaissance in the 14th century, the Humanist scholars (the first of whom is traditionally claimed to be Petrarch) went back to the manuscripts of the 9th, 10th and 11th centuries written in Carolingian minuscule and adapted their Gothic hands to produce the fine Roman writing we call *humanistic*. In just the same way as the Gothic script gave way in Italy to the humanistic, so Gothic types gave way to Roman. In 1465, only ten years after Gutenberg's black-letter Bible, a type approximating to the humanistic script was used by Sweynheim and Pannartz in Subiaco for printing Cicero's *De oratore*. By the end of the 15th century Roman type, as we know it today, had arrived, largely as a result of the publications of the Venetian printer, Aldus Manutius. Aldus was also the first printer to use an *italic* type, based on the *chancery* script which was used for writing letters, commercial notes and documents of minor importance.

**NUMERALS** The classical Roman system of numbering is still used to indicate the serial numbers of certain periodicals, to mark the hours on a clock, and to denote the dates of buildings. The modern so-called "Arabic" numerals are probably of Indian origin. (There is some doubt as to whether they were actually invented in India or introduced there by Greek or other traders from the west.) They were brought from the east by Arabian traders and introduced by them into Spain in the Middle Ages. From there they spread over Europe, coming into use in England during the 11th century.

#### European languages

Most European languages, and those languages of European origin, make use of the Latin alphabet. The normal 26 letters are sometimes inadequate to indicate the pronunciation of a letter or to distinguish one sound from another. To meet this difficulty letters are often modified by diacritical marks. A diacritical mark may be defined as a sign placed above, below or across a letter to indicate the correct sound represented by that letter in the given word.

Of the accents, only the *acute* was known during the Middle Ages, and even that was not used for its present purpose of distinguishing one sound of *e* from another until Robert Estienne, the famous printer, employed it for that purpose in his *Dictionarium* (1530). The apostrophe to indicate elision was introduced by another French printer, Geoffroy Tory, in 1592, and the *cedilla* (originally a small *z*) was borrowed by him from Spanish but its use did not become general until much later. The *grave* accent was sparingly used; Pierre Corneille was the first to advocate its systematic use to indicate the open quality of *e* (*è*), but his example was not followed for a long time. The *tréma* (to indicate a dieresis) and the *circumflex* (to indicate a contraction) were introduced by Sylvius in 1532, and the hyphen by Montflory in 1538. Quotation marks are relics of small vertical lines used to denote the sacred words of Holy Writ in mediaeval times. The general adoption of these signs took place gradually in the course of the 16th century. The Dutch printers were largely responsible for their extensive and regular application.

### Africa script

In Africa there are some 1,400 different and distinct languages spoken. Many of these languages have never been written down, and of those that have most owe their existence to the efforts of missionaries and administrators.

Missionaries came from many countries and naturally transliterated native sounds into the nearest equivalent in their own languages. A French monk might represent a sound in one way and an Italian friar in another, while two British officials might hit on two quite different ways of representing the same sounds. As a result a number of methods of transliterating native sounds grew up in various parts, all using the Latin alphabet, some with and some without diacritical marks, with no essential basis in common. Even locally there was no consistency, and the same vowel in the same language might be represented in as many as four different ways.

The resulting confusion was tackled by the International African Institute in 1927, when it set out to formulate principles for a simple and uniform system of writing African languages. The result of its efforts was published in 1930 as *A Practical Orthography of African Languages*. The new system there laid down is based on the alphabet of the Association Phonétique Internationale. It avoids the excessive use of diacritical marks or digraphs, and for sounds for which the Latin alphabet provides no symbols it uses new letters taken from the alphabet of the Association.

In the resulting alphabet the following consonants have their usual English values: *b d f h k l m n p s t v w z*. The other consonants, which have different values in different words when used in English, are used with a constant value: *c g j r y*. The following special consonants have been added: *β ḍ ɲ fɔ̣ ʒ ʎ*. The normal vowels are used, plus the following: *ɛ ə ə*. A full list of the characters used in the Africa script is shown on the right. By far the most widely used letters are *ε ə ɲ*, and these are sufficient for many languages.

Two of the leading figures in promoting the Africa script were Professor D. Westermann and Professor A. N. Tucker. Dr. Westermann visited West Africa in the years from 1927 to 1929 to investigate local languages and to explain their representation by the new script. As a result of his visit several territories adopted the orthography he suggested, but some of these are now having second thoughts and the present situation is far from settled.

A	a	H	h	R	r
B	b	X	x	S	s
β	β	I	i	Σ	ʃ
C	c	J	j	T	t
D	d	K	k	U	u
Ð	ḍ	K̄	k̄	V	v
E	e	L	l	U	ɔ
ε	ε	M	m	W	w
ə	ə	N	n	Y	y
F	f	D	ɲ	Z	z
F̣	f̣	O	o		ʒ
G	g	ɔ̣	ɔ̣		
Y	ʎ	P	p		

### The All India Alphabet

India is a country of many languages. The *Linguistic Survey* records 179 languages and 544 dialects, and these include 16 major literary languages which are written in a dozen different scripts.

In 1941 two well-known phoneticians, Professor Daniel Jones and Dr. (subsequently Professor) J. R. Firth put forward a scheme for an All India Alphabet, by which the various languages of India could be represented by the normal Roman letters with the addition of a few phonetic signs. This Alphabet has been applied to Hindi, Urdu, Bengali, Gujerathi, Marathi, Sinhalese, Tamil and Telugu, and quite a number of Hindi and Urdu text-books have been set in it.

The full Latin alphabet is used, and the additional letters taken from the phonetic alphabet are as follows: *ɖ ε ə ɣ l ɳ ɲ ɳ ə ʀ ʃ ʂ t ʒ*.

### American Indian languages

John Eliot, of Jesus College, Cambridge, arrived in Massachusetts in 1631 and began to study the language of the native Indians for the

purpose of preaching to them. His translation of the Bible into their language was completed and printed in 1663.

He adapted Latin characters to record the native speech, and he was the first of many scholars to do so. In the subsequent centuries the following Indian languages have been reduced to Latin script, usually with the addition of diacritical marks and sometimes of special characters: Cherokee, Chinook, Chippewa, Choctaw, Cree, Dakota (Sioux), Iroquoian, Kwakiutl, Maidu, Muskogee, Navaho, Osage, Pima, Takelina, Tsimshian; also the Caddoan language family comprising Caddo, Kitsai, Pawnee and Wichita.

#### Central and South American languages

According to one expert opinion, there are 96 known languages in Mexico and Central America and 783 languages in South America and the Caribbean. Some of these languages have been written down, chiefly by missionaries.

In Mexico over half a million Indians speak a form of Aztec that is probably derived from the language of the ancient Mexican civilization (see *Aztec*). Similarly in Yucatan, Mexico, the language of Maya is still spoken by Indians. The language of the ancient Inca civilization of Peru lives on in the various forms of Quechua, which are spoken by some four million Indians in central South America. Carib, the original language of the West Indians which gave its name to the Caribbean, is still spoken by Indians along the shores of British Honduras. Another widespread language was Arawak, once extensively spoken in the Caribbean and South America, and now strongest in British and Dutch Guiana.

In Central America Cakchiquel and Quiche in Guatemala, and Bribrí in Costa Rica, have been written down. So too have Makuchi in British Guiana, Aguaruna in northern Peru, Aymara in southern Peru and Paraguay, Mataco in Argentina, and in the very south of the continent Yangan in Tierra del Fuego.

#### Australian Aboriginal languages

There were probably about five hundred languages spoken in Australia when the first white men settled there. Many of these languages became extinct before they were written down. Of those that remain the best known is the Aranda, which is spoken in central Australia. Others that have been studied and written down are Djadjala in Victoria, Wiradjuri, Kamilaroi, Kurnu and Wongaibon in New South Wales, and Dieri in South Australia. Many of these languages may be of common origin, for a considerable amount of similar vocabulary occurs over large tracts of Australia (but not including Aranda). A few Australian words have even found their way into English. They include kangaroo (recorded by Cook in 1770), boomerang, corroboree, dingo, nulla-nulla, waddy, koala, wallaby.

#### Artificial languages

Numerous attempts have been made to bridge the language gap between countries and peoples by the invention of a universal artificial language. Amongst those projects that have met with some success are Anglic, Basic English (see *English: Part Two*), Esperanto (see *Part Two*), Ido, Interglossa, Interlingua (see *Part Two*), Novial and Volapuk, all of which use the Latin alphabet.

There is an International Auxiliary Language Association which examines all proposed artificial languages.

## Devanagari

Nothing is known with certainty about the origin of writing in India, but one theory is that an alphabet known as the Brahmi letters, which was derived from some primitive form of the Semitic alphabet, was introduced into Western India by traders about 500 B.C., and that from it most forms of writing now used in India are descended.

Another theory would make Indian writing entirely Indian in origin. Five thousand years ago there was a highly developed civilization in the Indus Valley which possessed a form of writing. This survives on ancient seals to the number of five hundred, but it has never been satisfactorily deciphered. In this, some Indian scholars see the primitive form from which the Brahmi alphabet was developed.

In northern India this original Brahmi script gradually evolved into what is now known as Devanagari. With the introduction of printing this alphabet was generally used for the ancient Sanskrit literature (see *Part Two*). It is also employed for six modern languages of northern India, namely Bihari, Hindi, Kashmiri, Marathi, Nepali and Rajasthani (see *Part Two*). The alphabet used in printing Gujarathi is a modified form of Devanagari, and the alphabets used for Bengali and Punjabi are also related to it, though not so closely.

All of these languages are written from left to right, and the words of a sentence are not necessarily divided. There are no capital letters.

The Devanagari alphabet is traditionally and conveniently arranged under the different classes of letters. First of all there are 14 vowel and diphthong characters, among which the ancient Hindu grammarians included *r̄* and *l̄*. Each vowel occurs in a short and a long form. Both forms of these vowels are used in Hindi, and the four alternative forms shown in the list on the right are preferred for Marathi and Sanskrit.

The forms of the first lot of vowels in the list are used only when the vowel forms a syllable by itself. The vowels and diphthongs, if preceded by a consonant, are represented by ordinary, superior or inferior symbols, as shown below. In addition the two superior signs ' and ^ indicate a nasal sound. : indicates an aspirate.

The consonants as normally written represent the consonant plus the vowel *a*. Where a consonant is to be pronounced without a vowel following it, it is printed with a sloping stroke below it, which is known as a *virama*, thus क्. Consonants can be combined with each other either side by side or one above the other. Most consonants have shortened forms that are used when they are combined with other consonants, thus क्क, क्ख, क्ग, क्घ, and so on.

Three compound letters are often given as part of the alphabet, because they differ markedly from the letters from which they are derived: क्श्cha, क्ग्या, क्त्रa.

In Hindi a few letters are printed with a dot beneath them to represent sounds that do not occur in Sanskrit, thus क्ड़a, क्ढ़a, क्फ. The three alternative forms in the list of consonants on the right are preferred for Marathi and Sanskrit.

Letters with a dot beneath them (ka, da, etc.) are known as cerebrates and are common in all Indian languages. They are pronounced with the lower side of the tip of the tongue touching the roof of the mouth.

A single vertical line is used in poetry to mark the end of the first line of a couplet, and two lines to mark the end of the couplet. In prose, the double lines correspond to a full stop, but modern European punctuation is frequently used.

There are two 'Monotype' faces available: Devanagari (155) and Devanagari Bold (346). Both are also available for filmsetting.

### VOWELS AND DIPHTHONGS

#### separate form

अ, अ	a	ऋ	ॠ	ओ, ओ	o
आ, आ	ā	ॠ	ॡ	औ, औ	au
इ	i	ऋ	ॠ		
ई	ī	ॠ	ॡ		
उ	u	ॠ	ॡ		
ऊ	ū	ॠ	ॡ		

#### combination form

		superior		inferior	
ॠ	ā	ˆ	e	ˆ	u
ॡ	ī	ˆ	ē	ˆ	ū
ॢ	ī	ˆ	ṛ	ˆ	ṛī
ॣ	o	ˆ	ṛē	ˆ	ṛī
।	au	ˆ	ṛē	ˆ	virama
॥	ṛī	ˆ	nasal		
०	ṛo	ˆ	nasal		

#### CONSONANTS

क	ka	ख	ḍa	म	ma
ख	kha	ḍ	ḍha	य	ya
ग	ga	ण, ण	ṇa	र	ra
घ	gha	त	ta	ल, ल	la
ङ	ṅa	थ	tha	व	va
च	ca	द	da	श	śa
छ	cha	ध	dha	ष	ṣa
ज	ja	न	na	स	sa
झ, झ	jha	प	pa	ह	ha
ञ	ña	फ	pha	ळ	ḷa
ट	ṭa	ब	ba		
ठ	ṭha	भ	bha		

#### NUMERALS

१	२	३	४	५	६	७	८	९	०
1	2	3	4	5	6	7	8	9	0

## Gujerathi

The script employed for Gujerathi, the language of Gujerath (see *Part Two*) in the Bombay State, resembles Devanagari to which it is closely related. The chief difference is that Gujerathi has discarded the top lines of Devanagari and favours rounded forms in place of straight bars.

As in Devanagari, the vowels have two forms, one for independent use and the other a shortened form for use in conjunction with a consonant. (These forms are combined as follows: on the right of the consonant - ā, ī; on the left - i; below - u, ū; above - e, ai; above and on the right - o, au.) Also, where no vowel is indicated, the first vowel *a* is to be assumed, and the absence of any vowel is indicated by the *virama*.

Consonants may be combined, as in Devanagari, to form compound characters. The letters may be abbreviated (६ for ५, ७ for ५, ८ for ५, and so on), or they may assume special forms. In particular, the letter *r* at the beginning of a word can be indicated by ५, thus ५ rpa. Other examples of conjunct consonants are: ५ = ५ + ५, ६ = ६ + ६, ७ = ७ + ५.

A dot above a character indicates a nasal pronunciation, thus ५̣.

In prose the same system of punctuation is used as for English. To punctuate verse the same system is used as for Sanskrit verse. As in other Indian scripts, there are no capital letters, and the script reads from left to right.

There are two 'Monotype' faces available: Gujerathi (460) and Gujerathi Bold (518).

separate form	combination form	transliteration	separate form	combination form	transliteration				
अ		a	ख		ḥi				
आ	।	ā	ख	।	e				
इ	।	i	ख	।	ai				
उ	।	u	ख	।	au				
ऊ	।	ū							
६	ka	६	ta	५	va				
५	kha	५	tha	५	ṣa				
७	ga	६	da	५	ṣa				
५	gha	५	dha	५	sa				
५	ca	५	na	५	ha				
७	cha	५	pa	७	la				
७	ja	६	pha		special forms				
७	jha	७	ba	५	ksha				
७	ṭa	७	bha	५	gn				
७	ṭha	५	ma	५	tra				
७	ḍa	५	ya	५	ja				
७	ḍha	५	ra						
७	ṇa	५	la						
५	२	३	४	५	६	७	८	९	०
1	2	3	4	5	6	7	8	9	0

## Bengali

The script in which Bengali (see *Part Two*) is written is derived from Devanagari, which it closely resembles. It has the same order of letters, the same alternative forms of vowels, and a similar manner of joining consonants together.

The vowel forms are combined as follows: on the right of the consonant - ā, ī; on the left - e, ē, i; on both sides - o, au; below - u, ū, ī.

Sometimes the second consonant is written beneath the first, thus ५ = ५ + ५, ७ = ७ + ७, and so on. Sometimes they are joined laterally, thus ५ = ७ + ५, ५ = ६ + ५. Either the first or the second character may be modified in form to facilitate the joining of two characters, thus ५ = ७ + ७, ५ = ७ + ५. The three consonants, ७, ५ and ५ have special forms that are used in combinations: ७ = ७ + ५; ५ = ५ + ५; ५ = ५ + ५. Some consonants have special forms when combined with vowels, thus ५ ru, ५ rā, ५ śu, ५ hu, ५ hri.

Three modifying signs are used. The mark \* above a character indicates a nasal sound. The mark † is used finally or medially as an alternative form of ७. The sign ‡ after a character represents an aspirate.

The absence of a vowel after a consonant is indicated by ५.

In prose, the end of a sentence is marked by an upright stroke. Other punctuation marks are used as in English. In modern verse, the same usage is often observed, although formerly it was the practice to follow the punctuation used in Sanskrit.

As in other Indian scripts, there are no capital letters, and the script reads from left to right.

There is one 'Monotype' face available: Bengali (470).

separate form	combination form	transliteration	separate form	combination form	transliteration				
अ		a	ख		ḥi				
आ	।	ā	ख	।	e				
इ	।	i	ख	।	ai				
उ	।	u	ख	।	au				
ऊ	।	ū							
क	ka	उ, ङ	ḍa	न	ma				
ख	kha	५, ५	ḍha	ब	ya				
ग	ga	५	ṇa	र	ra				
घ	gha	त	ta	ल	la				
ङ	ṅa	प	tha	श	ṣa				
च	ca	द	da	ष	ṣa				
छ	cha	ध	dha	स	sa				
ज	ja	न	na	ह	ha				
झ	jha	प	pa		special forms				
ञ	ña	क	pha	क	ksha				
ट	ṭa	ब	bha	ञ	jna				
ठ	ṭha	त	bha						
५	२	३	४	५	६	७	८	९	०
1	2	3	4	5	6	7	8	9	10

## Gurmukhi

This script derives its name from the second Sikh Guru or religious leader, Angad by name, who was head of the sect from 1538 to 1552. He is said to have found the script in use unsatisfactory, and so he devised a new alphabet, borrowing part of it from Devanagari. The new alphabet was then called Guru-mukhi, the alphabet proceeding from the mouth of Guru. This script has been used ever since for writing the Sikh scriptures and has been applied to other purposes, chiefly among members of that sect (see *Punjabi: Part Two* for language specimen).

The representation of the vowels, shown in the alphabet on the right, is somewhat complicated. As in other Indian scripts, the short vowel *a* (which is not distinguished) is implied with consonants which have no other vowel attached to them. Besides the four independent vowel forms shown in the alphabet on the right, there are further combination vowel forms which are attached to an independent vowel when it occurs initially or to a consonant when the vowel follows a consonant.

The script reads from left to right and there are no capital letters. A 'Monotype' Duplex Mould is used for casting Gurmukhi, and there are two 'Monotype' faces available: Gurmukhi Bold (601) and Gurmukhi (604).

## Sinhalese

Sinhalese (see *Part Two*) is a language of the Indo-European family that is spoken in Ceylon. It was introduced there by colonists who sailed from Bengal in the 5th century B.C. The script in which it is written is a development of the Brahmi script that was introduced from the mainland about the 3rd century B.C. It has been greatly influenced in its development by the fact that it was for a long time written on palm leaves. The interest of the writers of Sinhalese in calligraphy has given their script an air of complex elegance.

The vowel *a* is inherent in every consonant unless displaced by some other vowel or suppressed by the appropriate sign. Absence of a vowel is indicated by one of two additions to a consonant, thus  $\zeta$  da  $\zeta'$  d,  $\ominus$  ta  $\ominus$  t.

All vowels except *a* have two forms, the independent one shown in the alphabet on the right and a shortened one. In the following cases the shortened form can be shown separately:  $\circ$  ā, a r̄, aa r̄,  $\circ$  e,  $\circ$  ai,  $\circ$  - $\circ$  o,  $\circ$  - $\circ$  ō,  $\circ$  - $\circ$  au,  $\text{r}$  æ,  $\text{r}$  ǣ. In the remaining cases the shortened form is attached to the preceding consonant, thus:  $\text{z}$  ki,  $\text{z}$  ki,  $\text{z}$  ku,  $\text{z}$  kũ,  $\text{z}$  kē.

There are some forms that may be regarded as exceptional, thus  $\text{r}$  ra,  $\text{r}$  rā,  $\zeta$  da,  $\zeta'$  da. In combinations  $\text{c}$  often assumes the form  $\text{c}$ .  $\text{c}$  takes the form of a large loop below the consonant it follows, thus  $\zeta$  gra,  $\text{c}$  gra. The nasal sign  $\circ$  takes the form of an initial curlicue, thus  $\text{c}$  n̄ga.

Consonants can be joined together in pairs, and even in threes, but these combinations are generally used only in words of Sanskrit origin.

The ancient numerals of Sinhalese are no longer used by modern writers, who follow European usage in numbering as they do also for punctuation. There are no capital letters, and the script reads from left to right.

There are two 'Monotype' faces available: Sinhalese (557) and Sinhalese Bold (657). Both these faces are also available for filmsetting.

separate form	combination form	separate form	combination form
a	ā	i	ī
or	ai	or	ī
e	au	ē	ay
or	ū	o	o
sa	ñā	pa	
ha	ṭṭa	pha	
ka	ṭṭha	ba	
kha	ṭṭḍa	bha	
ga	ṭṭḍha	ma	
gha	ṭṭḍha	ma	
ṅa	ṭṭa	ra	
cha	ṭṭha	la	
chha	ṭṭda	va	
ja	ṭṭdha	ra	
jha	ṭṭna		
१	२	३	४
५	६	७	८
९	०		

අ	a	සා	}	එ	ai
ආ	ā	සා	}	ඔ	o
ඉ	i	සා	}	ඔ	ō
ඊ	ī	සා	}	ඔ	au
උ	u	ඒ	e	ඇ	æ
ඌ	ū	ඒ	ē	ඇ	ǣ
ක	ka	ධ	ḍa	ම	ma
ඛ	kha	ඨ	ḍha	ය	ya
ග	ga	ණ	ṇa	ර	ra
ඝ	gha	ඨ	ḍa	ල	la
ඞ	ṅa	ඨ	ḍha	ව	va
ච	ca	ඤ	ḍa	ඹ	śa
ඡ	chā	ධ	dha	ඹ	ṣa
ජ	ja	න	na	ස	sa
ඤ	jha	ඨ	ḍa	හ	ha
ඞ	ṅa	ඨ	ḍha	ල	la
ට	ta	බ	ba		
ඨ	ṭṭha	භ	bha		

# Kannada

The Kannada language (see *Part Two*), also known as Canarese, is one of the Dravidian languages spoken by 15 million people in the south of India, and it is related to Telugu, Tamil and Malayalam.

The earliest use of the alphabet in which Kannada is written has been found in inscriptions of the 5th century A.D., and there are literary documents of the 9th century. Like most other Indian scripts, it stems from the Brahmi alphabet which was written from left to right. The structural principles of the various regional alphabets which developed from Brahmi remained identical (e.g. long and short forms of vowels, different forms of vowels for initials, etc.), but the actual shapes of the letters were subject to endless change. Consequently, there is very little visual similarity between Kannada and Devanagari, which is the form that the Brahmi alphabet took in north India.

In the alphabet on the right, the two forms of the vowels are shown. The initial forms are used when a vowel occurs at the beginning of a word or on its own. Where a vowel other than  $\text{a}$  follows a consonant it is indicated by the secondary form, which is placed above and/or to the right of the consonant. In the normal forms of the consonants the vowel  $\text{a}$  is included, being implied by the top flourish of what are called the "crested letters" -  $\text{ka}$ ,  $\text{ra}$ , etc.

Vowel sounds that follow a consonant can be added thus:  $\text{ka} + \text{a} = \text{ka}$ . Those that are placed above the consonant are replaced by special composite characters, thus:  $\text{ka} + \text{a} = \text{ka}$ ,  $\text{ka} + \text{a} = \text{ka}$ ,  $\text{ka} + \text{a} = \text{ka}$ , and so on. The absence of a vowel after a consonant is indicated by  $\text{r}$  to the right.

When a consonant is followed by another consonant the second used to be written below the first, sometimes in a special form. This usage is still found, thus:  $\text{ka}$ . But in general use it is avoided in printing.

The traditional punctuation is  $\text{!}$  and  $\text{||}$  which correspond roughly to a comma and a full stop, but European punctuation (as well as numerals) are increasingly used, particularly in periodicals.

There is one 'Monotype' face available: Kannada (588).

# Telugu

The Telugu script shares many features with other Dravidian forms of writing used in southern India, such as Tamil and Malayalam. In particular, it has many points in common with Kannada, for the two had a similar origin and from 1000 to 1300 A.D. the same script was used for both. (For language specimen see *Part Two*.)

As with Kannada there are two forms of the vowels, one for independent use and one for use when attached to a consonant. In some cases the vowel sign is placed above the consonant, as in  $\text{ka}$ ,  $\text{ka}$ ,  $\text{ka}$ ,  $\text{ka}$ ; others to the right of it as in  $\text{ka}$ , and  $\text{ka}$ ; and in the case of  $\text{ka}$  a part is placed below and a part above. Where no vowel is indicated, the vowel  $\text{a}$  is implied. Absence of a vowel is indicated by  $\text{r}$  above a letter, thus  $\text{ka}$ .

There are many combined consonants, formed on obvious principles. Some letters assume special forms in combination:  $\text{ka}$  can become  $\text{ka}$  as in  $\text{ka}$ ,  $\text{ka}$  can become  $\text{ka}$  as in  $\text{ka}$ .

There are no capitals, and the language is written from left to right. European numerals and punctuation are now generally used. The earlier punctuation signs were one and two short vertical strokes.

Two 'Monotype' faces are available: Telugu (426) and Telugu Medium (626).

separate form	combination form	transliteration	separate form	combination form	transliteration
ಅ		a	ಋ		ri
ಆ	ಃ	ā	ೠ		ē
ಇ		i	ಋ	ಃ	ē
ಊ	ಃ	ū	ಋ	ಃ	ai
ಋ		u	ಋ	ಃ	o
ೠ		ū	ಋ	ಃ	ō
ಕ		ka	ಋ		au
ಕ	ಃ	kha	ಋ	ಃ	la
ಗ		ga	ಋ	ಃ	wa
ಗ	ಃ	gha	ಋ	ಃ	śa
ಙ		ṅa	ಋ	ಃ	śa
ಚ		ca	ಋ	ಃ	sa
ಚ	ಃ	cha	ಋ	ಃ	ha
ಜ		ja	ಋ	ಃ	la
ಜ	ಃ	jha	ಋ	ಃ	la
ಬ		ba	ಋ	ಃ	special letters
ಬ	ಃ	bha	ಋ	ಃ	o
ಮ		ma	ಋ	ಃ	ṁ
ಮ	ಃ	ma	ಋ	ಃ	h
ಯ		ya	ಋ	ಃ	
ಯ	ಃ	ya	ಋ	ಃ	
ರ		ra	ಋ	ಃ	
ರ	ಃ	ra	ಋ	ಃ	
ೠ		ṛ	ಋ	ಃ	
ೡ		ṙ	ಋ	ಃ	
ೢ		ṛ	ಋ	ಃ	
ೣ		ṛ	ಋ	ಃ	
೤		ṛ	ಋ	ಃ	
೥		ṛ	ಋ	ಃ	
೦	೧	೨	೩	೪	೫
1	2	3	4	5	6

separate form	combination form	transliteration	separate form	combination form	transliteration
ಅ		a	ಋ		e
ಆ	ಃ	ā	ಋ	ಃ	ē
ಇ		i	ಋ	ಃ	ai
ಊ	ಃ	ū	ಋ	ಃ	o
ಋ		u	ಋ	ಃ	ō
ೠ		ū	ಋ	ಃ	au
ಕ		ka	ಋ		lu
ಕ	ಃ	kha	ಋ	ಃ	lū
ಗ		ga	ಋ	ಃ	
ಗ	ಃ	gha	ಋ	ಃ	
ಙ		ṅa	ಋ	ಃ	
ಚ		ca	ಋ	ಃ	sha
ಚ	ಃ	cha	ಋ	ಃ	sa
ಜ		ja	ಋ	ಃ	ha
ಜ	ಃ	jha	ಋ	ಃ	ja
ಬ		ba	ಋ	ಃ	special letters
ಬ	ಃ	bha	ಋ	ಃ	o
ಮ		ma	ಋ	ಃ	ṁ
ಮ	ಃ	ma	ಋ	ಃ	h
ಯ		ya	ಋ	ಃ	
ಯ	ಃ	ya	ಋ	ಃ	
ೠ		ṛ	ಋ	ಃ	
ೡ		ṙ	ಋ	ಃ	
ೢ		ṛ	ಋ	ಃ	
ೣ		ṛ	ಋ	ಃ	
೤		ṛ	ಋ	ಃ	
೥		ṛ	ಋ	ಃ	
೦	೧	೨	೩	೪	೫
1	2	3	4	5	6





# Thai

The Thai system of writing (formerly known as Siamese) is derived, like several others, from some form of writing which came from India along with the Buddhist religion, and it was introduced into Thailand by way of Cambodia. Tradition has it that the Thai alphabet was the work of Ramkhamhaeng the Great of Sukhothai. The earliest record of it, a monument known as the Inscription of King Ramkhamhaeng, dates from the year 1283 A.D.

The Thai alphabet retains the basic arrangements of consonants found in Devanagari, in which the letters are grouped according to their phonetic classification. There are no capitals, and the script is written from left to right. (For language specimen see *Part Two*.)

The vowels in Thai are indicated by symbols which are either simple (composed of one part) or complex (composed of more than one part). Of the simple vowel symbols, some appear before, some after, some above and some below the consonant that follows. The complex vowel symbols are built up of two or more simple vowel and/or consonant symbols.

Most vowel symbols occur in pairs of short and long. In the list on the right, the dash shows the position of the consonant preceding the vowel. The last four in the list represent *a* plus *m*, *j* and *w*.

In addition there are tonal marks, which are placed over the first or second consonant of the syllable they mark. The tone marks are as follows: ' ° ° ° °. If a syllable already carries a vowel symbol above it, the tone mark is placed above the vowel, thus ' ° ° ° ° and so on. Similarly, the tone mark may appear above a consonant and the vowel below it, thus ' ° ° ° and so on.

European numerals are sometimes used in Thai books and newspapers, and Thai schoolchildren are normally taught both systems.

The following signs are peculiar to Thai: ◊ indicates the omission of a familiar word which should be supplied in reading; ◊ indicates that a word or expression is to be repeated; † indicates the suppression of the vowel or consonant beneath it in reading.

In printing Thai, words are not separated by spaces but are run together, as with Sanskrit and languages using Devanagari script and its derivatives. Spaces indicate the end of a phrase, clause or sentence, and thus play the part of a comma or period. Hyphens and other punctuation marks.

A 'Monotype' Duplex Mould is used for casting Thai, and there are five 'Monotype' faces available: Thai (577), Thai Bold (611), Thai Light (608), Thai Light Italic (628) and Thai Medium (621).

short	long			
— ◊ —	— ◊ —	a	— ◊	um
◊	◊	i	◊	i
◊	◊	y	◊	i
◊	◊	u	◊	ow
◊	◊	e	tone marks	
◊	◊	ē	◊	◊
◊	◊	o	◊	◊
◊	◊	o	◊	◊
◊	◊	ua		

dashes show the position of the consonants which precede the vowels in sound.

◊	k	◊	th	◊	m
◊	kh, c	◊	th	◊	y, j
◊	kh, c	◊	n	◊	r
◊	kh, c	◊	t	◊	l
◊	kh, c	◊	d	◊	w, v
◊	kh, c	◊	th	◊	s
◊	ng	◊	t	◊	s
◊	ch	◊	th	◊	s
◊	ch	◊	n	◊	h
◊	z	◊	b	◊	l
◊	sh	◊	p	◊	a
◊	y, j	◊	f, ph	◊	ryy
◊	d	◊	p	◊	lyy
◊	t	◊	f, ph		
◊	th	◊	p		

◊ Obsolete      † used as vowel

◊	◊	◊	◊	◊	◊	◊	◊	◊	◊
1	2	3	4	5	6	7	8	9	0

## PART TWO : LANGUAGES

### AFGHAN (Persi-Arabic script)

Afghan or Pashto is an Indo-European language, spoken by some 12 million people, and has been the official language of Afghanistan since 1936. It is also spoken in the North-West frontier region of Pakistan and in parts of Iran.

The written literature of the language, which goes back to the 16th century, has always been greatly influenced by Persian, but there is also a national oral literature, chiefly represented by popular songs. (See *Arabic in Part One.*)

SPECIMEN (I have seen a nightingale): بلبله م ليبل ده .

### AFRIKAANS (Latin script)

Afrikaans, in general a modified form of Dutch (with English and Malay additions), is an official language of South Africa.

The Latin alphabet is used, but *c q x z* occur only in foreign words and proper names. The following accented letters are used: á â ã ä è é ê ë ì í î ï ò ó ô õ ü ú û ü.

SPECIMEN This is the beginning of an account of the Portuguese discovery of the great ruins at Zimbabwe:

Die Portuguese het aan die begin van die sestiende eeu 'n neersetting gestig by Sofala, en daarvandaan het hul vrugtelose soektoegte onderneem na die legendariese koninkryk wat êrens in die binnelande van Afrika sou bestaan.

### ALBANIAN (Latin script)

Albanian is one of the oldest languages in Europe. It dates from pre-Hellenic times, but it has been considerably changed through the centuries. It is now spoken by over a million people in Albania and by scattered minorities elsewhere.

The language has been written in twenty different ways at different times, but by a decision of a special committee of 1908 it is now written in the Latin alphabet but with the following accented letters: â ç ë.

SPECIMEN Here are some lines sung by Highland shepherds:

Neve qytetarë a'jemi,  
Malësor emërin e kemi.  
Neve jemi bij bariu,  
S'na trëmp as 'bora as shiu.  
Puna jon' është vetiu,  
Natrya si desh na shiu.

### ALGERIAN (Arabic script)

See *North-West Africa in Arabic (Part One)*

SPECIMEN This passage, in normal unpointed Arabic, is from a story about a king and a madman:

ان بعض الملوك قصد التفرج على المجانين فلما دخل عليهم راي قبيم شابا حسن الهيئة نظيف الصورة.

### ALUR (Latin script)

Alur is spoken by people of Uganda living to the north of Lake Albert.

It is written in Latin script, with one special letter: η. SPECIMEN This is a passage from a children's story:

Akuru githeyo rimo ku katalango. I oro accel ceg oryeny mandha. Nam ocwiny ceke.

### AMHARIC (Amharic script)

The Amharic language is of Semitic origin and is distantly related to Arabic. It takes its name from the district of Amhara in Ethiopia, where it was originally spoken. With the ascendancy of the Shoan race, whose mother tongue it is, it has become the most widely used language of Ethiopia. The number of speakers of Amharic is variously estimated at from three to five millions.

SPECIMEN These sentences are from the invocation at the beginning of the life of a saint:

በሰሙ ፡ እግዚአብሔር ፡ ሥሉሰ ፡ በኢተሐልዮ ፤ ዋሕድ ፡ በኢተሐልዮ ፤ ከድኑ ፡ ሙለኮት ፡ እምተሐልዮ ፤ ወሃሩሙ ፡ ሥልጣን ፡ እምተሙኣሰዮ ፤

### ANGLO-NORMAN (Latin script)

Anglo-Norman was the language brought into England by the Normans and there so modified as to need a new name to distinguish it from the Norman spoken on the Continent.

The alphabet used was the same as for French, but without accents.

SPECIMEN These are the opening lines from a poem by Jordan Fantosme, a schoolmaster at Westminster, early in the 13th century:

Li sires d'Engleterre aq en sun cuer pesance  
Quant sun fiz le guerreic, qu'il nurri ad d'enfance,  
E veit que cil de Flandres l'unt mis en errance.

### ANGLO-SAXON (Latin script)

Anglo-Saxon is the name given to the vernacular Germanic language in use in England up to the Norman

Conquest. It is the ancestor of modern English and for this reason is also called Old English.

The modern practice is to use the Latin alphabet for setting Anglo-Saxon, with the addition of two special letters þ and ð and the diphthongs æ œ. It is also customary to indicate long vowels by a horizontal line above them: ā ē ī ō ū æ œ.

**SPECIMEN** This is part of a description of England from Bede's *Historia Ecclesiastica*:

Breoton is gārseges ealand, ðæt wæs iūgeāra Albion hāten: is geseted betwyh norðdæle and westdæle, Germānie and Gallie and Hispānie, þam mæstum dælum Eurōpe, mycele fære ongegen.

### ARABIC (Arabic script)

Arabic is a Semitic language, related to Hebrew. It is the sacred tongue of Islam and is spoken, in a great variety of dialects, by some 37 million people in North Africa and the Near East.

In spite of the differences that have arisen in the spoken forms of Arabic, the literary language is the same throughout the Mohammedan countries. Thus, though a man from Iraq might have some difficulty in speaking to an Egyptian, their written language would be the same.

The Arabic language is unique in that it was, from its first recorded use, fully developed and richly endowed. The language of the poets of the 6th century and of the Koran, which was composed in the 7th century, is substantially that used today. From the first it showed its characteristic wealth of expression and flexibility of form, and from that day it has undergone no important change beyond essential additions to its vocabulary.

Its resources have proved fully adequate for such diverse purposes as the wonderful stories of *The Arabian Nights*, the philosophy of Averroes, the geographical work of Ibn Batuta, the history of Idrisi, translations from Aristotle, and the modern journals of Cairo and Baghdad.

**SPECIMEN** This is the opening *surah* or chapter of the Koran, in the pointed form in which it is always printed.

اَلْحَمْدُ لِلّٰهِ رَبِّ الْعَالَمِينَ ۝ اَلرَّحْمٰنِ الرَّحِيْمِ ۝ مَالِكِ يَوْمِ  
الْحِسَابِ ۝ اِنَّ اِنَّا لَنَرٰكَ فَنسُوْنَهُ ۝ اَفَدَلًا الْفَرِیْطَ ۝  
اَلَسُنَّتِیْمِ ۝ صِرَاطًا الَّذِیْنَ اَنْشَأْتَ عَلَیْهِمْ ۝ غَیْرَ  
الْمَغْضُوْبِ عَلَیْهِمْ ۝ وَلَا الضَّالِّیْنَ ۝

### ARAMAIC (Hebrew script)

Hebrew script has been used to write various Semitic languages, including Aramaic, which is closely related to Hebrew. Aramaic in various forms was widely used as a *lingua franca* throughout the Near East from the 5th century B.C. to the 7th century A.D., and it is still spoken by small communities in the Lebanon and Mesopotamia. Biblical Aramaic, in which some parts of the Bible were originally written, is sometimes called Chaldean.

**SPECIMEN** Here is a message of greeting, written in the 5th century (possibly in Babylon) from a prince to his king, with the hope that the king may enjoy peace:

ארשם על נתחזור כנוסרם וכוותה וכעת  
פטוסרי שמה ורשבר עלים

### ARMENIAN (Armenian script)

See *Armenian (Part One)*

**SPECIMEN** This is part of a folk-tale of Abu-Jasana:

Մի օր Աբու-Շասանը երկրոր ու բարակ մտածելուց  
Հետ վերջապես վճոց բարկիք գնալ:

— Համ շատ ժամանակ է չես երել, Համ էլ էժամ  
է: Կարծեմ մի քիչ էլ կկտտա պիտի լինես:

Հանդիլիս ծանօթեցրից մեկը նկատեց ու սասմլ.

### ARMORICAN (Latin script)

See *Breton*

### ASSAMESE (Bengali script)

ASSAMESE, the language of Assam, is spoken by some five million people. It is related to Bengali and is commonly written in Bengali script.

Literary work has survived from the 13th century, but Assamese literature came into its own under the leadership of Sankardev in the 15th and 16th centuries. During the second half of the 19th century there was a literary renaissance in which Hemchandra Barua and Guna-bhiram Barua were outstanding figures. There is now a flourishing literary output and a high level of journalism. As elsewhere, the older writers have shown great interest in folk literature, while younger writers strive for greater realism and social consciousness.

**SPECIMEN** This extract from court records (taken from Grierson's *Linguistic Survey of India*) is part of the statement of an accused person:

এই বকর্দ্বা মিচা । মই তাব যবত কোনো বস্ত চোব কবিবনৈ  
যোবা নাহিলৌ । কথা হযে ।

### AZTEC (Latin script)

Aztec, also called Nahuatl, was the language of the ancient civilization of Mexico. It is now extinct, but there are extensive remains in existence, largely due to Bernardino de Sahagun, a Spanish monk of the 16th century who wrote down long recitations of both prose and verse in the Latin alphabet.

**SPECIMEN** Here is a short poem (The river passes, passes and never ceases; the wind passes, passes and never ceases; life passes and never returns):

Danthe togui togui  
hin hambi tegui.  
Ndahi togui togui  
hin hambi tegui.  
Nbui togui. . . .  
hin hambi pengui.

**BAMBARA** (Latin script)

Bambara is an important member of the Mandingo group of languages. It is spoken in former French territories in West Africa.

It is written in Latin script with the following accented letters: ä ë ì ö ü.

**SPECIMEN** (The sea was getting rough because of a strong wind that was blowing):

Dyi kuru bau wulila flye ba fe.

**BAMILEKE** (Latin script)

Bamileke is classified as a semi-Bantu language. It has affinities with that language family but does not wholly belong to it.

It is spoken in the West Cameroons and is written in Latin script with the following extra letters: á à ä é ê ë ò ù ù ñ.

**SPECIMEN** This passage is from a translation of St. Mark's Gospel:

A bë tum ntse, njen kamenyam fagte, Ywet swê\* num i z'a na' mfute boñsañ na.

**BASQUE** (Latin script)

Basque is a unique language in that it has no resemblance to or relation with any known language. It is spoken by about 600,000 people in Spain and France, on both sides of the Pyrenees.

There seems little agreement or uniformity about the way in which it is written. The Latin alphabet is used, and the following accented letters have been employed: á ä é ê ì ñ ö ù ñ ü.

**SPECIMEN** The following sentence means that, though you say many things, you do not speak the truth:

Zuk gauz asko esaten dituzu; baña ez dezu egi bat esaten.

**BENGALI** (Bengali script)

Sanskrit was the language of scholars and priests. In the mouths of men in the street is assumed a less elaborate form known as Prakrit. Prakrit took on different characteristics in different regions, and in Bengal it became the source of the Bengali language, now the mother tongue of some sixty million people in India and East Pakistan.

The Bengali language has had an independent existence for a thousand years. Some verses have been discovered which were written in the 10th century, and these represent the earliest known specimens of Bengali writing.

With the development of Bengal, and in particular the city of Calcutta, as a centre of culture, Bengali has become one of the most important languages in India. It owes much of its present eminence to the great poet Rabindranath Tagore, the first Indian writer of modern times to achieve international fame.

**SPECIMEN** This is part of an address on university education by Rabindranath Tagore:

একদিন অপেক্ষাকৃত অল্পবয়সে যখন আমার শক্তি ছিল তখন কখনো কখনো ইংরেজি সাহিত্য নুখে মুখে বাংলা করে জনিয়েছি আমার পুঁজতার ইংরেজি জানতেন সবাই।

**BIHARI** (Devanagari script)

Bihari, the language of the province of Bihar in northern India, is the mother tongue of more than 37 million people. It is an Indo-European language, descended from Sanskrit by way of Prakrit, and is related to most European languages. It is written in Devanagari script.

**SPECIMEN** There are various dialects of Bihari with didyent names. This specimen is from a poem by Vidyapati Thakur and is in the Maithili dialect:

कामिनि करएँ मिनान । हेरते हूद्व हएँ यचनाने ॥

**BINI** (Latin script)

Bini, or Edo, is spoken in the west of Nigeria.

It is written in Latin script, with two extra letters: e ɔ.

**SPECIMEN** Here is a sentence from a children's story:

Okhuo ɔman-ɔn ɔkpa nɔn ɔfɛse ke gha ri che.

**BIRWA** (Latin script)

Birwa is a member of the Sotho group of the Bantu family of languages and is spoken in Southern Rhodesia.

It is written in Latin script with the following extra letters: ú û ß ɳ ʒ ʒ.

**SPECIMEN** This is a verse from a hymn:

Kiuemula vɛkati tsotha  
Mata ahaɔ ahenya Satane;  
Haho muthu yatswanaho wena;  
Thavoŋ leuvthukuj, lala hunna!

**BRAZILIAN** (Latin script)

See *Portuguese*

**BRETON** (Latin script)

Breton is a Celtic language, also known as Armorican, spoken by over a million inhabitants of Brittany. It has a considerable literature.

The Latin alphabet is used. There are no fixed rules about accents, but the following are found: ä é ê ì ñ ö ü.

**SPECIMEN** In the following lines Prosper Proux, the leading Breton poet, warns young maids that his book is not for them:

Seurzet fur a zantel, a voar quer brao ruya,  
Evit o silvidiguës, serrit prest al levr-ma.

**BULGARIAN** (Cyrillic script)

The Bulgarian language has a long and interesting history, since it is closely related to the language in which the earliest Slavonic documents were written in the 10th and 11th centuries A.D. The latest event in its development was the spelling reform introduced in 1945. It is now the language of some seven million people.

The Bulgarian alphabet differs from the Russian (see *Cyrillic [Part One]*) in omitting ъ, ѣ and ь. The following

additional characters are found in books printed prior to 1945: ж, ъ.

**SPECIMEN** These lines from a poem by Yavorov are in the new orthography:

От заник-слънце озарени,  
Алеят морски ширини;  
В игра стихийна уморени,  
Почиват яростни вълни...  
И корабът се носи леко  
С попълни тихи ветрове —  
И чезнете в мъгли далеко  
Вий родни брегове.

### BUNYORO (Latin script)

Bunyoro is the language of a Bantu tribe which is widespread in Uganda.

It is written in Latin script without additions.

**SPECIMEN** This passage is taken from a description of the weather:

Ebucu bihindukahinduka muno. Omu kasumi kake muno, okusisana kwabyo oku obaire noorolerra kuba kuhingsiibwe, noosisa nk'arukurora ekisisani ekindi.

### BURMESE (Burmese script)

The Burmese language is spoken by some ten million people. It is related to Tibetan and Chinese, and like those languages it is largely monosyllabic and depends upon tone to distinguish meaning.

There is a considerable literature in the language, much of it influenced by the long association with Buddhism.

**SPECIMEN** (When Burmese elders talk together, they like to quote proverbs):

မြန်မာ လူကြီးချင်း စကား ပြောသည့် အခါမှာ စကားပုံများကို ထည့် ပြောတတ်သည်။

### BYELORUSSIAN (Cyrillic script)

The Byelorussian language, spoken in the region of Vitebsk, Mogilev, Minsk and Mozir, is closely akin to Ukrainian and Russian. As with Ukrainian, the language took on a fresh lease of life after 1905, and since then many books have been written in the language. It is now the native tongue of 6½ million people.

For the Byelorussian alphabet, see *Cyrillic (Part One)*.

**SPECIMEN** This passage is taken from a folk-tale:

У нашей вѣсны жмѹ адзін вядзьмар. Яго зналі не толькі ў суседніх сѣлах, але мо' ў цэлай акрузе. Гэта быў такі вялікі вядзьмар, што ён усё ведаў.

### CANARESE (Kannada script)

See *Kannada (Parts One and Two)*

### CATALAN (Latin script)

Catalan is a Romance language, akin to Provençal, spoken in Catalonia, Valencia, Andorra, the Balearic Islands, and Corbières in southern France, by approximately six million people.

The Latin alphabet is used with the following accented letters: à ç è é í ò ó ú.

**SPECIMEN** Here is a song to the lady of Aragon:

A Aragó n'hi ha una dama  
que és bonica com un sol.  
Té la cabellera rossa,  
li arriba fins als talons.

### CHIBEMBA (Latin script)

More than half the people of Northern Rhodesia understand Chibemba, the language of the Bemba tribe and its neighbours. North of Lusaka it is spoken by most Africans in contact with Europeans.

It is written in Latin script without additions.

**SPECIMEN** This is a sentence from a story about a young boy's adventures:

Wongo ubulanda bwamwikata sana, pantu Anatole alimwishiba nokuti muntu umusuma sana.

### CHINESE (ROMANIZED) (Latin script)

There have been many attempts to reduce the Chinese language to Latin characters. The most widely known is that of Sir Thomas Wade (1819-95), who was the first Professor of Chinese at Cambridge University. His system uses the Latin alphabet, with superior numerals after each syllable to indicate the tones.

A much more recent system is known as the Yale Romanization, which was created in 1943 largely through the efforts of Professor George A. Kennedy of Yale University. Two further attempts at romanization were made by the Chinese themselves, namely the *Gwoyeu Romatzyh* (1926) and *Sin Wenz* (1931).

**SPECIMEN** The following sentence, in each of the four systems, records the fact that in the city of Bilbao two places have caught fire:

Pi<sup>4</sup>-erh<sup>2</sup>-pa<sup>1</sup>-o<sup>4</sup> ch'eng<sup>2</sup>-li<sup>2</sup> yu<sup>2</sup> liang<sup>2</sup>-ch'u<sup>4</sup> shih<sup>1</sup>-huo<sup>2</sup>.

(*Wade Romanization*)

Bilbao chernglii yeou leangchuh shyhuoo. (*Gwoyeu Romatzyh*)

Bilbao chengli yōu lyāngchū shrhūō. (*Yale Romanization*)

Bilbao chengli iou liangchu shxuo. (*Sin Wenz*)

### CINYANJA (NYANJA) (Latin script)

Cinyanja, or Nyanja, is spoken in Northern Rhodesia and throughout the greater part of Nyasaland.

It is written in Latin script without additions.

**SPECIMEN** This is a sentence from a story about a boy:  
Anthu onse pamodzi ndi ana ao anapita kukaceka nyamaayo.

Among the Mananja or Southern Nyanja people of

Nyasaland, the language is usually written with one extra letter: ŋ.

**SPECIMEN** From a description of the initiation ceremony:

Mwana wamwamuna wazaka khumi wafikila pamsinkhu wobvinidwa.  
M'mudzi mukakhala ana osabvinidwa, makolo ao amapangana zakuitana mmichila kudzabvinila ana.

#### CORNISH (Latin script)

Cornish is an extinct Celtic language of the Brythonic group, related to Welsh and Breton.

Spelling has never been standardized, but one suggestion is that the complete Latin alphabet be used with the following accented letters: ã ë ì ö ü ŷ.

**SPECIMEN** In this sentence John Bosc (c. 1710) regrets that the ancient language of Cornwall should be lost:

Ny a-ŷl gwelias hag aswon bōs an Tavas Kernewek tavas cōth, ha trūeth yu y-vōs kellys.

#### CREOLE (Latin script)

For mention of Spanish, French, Portuguese, Dutch and English Creole, see section in *English*.

#### CZECH (Latin script)

The Czech language, the westernmost member of the Slavonic linguistic family, is spoken and read by some 8½ million people, mainly in the western and central parts of Czechoslovakia. The Czechs possess a rich literature which had its beginnings in the 14th century. The diacritical marks, which are a feature of the Czech alphabet, were introduced to represent sounds not existing in Latin by the reformer Jan Hus in the year 1441.

The Latin alphabet is used, but *f g q w x* occur only in words of foreign origin. The following accented and special letters are used: á ě é é ě í ě ň ó ř š ť ú ů ý ž; in capital letters *Ď* is used for *d* and *Ť* for *t*.

**SPECIMEN** This passage is from an account of Czech music:

Je málo zemí, kde by lid měl tak přirozenou náklonnost k hudbě jako v Čechách. Češové měli vždy pověst dobrých hudebníků, a velcí skladatelé jako Mozart a Weber rádi pobývali v Praze.

#### DAGBANI (Latin script)

The Dagbani or Dagbane language is spoken in the Upper Volta territory.

It is written in Latin script with the following extra letters: ɛ ɔ ɲ ʃ ʎ ʒ.

**SPECIMEN** (And he went away from there and came into his own country):

O yĩsĩ nimaani, n-kulĩ omanɲaɲa ya.

#### DANISH (Latin script)

Danish is a Scandinavian language, closely akin to Norwegian and Swedish, and is spoken by some four million people in Denmark. Until 1905 it was also the official language of Norway, but when Norway seceded from Sweden in that year it also adopted its own language, based upon the spoken form or Landsmål. Many early Danish books were printed in Gothic type.

The Danes have preserved a wealth of oral literature from early times and of mediaeval ballads from the 12th century. Popular literature began in the 18th century with Ludwig Holberg. In the 19th century two writers of widely differing outlooks won world-wide attention, Hans Christian Andersen with his fairy-tales and Sören Kierkegaard with his works on religion and philosophy.

The following accented letters and diphthong are used: å æ ø.

**SPECIMEN** Hans Christian Andersen visited Charles Dickens in 1857 and wrote a description of his visit:

Uden al øvelse i tidligere at tale engelsk og høre det tale, forstod jeg fra første øjeblik næsten alt, når Dickens talte til mig; kom der mig noget vanskeligt, da gengav han det i en ny sætning; ingen var hurtigere til at forstå mig end han.

#### DINKA (Latin script)

The Dinka people live in Sudan, on the upper Nile, where they number almost a million.

Their language, which has been written in Arabic script, is now usually written in Latin script, using the special letters: ŋ ɔ.

**SPECIMEN** Here is a sentence from a folk-tale:  
Ne akol cien, ke koc kedhia aabi ceɲden war.

#### DUALA (Latin script)

Duala is a Bantu language, spoken in the East Cameroons.

It is written in Latin script with the following extra letters: á ä ë ɛ ɪ ɔ ɔ̃ ń ñ.

**SPECIMEN** This extract is from a book of Bible stories:  
Nde mu m'bwaa mabu mɛng mu si ta mu kɔbɛ o manglea ma bwaa; mandabo mabu ma ta nde ma longabɛ na mɔni ba dumbannɔ batakedi o ekombo.

#### DUTCH (Latin script)

Dutch is the language of the eleven million inhabitants of the Netherlands, and of at least a further million Dutchmen elsewhere. Dutch literature has flourished since the 13th century. Many early works were translations, but since the time of Pieter Hooft and Joost van den Vondel, both younger contemporaries of Shakespeare and Cervantes, the native genius has blossomed. From the time of Erasmus the climate of Holland has been friendly

to scholars, and many important works of scholarship have been written in Dutch.

The Latin alphabet is used, but *g x y* appear only in foreign words, though *y* used at one time to be written for *ij*. Accents are normally used as little as possible in modern Dutch, but acute and grave accents are sometimes employed as indications of stress where wrong pronunciation might result. The diacresis is also used, as in *industriële*.

**SPECIMEN** This is the beginning of a poem on the Dutch language by J. P. Heye:

Neerland! was uw arm van staal,  
't Hart was zacht en mild en goedig;  
Zo ook huwt zich, vroom en moedig,  
Kraacht en fierheid in uw Taal!

### EDO (Latin script)

See *Bini*

### EFIK (Latin script)

Efik, of which Ibibio is a dialect, is spoken in the Calabar district of southern Nigeria.

It is written in Latin script with the following extra letters: *ɔ ŋ*.

**SPECIMEN** (It is thus that in Efik land they say: "He who hates another man's children has none of his own."):

Do ke Efik edoho ete: Asua eyen-owo inyeneke esie.

### ENGLISH (Latin script)

The English language is Germanic in origin and is derived chiefly from the Jutes, Angles and Saxons who invaded England in the 5th century A.D. The earlier Celtic-speaking inhabitants were driven to the western fastnesses of Cornwall and Wales. (See *Cornish* and *Welsh*.) The Jutes, who came first, settled in Kent, southern Hampshire and the Isle of Wight; the Saxons occupied the rest of England south of the Thames; and the Angles settled in regions north of the Thames, which were divided by the Humber into Mercia and Northumbria. A different form of language was spoken by each, the differences being preserved in dialects and place-names. The invasions of the Vikings in the 8th and 9th centuries added a further strain to the country's inhabitants and their language. In the south King Alfred (871-899) repelled the invaders from his kingdom of Wessex, and in the *Anglo-Saxon Chronicle* contributed to a vernacular prose literature that has had a continuous history for over a thousand years. (See *Anglo-Saxon*.)

Under Alfred's grandson, Aedred, England was first united, and under Canute it achieved a greater degree of political unity. This in turn led to the formation of a national written language.

The Norman Conquest (1066) introduced a further linguistic strain. For the next three centuries the kings of

England spoke French, and this was the official language of the law until 1731. (See *Anglo-Norman*.) One result was a tremendous enrichment of the English vocabulary by the introduction of numerous words from French.

Meanwhile, the English language had made progress. In 1362 Parliament was first opened in English, and about the same time Chaucer, then a youth of 21, was beginning to write the poems that were to put England on the literary map of Europe. He used a vocabulary of just over eight thousand words, of which about half were of Romance origin derived from French and Latin. In this he set a pattern for the future, for the present-day vocabulary of English is roughly half Germanic and half Romance. Chaucer established the tradition of applying English to literary purposes, in which he was followed by Tyndale and Wycliffe, Spenser and Marlowe. In the hands of William Shakespeare (1564-1616) it attained a power and beauty never excelled by any author in any language.

When Shakespeare wrote, English was the native language of some five million people. As a result of the discovery of the New World and its subsequent settlement by Europeans, several European languages took the first steps towards becoming world-wide. As Spanish and Portuguese established themselves as the main languages of South America, so English became the main language of North America. At a later date, explorers and settlers spread the use of English to other parts of the world, notably to Australia and New Zealand. Today, as a result, it is spoken by over 250 million people and is understood in some measure by a further 350 million, representing nearly one in four of the inhabitants of the earth.

### Basic English

Basic English, the invention of C. K. Ogden, is a simplified form of the English language. In it the vocabulary is reduced to 850 essential words which are able to give the sense of anything that may be said in English. Its purpose is to provide a form of the language which would at once be understood by all who know English and could be easily learned as an auxiliary language by all who do not already speak English. It is intended, in fact, as an international means of communication and as a step towards the possible adoption of English as the universal language of the world.

**SPECIMEN** The following familiar passage from the New Testament is in Basic English:

Happy are those who are sad: for they will be comforted.

Happy are the gentle: for the earth will be their heritage.

Happy are those whose heart's desire is for righteousness: for they will have their desire.

Happy are those who have mercy: for they will be given mercy.

**Dialects**

A dialect is a form of language spoken in a certain locality showing sufficient differences to be regarded as a distinct entity, yet not enough to be regarded as a distinct language. There are many dialects of English. The poems of Burns in the Scottish, of Tennyson in the Lincolnshire, and of Barnes in the Dorset dialect are well-known literary examples. There is a distinct Australian dialect, a New Zealand dialect and a variety of American dialects.

**SPECIMEN** Here is part of a poem by William Barnes:

Sweet Linda Deïne do match the skies  
 Wit' sheenèn blue o' glisnèn eyes,  
 An' fealrest blossoms do but show  
 Her forehead's white, an' feâce's glow.

**Creole**

The word *Creole* was originally used in the 16th century to denote persons in the West Indies of Spanish parentage. The word was later applied to languages founded on French and Spanish usually spoken by natives in tropical countries, and it is now used for debased forms of other languages, including English.

The original French Creole is spoken in Haiti and Mauritius, and three forms of Spanish Creole are spoken in the Philippines. Dutch Creole is heard in Surinam, and Portuguese Creole in the Cape Verde Islands.

There are various forms of English Creole. One is spoken by the natives of Jamaica. Another is known as Krio, which is the mother tongue of some 25,000 inhabitants of Freetown and of the surrounding villages in Sierra Leone. It is not recognized officially for educational purposes and has no standard written form. As with other Creole languages, when its speakers become literate, they are literate in the mother tongue, in this case English.

**Pidgin English**

There is an important difference between Creole and Pidgin English. Creole is the native language, and often the only language, of those who speak it, whereas Pidgin English is a secondary language acquired for the purpose of communication with speakers of other languages, not necessarily English. The word *pidgin* itself is said to be a Chinese corruption of the word *business*.

**SPECIMEN** This specimen of Beach-la-Mar (also known as Beach-le-mar, Beche de mer), the trade speech of the Western Pacific, is fairly typical:

Altogether you boy belong Solomon you no savvy  
 white man. Me fella me savvy him. Me savvy talk along  
 white man.

**ESKIMO (Latin and Cyrillic scripts)**

The different Eskimo dialects that occur in Greenland, Alaska, Labrador and Northern Canada have been

recorded in a Latin alphabet that does not use *c q x z*, but requires the following accented letters: *â ë î ö ö î*.

A much more elaborate phonetic alphabet has also been compiled for these dialects by the Bureau of American Ethnology.

Russian scholars have applied the Cyrillic alphabet to recording the language of the Eskimos in the Aleutian Islands.

**SPECIMEN (Do you speak Eskimo?):**

Inuk-tut-okar-rung-narpunga?

**ESPERANTO (Latin script)**

Esperanto is an artificial language, devised by Dr. L. L. Zamenhof in 1887 and claimed to be spoken by several million people throughout the world.

It uses the Latin alphabet, with the exception of *q w x y*, and also the following accented letters: *ĉ ĝ ĵ ŝ ŭ*. In addition *h* is sometimes used but is generally replaced by *k*.

**SPECIMEN**

La inteligenta persono lernas la interlingvon Esperanto rapide kaj facile. Esperanto estas la moderna, kultura, neŭtrala lingvo por ĝenerala interkomunikado. La interlingvo estas simpla, fleksebla, praktika solvo de la problemoj de globa interkompreno.

**ESTONIAN (Latin script)**

The Estonian language, which is akin to Finnish, is today spoken by over a million people in the Estonian Soviet Socialist Republic and by about 200,000 people in other parts of the U.S.S.R.

The Latin alphabet is used but *c f q w x y z* occur only in foreign words. The following accented letters are used: *ä ö õ š ž*.

**SPECIMEN** This is the beginning of a tale of the year 1343:

Kevade tuli, 1343-nda aasta kevade. Jürrikuu näitas, nagu kõik jürrikuud, ka sel aastal oma sagedasti muutuvat nägu.

**EWE (Latin script)**

Ewe (or Ewe) is one of the most important of the languages spoken in Ghana, where it is widely used and understood.

It is written in Latin script with the following special letters: *â ð è é f ŋ ó ò ð ŋ ó*.

**SPECIMEN** This is part of a song which begins with the line, "What happened to a whale is best known to the whalers":

Afiké boso tsi, asabudaduawo le nyua lo!  
 Xe-dze-u-dzi-novi, Agamado  
 Sese-mekana-novi, didi-mekaláwo fe vi,  
 Nu-na-'mea lano! menyé lá nœ yeno  
 a agbe va dzia lá tsui o dâ?  
 Menyé lá tsuiwo ta dzo gladza nao o dzâ?

**FANAGALO** (Latin script)

Fanagalo is an artificial language widely spoken in South Africa where peoples of different nationalities and tongues have to work together. It is, for example, used as a means of intercommunication among the many tribes and races in the Johannesburg goldfields. It is said to be the most easily learned language in the world, and it is made up of elements from Zulu, Xhosa and other Bantu languages and simplified terms from English and Afrikaans. Its name is derived from the phrase "Kuluma fana ga lo" meaning "talk like this". It is also called Kitchen Kafir, Mine Kafir and Basic Bantu.

It is written in Latin script without additions.

**SPECIMEN** (We shall go to Durban by car):

Tina zo hamba nga lo motokali lapa Tekweni.

**FANG** (Latin script)

Fang is a Bantu language, widely spoken in the territories which used to constitute French Equatorial Africa.

It is written in Latin script with the following extra letters: é è í ò ñ ŋ ŵ ž.

**SPECIMEN** This extract is from a school reader:

Bör abí be ñga buné Jésus ve nyeghe nye. Bevokh kí be ñga vine nye, toge be be ye minlem mibi.

**FANTE** (Latin script)

Fante is a dialect of Twi and is spoken in Ghana, where it is the medium of an important vernacular literature.

It is written in Latin script, using the special letters: ñ é ó.

**SPECIMEN** This is the beginning of the story of the hare and the tortoise:

Annkyere na asoaso nyanee. Otuu mbirka de mbre obotum kor Egya Sisi ne haban mu ha.

**FAROESE** (Latin script)

Faroese (Faeröese) is a Scandinavian language, akin to Icelandic, spoken by the inhabitants of the Faroe Islands and written only since the middle of the 19th century.

The Latin alphabet is used without *c q w x z*, but with the following accented and special letters: á ð í ó ö ú ý and æ.

**SPECIMEN** The following is a translation of *Auld Lang Syne*:

Hvær skuldi gamlar gøtur gloymt  
og gamalt vinalag,  
hvor skuldi gamlar gøtur gloymt  
og mangan góðan dag.

**FINNISH** (Latin script)

Finnish, akin to Estonian and Magyar, is the national language of Finland, where it is spoken by some four million people. (Swedish was the official language of Finland until the latter part of the 19th century.) There is also the Karelian (or Carelian) dialect which is sometimes treated as a separate language.

The Latin alphabet is used, but *b c f q x z* occur only in foreign words. The following accented letters are used: ä ö; å occurs in Swedish words.

**SPECIMEN** The great epic of the Finnish people is the *Kalevala*, of which a verse is given below. It was the *Kalevala* that suggested the metre of *Hiawatha* to Longfellow.

Vaka vanha Väinämöinen  
elelevi aikojansa  
noilla Väinölään ahoilla,  
Kalevalan kankahilla,  
laulelevi virsiänsä,  
laulelevi, taitelevi.

**FLEMISH** (Latin script)

Flemish, a language closely similar to Dutch, is spoken by about 4½ million people in Belgium. In mediaeval times Flemish was the language of the richest literature of the Netherlands, and more recently the rediscovery of mediaeval masterpieces has been an inspiration to authors. Hendrik Conscience created the Flemish novel in the first half of the 19th century, and by the end of the century Flemish literature was of European importance.

It is written with the Latin alphabet and uses the two accented vowels *ë* and *ó*, and *ij*.

**SPECIMEN** Here is a passage about the battle of Waterloo:

Het congres van Wenen was in 1815 nog vergaderd, toen de tijdging kwam dat Napoleon op 26 februari uit het eiland Elba ontsnapt was. De bloedige slag, die op 18 juni 1815 te Waterloo geleverd werd, besliste over het lot van Europa; Napoleon, die een volkmen nederlaag geleden had, vluchtte naar Parijs, waar hij op 22 juni een tweede maal troonsafstond deed.

**FRENCH** (Latin script)

The French language, like other Romance languages, is descended from Latin. Its earliest recorded use as a distinct language was in 842 A.D. It was used during the Middle Ages, without any efforts at standardization, for works of devotion, national epics and courtly romances.

In the 15th century François Villon demonstrated its poetic capabilities and Froissart its use for historical narrative. In the following century Calvin and Rabelais, in their rather different ways, developed it as a medium for prose, and Ronsard for poetry. Montaigne in his

Essays made his individual contribution, and Descartes proclaimed the sovereign authority of reason.

The 16th century saw the beginning of that conscious preoccupation with the language which was to become a permanent feature. Frenchmen still argue about their language in a way no Englishman or German would.

French is the native language of 48 million people in France, a further four million in Belgium and Switzerland and some four million in overseas territories.

Canada was ruled by France from 1608 to 1759, by which time there were some 60,000 French people in Canada, chiefly in the valley of the St. Lawrence. Their descendants have proudly retained their native language, which largely accounts for the fact that close on three million Canadians speak French as their native language. In Quebec and to a lesser extent in other provinces of Canada, there are numerous French schools where all the teaching is in French, and English is studied as a second language. French is one of the two official languages of the country. Deputies to the Federal Parliament in Ottawa may speak in either language, and government publications are issued in both languages. Many periodicals, both commercial and cultural, are published in French.

The following diphthong and accented letters are used: œ à â ç é è ê ë ï î ó ô û ü.

**SPECIMENS** This extract from Descartes illustrates not only the language, but something of the wit and irony that is part of the national character:

Le bon sens est la chose du monde la mieux partagée: car chacun pense en être si bien pourvu, que ceux même qui sont les plus difficiles à contenter en toute autre chose n'ont point coutume d'en désirer plus qu'ils en ont. En quoi il n'est pas vraisemblable que tous se trompent; mais plutôt cela témoigne que la puissance de bien juger, et distinguer le vrai d'avec le faux, qui est proprement ce qu'on nomme le bon sens ou la raison, est naturellement égale en tous les hommes; et ainsi que la diversité de nos opinions ne vient pas de ce que les uns sont plus raisonnables que les autres, mais seulement de ce que nous conduisons nos pensées par diverses voies, et ne considérons pas les mêmes choses.

These "Stanzas to a Marquis" are by Corneille:

Marquis, si mon visage  
A quelques traits un peu vieux,  
Souvenez-vous qu'à mon âge  
Vous ne vaudrez guère mieux.

Le temps aux plus belles choses  
Se plaît à faire affront,  
Et saura faner vos roses  
Comme il a ridé mon front.

Le même cours des planètes  
Règle nos jours et nos nuits,  
On m'a vu ce que vous êtes;  
Vous serez ce que je suis.

### FRIESIAN (Latin script)

Friesian is the language of some 3½ million people in Friesland, which is north of the Zuyder Zee in Holland. It is of special interest as the language closest to English.

The following accented letters are used: â ë ù.

**SPECIMEN** These lines are from a poem called *Tsjuster* (Darkness) by P. Sipma:

Swart, aeklik swart is nou de nacht: gijn brân  
Fen sinnefûr, noch 't sêfte moantsje jiet  
In wit striel oer 't stille gea; men wit  
Fen bou noch greide of mar, fen sê noch strân.

### FULANI (Latin script)

The Fulani are found all over the area from Senegal to the upper Nile. In West Africa they number nearly five million, of whom more than half live in Nigeria. The people call themselves Pullo (in the singular) and Fulbe (in the plural), and their language Fulfulde. It is often called Fula by outsiders.

It is written in Latin script with the following extra letters: ɓ ɗ ǰ ń w ǎ ẽ õ ɛ ɔ.

**SPECIMEN** This extract is from a book about the language:

Fulbe lewɓe wari fã d'õga 'e wẽnu nu, ɓe ta'wi puldebbõ  
'õn mãyi, ɓingẽn mum na wulla 'e da'adã; ɓe 'adi  
ɓijidõ 'õn, ɓe hõri wuro, ɓe hokko puldebbõ na  
wi'ẽ Asẽtu.

### GA (Latin script)

Ga is one of the languages of Ghana, a country which has six major languages and 65 dialects.

It is written in Latin script, using the special letters: á â ɛ ê ɳ ɔ ð ʒ.

**SPECIMEN** This is the beginning of the story of the hare and the tortoise:

Kpenkple hie tɔf. Edzo foi waa ke ba Ataa Ofɔfɔfɔfɔ  
ɳma le fi. Enãa akpokplonto ye gbe le no.

### GAELIC (Latin script)

Gaelic is a Celtic language, sometimes called Scottish Gaelic to distinguish it from Irish Gaelic, to which it is closely allied. It is spoken by perhaps a hundred thousand people in the highlands and islands of Scotland, and it has a considerable literature.

The Latin alphabet is used, with the following accented letters: à é è ì ó ù.

**SPECIMEN** This is part of a fishing story:

Chaidh mi-fhèin agus Iain a dh' iasgach air an abhainn. Dh' èrrich sinn moch anns a' mhadaunn agus ghabh sinn ar tràth-maidne. Bha an latha ciùin le smodan uisge. Chuir sinn na slatan an òrdugh agus dh' iasgach sinn leis an t-sruth.

**GEORGIAN** (Georgian script)

Georgian is the most important of the Caucasian group of languages. It has a literature that goes back to the 10th century, and it is now spoken by over two million people in the Soviet Republic of Georgia.

For the Georgian alphabet, see *Georgian (Part One)*.

**SPECIMEN** This is a folk-tale about a thieving tailor:

ოყო ერთი ქუჩლი მვერვალი. რაც უნდა გამოეცრა,  
უთუოთ ნაკერი უნდა მოეპარნა. ერთ ღამეს  
ნახა სიზმარი, ვითომც ჰირში ამოსლოდა დიდი  
ხე და მას ტოტებზე ევიდა ყველა მისგან  
ნაპარევი ნაკრები.

**GERMAN** (Latin script and Fraktur)

German, a Teutonic language, is the native language of Germany and Austria and part of Switzerland, and it is spoken by some hundred million people.

The first literary document was written about 750 A.D. Mediaeval times saw the beginning of a strong literary tradition, both oral and written. Heroic lays are represented by the *Nibelungenlied*, romantic songs by Walther von der Vogelweide and other noblemen, and traditional stories by the great collection about *Till Eulenspiegel*, which was popular long before it was printed in 1515. Martin Luther, by his translation of the Bible (1522-34), contributed greatly to the development of the German language by imposing upon it a measure of uniformity and by lending impetus to the movement for modernizing the language.

German is now commonly printed in the Latin script (called *Antiqua* by German writers), with the additional letter ß and the *umlaut* over *a o u*, thus *ä ö ü*. (The *umlaut* represents a missing *e* which is occasionally retained, e.g. Mueller and Müller.)

In the past, letter-spacing was used for words to be emphasized (e.g. Alte Liebe rostet nicht!), but there is an increasing tendency to use italic.

**SPECIMEN** This passage is by Wilhelm Grimm and is taken from the prospectus of his German dictionary (1847); the Lighter Capitals of Times New Roman (727) are used:

Das Wörterbuch soll die deutsche Sprache umfassen, wie sie sich in drei Jahrhunderten ausgebildet hat; es beginnt mit Luther und schließt mit Goethe. Zwei solche Männer, welche, wie die Sonne dieses Jahres den edlen Wein, die deutsche Sprache beides feurig und lieblich gemacht haben, stehen mit Recht an dem Eingang und Ausgang. . . In Luther gewann die deutsche Sprache, nachdem sie von der früheren kaum wieder erreichbaren Höhe herabgestiegen war, wieder das Gefühl ihrer angeborenen Kraft. . . Der Stab, mit dem Goethe an den Felsen schlug, ließ eine frische Quelle über die dürren Triften strömen; sie begannen wieder zu grünen, und die Frühlingsblumen der

Dichtung zeigten sich aufs neue. Es ist nicht zu erschöpfen, was er für die Erhebung und Läuterung der Sprache getan hat, nicht mühsam suchend, sondern dem unmittelbaren Drange folgend; der Geist des deutschen Volkes, der sich am klarsten in der Sprache bewährt, hatte bei ihm seine volle Freiheit wieder gefunden.

German is now rarely set in Fraktur, a form of Gothic letter which was in general use until quite recently. There are over sixty Fraktur series available in the range of 'Monotype' faces. Alphabets of one of these (Series 37) are shown below:

ა ბ გ დ ე ვ ზ თ ი კ ლ მ ნ ი პ რ ს თ ვ ზ ა ბ გ დ ე ვ ზ  
ჟ კ ა ბ ც დ ე ფ გ ხ ი ჟ მ ნ ი პ რ ს თ ვ ზ ა ბ გ დ ე ვ ზ

There are two forms of small *s*, *f* and *ß*, the latter being used mainly at the end of words. There are also four ligatures: *th*, *ff*, *ß*, *ß*. As with the Latin script, the *umlaut* is used.

**SPECIMEN** Here is part of a speech from Goethe's *Faust*:

ღ გლძიფ, სერ ნოფ ჰოფენ ზან,  
სუს ძიეფმ შლერ ბეს ჴრტჳუმს აფჳტაუდენ!  
სჳას მან ნიქჳ უეიფ, ბად ებენ ბრანფე მან,  
ლჳბ მან მან უეიფ, ზანნ მან ნიქჳ ბრანფენ.  
დოქჳ სეფ ანს ძიეფრ შტუნდე სჳინეს ჳუტ  
დურქჳ სოქჳენ შრბჳჳინ ნიქჳ ვერჳამმერნ!

**GREEK** (Greek script)

See *Greek in Part One*

Greek is in many ways the most important of all European languages, since we owe so much to it. Great achievements in the fields of literature and philosophy, the arts and the sciences, politics and government, have affected all subsequent European civilization. The New Testament was written in Greek, and the Old Testament in part owes its survival to the Greek translation known as the Septuagint.

Modern Greek differs considerably from classical Greek. It has two written forms, *katharevousa* or purified form which is used for official documents, and *demotiki* or colloquial form which is favoured by poets and modern writers. Modern Greek is spoken by some eight million people, most of whom live in Greece and Cyprus.

**SPECIMEN** The following are the opening lines of Homer's *Iliad* (Achilles' baneful wrath, resounding, O Goddess, that imposes/Infinite sorrows on the Greeks, and many brave souls' d), which are generally considered to be the beginnings of European literature. They may have been composed as early as 1200 B.C. and written down about 750 B.C. They are set in 'Monotype' New Hellenic (192).

Μῆνιν ἄειδε, θεῆς, Πηληϊάδεω Ἀχιλῆος  
οὐλομένην, ἣ μυρ' Ἀχαιοῖς ἄλγε' ἔθηκεν.

Here is a translation into modern Greek, set in 'Monotype' Greek Times Upright (565):

Μοῖσα, τραγοῦδα τὸ θυμὸ τοῦ ἔπακουστοῦ Ἀχιλλῆα,  
τὸν ἔρημο! π' ἔδλους πότισε τοὺς Ἀχαιοὺς φαρμάκια.

**GUJERATHI** (Gujerathi script)

The Gujerathi language, which is one of the Indo-European group of languages, is the mother tongue of the 15 million inhabitants of the extensive region, from which it takes its name, on the west coast of India in the State of Bombay. It is also spoken fairly extensively in West Pakistan (especially in Karachi) by Moslems and Parsees who recently came there from the area around Bombay.

The language has been put to considerable literary use from the 10th century down to our own day. Tarunaprabha, a prose novelist who was contemporary with Chaucer, is only one of the earliest in a long line of accomplished writers.

The modern era may be said to have started with Narmad (1833-86), who attempted and embellished many literary forms, including the first dictionary of the language. During his lifetime and the years that followed many Western influences made themselves felt among Gujerathi writers. The most outstanding work of this period was *Sarasvatichandra* by Govardhanram Tripathi (1855-1907), an epic novel published in four volumes.

More recent literature has been influenced by political events. Mahatma Gandhi wrote his autobiography in Gujerathi, and his influence was widely felt in the literary as well as the political field. More recently still, national independence has proved an inspiration to writers.

**SPECIMEN** This is a passage from a sea story and tells how the narrator left Borneo and ran into a tempest:

સારો પવન ચાલેને પડ્યો અને બોલનારો બેઠાંબી નીકળ્યો  
 વધુ આગળ ચાલનાં તોડાન લાગવા માંડ્યું. કેટલાક દહાણુ લગી  
 તે આમાંથી વહાણુ ઊછાડવાનું સમજ્યું. ઇચ્છે પવન નરમ પડ્યો અને  
 દરિયા સ્તર ધીમો થયો એટલે બહુ ઊંઠે અને એક મછાલે ડબકાં  
 ખાંડો દીધાં.

**HAUSA** (Latin script)

Hausa is the second most widely spoken language in Africa. In Nigeria alone it is probably spoken and understood by ten million people.

It is written in Latin script, using the following special letters: Ɔ ɗ K.

**SPECIMEN** This is part of a Hausa folk-tale:

Wannan wani mutum ke nan, raƙumara ta faɗe. Sai ta kama hanya, tana tafiya. Tana tafe tana figar ganyen itace, tana ci.

**HEBREW** (Hebrew script)

See *Hebrew in Part One*

**SPECIMENS** These are the opening verses of the *Book of Genesis*, set in 'Monotype' Hebrew Sonzino (218):

בְּרֵאשִׁית בְּרָא אֱלֹהִים אֶת הַשָּׁמַיִם וְאֶת הָאָרֶץ׃  
 וְהָאָרֶץ הָיְתָה תוֹהוּ וָבֹהוּ וְחָשֶׁךְ צִלְפָּנִי תְהוֹם׃  
 וְרוּחַ אֱלֹהִים מְרַחֵף עַל-פְּנֵי תְהוֹם׃

Here, in modern Hebrew, which is written with the same characters and vowel signs, is a passage about the State of Israel, set in 'Monotype' Hebrew Sonzino (218):

מְרִיבֵית יִשְׂרָאֵל  
 בְּחֹדֶשׁ מֵאֵי שָׁנָה 1948 חֲסָלָה בְּרִישׁוּמֶיהָ אֶת הַמְּעִרָה  
 שֶׁהָיוּקָה עַל פְּלִשְׁתִּינְיָה. בְּתוֹצְאָהּ מִנְּה הִוּקְמָה  
 הַמְּדִינָה הַיְהוּדִית הַשְּׂלִישִׁית. זֶה הָיָה הַמְּאוֹרֵעַ  
 הַהִיסְטוֹרִי הַגָּדוֹל בְּיוֹמֵר מִזְמַן חֲרָבֵן תְּפִירַת הַשָּׁמַיִם.

The two specimens shown above have all the vowels marked. Hebrew was originally written without vowel marks, and it can still be printed in that way, although it calls for greater familiarity with the language to supply the missing vowels.

The following passage of unpointed Hebrew is taken from a description of the War of the Sons of Light with the Sons of Darkness, found in one of the Dead Sea Scrolls:

אבות הערה שנים והמשים ואת ראשי הכורנים  
 יסרכו אחר כוהן הראש ומשנהו ראשים שנים  
 עשר לדיות משרתים בתמיד לפני אל וראשי  
 מוהשמורת

**HERERO** (Latin script)

Herero is a member of the Bantu language family and is spoken in Damaraland in South-West Africa.

It is written in Latin script without additions.

**SPECIMEN** (Him that cometh to me, I shall in no wise cast out):

Ngu mee ja ku ami, hi mee mu rambere pendje.

**HINDI** (Devanagari script)

Hindi, of which several dialects are found, is a member of the Indo-European group of languages. It is the mother tongue of some two hundred million people and, under the Constitution of India, is its official language.

Its literary beginnings are said to date from the 8th century, and the earliest manuscripts are from the 12th. Much of the earlier literature consists of hymns and religious epics. One work of the 16th century, a *Life of Rama*, is read and recited throughout the north of India, and it has been the Bible of countless millions of people through three centuries.

The modern era of Hindi literature began late in the 18th century. Since then it has expanded and embraced new forms of expression. It received a further impetus upon the achievement of national independence and is

now being developed to meet all the cultural and technical purposes expected of a great modern language.

The modern tendency is to purify the language of Arabic and Persian words and to rely more and more upon Sanskrit sources.

**SPECIMEN** This couplet from a Hindi poem compares worry to a forest fire. Though no smoke is seen it smoulders within the breast:

चिंता ज्वान सरोर बन दावा लगि लगि जाय । प्रयत्न धुषी नहि  
देखिए उर धंतर धुषुवाय ॥

#### HOVA (Latin script)

See *Malagasy*

#### HUNGARIAN (Latin script)

The Hungarian or Magyar language has no obvious connection with any other language in Europe, although it is believed to be related to Finnish. It is spoken by over 13 million people in Hungary and neighbouring countries.

There are works in Hungarian dating back to the 13th century, but the national literature begins in the 16th. An outstanding figure was the poet Bálint Balassi, who was contemporary with Shakespeare. His poetry, and that of other poets down to our own time, reflects the life of action they were forced to follow. Two of the greatest literary figures were two friends, Sándor Petöfi (1823-49) and János Arany (1817-82).

The Latin alphabet is used with the following accented letters: á é í ó ö ő ü ú ű.

**SPECIMEN** In this poem a Hungarian poet, Sándor Petöfi, sings about his native land and its plains:

Itt születtem én ezen a tájon,  
Az alföldi szép nagy rónaságon:  
Ez a város születésem helye,  
Mintha dajkám dalával vón' tele.

#### ICELANDIC (Latin script)

Icelandic, the native language of over 150,000 inhabitants of Iceland, is one of the oldest languages in Europe, for it was in use before the year 1000 A.D. In it were written the great prose-tales, the *sagas*, which in their original form, and in translations by William Morris and others, have been a source of pleasure and inspiration to countless readers.

The Latin alphabet is used with the addition of two special letters Ð ð and Þ þ, the diphthongs æ œ, and the following accented letters: á é í ó ö ú ý. The modified vowel ǿ is used in some old texts.

**SPECIMEN** Here is a passage in modern Icelandic about early Icelandic poems:

Hver sá, er rita vill til hlitar sögu islenzks skáldskapar  
frá upphafi til þessa dags, verður að horfast í augu

við þann ördöguleika, sem á því er, að brúa samleiga  
hið breiða sund á milli fornaldarinnar og nýja tímans -  
sund nálega fjögurra alda.

#### IDOMA (Latin script)

Idoma is spoken in northern Nigeria and is written in Latin script with the following extra characters: ɛ ɔ ɲ.

**SPECIMEN** (The woman said to him: Sir, I see that you are a prophet):

Ọnya aa da ɔ ka: N le je ka a we ɔce omaebe.

#### IGBO (Latin script)

Igbo, or Ibo, is the language of the Ibo people, who number over five million in Nigeria, where they inhabit the rolling forest plains of the south-east.

It is written in Latin script, sometimes with the following special letters: Ọ ọ; and sometimes with these: ɛ ɔ ɲ.

**SPECIMEN** Here is a passage in both orthographies:

Ọ nwere umu ntakiri abuo. Nne umu ntakiri abuo ndi  
a bu nwanyni nke no na-ebi nime osimiri Bingo.

O nwere omọ ntakere aboo. Nne omọ ntakere aboo ndi  
a be nwanyni nke na na-ebi nime osimiri Bingo.

#### INDONESIAN (Latin script)

Since Indonesia became an independent republic in 1954, its official language has been Indonesian or *Bahasa Indonesia*. This is related to Malay on the one hand and to the Polynesian languages on the other.

The republic of Indonesia, which has a total population of 90 million, is made up of a number of islands, of which the largest are Java, Sumatra, Celebes, Borneo, Bali and the Moluccas. These islands have their own dialects (see *Javanese*), but Indonesian is everywhere the official language.

It is written with the Latin script without accents, although the *acute* accent and the inverted *circumflex* are used in words and names of foreign derivation. It has one feature not found in any other language: the plural is often indicated by repeating a word ("kota besar-besar lainnja" means *other big towns*). The repetition is often indicated in printing by the numeral 2, either as a superior or a lining figure ("kota besar<sup>2</sup> lainnja": similarly "kadangz<sup>2</sup>" meaning *from time to time*).

**SPECIMEN** The following passage is from a story in the collection known as *Bahan budiman* (The Clever Parrot):

Ada seorang perempuan anak saudagar didalam negeri Kasam, terlalu baik parasnja. Maka terlalu amat dikasih oleh sauminja, tiada bagainja didalam negeri itu. Maka diambil Allah akan njawanja. Maka sauminja pun daripada sangat kasih tiada dapat bertjerai dengan mat insterinja.

**INTERLINGUA** (Latin script)

Interlingua is an artificial language fostered by the International Auxiliary Language Association, chiefly for the purpose of promoting scientific progress. It uses the Latin alphabet without accents.

**SPECIMEN**

Le modifications de aeroconstruction que has rendite possible le velocitates supersonic recentements attingite per pilotas stato-unitesse esseva dictate per le resultatos de analyses a methodos photographie del undas de compression que se manifesta in le vicinitate del "barriera" sonic.

**IRISH** (Latin script and Gaelic)

Irish (sometimes called Erse) is a Celtic language, from which stem also Scottish Gaelic and Manx. It is spoken by about 250,000 people mostly resident in the western counties of Ireland from Donegal to Kerry. It is also taught in all the schools of the Republic of Ireland. All government notices, bills and acts are published in Gaelic as well as in English, both being recognized as official languages.

The distinctive Gaelic script, of which there are four 'Monotype' faces, is gradually being abandoned in favour of the Latin script.

The Gaelic alphabet comprises 18 letters, no use being made of *j k q r v w x y z* except in foreign words.

ABCDEF GHIJLMNOPRSTU  
abcde fghilmnoprrtu

The following accents may be used: *á ß é ó é ð ß í ñ ó ð ð ú*. The acute accent denotes length, and the superior dot an aspirate value to the consonant. (In the Latin script, *h* must be inserted after the consonant to replace the dot.)

**SPECIMENS** Here is a translation of "This is a Printing Office":

OIFIS CLÓDÓRA AN ÁIC SEO

Croif-bóthar na Sibialcaicta

Dor óim na n-Calaion ar ghéar-ruaict amhrpe

Sgiaí éorcanca na ríinne calma

ar phormhaic bréige

Buan-ruoc na Ceannáicicta

Ar an áic seo amaic ar lán-eice míonn uplaíra

*San baol a báitec rooi tuinn corainn*

*San fail a malouice rooi láim ríribneóra*

*Áé Suícté so buan ar na ríomáid i zcló*

Ar rós éoríreacicta, a éara, acá no íearain

Oirg ClóDóra on áic seo

'Monotype' Colm Cille (121) is equipped with all 26 letters (some in alternative forms) in both roman and

italic and can be used for setting other languages. Here are four lines from Oliver Goldsmith's *The Deserted Village*:

Sunk ere thy bowers in shapeless ruin all,  
And the long grass o'ertops the mouldering wall;  
And trembling, shrinking from the spoiler's hand,  
Far, far away, thy children leave the land.

**ITALIAN** (Latin script)

Italian is a Romance language, descended directly from Latin. Besides being the native language of Italy, it is widely spoken in her former overseas territories, in southern Switzerland and amongst immigrants in various parts of the world, amounting to a total of over sixty million speakers.

In the 14th century Dante and Petrarch in poetry and Boccaccio in prose demonstrated the use of a modern language for literary purpose and so led the way to the development of modern European literature. Italian has ever since been one of the richest literatures in the world.

The Latin alphabet is used, but it should be noted that *j* is no longer used and is replaced by *i*, while *k w x y* occur only in foreign words. There is no established uniformity in the use of accents. The following accented letters are found: à è é ì í î ò ó ú.

**SPECIMENS** In this passage from Dante's *Inferno*, Virgil tells of his meeting with Beatrice:

Io era tra color che son sospesi,  
e donna mi chiamò beata e bella,  
tal che di comandare io la richiesi.  
Lucevan gli occhi suoi più che la stella;  
e comincionmi a dir soave e piana,  
con angelica voce, in sua favella:

This passage is by the great Italian critic, Benedetto Croce:

È fuori dubbio che il D'Annunzio occupa un gran posto nell'anima moderna e che lo occuperà di conseguenza nelle storie che si scriveranno della vita spirituale dei nostri tempi. Intorno a ciò mi parrebbe ozioso disputare, nè so rispondere senza impazienza alla domanda, che così spesso si ode risonare nelle conversazioni: "Credete che il D'Annunzio sia davvero artista?" Credete? E come, di grazia, si farebbe a credere altrimenti?

**JAPANESE (ROMANIZED)** (Latin script)

For a long time it has been possible to write the Japanese language with Latin characters, without the use of symbols or accents. An early form of transcription, known as *Romaji*, was popular among teachers and others, and a number of books were published in it.

In 1937 a more modern system was introduced by the government, and it is known as *kunrei-siki*. Like *Romaji*,

now being developed to meet all the cultural and technical purposes expected of a great modern language.

The modern tendency is to purify the language of Arabic and Persian words and to rely more and more upon Sanskrit sources.

**SPECIMEN** This couplet from a Hindi poem compares worry to a forest fire. Though no smoke is seen it smoulders within the breast:

चिंता ज्वाल सरोर बन दावा लगि लगि जाय । प्रगट धुवाँ नहिं  
देखिए उर अंतर चुपचाप ॥

#### HOVA (Latin script)

See *Malagasy*

#### HUNGARIAN (Latin script)

The Hungarian or Magyar language has no obvious connection with any other language in Europe, although it is believed to be related to Finnish. It is spoken by over 13 million people in Hungary and neighbouring countries.

There are works in Hungarian dating back to the 13th century, but the national literature begins in the 16th. An outstanding figure was the poet Bálint Balassi, who was contemporary with Shakespeare. His poetry, and that of other poets down to our own time, reflects the life of action they were forced to follow. Two of the greatest literary figures were two friends, Sándor Petöfi (1823-49) and János Arany (1817-82).

The Latin alphabet is used with the following accented letters: á é í ó ö ő ú ü ű.

**SPECIMEN** In this poem a Hungarian poet, Sándor Petöfi, sings about his native land and its plains:

Itt születtem én ezen a tájon,  
Az alföldi szép nagy rónaságon:  
Ez a város születésem helye,  
Mintha dajkám dalával vón' tele.

#### ICELANDIC (Latin script)

Icelandic, the native language of over 1,500,000 inhabitants of Iceland, is one of the oldest languages in Europe, for it was in use before the year 1000 A.D. In it were written the great prose-tales, the *sagas*, which in their original form, and in translations by William Morris and others, have been a source of pleasure and inspiration to countless readers.

The Latin alphabet is used with the addition of two special letters Ð ð and Þ þ, the diphthongs æ œ, and the following accented letters: á é í ó ö ú ý. The modified vowel ǰ is used in some old texts.

**SPECIMEN** Here is a passage in modern Icelandic about early Icelandic poems:

Hver sá, er rita vill til hliatar sögu íslenzks skáldskapar  
frá upphafi til þessa dags, verður að horfast í augu

við þann örðugleika, sem á því er, að brúa sémilega  
hið breiða sund á milli fornaldarinnar og nýja tímans –  
sund nálega fjögurra alda.

#### IDOMA (Latin script)

Idoma is spoken in northern Nigeria and is written in Latin script with the following extra characters: ɛ a ŋ.

**SPECIMEN** (The woman said to him: Sir, I see that you are a prophet):

Ŋnya aa da o ka: N le je ka a we œe omaœe.

#### IGBO (Latin script)

Igbo, or Ibo, is the language of the Ibo people, who number over five million in Nigeria, where they inhabit the rolling forest plains of the south-east.

It is written in Latin script, sometimes with the following special letters: Ọ ọ; and sometimes with these: ɛ ɔ ɔ.

**SPECIMEN** Here is a passage in both orthographies:

Ọ nwere umu ntakiri abuọ. Nne umu ntakiri abuọ ndi  
a bu nwanýi nke nọ na-ebi nime osimiri Bingu.

O nwere omo ntakere aboo. Nne omo ntakere aboo ndi  
a be nwanýe nke no na-ebi nime osimiri Bingu.

#### INDONESIAN (Latin script)

Since Indonesia became an independent republic in 1954, its official language has been Indonesian or *Bahasa Indonesia*. This is related to Malay on the one hand and to the Polynesian languages on the other.

The republic of Indonesia, which has a total population of 90 million, is made up of a number of islands, of which the largest are Java, Sumatra, Celebes, Borneo, Bali and the Moluccas. These islands have their own dialects (see *Javanese*), but Indonesian is everywhere the official language.

It is written with the Latin script without accents, although the *acute* accent and the inverted *circumflex* are used in words and names of foreign origin. It has one feature not found in any other language: the plural is often indicated by repeating a word ("kota besar-besar lainnja" means *other big towns*). The repetition is often indicated in printing by the numeral 2, either as a superior or a lining figure ("kota besar<sup>2</sup> lainnja": similarly "kadang<sup>2</sup>" meaning *from time to time*).

**SPECIMEN** The following passage is from a story in the collection known as *Baján budiman* (The Clever Parrot):

Ada seorang perempuan anak saudagar didalam negeri Kasam, terlalu baik parasnja. Maka terlalu amat dikasih oleh sauminja, tiada bagainja didalam negeri itu. Maka diambil Allah akan njawanja. Maka sua-minja pun daripada sangat kasih tiada dapat bertjerai dengan amat isterinja.

**INTERLINGUA** (Latin script)

Interlingua is an artificial language fostered by the International Auxiliary Language Association, chiefly for the purpose of promoting scientific progress. It uses the Latin alphabet without accents.

**SPECIMEN**

Le modificaciones de aeroconstruction que has rendite possibile le velocitates supersonic recentemente attingite per pilotas stato-unitesse esseva dictate per le resultados de analyses a methodos photographie del ondas de compression que se manifesta in le vicinitate del "barriera" sonic.

**IRISH** (Latin script and Gaelic)

Irish (sometimes called Erse) is a Celtic language, from which stem also Scottish Gaelic and Manx. It is spoken by about 250,000 people mostly resident in the western counties of Ireland from Donegal to Kerry. It is also taught in all the schools of the Republic of Ireland. All government notices, bills and acts are published in Gaelic as well as in English, both being recognized as official languages.

The distinctive Gaelic script, of which there are four 'Monotype' faces, is gradually being abandoned in favour of the Latin script.

The Gaelic alphabet comprises 18 letters, no use being made of *j k q v w x y z* except in foreign words.

ABCDEFGHIJKLMNOPQRSTU  
 abcdefghilmnoprrtu

The following accents may be used: *á b é ó é f é í h í ó p r t ú*. The acute accent denotes length, and the superior dot an aspirate value to the consonant. (In the Latin script, *h* must be inserted after the consonant to replace the dot.)

**SPECIMENS** Here is a translation of "This is a Printing Office":

OIFIS CLÓDÓRA AN ÁIT SEO

Cnór-bócar na Sibialtaíca

Dof b'áin na n-ealaíon ar édar-ímaíct amhré

Sgíac éaranta na írinnne calma

ar íoríarínac bhéige

Buan-íoc na Ceannóíca

Ar an áit seo amáe ar lán-eice imíonn uplabra

Son baol a báice faoi éinn copáin

Son faill a malairce faoi láim íarínéara

Áé Suidce go buan ar na íríomáe í gcló

Ar íos éaríreaca, a éara, acá só íearáin

Oirg Clódóra an áit seo

'Monotype' Colm Cille (121) is equipped with all 26 letters (some in alternative forms) in both roman and

italic and can be used for setting other languages. Here are four lines from Oliver Goldsmith's *The Deserted Village*:

Sunk are thy bowers in shapeless ruin all,  
 And the long grass o'ertraps the mouldering wall;  
 And trembling, shrinking from the spoiler's hand,  
 Far, far away, thy children leave the land.

**ITALIAN** (Latin script)

Italian is a Romance language, descended directly from Latin. Besides being the native language of Italy, it is widely spoken in her former overseas territories, in southern Switzerland and amongst immigrants in various parts of the world, amounting to a total of over sixty million speakers.

In the 14th century Dante and Petrarch in poetry and Boccaccio in prose demonstrated the use of a modern language for literary purpose and so led the way to the development of modern European literature. Italian has ever since been one of the richest literatures in the world.

The Latin alphabet is used, but it should be noted that *j* is no longer used and is replaced by *i*, while *k w x y* occur only in foreign words. There is no established uniformity in the use of accents. The following accented letters are found: *à è é í î ì ò ó ù ú*.

**SPECIMENS** In this passage from Dante's *Inferno*, Virgil tells of his meeting with Beatrice:

Io era tra color che son sospesi,  
 e donna mi chiamò beata e bella,  
 tal che di comandare io la richiesi.  
 Lucevan gli occhi suoi più che la stella;  
 e cominciommi a dir soave e piana,  
 con angelica voce, in sua favella:

This passage is by the great Italian critic, Benedetto Croce:

È fuori dubbio che il D'Annunzio occupa un gran posto nell'anima moderna e che lo occuperà di conseguenza nelle storie che si scriveranno della vita spirituale dei nostri tempi. Intorno a ciò mi parrebbe ozioso disputare, nè so rispondere senza impazienza alla domanda, che così spesso si ode risuonare nelle conversazioni: "Credete che il D'Annunzio sia davvero artista?" Credete? E come, di grazia, si farebbe a credere altrimenti?

**JAPANESE (ROMANIZED)** (Latin script)

For a long time it has been possible to write the Japanese language with Latin characters, without the use of symbols or accents. An early form of transcription, known as *Romaji*, was popular among teachers and others, and a number of books were published in it.

In 1937 a more modern system was introduced by the government, and it is known as *kunrei-siki*. Like *Romaji*,

it uses the complete Latin alphabet without accents, and the two systems are very similar.

It is interesting to note that all school children in Japan are taught the Latin alphabet, although they still write their language in the traditional script.

**SPECIMEN** Here, in *Romaji*, is a sentence about volcanoes:

Nihon ni kiken na kwazan ga daibu arimasu.

And here, in *kunrei-siki*, is a sentence about a great earthquake and mountains:

Mukasi oo-zisin ga atte tatimati ano yama ga dekita to yuu kanasi desu.

### JAVANESE (Latin script)

The Javanese language is spoken in Java by some thirty million people. It is more closely related to Indonesian, which is the official language of Java as part of Indonesia, than to the Polynesian languages.

It is written both in Javanese characters and with the Latin alphabet, and both forms of writing are now taught in schools. The following accented and special letters are used in writing Javanese with Latin script: é ê ð (or ð) † (or †).

**SPECIMEN** The following passage tells how the kingdom of Giling Wesi was visited by catastrophe:

Kala semanten nagari ing Giling Wesi kadatengan geger ing ageng. Kajah tjang alit ingkang sami risak, sarta awis teġa, asring wonten grahana srengéngé octawi grahana remboelan, djawah salah mangsa, lindöe kaping pitoe sadinten.

### KANNADA (Kannada script)

See *Kannada (Part One)*

**SPECIMEN** The following comes from an introduction to a book on the grammar of Kannada:

ಸೂಚನೆ—ಹಿಂದಿನ ಭಾವದಲ್ಲಿ ಕ್ರಿಯಾ ಪದದ ಮೂರು ಕಾಲಗಳನ್ನು ಕುರಿತು ಹೇಳಿರುವುದಿಲ್ಲ. ಅವುಗಳ ರೂಪ ಭೇದಗಳನ್ನು ಸಹ ಹೇಳಿರುವುದಿಲ್ಲ. ಇವುಗಳನ್ನು ಚಿವುಟಗ್ರಹಿಸಿದ ಮೇಲೆ ಇಲ್ಲಿ ಬರೆದಿರುವ ವಿಷಯಗಳನ್ನು ಮತ್ತು ಸಾಧನೆಗಳನ್ನೂ ಗ್ರಹಿಸಬೇಕು.

### KARANGA (Latin script)

Karanga is a Bantu language, spoken in Mashonaland.

It is written in Latin script, using the following extra letters: ß ð ŋ ʒ v z.

**SPECIMEN** These lines are from a hymn:

Mukai, mukai, vatendi!  
She wenyu mirirai;  
Nesimba rake roga  
Uarume, simukai!

### KARELIAN (Latin script)

See *Finnish*

### KASHMIRI (Persi-Arabic and Devanagari scripts)

Kashmiri is the language of the valley of the Kashmir, where it is spoken by about a million people. It is Indo-European in origin. There is a scanty literature but a wealth of proverbs and folk-tales.

It is written in three scripts: Persi-Arabic, Devanagari and a local script called Sarada.

**SPECIMENS** These lines from a poem by Mahmud Gami tell how Jacob yearned for Joseph after the death of Rebecca:

قوت كاتياہ كو يقويس — اوسو مشاق تس كن تيس بس

This is part of a folk-tale in Devanagari script:

अकिस महविस् श्रोसि बृह न्यनिवि ।

### KIKUYU (Latin script)

Kikuyu is a Bantu language, spoken by the tribe of that name who inhabit the slopes of Kilimanjaro in Kenya.

It is written in Latin script, using the following special letters: ū.

**SPECIMEN** Here is part of the story of the goat and the wolf:

Mbwe o na ingikurũ ni njũgi-rĩ, kwĩ mahinda mamwe na mamwe cionaga thĩna.

### LAPPISH (Latin and Cyrillic scripts)

Lappish is related to Finnish and is spoken by some 30,000 people mainly in the northern parts of Scandinavia and Finland, and to a much smaller extent in the U.S.S.R.

It is usually written in Latin script, but Russian scholars sometimes transcribe it into Cyrillic.

**SPECIMEN** This Biblical extract is from St. John's Gospel:

Tastko nũ rakisti Ibmil mailmi, atti son addi aidnu riegadam Parnis, amas oktage, kutte sudnji osku, kadutussi šaddat, muttu vai son oũži agalaš aellim.

### LATIN (Latin script)

Latin, though extinct as a spoken language, remains of great importance for several reasons. It gave rise to one of the richest literatures: it was the parent of the Romance languages (Italian, Spanish, Portuguese, French, Roumanian); and it supplied a great number of words to the vocabularies of English and other languages. It is still used as a liturgical language and as a means of communication in the Roman Catholic Church, and it is also widely used for medical, biological and other scientific terminology.

**SPECIMEN** These are the opening lines of Virgil's first eclogue:

Tityre, tu patulae recubans sub tegmine fagi  
silvestrem tenui musam meditaris avena:  
nos patriae finis et dulcia linquimus arva.  
Nos patriam fugimus: tu, Tityre, lentus in umbra  
formosam resonare doces Amaryllida silvas.

**LATVIAN** (Latin script)

Latvian is a Baltic language, also known as Lettish, which is a member of the Indo-European family of languages and thus has affinities with English and other European languages. It is spoken by about two million people in the Latvian Soviet Socialist Republic.

The Latin alphabet is used, but *qxy* occur only in foreign words. The following accented letters are used: ā ē š ģ ī ķ ļ ņ ū ţ ž š ū ž.

**SPECIMEN** Here is a Latvian folk-song, "Singing I was born":

Dziedot dzimu, dziedot augu,  
Dziedot mūžu nodzivoju;  
Dziedot gāja dvēselīte  
Dieva dēla dārzīpā.

**LETTISH** (Latin script)

See *Latvian*

**LIBYAN** (Arabic script)

**SPECIMEN** This is the beginning of a folk-tale:

فيه واحد اسمه مرّا تجيب فيه في التباويت

**LITHUANIAN** (Latin script)

Lithuanian is a Baltic language, similar to Latvian, and like Latvian is a member of the Indo-European family of languages. It is spoken by some three million people in the Lithuanian Soviet Socialist Republic.

The Latin alphabet is used, but *qwx* occur only in words of foreign origin. The following accented letters are used: ą č ė ę š ū ū ž.

**SPECIMEN** This passage is taken from an account of the Lithuanian language:

Dažniausiai dabartinėje lietuvių kalboje pasitaiko inesyvas, reiškias vietą, kurios viduje kas nors yra arba vyksta. Kitaip sakant, toks inesyvas rodo, kad veikėjas arba veiksmas yra ar vyksta tam tikroje, šiuo linksniu reiškiamoje vietoje.

**LOZI** (Latin script)

Lozi is a Bantu language of the Sotho group, spoken in Northern Rhodesia.

It is written in Latin script with the additional letter: ñ.

**SPECIMEN** This passage is from a book of stories:

Zazi le liñwi ne i li mulisezo wa mushimani ya n'a s'a talifile, wa lilimo ze ka ba 15, kamba 16, i li mulisezo wa manamani.

**LUGANDA** (Latin script)

Luganda is spoken by people inhabiting the central part of Uganda between Lake Victoria and Lake Albert.

It is written in Latin script without additions.

**SPECIMEN** Here is a sentence from a folk-tale:

Awo olwatuuka, mu mirembe egy'edda ennyo, ne wabaawo abavubuka bataan abaabeeranga awamu ennaku zonna.

**LUO** (Latin script)

The Luo language is spoken in Kenya and is written in Latin script without additions.

**SPECIMEN** (And five of them were wise, and five were foolish):

To jabich kuomgi ne jo mofuwo, to jabich ne jo mariek.

**LUSATIAN (WENDISH)** (Latin script)

The Lusatians, who are also known as the Lusatian Serbs or Wends, number between one and two hundred thousand. They inhabit a corner of Germany, bordering on Czechoslovakia and Poland on the upper reaches of the River Spree, and they treasure their language which is Slavonic in origin.

The Latin alphabet is used but *qwx* occur only in foreign words. The following accented letters are used: ǒ é č ě ĺ ř ů ó ǖ í ř š ŵ ž ž.

**SPECIMEN** This is part of a folk tale:

Borkač pŕiwjeđe sebi kmótra Hansa na kermušu. Jeho za blido sadžiwiš, počina nože rjeđić a točić. Žona pak pječe w kuchini kurjatko, a po kuchařsku woptawa a woptawa, hač běchu so kurjatka minyle.

**MACEDONIAN** (Cyrillic script)

The Macedonian people have formed part of several countries in the course of their long history, and partly no doubt as a result of this their national identity and their cultural and artistic individuality have often been overlooked. In fact it was not until 1943 that the Macedonian people were recognized as a separate nationality with a language of their own. Since then, many books and periodicals have been published in Macedonian, and the wealth and beauty of their folk-songs and folk-tales have become more widely known.

For the Macedonian alphabet, see *Cyrillic (Part One)*.

**SPECIMEN** These lines are from a folk-song:

Рачај, порачај, бела Бојано:  
сега сум овде, утре ќе одам,  
утре ќе одам туѓа туѓина.  
Рачај, порачај, што да ти пратам,  
ал пари сакам, ал книга сакам?

**MAGYAR** (Latin script)

See *Hungarian*

**MALAGASY** (Latin script)

Malagasy is the name given to the language spoken by the various tribes who inhabit Madagascar. The leading tribe is the Hova, from which the language is sometimes known as the Hova language. Malagasy belongs to the Malay-Polynesian group and has no affinity with the languages of the African mainland. At the last census there were over four million natives speaking Malagasy in the island.

The language was first reduced to writing by English missionaries early in the 19th century. They used the Latin alphabet, with the exceptions of *cquwx*, which

have no corresponding sounds in Malagasy. The following accented letters are used: á é í ó ô ö ý.

**SPECIMEN** This is part of a story about a baby and a water-serpent with seven heads:

Nandéha, hóna, itý izý télo miánaka nankány ány ny ráno lálina, ja ny zaza-kély kósa, hóno, napétrany téo amóron-dráno. Dia niántso ny bíby, hóno, ny zaza ka nanáo hoe: "Izáho natérin'ikáky sy inény hoháninaréo."

#### MALAY (Persi-Arabic and Romanized scripts)

Malay is the mother tongue of the Malay peoples of Malaya and of many inhabitants of neighbouring parts of the mainland, the Riau archipelago south of Singapore, and parts of Sumatra and Borneo. It is also used and understood in the mainland of Indo-China and the Indonesian archipelago. The total number of people speaking it as their mother tongue is about three million, but the number who employ it as a medium of communication probably exceeds ninety million.

It has been employed as a literary language since the 15th century. The Buddhist scriptures have been translated from the Indian languages, Moslem theology from Arabic and Persian, romances from Javanese. In fact, so far it has been largely a literature of adaptation, although there is a considerable heritage of traditional folk tales, poetry and historical works, including the popular verses known as *pantun*.

For centuries Malay has been written in Persi-Arabic characters. Both Arabic and Romanized scripts are now taught in schools, and in time Romanized may displace Arabic in general use.

**SPECIMEN** This is an example of the popular verse known as *pantun*. In translation it runs: Do not handle unripe rice or the brittle stalk will break. Do not follow the impulses of a too young heart or it will long be broken.

جاڤن لوروت فادي مودا  
فجه بانغ كئا تانڤن.  
جاڤن دتورت هاتي مودا  
روسق بناس فتبعق ٢ غن.

Jangan lurut padi muda,  
Pechah batang kena tangan.  
Jangan di-turut hati muda,  
Rosak binasa panjang-panjang.

#### MALAYALAM (Malayalam script)

Malayalam is one of the four main cultivated languages of the Dravidian family, the others being Tamil, Telugu and Kannada. It is the mother tongue of over 13 million people on the west coast of South India in the recently formed state of Kerala. The language is very similar to Tamil, and in fact the two were not distinguished as separate languages until about the 9th century A.D.

Malayalam literature was always more influenced by Sanskrit than was Tamil, but in the 18th century there was a movement, led by Kunchan Nambiyar, to write

poetry in the spoken language rather than in the traditional literary style. In the 20th century, the same trend has been pursued by Venmani Nambudiri. Today there is an extensive literature, in which poetry and realistic fiction predominate.

**SPECIMEN** Here is the opening sentence of a passage about students and their studies:

കളിയിൽമാത്രമല്ല, കാര്യത്തിലും അവരൊട്ടും  
പിന്നോക്കമായിത്തന്നിട്ടു.

#### MALTESE (Latin script)

Maltese is a Semitic language, related to Arabic, with a strong admixture of Italian. It is the official language of the island of Malta, which has a population of over 300,000.

Leading Maltese writers use the Latin alphabet (without *y*) and the following accented letters: ċ ġ ħ ž. Sometimes vowels are accented with the *grave*, *acute* and *circumflex*.

**SPECIMEN** Here, in Maltese, is the injunction to love thy neighbour as thyself:

Hobb il-ġhajrek bhalek innifsek.  
And here is the proverb about the barking dog:  
Kelb nebbieħ ta tibħax minnu.

#### MANX (Latin script)

Manx, the language once spoken on the Isle of Man, is of Celtic origin and related to Irish. It is now almost extinct, but vigorous efforts are being made to revive it.

The Latin alphabet is used, generally without accents.

**SPECIMEN** This verse from an old ballad tells how St. Patrick came to the island:

Eisht haik ayn Pharick Noo nyn mean;  
She dooinney-noo v'ch lane dy artue;  
Dimman eh Mannanan er y tonn,  
As e gh voojnjer lieh-chiart.

#### MAORI (Latin script)

Maori is a Polynesian language, akin to Samoan and Tahitian, and is spoken by descendants of the original inhabitants of New Zealand.

The alphabet consists of 13 letters: the five vowels *a e i o u* and eight consonants *h k m n p r t w*.

**SPECIMEN** This song tells how Hauhao-tu-puni sets off to sea in good weather:

E hora te marino,  
Horahia i waho ra;  
Hei paki haerenga  
Mo Huahao-tu-puni  
Noku te wairangi  
Te whai rangi au  
Te hukanga wai hoe  
Nou a Ahu-rei  
Kai tonu ki te rao  
Ki Ko-hi ra ia  
Marama te titiro.



people. Norwegian literature achieved world-wide recognition with the plays of Ibsen in the 19th century.

The Latin alphabet is used, but *qwxz* occur only in foreign words. The following accented letters are used: á ø; and the diphthong æ.

**SPECIMEN** Here is part of the stage directions at the opening of Act Two of Ibsen's *John Gabriel Borkman*:

Borkman er en middelhøj, fast og kraftig bygget mand oppe i sekstiårene. Fornemt udsende, fint skåret profil, hvasse øjne og gråhvidt, kruset hår og skæg. Han er klædt i en sort, ikke ganske moderne dragt og bærer hvidt halstørklæde. Frida Foldal er en smuk, bleg, femtenårig pige med et noget træet, anstrængt udtryk. Tarvelig pyntet i lys påklædning.

## OLD CHURCH SLAVONIC

(Cyrillic and Glagolitic scripts)

**SPECIMEN** Here is a scriptural passage set in Glagolitic (598):

ДѢНЪ СЪВЪЩЕНОУ СЪВЪЩЕНОУ  
 ДѢНЪ СЪВЪЩЕНОУ СЪВЪЩЕНОУ  
 ДѢНЪ СЪВЪЩЕНОУ СЪВЪЩЕНОУ

**SPECIMEN** Here is a scriptural passage set in Old Bulgarian (597):

Еъ оно вѣѣма прѣде ноанъ крѣститель  
 проповѣдалъ въ поустыни недѣкѣнѣн и  
 гадѣла: Покантеса, прибанжи во са  
 царѣство небесѣне.

## PASHTO (Persi-Arabic script)

See *Afghan (Part Two)*

## PERSIAN (Persi-Arabic script)

Persian literature began with an outburst of poetry in the 9th century. It reached its highest achievement in the work of Firdousi, whose real name was Abul Kasim Hasan. He lived from about 930 to 1020 A.D., and spent some 25 years polishing the couplets of his epic poem *Shah-nama*, celebrating the old Persian rulers. For a thousand years this poem has been cherished as the supreme expression of the national genius.

Second among Persian poets was Hafiz, who died in 1389, but undoubtedly the most celebrated is Omar Khayyam (1050-1132), who was better known in his lifetime as a mathematician. His celebrity amongst English-speaking peoples depends upon Edward FitzGerald's famous paraphrase of the *Rubaiyat*.

In modern times the Press has played an important part in Persian literary history by keeping the language pliant, simple and popular.

**SPECIMEN** This is part of the story of Ali Baba:

و چنينكه على بابا تمام زر جمع كرده بود خيلى ميخواست كه وزنا  
 و بيك كينف با خودش آورده باشد. براى اينكه بتواند آن را بكشد. بعد از

اين كه راجع باين چنين وقت فكر كرده بود بنظرش خوب آمد بيزنل  
 برادرزن خودش بريد تا وزنا بكشد. مصطلق (كه اسم برادر زنش  
 بود) مايل بود باوكينف و وزناها را بنده و على راه رفت زر خود را  
 بكشد.

## POLISH (Latin script)

Polish is a Slavonic language, akin to Russian, and is spoken by over 31 million people in Poland and ten million elsewhere. It has a long literary history, and its records go back to the 13th century. When it came to be printed at the end of the 15th century, most of the accented letters still employed were introduced by the early printers.

The novelist Sienkiewicz (the author of *Quo Vadis*) and the poet Mickiewicz are the best-known Polish authors, but there have been many others.

The Latin alphabet is used, but *qvx* occur only in foreign words. The following accented letters are used: ą ć ł ń ó ś ź ż.

**SPECIMEN** This passage is from a letter by H. Sienkiewicz:

Idę sobie wieczorem nad Wisłą, patrzę: rozbił się galar z jabłkami. Andrusy wylawiają jabłka z wody, a nad brzegami siedzi cała rodzina żydowska w takiej rozpaczy, że nawet nie lamentują, tylko pozalamywali ręce i patrzą na wodę jak posagi.

## PORTUGUESE (Latin script)

Portuguese is a Romance language, being descended from Latin and related to Spanish, French, Italian and Roumanian.

Besides being the native tongue of the eight million inhabitants of Portugal and of over 12 million in Portuguese overseas territories, it is also the language of 72 million people in Brazil. By an agreement signed on 29th December 1943, Portugal and Brazil undertook to use the same orthography. Such differences as exist are in vocabulary and punctuation.

The outstanding figure in Portuguese letters is Camoens, a contemporary of Shakespeare, who wrote *Os Lusíadas*. Brazil is the home of many poets and novelists.

The Latin alphabet is used, but *kwy* occur only in foreign words. The following accented letters are used: à á â ã ç è é ê ì í ð ó ô õ ù.

**SPECIMEN** This is part of the announcement of the agreement on spelling between Portugal and Brazil:

O Governo brasileiro, pela voz do seu Ministro de Educação, Dr. Gustavo Capanema, acaba de declarar, em sessão da Academia Brasileira de Letras de 29 de Janeiro de 1942, que há uma só língua portuguesa no Mundo; que essa é, na sua esplêndida unidade, a língua de Portugal e do Brasil; e que o Governo brasileiro aceita, como cânone ortográfico do idioma único e imortal, o Vocabulário da Academia das Ciências de Lisboa, recentemente publicado.

**PROVENÇAL** (Latin script)

Provençal is a Romance language which was used throughout southern France in the Middle Ages as a unified literary language. It is now represented by a number of local dialects. About 1900 the poet Mistral attempted to reconstruct a literary Provençal.

In its original form the language was written in Latin script without accents.

**SPECIMEN** This extract from the 12th-century *Lives of the Troubadours* tells of William of Gabestaing:

Guillems de Cabestaing si fo us cavalliers de l'encontrada de Rossillon que confina ab Cataloigna et ab Narbones. Mout fo avinens hom de la persona e prezzat d'armas e de cortesia e de servir. Et en la soa encontrada avia una dompna que avia nom ma dompna Soremonda, moiller d'en Raimon de Castel-Roussillon, que era mout gentils e mals e braus e fers e rics et orgoillos.

**PUNJABI** (Gurmukhi, Persi-Arabic and Devanagari scripts)

The Punjabi language is spoken by some 18 million people in the north-western regions of India and Pakistan. It is an Indo-European language, descended from Sanskrit by way of Prakrit. As a language it has been receptive to innovations and has imported more Arabic and Persian words than any other Indian language. Greeks, Turks, Afghans and English have also added to its resources. It is the language of the Sikh religion, as well as being spoken by many Moslems and Hindus.

Among its literary figures have been Baba Farid, who converted millions of Hindus to Islam, and Nanak, the founder of the Sikh religion. More recently, Bhai Vir Singh brought several forms of literature to a high pitch of achievement. Writers like Puran Singh, Daniram Chatrik and Mohan Singh have, on the one hand, imported ideas from other cultures and, on the other, enriched contemporary literature with native traditions.

Punjabi can be written in three scripts: Gurmukhi, Persi-Arabic and Devanagari.

**SPECIMEN** This is part of a traditional poem, set in both Gurmukhi and Persi-Arabic characters (The month of Chait has begun, and showers have fallen. My friends, great is the might of the Sikh government. Kabul and Kandahar trembled before it, and its tents are pitched beyond the Indus):

ਚੜਿਆ ਚੌਤ ਪੁਹਾਰਾ ਯਾਰੋ ਵਡੀ ਸਰਕਾਰ ਯਾਰੋ ਕਾਬੁਲ ਤੋ  
ਕੰਧਾਰ ਤੋਰੇ ਖਤੋ ਅਟਕੇ ਪਾਰ।

چڑھیا چہتر پئی پُہار -  
یارو وڑی ہو اے سؤر -  
دھسکی کاہد ت تفتدار -  
ژیری کھنک انکون پار \*

**RAJASTHANI** (Devanagari script)

The name of Rajasthani was invented to distinguish a group of kindred dialects from their neighbours, Western Hindi on the one hand and Gujerathi on the other. It means the language of Rajasthan, the country of the Rajputs, and it is spoken by 18 million people who inhabit an area of 180,000 square miles in north-western India.

Books printed in this language make use of Devanagari script.

**SPECIMEN** This refrain from a folksong runs, "You are charming when full of opium, my lord. Do drink some wine."

अमलौ-मैँ बाबा लागो म्हारा राज । पीबोमो दादुड़ी ॥

**ROLONG** (Latin script)

Rolong is a Bantu language, spoken in the Orange Free State.

It is written in Latin script with one extra letter: ñ.

**SPECIMEN** This passage is from a translation of St. John's Gospel:

Batho botlhe ba tla itse ha lo le barutwa ba me ka mo, ha lo ratana.

**ROMANY** (Latin script)

Romany (Romani), the language of the gipsies, has been proved to be of Indian origin.

It has been written with the Latin alphabet without accents.

**SPECIMEN** The following lines reported by George Borrow concern a meeting between two strange gipsies:

Coin si deya, coin se dado?  
Pukker mande drey Romanes,  
Te mande pukkeravava tute.

**ROMAUNTSCH** (Latin script)

Romauntsch (Romansch) is a language of Romance origin which occurs in a number of closely related dialects; it is also known as Rhetian, Rhaeto-Romanic and Ladin. It is recognized as a national but not as an official language of Switzerland, where it is spoken by some 50,000 people, and it is also the native tongue of 820,000 people in nearby parts of Italy.

It is written with the Latin alphabet, without accents except the diaeresis.

**SPECIMEN** This is a verse from a national song:

In guerra valurus  
In pasch religius  
In dits e fats,  
Usche onur rendain  
Al nom cha nus portain  
Liberans cunservain  
Sco'ls antenats.

## RONGA (Latin script)

Ronga is a Bantu language, spoken in Portuguese East Africa.

It is written in Latin script without additions.

**SPECIMEN** This passage is from a children's primer:

Sofia a loloha, a nga tiri. A tota mombo hi powa ja ntlohe, a bihile. Tatana wa Sofia a xavisi role, Sofia a kombela mali ku tatana.

## ROUMANIAN (Latin script)

Roumanian (Romanian, Rumanian) is a Romance language, derived originally from the speech of Roman soldiers who were posted to what was then the province of Dacia in the reign of the Emperor Trajan. It is now the native tongue of some 15 million people in Roumania.

The country has always been rich in folk literature, in the form of epics, lyrics, and didactic and dramatic literature. Much of its best literature is inspired by rural life, including the work of Mihail Eminescu, the country's greatest poet, and his friend, Ion Creangă, a master of narrative prose.

The Latin alphabet is used, but *k q w y* occur only in foreign words. The following accented letters are used: *ă â ä è î î ș ț ü*.

**SPECIMEN** In this verse a poet praises his native language:

Mult e dulce și frumoasă  
Limba, ce vorbim!  
Altă limbă armonioasă  
Ca ea nu găsim!  
Saltă inima 'n plăcere,  
Când o ascultăm,  
Și pe buze aduce miere,  
Când o fine-cuvântăm.

## RUSSIAN (Cyrillic script)

The Russian language, which is spoken by at least 124 million people, is one of the most important in the world today. This importance is of fairly recent growth, for before the time of Peter the Great the language was largely dominated by the vocabulary and style of Church Slavonic, and Russian was of little literary significance until the 19th century. Since then, however, writers like Pushkin, Dostoyevsky, Tolstoy, Chekov, Gogol and Gorky have developed its wonderful potentialities and made it into one of the great languages of the world. Besides its literary value, it is of ever increasing importance to students of political thought and culture, and to all who wish to keep abreast of developments in science and technology.

For the Russian alphabet see *Cyrillic (Part One)*.

**SPECIMEN** This famous passage from Turgenev tells how he felt towards his native language. (In days of doubt, in days of sad reflection on the fate of my country, you alone succour and support me, O great, mighty, truthful and free Russian tongue. Were it not for you, how could

## THE MONOTYPE RECORDER

one not fall into despair at the prospect of all that is happening at home? But it is impossible to believe that such a language was not given to a great people.)

Во дни сомнений, во дни тягостных раздумий о судьбах моей родины, — ты один мне поддержка и опора, о великий, могучий, правдивый и свободный русский язык. Не будь тебя — как не впасть в отчаяние, при виде всего что совершается дома? Но нельзя верить, чтобы такой язык не был дан великому народу.

## SAMOAN (Latin script)

Samoa is a Polynesian language, spoken on the island of Samoa. It is akin to Tahitian and Maori.

A limited alphabet is required, since *b c d g w x y z* are never used and *h k r* occur only rarely. The only accented letter is *ā*.

**SPECIMEN** The following is part of a description of the fire-rubbing technique:

O le māsani o Samoa mai le vavau e leai ni a lātou afitusi po'o ni afitā e tusa ma afitā māsani ai alii papalagi.

## SANSKRIT (Devanagari script)

Sanskrit is the language of the ancient Hindu scriptures. The earliest of these, the *Rigveda*, was in existence in the period 1200–1000 B.C. and may be much older. From that time onwards the literary tradition of the language is continuous and uninterrupted. Its early literature is second only to that of Greece in the ancient world for the wealth and range of its achievement.

For many centuries the Vedic hymns were retained in the memory of Hindu priests before they were written down in native scripts. The Brahmins, whose duty it was to protect this literature, made it their concern to preserve the language against change. To this end they devised a system of grammar that covers the whole language in minute detail. As a result it has been retained, as Latin was in Europe, as a medium for writing on learned and religious topics long after it had ceased to be spoken.

Several English scholars were attracted to the study of Sanskrit in the 18th century. The greatest of them was Sir William Jones, who pointed out that Sanskrit bore a resemblance to Latin and Greek that was too close to be due to chance and suggested that all three "have sprung from some common source which, perhaps, no longer exists", and the Germanic and Celtic languages probably had the same origin. Later scholarship has established the truth of these suggestions. From them the concept of the Indo-European group of languages and the whole science of comparative philology originated.

At first, Sanskrit literature was written down in the script of the local language, Bengali in Bengal, Telugu in South India, Sinhalese in Ceylon, and so on. But from the beginning of the 19th century Sanskrit has been more

and more exclusively printed in Devanagari, the script of the sacred city of Benares (Varanasi) and of the local language, Hindi.

**SPECIMEN** One of the greatest Sanskrit authors was Kalidasa, who is thought to have flourished in the 5th century A.D. In this couplet from his poetic drama, *Abhijñana Shakuntalam*, good wishes are expressed for a young woman setting out on a journey:

रम्यान्तरः कमलिनोहरितैः सरोभि-  
रछायाद्गर्भैर्निर्मिताकैर्ममूखात्पुनः।  
मृत्वात्कुचोद्ययरजीभुङ्गोष्णरस्याः  
शान्तामृकूलपवनश्च शिवश्च पत्याः॥२३॥

### SARDINIAN (Latin script)

There are a number of dialects spoken in the island of Sardinia which are generally known collectively as Sardinian.

The Latin alphabet is used with *grave* accented vowels. **SPECIMEN** This verse about a dove shows how widely Sardinian differs from Italian:

Unu culumbu biancu  
Lu tenz' in domo mia,  
Chi lu tent' e recreo.  
Unu culumbu biancu.  
S'amore tu' e mia,  
S'er posta dae Deus  
Mai non dà benner mancu.

### SCOTTISH GAELIC (Latin script)

See *Gaelic*

### SEA DAYAK (Latin script)

In the British territory of Sarawak in North Borneo, one of the native languages is Sea Dayak.

It is written in the Latin script without extra letters.

**SPECIMEN** Here is part of a lesson about telling the time: Gaga gambar jam ba keretas tau ka batu slate nuan, lalu ambi rangkang dua iti ka jarum ia.

### SERBO-CROATIAN (Cyrillic and Latin scripts)

Serbo-Croatian is the principal language of Yugoslavia, where it is spoken by some ten million people. Linguistically Serbian and Croatian are identical, the only distinction being that Serbian is written in Cyrillic characters, according to a reformed orthography of 1840, and Croatian is written in Latin characters, with the following accented letters: č é đ š ž.

For the Serbian-Croatian alphabets, see *Cyrillic (Part One)*.

**SPECIMEN** This is a passage from a folk-tale, in both Latin and Cyrillic scripts:

Tako najprije uzjaše đavo na Švaba i započevši najđužu pjesmu koju je znao, stane pjevati, a kad je svrši, onda Švabo uzjaše na đavola, ali mjesto kake prave pjesme zaintači pjevati *tralala*, a tako Švabo

prevario đavola, te ga je morao nositi čitav dan, a pjesmi ni kraja ni konca.

Tako najprije uzjaše đavo na Švaba i započevši najđužu pjesmu koju je znao, stane pjevati, a kad je svrši, onda Švabo uzjaše na đavola, ali mjesto kake prave pjesme zaintači pjevati *tralala*, i tako Švabo prevario đavola, te ga je morao nositi čitav dan, a pjesmi ni kraja ni konca.

### SHONA (Latin script)

The name Shona is now given to a number of dialects in Southern Rhodesia (e.g. Karanga, Nda, Zezuru) that were once thought to be separate languages.

In 1932 a system of writing the language was adopted which incorporated six special letters: ɓ ɗ ɗɗ ɗɗɗ ɗɗɗɗ. In 1955 it was decided to discard these letters and return to the Latin alphabet without additions.

**SPECIMEN** This sentence, in the 1932 spelling, says that the people of old used to think that Europeans had no knees:

Uanhu vakare vaifungga kuti varungu hanana-mabvi.  
And here is a proverb in the modern spelling (Do what your heart wishes. The tortoise likes his shell):  
Chida moyo. Hamba ya ka da makwande.

### SIAMESE (Thai script)

See *Thai (Parts One and Two)*

### SINDEBELE (NDEBELE) (Latin script)

Sindebele or Ndebele is the language of the Matabele and has been adopted by other tribes in the central area of Southern Rhodesia.

It is written in Latin script, to which the special letter ɓ is sometimes added.

**SPECIMEN** Two proverbs tell us that spilt milk cannot be picked up and that the lingering bush-pig is caught by dogs:

Amanzi citegileyo ka wa butwa'muntu.  
Ingulube e sala ngemuva i banjwa yizinja.

### SINHALESE (Sinhalese script)

The Sinhalese language is the mother tongue of some nine million people in Ceylon. Dialects of it are also spoken in the Laccadive and Maldivé Islands. It has existed in the form of inscriptions or of manuscripts since the 3rd century B.C., and it has a longer continuous written history than any other Indo-European language with the exception of Greek. Over the centuries many of the literary efforts of the Sinhalese, both in prose and in verse, have been devoted to religious subjects.

**SPECIMEN** Here is a version of Aesop's fable about the dog and his shadow:

ඔලලක් මිස් කැටියකින් මස් කැටියක් ගෙනා ගත්තේ ය.  
එදෙනෙකින් උඩින් යන අතර, උර උගේ මි ගෙඩිකැල්ල වතුර  
දැක්වේ ය. ගෙඩිකැල්ලෙන් ගෙනාගත් ඔලලා, ටෙන මස්  
කැටියක් ගත් තඩින් ඔලලෙකැයි සිතා ඒ මස්කැටිය ද  
ගැනීමට උර සිතාගත්තේ ය.





**TELUGU** (Telugu script)

Telugu is one of the four main Dravidian languages of India, the others being Tamil, Malayalam and Kannada. It is spoken by some thirty million people in southern India.

Inscriptions prove that Telugu had a separate existence as a language at least as early as the 5th century, and it has been in use as a literary vehicle since the 11th century. For many years Sanskrit was almost the only source of inspiration, but, within the last century, Western influence on the literature can be seen in the work of K. Viresalingam and those who followed him. In recent times Viswanatha Satyanarayana and Adivi Bapiraju are acknowledged as two great novelists. As elsewhere in India, the short story is a popular form of literature.

**SPECIMEN** This passage comes from a Telugu story:

అంజను మర్కటరామమ్మ బాతా కోమలిలో చెయ్యగా, అంజను  
మిలచిపోయినా గుండక తలచినా మర్కటరామమ్మ -వచ్చి పోయిన  
యేమనడమున్న మనోపోయిన కంఠంబున్న ఉప్పించడము పొద్దుము కాని  
వని గుండక మనోపోయిన కంఠంబున్న వచ్చి పోయిన యేమనడమున్న  
చెయ్య అని తీర్పుచేసెను.

**TEMNE** (Latin script)

Temne, a language of Sierra Leone, is written in Latin script with the following extra letters:  $\varepsilon$   $\eta$ .

**SPECIMEN** (His mother said to the servant, whatever he says to you, do it):

Okara kaj  $\varepsilon$  pa ka eboi, Tai o tai be ata o kans nu-  
 $\varepsilon$  yo ti nu.

**THAI** (Thai script)

The Thai language, which used to be known as Siamese, is spoken by some 27 million people throughout Thailand. The language is related to the Chinese-Tibetan group and, like the others in this group, possesses a system of tones.

The written history of the language goes back to the 13th century. At first and until quite recently, Buddhist monks were the leading teachers, and monasteries were the only important seats of learning and the only repositories of written documents. The introduction of printing, however, and the improvement of communications have had a widespread effect upon a people who are already highly civilized and intelligent.

**SPECIMEN** This is part of a recent article on the Thai language:

ไมทานาไทยที่มีอยู่ขณะนี้ทั้งหมด ๔๔ คำ คำที่อักษรตัวแรก คือ ก ไก่  
และอักษรตัวสุดท้ายเป็นสระกลุ่ก แต่ทว่า คำที่อักษรกลางตัวนี้ไม่มี  
น้อยมาก และบางคำกลุ่กใช้แล้ว ถ้ามีการออกเสียงวรรณยุกต์ขึ้น  
หรือขุ่นลงมี ๔๔ คำ ถูกอ่านออกอกเป็นสามทวน คือ

**TIV** (Latin script)

Tiv, also called Munchi, is spoken by nearly a million people in northern Nigeria.

It is written in Latin script with one extra letter:  $\delta$ .  
**SPECIMEN** (When they went in they saw a young man seated on the right side, dressed in a white robe: and they were full of wonder):

Mba nyer ken uwar la y $\delta$ , ve nenge wanyekwaor tema  
ken uwegh ku yanegh, haa g $\delta$ g $\delta$ riga u pupuu; cier  
ve iyol.

**TONGA** (Latin script)

Tonga, like Cinyanja, is spoken in Northern Rhodesia and Nyasaland.

It is written in Latin script, without additions.

**SPECIMEN** (This is your spear, take it):

Ndeeli isumo lyako, litole.

**TSONGA** (Latin script)

Tsonga is a Bantu language, widely spoken in Portuguese East Africa.

It is written in Latin script without additions.

**SPECIMEN** This sentence is from a book of prayers:

Tatana Xikwembu, a ndzi lavi ku va mukriste hi  
vito tnsena.

**TSWANA** (Latin script)

Tswana belongs to the Sotho group of the Bantu language family and is spoken in Bechuanaland.

It is written in Latin script with the following extra letters:  $\xi$   $\delta$   $\xi$ .

**SPECIMEN** This is a verse of a poem by a local poet:

Tum $\delta$  kgolo kwa borwa, tum $\delta$  ya kgosi Mmanaka,  
Kg $\delta$ sana ya mariga e palame ka boganka;  
E it $\delta$ kwa thata ke thakana ya makawana,  
Ga ba e rate ka e pitl $\delta$ lela dirathana.

**TUMBUKA** (Latin script)

Tumbuka is a Bantu language, spoken in northern Nyasaland.

It is written in Latin script with the following extra letters:  $\eta$   $\upsilon$ .

**SPECIMEN** Here is a sentence from a folk-tale:

Nteura munyake wakayezga kulipa zizombe, simbuzi,  
zimerere, ndipo na vantu, kweni iyo wakakana  
wakati, "Nkukhumba muvwi wane ndipera".

**TURKISH** (Latin script)

Turkish is the native language of some twenty million people in Turkey, and it is closely related to other languages that are spread from Macedonia to Siberia. It is of great historical importance, since it was the official language of the Ottoman Empire, which for several centuries comprised the greater part of Asia Minor and part of Europe. During that time Turkish was written with Arabic characters, but since 1928 the Latin alphabet has been used.

The letters *q w x* occur only in foreign words, but the following accented and special letters are used: â ç ğ ı î ö ş ü ū. (Note: there are two versions of *i* – one in which both the upper and lower case are dotted and one in which neither is dotted.)

**SPECIMEN** This is part of a folk-tale:

Son günlerde sık sık hatırlanan bir orman masalı da hayvanların silahlarını bırakma hikâyesidir. Belki bilirsiniz: Hayvanlar arasında bir silâhsızlanma kongresi yapılmıştı. Sebabi malum: Bir gün arslan havada heybetle uçan kartala bakmış ve "Tırnakları kaldırmalıyız!" demişti.

**TWI** (Latin script)

Twi is the name given to a group of dialects found in Ghana between the rivers Tano and Volta. It is related to the Sudanic languages.

It is written in Latin script, using the following special letters: ɛ ɛ̃ ɔ.

**SPECIMEN** This is part of a tale of a monkey:

Dkontromfi abieg bi a woye anuanom traa kwac kwee bi mu. Əpanyij no de Kwagyadu na akumaa no de Kakaweadwe.

**UKRAINIAN** (Cyrillic script)

Ukrainian is of a common origin with Russian and was for centuries regarded as a dialect of Russian. It was prohibited for many years, and only with the lifting of the ban in 1905 was it possible for a distinctive Ukrainian literature and culture to develop, with its centre at Kiev. It is now the native language of some 33 million people.

The Ukrainian alphabet differs from the Russian (see *Cyrillic, Part One*) in omitting ъ ы э є and in adding є і.

**SPECIMEN** These lines are from a poem by Shevchenko (1814–69), a founder of the national literature who was banished by the Czar:

Утоптала стежечку  
Через яр,  
Через гору, серденько,  
На базар.  
Продавала бублики  
Козакам,  
Вторгувала, серденько,  
П'ятака.

**URDU** (Persi-Arabic script)

Urdu, which is related to Persian, the Indo-European languages and Sanskrit, is one of the official languages of Pakistan and is spoken by some seventy million people. Its script is used for one of the two literary forms of Hindustani, the other being Hindi.

The literary history of the language begins in the 11th century with the invasion of the Indian sub-continent by Persian-speaking Moslems, who developed the Urdu language and wrote it in Arabic characters as a means of communication with the inhabitants. From about 1350 it was used to translate Islamic religious tracts and poems.

The first considerable Urdu poet was Wali, who early in the 18th century caught the popular fancy with his *ghazals* or love poems and broke the monopoly hitherto enjoyed by poets writing Persian under court patronage. Wali was soon followed by two outstanding but strongly contrasted poets, Sauda and Mir Taqi, who are still honoured among the great poets of the language. They were followed by Anis and Dabir, two writers of elegiac epics, and Ghalib, who was an inspiration to the generation that came after the Mutiny. The new outlook, created by nationalist ambitions, was well illustrated by Iqbal, who died in 1938 a few years before national independence.

Urdu prose literature first attained an acknowledged literary standard in 1801 with the publication of Mir Amman's *Bagh o Bahar*. The literary impetus thus generated has now invaded all fields of human interest. The novels of Nazir Ahmad served as a model for several generations of youthful writers. But in Urdu, as in other vernacular literatures of India, the short story has become more popular than the novel and is often employed as a means of ventilating social problems. Prem Chand and Rashid ul Khairi were among the writers who were inspired to write as a result of social injustices.

**SPECIMEN** These specimens were set on a 'Monophoto' Filmsetter. They mean: My eyes are anxious to see her again, and the thought of any difficulty arising has driven me out of my mind; difficulties in love give birth to love itself, and the shortcomings of the loved one help to develop the fullest intensity of love in the lover.

پہرامں کے دیکھنے کو آنکھیں ترس رہی ہیں۔ یادش بہتر جس نے دیوانہ کر دیا ہے  
مجاہزی رنگ میں فراق عشق کو زندگی عطا کرتا ہے اور عاشق اپنے محبوب کے تیر  
تغافل ہی سے سوز باطن اور دردِ محبت کے مزے لیتا رہتا ہے۔

**VA-NYANEKA** (Latin script)

Va-Nyaneka is a Bantu language, spoken in Angola.

It is written in Latin script without additions.

**SPECIMEN** This passage is from a poem about maize:

Ovi lia vi ou-hala ng' ovi ta vi ou m belesoko ng'  
ova n tu.

**VENDA** (Latin script)

Venda is a Bantu language, spoken in northern Transvaal.

It is written in Latin script with the following extra letters:  $\text{ǃ} \text{ǀ} \text{ǂ} \text{ǃ} \text{Ǆ}$ .

**SPECIMEN** This is a passage from a school reader:

Hone Dzindi ho vha hu na maḁambuwo o vhiḁhaho;  
ngauri ho vha hu sathu u vha na magondo ane a  
nga a zwino.

**VIETNAMESE** (Latin script)

Vietnamese, formerly called Annamese, is the language spoken by most of the 25 million inhabitants of the republics of North- and South-Vietnam. Since the time of the first missionaries to Indo-China, the language has been written in Latin characters.

The only special consonant used is  $\text{Đ đ}$ , but numerous diacritical marks are employed to denote both the length of the vowels and their intonation, on which (as in Chinese) the meaning of the word depends. The following are the signs affecting the length of vowels:  $\text{â} \text{ã} \text{ê} \text{ô} \text{u}'$  (the last two are sometimes combined with the vowel). The following are the signs affecting intonation:  $\text{â} \text{ã} \text{ê} \text{ô} \text{u} \grave{\text{a}}$  etc. These two systems of marking the vowel are often combined, with the intonation marks above the marks denoting length or alongside each other.

**SPECIMEN** This is the beginning of a folk-tale, which is set in Italian Old Style (108):

Ngày xưa ông giời có một người cháu gái dẹt cưỡi  
bên bờ sông Ngân, tên là Chức nữ. Chỉ yêu một anh  
chăn trâu tên là Nguru Lang. Biết chuyện đó, giời giận  
lâm vì cháu giời lại yêu một anh chăn trâu.

**WALLOON** (Latin script)

Walloon is a language closely related to French. It is spoken in southern Belgium and in the adjoining region of north-eastern France.

It is written with the same accented letters as French.

**SPECIMEN** This is a verse from a poem by Joseph Demoulin (1825-1879):

Si nos vix pére  
Riv'nit à s'pére  
I ririt bin di nos oy;  
Lès tièsse di hoie  
Jurit-leu cole,  
C'êstèu por z'èll li vix spot dé pays!

**WELSH** (Latin script)

Welsh is a language of the Brythonic branch of the Celtic languages and is spoken in Wales by about a million people. It is also called *Cymric* or *Cymraeg*. There is a considerable literature in this language dating from the Arthurian legends of mediæval times, and it is still in active use as the living tongue of a highly literate people.

The method of writing Welsh has long been agreed, but it is still liable to change. Until recently *j* and *x* were

regarded as non-Welsh, but both have been in use since 1955. The following accented letters are sometimes used, though many texts are printed without them:  $\text{â} \text{ã} \text{ê} \text{ô} \text{u}'$   $\text{ò} \text{ô} \text{w} \text{y} \text{y}$ .

**SPECIMEN** In this verse the poet Goronwy Owen pines for his native Anglesey:

Y lle bum yn gware gyt  
Mae dynion na'm hadwaenynt;  
Cyfaill neu ddau a'm cofiant,  
Prin ddau, lle'r oedd gynau gant;  
Dyn didol dinod ydwyf,  
Ac i dir Mon estron wyf.

**WENDISH** (Latin script)

See *Lusatian*

**WOLOF** (Latin script)

Wolof, also called Jolof, is the *lingua franca* of Senegal and is spoken by some 700,000 people.

It is written in Latin script with one extra letter:  $\text{ø}$ .

**SPECIMEN** These are some lines from a rhyming game:

Ai yoo, neene tuuti,  
tuuti lu mo jooi?  
hana reenu la jooi,  
reuu mutul a jooi,  
dunda mut a jooi,  
so dunde am sa muur . . .

**XHOSA** (Latin script)

Xhosa is a Bantu language, spoken in the south of Africa between East London and Durban.

It is written in Latin script, using the special letters:  $\text{ǃ}$   $\text{ǂ}$ .

**SPECIMEN** (They enquired who were present):

faɓuzile ukufa nguɓani na okhoyo.

**YIDDISH** (Hebrew script)

Hebrew script is used to write Yiddish, a language based on 14th-century German with a background of Hebrew and additions from the language of the countries where it is spoken, e.g. Polish, Russian, French, Spanish, etc. It is widespread, and the total number of people speaking it, chiefly in Europe and America, has been estimated at 12 million.

The earliest dated document in Yiddish is in the Cologne records and goes back to 1396 A.D. There is a considerable literature, and several newspapers are published in Yiddish.

As it employs characters to indicate vowels, few diacritical marks are required.

**SPECIMEN** Here is part of a dialogue in a bank:

וואו איז דער נאָענטסטער באַנק?  
זײַט אַזוי נעבן, קאָסירט מיר אײַן דעם טשעק.

**YORUBA** (Latin script)

The territory of the Yoruba covers both south-western Nigeria and parts of Dahomey. In Nigeria they number over five million.

Attempts have been made to apply Africa Script to Yoruba, but in general it is still written according to the system laid down by Archbishop Crowther, who was himself of Yoruba descent. According to this system, the normal Latin alphabet is used, supplemented by the special letters: à ã ẹ è é í ọ ó ò ñ ù.

**SPECIMEN** This passage is from a story book for children:  
 Ní kùkuru Ọ̀gbóní nì o wá lárín Ọ̀bá atí awọ̀n eníà rẹ̀,  
 tí nwọ̀n nri pe ọ̀bá kò ẹ̀ ohun tí o wu u, tí nwọ̀n nri pe  
 awọ̀n eníà nẹ̀ ẹ̀tọ̀ pẹ̀lu ọ̀bá.

**ZANDE** (Latin script)

Zande is spoken in the very south of Sudan. It is written in Latin script.

**SPECIMEN** Here is a sentence from a children's story:  
 Zabía agi he ya ambara adu bire yo mbembei na gako bambu.

**ZULU** (Latin script)

The Zulus, the best known of the Bantu peoples, inhabit the territory north of Durban. Their language is written in Latin script, using the special letter: ǃ.

**SPECIMEN** (When the beer was boiling, it got spoilt):  
 Utshwala ǃathi ǃuǃila, ǃachitheka ǃonke.

# The Monotype Corporation Limited

Registered Office Monotype House, 43 Fetter Lane, London E.C.4, Fleet Street 8351  
 Head Office Salfords, Redhill, Surrey, England. Redhill 4641  
 Works Salfords, Redhill, Surrey, England. Redhill 4641  
 Halbeath Road, Dunfermline, Scotland. Dunfermline 3121

## BRANCHES

Birmingham The Monotype Corporation Limited  
 Bristol  
 Dublin  
 Edinburgh  
 Harrogate

Legal & General Buildings, 8 Waterloo Street, *Milank 1968*  
 West India House, 54 Baldwin Street, *Bristol 2462*  
 39 Lower Ormond Quay, *Dublin 74667*  
 45 Hanover Street, *Calcutta 2909*  
 12a-14 West Park, *Harrogate 6459*

## OVERSEAS ADDRESSES

Algeria C. Riahi  
 Australia The Monotype Corporation of  
 Australia Pty. Limited  
 Brazil Companhia Lanston do Brasil S.A.

23 Rue Mokhtar Attia, Tunis  
 100-114 Chalmers Street, Sydney, N.S.W.  
 567 Little Bourke Street, Melbourne, Victoria  
 Rua Santos Rodrigues 190, Rio de Janeiro  
 Rua Major Diego 877, São Paulo  
 550-2 Merchants Street, Rangoon  
 263 Adelaide Street West, Toronto 1, Ontario  
 53 Sir Ernest de Silva Mawatha, Colombo 7  
 Casilla 3328, Santiago  
 Apartado Aereo No. 5357, Bogota  
 23 Sh. Abdel Khalek Sarwat, Cairo  
 307-9 Marina House, Queen's Road Central, Hong Kong  
 Rotary Building, 4th Floor, 27, 1-chome, Kanda, Nishikicho, Chiyoda-ku, Tokyo  
 1-2 Magazine Road, Singapore  
 73 Ampang Road, Kuala Lumpur, Malaya  
 20 Rajendramath Mukherjee Road, Calcutta 1  
 India House, Opp. G.P.O., Bombay 1  
 18 Second Line Beach, Madras 1  
 Vohra House, 25 Asaf Ali Road, New Delhi 1  
 Medina House, 22 Residency Road, Bangalore  
 31 Edmonstone Road, Allahabad  
 Division II, Graphic Department, 127 Djalan Hajam Wuruk, Djakarta  
 Jafid Hassan Pasha, Baghdad  
 20 Hamashir Street, Tel Aviv  
 Sidq Qassim Building, Wadi Socr Street, Amman  
 B.P. 1119, Beyrouth  
 390 Bd. Mohammed V, Casablanca  
 208-10 Madras Street, Christchurch  
 13 Mumtaz Street, Off Durand Road, Lahore  
 Mandivilla Chambers, Wood Street, Karachi 2  
 8 Purana Paltan, Dacca  
 Apartado No. 159, Lima  
 Monotype House, 58 Siemert Road, New Doornfontein, Johannesburg  
 84-6 Bultengracht Street, Cape Town  
 599 Mitrichi Road, Bangkok  
 23 Rue Mokhtar Attia, Tunis  
 3620 G Street, Philadelphia 34, Pa.  
 630 - 3rd Avenue, New York 17, N.Y.  
 Room 1150, Transportation Building, 608 South Dearborn Street, Chicago, Ill.  
 1562 Spring Street N.W., Atlanta, Ga.  
 115 New Montgomery Street, San Francisco 5, Calif.  
 Ituzingo 1533, Montevideo  
 Monotype House, 60 Campbell Street, Lagos, Nigeria

Burma S. Oppenheimer & Company Limited  
 Monotype Company of Canada Limited  
 Canada The Monotype Corporation Limited  
 Chile Juan Markiewicz Munter  
 Colombia Amerex Ltda  
 Egypt Joseph Lindell  
 Far East The Monotype Corporation (Far East) Limited  
 London (Pacific) Limited  
 London Wiggins Teape Limited  
 Wiggins Teape (Malaya) Limited  
 The Monotype Corporation Limited

Indonesia Perusahaan Negara Tulus Bhakti  
 Iraq Kamal Ahmed Agha  
 Israel Palewco Limited  
 Jordan Rudolph Habesh & Sons  
 Lebanon R. Kaschun (Representative)  
 Morocco R. Simon  
 New Zealand Morrison & Morrison Limited  
 Pakistan The Monotype Corporation Limited

Pers Sadolin & Simman S.A.  
 South Africa Monotype Machinery (S.A.) Limited

Thailand Thai Watana Panich  
 Tunisia C. Riahi  
 United States of America Lanston Monotype Company

Uruguay José Castiglioni (h) S.A.  
 West Africa The Monotype Corporation Limited

## CONTINENTAL ADDRESSES

Austria L. A. Rode  
 Belgium and Luxembourg The Monotype Marketing Company Limited  
 Denmark Hugo V. Larsen  
 Finland Kirjainliikemiesseiniemistö Oskaryhtö  
 France Société Anonyme Monotype  
 Germany Setemaschinen-Fabrik Monotype G.m.b.H.  
 Greece Electrotyping and Technical Company  
 Limited  
 Holland The Monotype Corporation Limited  
 Iceland Snabjörn Jónsson  
 Italy Agenzia Generale Italiana Macchine Monotype  
 Norway Olaf Gulowen A.S.  
 Portugal Monotype Portuguesa Ltda  
 Spain F. Lozano Navarro  
 Sweden A. B. Maskinfirmen Monotype  
 Switzerland The Monotype Corporation Limited  
 Turkey Burkhard Gastenbein & Company

Hytgasse 30-2, Vienna XVI  
 54 Rue du Commerce, Brussels IV

Højtorvej 12, Copenhagen NV  
 Kallevalkanki 13, Helsinki  
 85 Avenue Denfert-Rochereau, Paris 14e  
 Arnburgerstrasse 68-70, Frankfurt am Main  
 10 Lahnstrasse, 1 Berlin 44  
 Kanigos Street 6, Athens

Kaizergracht 142, Amsterdam C  
 T Holtenans, Reykjavik  
 Corso Matteotti 19, Turin  
 Viale di Villa Pamphili 124, Rome  
 Grensen 5-7, Oslo  
 Rua dos Lusíadas 8a, Lisbon 3  
 Doctor Esquerdo 16, Madrid 2  
 Fogelvägen 108, Johannestun 4, Stockholm  
 Aarbergergasse 56, Bern  
 Ich Han No. 2, 23-8 (P.O.B. 176), Galata-Istanbul



