

**A Brief Introduction to the  
Semitic Languages**



**A Brief Introduction to the  
Semitic Languages**

**Gorgias Handbooks**

**19**

**Series Editor  
George Anton Kiraz**

**Aaron D. Rubin**

The Gorgias Handbooks series provides students and scholars with textbooks and reference books useful for the classroom and for research.



**gorgias press  
2010**

Gorgias Press LLC, 954 River Road, Piscataway, NJ, 08854, USA

www.gorgiaspress.com

Copyright © 2010 by Gorgias Press LLC

All rights reserved under International and Pan-American Copyright Conventions. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise without the prior written permission of Gorgias Press LLC.

2010

ISBN 978-1-61719-860-1

ISSN 1935-6838



## TABLE OF CONTENTS

Table of Contents.....	i
Preface .....	iii
Acknowledgements.....	v
Abbreviations.....	vii
Introduction.....	1
1 Languages and Classification.....	3
1.1 Akkadian.....	6
1.2 Eblaite.....	7
1.3 Modern South Arabian .....	7
1.4 Ethiopian Semitic .....	8
1.5 Arabic.....	10
1.6 Sayhadic (Old South Arabian).....	13
1.7 Ugaritic.....	15
1.8 Canaanite.....	15
1.9 Aramaic .....	18
2 Phonology.....	23
2.1 Semitic Consonants.....	23
2.2 Semitic Vowels .....	26
2.3 Roots and Root Integrity .....	26
3 Morphology .....	29
3.1 Pronouns.....	29
3.1.1 Independent Pronouns.....	29
3.1.2 Suffixed Pronouns.....	32
3.1.3 Reflexive and Reciprocal Pronouns.....	33
3.2 Nominal System.....	34
3.2.1 Gender and Number .....	34
3.2.2 Nominal Case.....	36
3.2.3 Nominal State .....	38
3.2.4 Articles.....	40
3.3 Numerals .....	41

### Library of Congress Cataloging-in-Publication Data

Rubin, Aaron D., 1976-

A brief introduction to the Semitic languages  
/ by Aaron D. Rubin.  
p. cm. -- (Gorgias handbooks; 19)

Includes bibliographical references.

1. Semitic Languages--Grammar, Comparative.

I. Title.

PS3021.R823 2010

492'.045--dc22

2010027121

Printed in the United States of America

3.4 Verbal System.....	43
3.4.1 Verbal Stems.....	44
3.4.2 Voice.....	47
3.4.3 Verbal Tense/Aspect.....	48
3.4.4 Mood.....	52
3.4.5 Phrasal Verbs.....	55
3.5 Adverbs.....	56
4 Syntax.....	59
4.1 Word Order.....	59
4.2 Positional Relations.....	60
4.3 Agreement.....	62
4.4 Comparison.....	65
4.5 Coordination.....	66
4.6 Copular Clauses.....	67
4.7 Existential Clauses.....	69
4.8 Possession.....	70
4.9 Interrogatives.....	71
4.10 Relative Clauses.....	72
4.11 Subordinate Clauses.....	75
5 Lexicon.....	77
6 Guide to Further Reading.....	79
Bibliography.....	85

## PREFACE

The contents of this book were originally intended to form a single chapter in a larger work on the Afroasiatic languages. When that project fell through, I decided that what I had written was worth publishing on its own. I had already been using it in seminars as an introduction to the Semitic family, and students found it useful. I expanded the original work, and this book is the result.

A comparative study of the Semitic languages is a very substantial undertaking, as the family comprises dozens of languages, spread out among the ancient, medieval, and modern periods. Numerous comprehensive studies have been made (see Chapter 6, below), and much more still needs to be written. But these larger works are much more than introductions to the Semitic family. There are also numerous brief sketches of the Semitic family, ranging from a few pages to a few dozen pages. These are more appropriate for the beginner or non-specialist, but they lack adequate detail to be very useful. This survey, therefore, is something of a middle ground between these two types. Its aim is to be both practical and manageable.

This brief introduction is intended to provide the student, general linguist, or amateur language enthusiast with an overview of the characteristic features of the Semitic languages, as well as some of the more interesting and unique developments that take place in the individual languages. Moreover, it is my goal to give the reader an idea of the diversity of the Semitic languages, as well as their similarities. All too often, sketches of this family focus mainly on the classical languages, and so in this book I have tried to incorporate at least as many examples from modern languages (particularly lesser known languages) as there are from ancient languages. I want the reader to know that

the Semitic family includes numerous vibrant and fascinating modern languages, worthy of all sorts of linguistic studies.

In order to make all of the examples as clear as possible, I have provided morpheme-by-morpheme glossing, following the Leipzig Glossing Rules. I assume that the reader is familiar with linguistic terminology, and has a basic understanding of historical linguistics. For those who do not have this background, I have provided some references to general linguistic works in the Guide to Further Reading (Chapter 6).

Given the length of this volume, many topics have necessarily been left unexplored (or underexplored), but I hope that the Guide to Further Reading (Chapter 6), as well as the many references scattered throughout the book, will encourage the reader to pursue the study of the Semitic languages further. With such a long attested history, such a diversity of languages—including languages that are important to several major religions and to some of history's most influential cultures—the study of the Semitic languages is exceptionally rewarding. And while I do not expect that everyone should choose a career devoted to the study of the Semitic languages, I do hope that everyone can appreciate their importance to both the past and present.

## ACKNOWLEDGEMENTS

My sincere thanks are due to my teacher and friend John Huehnergard, whose influence is obvious throughout this volume.

My thanks also to Gary Rendsburg and Rebecca Hasselbach, who provided me with very valuable comments on earlier drafts of this book. I am also grateful to Dr. George Kiraz, who accepted this book into the Gorgias Handbooks series.

My wife Kim has, once again, dedicated a large amount of time to editing this book in its various stages, and her assistance has been invaluable. I thank her, as always, with love and great appreciation.

## ABBREVIATIONS

1	first person
2	second person
3	third person
ACC	accusative
ACT	active
ADJ	adjective
BCE	before the common era
C	common gender
C	consonant
ca.	<i>circa</i>
CE	common era
COLL	collective
COMPAR	comparative
CONJ	conjunction
CONSTR	construct
DAT	dative
Deut.	Deuteronomy
DN	Divine Name
DU	dual
Exod.	Exodus
EXIST	existential particle
F	feminine
FUT	future
Gen.	Genesis
GEN	genitive
GER	gerund
IMPER	imperative
INDEF	indefinite
INF	infinitive
INTERROG	interrogative
JT.SS	jussive

lit.	literally
M	masculine
NEG	negative
NENA	Northeastern Neo-Aramaic
NOM	nominative
NONPAST	non-past (present or future)
Num.	Numbers
OBJ	object
PART	participle
PASS	passive
PAST	past tense
PL	plural
PREP	preposition
PRES	present
PROG	progressive
REL	relative pronoun
SG	singular
SUBORD	subordinating particle
V	verb; vowel
X > Y	X develops into Y.
X < Y	X derives from Y.
*	An asterisk marks a reconstructed or underlying form. For languages whose vocalization is not indicated in the script (e.g., Ugaritic), an asterisk also marks a theoretical vocalized form.
**	A double asterisk indicates a non-existent or ungrammatical form.

## INTRODUCTION

With a written history of nearly five thousand years, the Semitic languages comprise one of the world's earliest attested and longest attested language families. Most of the Semitic languages were or are spoken in the areas of the Levant, Mesopotamia, Arabia, and across the Red Sea in Ethiopia and Eritrea. Small pockets of Phoenician speakers settled in North Africa in the first millennium BCE (where their language is referred to as Punic), but it was not until the spread of Islam, and its language Arabic, that much of North Africa became Semitic speaking.

Today, Arabic is the most important of the Semitic languages, as it is the *lingua franca* of the Near East and North Africa. With roughly two hundred million speakers, it is also (by far) the Semitic language with the greatest number of speakers. Two thousand years ago, Aramaic was the *lingua franca* of the Near East, and a thousand years earlier, Akkadian had this distinction. The Semitic languages are thus inextricably linked to this region that is often called the cradle of civilization, and that still today is the subject of much international attention.

Two of the Semitic languages, Hebrew and Arabic, are the holy languages of major religions, Judaism and Islam, respectively. A third language, Aramaic, has played a significant role in the histories of both Christianity and Judaism. As such, these languages have been widely studied for many centuries. In fact, the relationship of Arabic, Hebrew, and Aramaic was recognized already in the 9th century by the Jewish grammarian Judah ibn Quraysh. Comparative Semitic studies has a long history indeed.

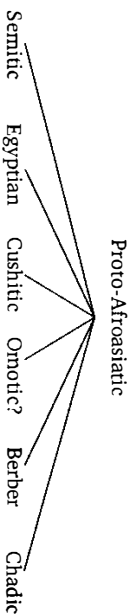
With the coming of the Reformation in Europe, and the consequent focus on the original languages of the Bible, the study of Hebrew became very popular among Christians. In the following centuries, with the rise in popularity of the study of Arabic, and with the European discovery of other Semitic lan-

guages like Syriac and Ge'ez, came a rise in linguistic comparison of these languages. When the modern field of historical linguistics was developed in the 19th century for Indo-European studies, it was natural then that the methods were quickly applied to the Semitic languages.

The Semitic languages, both ancient and modern, continue to be widely studied. They play a role in many areas of academic study, such as linguistics, religious studies, history, political science, comparative literature, and others. Some of the languages have died out, and others are spoken by tiny, vanishing communities, yet many continue to grow and thrive as languages of both literature and daily communication. One must keep in mind that the Semitic family is continually evolving, and that no study can ever be the last word.

## 1 LANGUAGES AND CLASSIFICATION

The Semitic language family is part of a larger family of languages called Afroasiatic, which in older literature is also called Hamito-Semitic. Afroasiatic can be divided into the following branches:



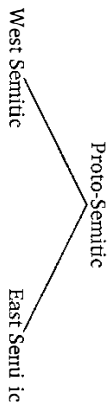
The existence of this macro-family has been recognized and studied since the 19th century, but given the great time depth that separates these branches, comparative Afroasiatic studies are tremendously difficult. Even the branches attested in antiquity, Semitic and Egyptian (both attested in the third millennium BCE), are highly divergent, suggesting a long period of separation. It does seem that some of the branches of Afroasiatic (e.g., Semitic and Egyptian, Semitic and Berber) share a greater number of similarities than do some other branches, but as yet it is impossible to demonstrate distinct sub-branches within the Afroasiatic family tree. Note that with the exception of Semitic, all of the branches of Afroasiatic are native to Africa: Egyptian, Cushitic, and Omotic in the east, and Chadic and Berber in the west. Also note that the position of the Omotic languages is still debated; some may really belong to the Cushitic family, and some of the languages may not even be Afroasiatic at all.

Within the Semitic family of languages, there is no consensus among scholars with regards to the proper subgrouping of the Semitic family, and probably there never will be. The following presentation reflects the subgrouping of the Semitic family as it is best understood given the facts available to date. For



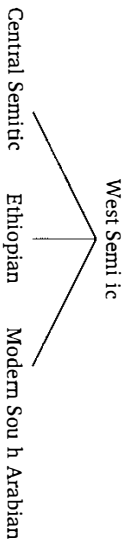
a detailed discussion of the internal classification of Semitic, see Rubin (2008) and the many sources cited therein.

There is a primary division between East and West Semitic, a division that has remained relatively uncontroversial for more than a century.



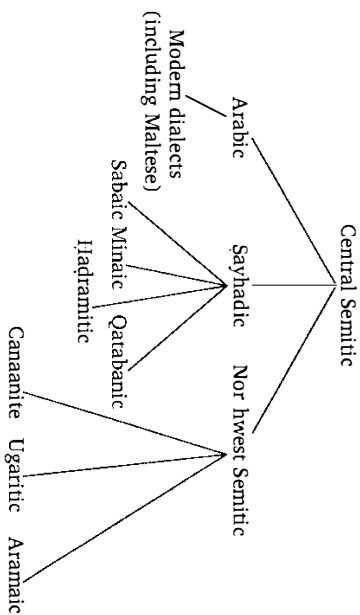
The foundation for this division is a major innovation that took place in the verbal system of West Semitic. In West Semitic, the inherited prefixed past tense (*yqtulu*) was replaced by the suffix-conjugation (*qatala*), which in East Semitic functions as a stative (or verbal adjective). This shared development of a stative into a past tense, and the replacement (or marginalization) of the inherited past tense, is a defining characteristic of West Semitic, though there are, of course, other innovations that are specific to each of these two branches; see Huehnergard (2006b) for further discussion of East Semitic innovations.

Among the West Semitic languages, there is a division between the Central Semitic languages, the Ethiopian Semitic languages, and the Modern South Arabian languages.



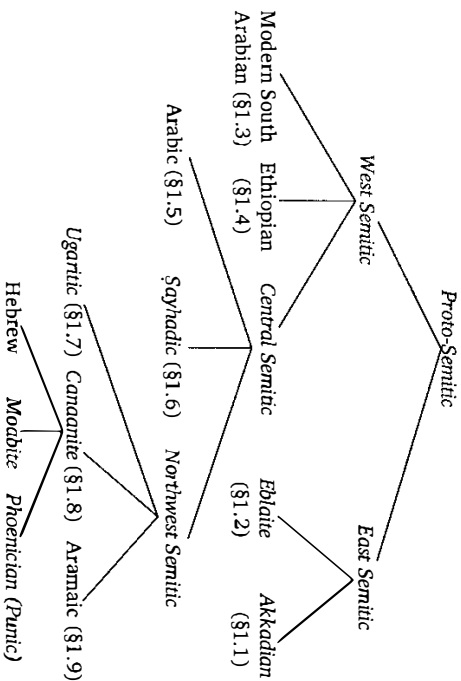
Central Semitic was first defined as such by Hetzron (1974, 1976), based largely on the development of a new imperfective verbal form (*yaqtulu*). This new form completely replaced the inherited imperfective form of the pattern *yqattul*, which survived in East Semitic, Ethiopian Semitic, and Modern South Arabian. The most detailed treatment of Central Semitic to date is that of Huehnergard (2005a). The Central Semitic family comprises Arabic, including Classical Arabic and other ancient North Arabian dialects, as well as a multitude of (often mutually unintelligible) modern Arabic dialects; the four Sayhadic languages,

also known as the Old South Arabian languages; and the Northwest Semitic languages.



Note that in addition to Ugaritic, Canaanite, and Aramaic, there are two other dialects of Northwest Semitic, attested in inscriptions from the first millennium BCE. One is called Sam'alian, known from just a few monumental inscriptions from ancient Sam'al, which is the site of modern Zincirli in south-central Turkey (about 75 miles northeast of Antakya). These texts date from about 820–730 BCE. The other Northwest Semitic dialect, which is given no special name, is known from a single inscription found at the site of Deir 'Allā, in northwestern Jordan (about 30 miles northwest of Amman). This inscription, which contains only about thirty broken lines of text, dates from about the same time period as the Sam'alian texts. The exact position of Sam'alian dialect and the dialect of the Deir 'Allā inscription within Northwest Semitic is unclear, but based on the scanty evidence that we possess, it seems that they cannot be called Canaanite, Ugaritic, or Aramaic.

Details regarding dates of attestation and location, as well as further information on the subgrouping of some of the individual branches of Central Semitic, will be given in the following sections. Following is a more complete Semitic family tree, combining what has been presented so far. In this tree, the numbers in parentheses indicate the subsections below in which the language or language group is treated. Languages (or groups) that are no longer spoken are indicated by italics.



### 1.1 AKKADIAN

Akkadian was spoken by the ancient Babylonians and Assyrians in Mesopotamia, more or less in the area of modern Iraq. It is the oldest attested Semitic language, with written records already around 2400 BCE. Akkadian is not a uniform language, but rather can be divided into multiple dialects, many of which achieved a standardized written form. The major division is between the Babylonian and Assyrian dialects, which in turn are distinguished chronologically. Following are the major dialects, along with their approximate dates of attestation:

Old Akkadian	(2400–2000 BCE)
Old Babylonian / Old Assyrian	(2000–1500 BCE)
Middle Babylonian / Middle Assyrian	(1500–1000 BCE)
Neo-Babylonian / Neo-Assyrian	(1000–600 BCE)
Late Babylonian	(600 BCE–100 CE)

Already in the Middle Babylonian period, Old Babylonian came to be considered the classical period of Akkadian. Authors in both Assyria and Babylonia developed a purely literary dialect based on the Old Babylonian model, known as Standard Babylonian. Standard Babylonian was used for literary and some

monumental texts in all later periods. Some time in the first half of the first millennium BCE, Akkadian died out as a spoken language, though Late and Standard Babylonian continued to be used in writing until about 100 CE.

There are also what are called “peripheral” dialects of Akkadian, which are essentially dialects attested outside of the Babylonian and Assyrian homelands, usually reflecting substrate influence from the local language. Notable sites where peripheral Akkadian texts have been found include Nuzi, Alalakh, Mari, Emar, Ugarit, and El-Amarna (see §1.8). These come mainly from the mid- to late second millennium BCE, when Akkadian was used as a *lingua franca* throughout the Near East.

Knowledge of the cuneiform system used to write Akkadian died out by the 2nd century CE, and cuneiform was deciphered only in the 19th century.

### 1.2 EBLAITE

Eblaitic designates the language of ancient Ebla, modern Tell Mardikh, which lies just south of Aleppo, Syria. The language was discovered only in the 1970s, when several thousand cuneiform tablets were excavated at the site. All of the texts can be dated to a period that spans less than a century, from the late 24th to the mid-23rd century BCE. Because of the nature of the cuneiform writing system used for Eblaitic—in particular the broad use of logograms and the ambiguity in the representation of nearly all consonants and vowels—knowledge of the language remains patchy. Still, it is clear that it is a close relative of Akkadian, but with enough differences to warrant placing it in its own branch of East Semitic (Huehnergard 2006b).

### 1.3 MODERN SOUTH ARABIAN

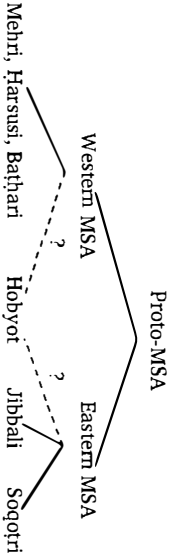
The Modern South Arabian family includes six languages: Mehri, Jibbali (or Šheri), Harsusi, Sogotri, Hobyot, and Baṭṭhari. All are spoken in eastern Yemen and western Oman, with the exception of Sogotri, which is spoken on the Yemeni-governed island of Soqatra, located in the Indian Ocean about 150 miles east of the Horn of Africa. Mehri, Jibbali, and Sogotri in turn have a number of dialects; in fact, Harsusi and Baṭṭhari are similar enough to

Mehri that they may also be considered dialects of that language. Each of these languages has a relatively small number of speakers, though exact figures are unknown. Mehri is the language with the greatest number of speakers: estimates on this number range from about 75,000 to 150,000. Jibbali has perhaps 30,000 speakers; Sogotri has perhaps 10,000 speakers, and the remaining three languages have probably less than a thousand speakers each.

None of these languages has a tradition of writing, and so our knowledge of these languages is quite recent. A couple are documented by Europeans as early as the 1830s, while others are known only from the 20th century. Hobyot and Baṭṭari remain especially poorly documented.

Despite the confusing terminology, the Modern South Arabian languages did not descend from the languages sometimes known as Old South Arabian (see below, §1.6), nor are they varieties of Arabic.

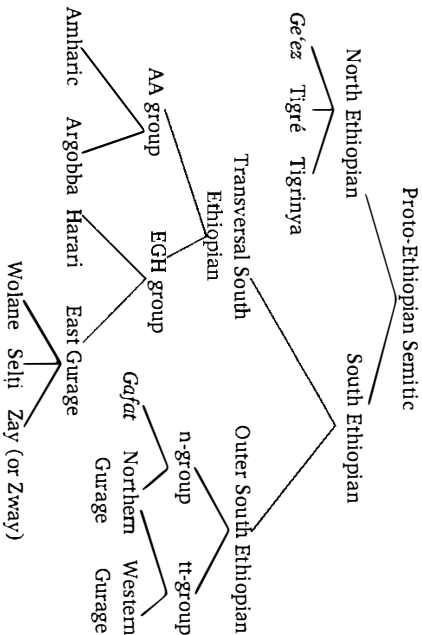
Little has been done with the classification of the Modern South Arabian group. Lomnet (2006) provides a tentative classification scheme:



1.4 ETHIOPIAN SEMITIC

The Ethiopian branch of Semitic contains a variety of languages, most of which are known only from modern times. The major exception is Ge'ez, the classical language of Ethiopia and still the liturgical language of the Ethiopian Church, which is attested in inscriptions already in the early 4th century CE, or perhaps even in the late 3rd century CE. Amharic is attested from the 14th century CE, but was not widely written until the 19th century. Though some of the languages still await comprehensive description, the Ethiopian Semitic family has on the whole been well studied, as has its internal classification. Following is

a simplified family tree of the languages, modelled after the work of Hetzron (1972). Those languages that are no longer spoken are indicated with italics:



Ge'ez, the language of the ancient Kingdom of Axum, in Northern Ethiopia, probably died out as a spoken language close to a thousand years ago, but remained the primary written language of Ethiopia up to the 20th century. Tigrinya is spoken in the northern parts of Ethiopia, as well as in Eritrea, where it is the national language. It has about six million speakers total. Tigre is also spoken in Eritrea, where it has about a million speakers. Dahalik, spoken by the several thousand inhabitants of the Dahlak Archipelago, off the coast of Eritrea, may be a Tigre dialect, but some have claimed that it should be considered an independent language (Simeone-Senelle 2006).

Amharic is the most widely spoken Ethiopian Semitic language, and the second most widely spoken (after Arabic) of any Semitic language today. It is the national language and *lingua franca* of Ethiopia, where it has about 20 million native speakers, though it is understood by many more non-native speakers. The closely-related Argobba, spoken mainly in the Ankober region, about 100 miles northeast of Addis Ababa, has about 10,000 to 15,000 speakers remaining. Harari has perhaps 30,000 speakers. It was historically spoken only in the walled

Muslim city of Harar, in southeastern Ethiopia, but as a consequence of the political upheavals of the 1970s, most Harari speakers now live in the vicinity of Addis Ababa.

The remaining Ethiopian Semitic languages, other than Gafar, which became extinct in the mid-20th century, are often called together the "Gurage Languages", though, as illustrated in the above tree, they do not form a single genetic group. The Northern Gurage group contains three languages/dialects, while the Western Gurage group (which can be further subdivided) contains about ten. Speakers of the Gurage languages, totalling between one and two million, all live in the same region, about 150 miles southwest of Addis Ababa, and most of them are also grouped together ethnically. Thus for geographical and ethnic reasons, it is convenient to speak of the Gurage languages as a group. The Gurage languages form a Semitic island surrounded by speakers of Cushitic languages.

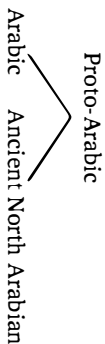
Those Ethiopian languages that are written all use a unique syllabary, based on the alphabet borrowed from the Sabaeans. The exception is Harari, which has traditionally been written with the Arabic script; since the 1990s, however, the Ethiopian script has been used to write Harari.

### 1.5 ARABIC

Arabic today has roughly 200 million speakers, whose domain stretches from Mauritania in the west to Oman in the east. In Africa, it is the main language of all the North African countries, from Mauritania to Egypt, as well as in the northern regions of Chad and Sudan, and is also widely spoken in Djibouti and Eritrea. It is the main language of all the countries of the Middle East, with the exception of Turkey, Iran, and Israel. Still, it is widely spoken in Israel and is spoken in small pockets of Turkey and Iran. There are also Arabic-speaking communities in Central Asia (Afghanistan and Uzbekistan). In the Mediterranean, we find Maltese (see below) and a tiny community of speakers on Cyprus. In the Middle Ages, there were thriving Arabic-speaking communities in Iberia and in Sicily. Arabic is also spoken today by large expatriate communities in Europe and the Americas,

and is the liturgical language of hundreds of millions of Muslims around the world.

Arabic is not a single linguistic entity, and we can distinguish different types of Arabic both chronologically and geographically. First, we must distinguish Arabic from what is known as Ancient North Arabian.

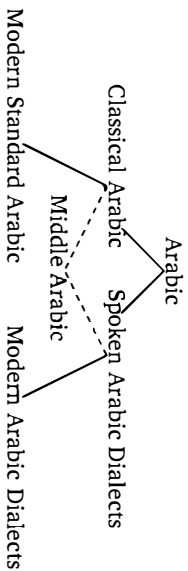


Ancient North Arabian is a cover term for several closely related dialects that are attested in inscriptions found mainly in the territories that are now Syria, Jordan, and Saudi Arabia, and that date from about the 8th century BCE to the 4th century CE. The names of these dialects include Taymanitic, Dadanitic, Dumaitic, Safaitic, Hismaic, Hasaitic, and Thamudic. Ancient North Arabian is attested about a millennium before Arabic, but none of these dialects are the ancestor of Arabic. Ancient North Arabian should be considered distinct from, though closely related to, the language that would become Classical Arabic.

For Arabic proper, we can distinguish several periods. Old Arabic designates the language of Arabic inscriptions, from about the 3rd to 7th centuries CE, that is, until the Islamic period. Old Arabic is relatively sparsely attested, and is found written in a variety of scripts. Classical Arabic refers to the variety of the written language that was standardized by the 8th century CE, based on the Qur'an and other pre-Islamic poetry. This has essentially remained the literary standard since, at least in terms of its grammar. Modern Standard Arabic (in Arabic, *fushā*) is essentially a modernized version of Classical Arabic that began to take shape in the 19th century. Not surprisingly, Modern Standard Arabic differs most from Classical Arabic in its lexicon, though there are also some minor differences in grammar. Modern Standard Arabic is the official language of the twenty-two Arab nations. It is the language of education, mass media, formal writing, and is used as a *lingua franca* across the Arab world.

While Modern Standard Arabic is the main written language of the Arab world, there exists a large number of spoken Arabic dialects, on which more will be said below. Many of the

differences that exist among the modern dialects certainly existed already a millennium ago, though in literary (Classical Arabic) works, this fact is largely concealed. However, there are a very large number of texts from the medieval and early modern periods in which post-classical, non-standard features abound. These kinds of texts are loosely called Middle Arabic, and essentially form a sort of middle ground between the literary language and the spoken dialects.



The spoken varieties of modern Arabic can be divided into several dialect bundles, which in turn contain a large number of dialects. Some of these dialects are so different from one another that, if we use mutual intelligibility as a distinguishing criterion, we should really speak of the modern Arabic *languages*, in plural. Only for cultural reasons, and because these dialects share a literary (and formal spoken) standard, do we usually find the term “dialect” used for the diverse forms of modern Arabic. Thus all Arabic speakers share a common written language, though their vernacular tongues can differ considerably.

Scholars usually divide the modern dialects into five major groups: Arabian peninsular; Mesopotamian, Syro-Palestinian (or Syro-Lebanese), Egyptian, and Maghrebi (or North African). This division is largely geographic, though there are indeed linguistic characteristics that pertain to each group. The individual dialects within each of these groups can vary considerably, not only with respect to location, but often also with respect to religious affiliation. For example, one cannot speak accurately of Baghdadi Arabic, but only of Muslim, Jewish, or Christian Baghdadi Arabic (Blanc 1964). There is also often a difference between the Arabic of the Bedouin (nomads) and that of sedentary Arabic speakers in the same region. Linguistic subgrouping of the Arabic dialects is an enormously difficult task and so scholars have

for the most part remained satisfied with this rough geographical division.

One variety of modern Arabic stands apart from the rest, in that it can safely be designated as a separate language. This is Maltese, spoken on Malta and its neighboring islands. Maltese has no tradition of written Arabic (in either its classical or modern standardized form), and instead has developed a written tradition of its own, using the Roman script. But historically, Maltese can be classified as a Maghrebi Arabic dialect, and despite the many phonological and lexical influences of Romance languages (and more recently, English), it is still easily recognizable as such.

Finally, it should be mentioned that there exist in Africa some creolized varieties of Arabic, notably Juba Arabic in southern Sudan and the closely related Nubi (or Ki-Nubi) in Uganda and Kenya. These are well known among creole specialists, but little studied by Semitists.

## 1.6 SAYHADIC (OLD SOUTH ARABIAN)

The Sayhadic group—also called Old South Arabian, Ancient South Arabian, or Epigraphic South Arabian—includes four languages (or dialects): Sabaic (or Sabean), Minaic (or Minian), Qatabanic (or Qatabanian), and Hadramitic. These designations are based on those used by the Greek scholar Eratosthenes in his *Geography* (late 3rd century BCE) for the four main peoples who inhabited the area of Southern Arabia corresponding roughly to what is now Yemen (see the edition of Roller 2010). The term Sayhadic derives from Sayhad, the name that medieval Arabic geographers gave to the Yemeni desert area now called Ramlat as-Sab'atayn, at the fringe of which the speakers of these languages had their major cities. Sayhadic is not yet a widely used term, but it is preferable to the other possibilities, since it avoids any misleading connections with the terms Arabic, Ancient North Arabian (§1.5), and Modern South Arabian (§1.3). It also allows the possibility of distinguishing Sabaic, Minaic, Qatabanic, and Hadramitic from the rare and very poorly understood other epigraphic languages that seem to have existed in ancient South Arabia (Beeston 1984; 1987). The approximate dates of written attestation for the Sayhadic languages are as follows:

Sabaic	(ca. 1000 BCE–ca. 560 CE)
Minaitic	(ca. 1000 BCE–ca. 120 BCE)
Qatabanic	(ca. 700 BCE–ca. 200 CE)
Hadramitic	(ca. 700 BCE–ca. 300 CE)

The Sayhadic languages are known almost exclusively from monumental texts, found mainly in modern Yemen and Saudi Arabia, but also in Ethiopia and Northern Arabia (e.g., Jordan, Syria, and Iraq). There are several thousands of these, though the great majority are in Sabaic. As can be seen from the list of dates above, Sabaic also has the longest period of attestation, and can be divided into different dialects (Stein 2004). Hadramitic is the most poorly attested. The limited content of these monumental texts, the almost complete lack of vocalization in the script, and the paucity of material for some of the languages mean that many features of the languages remain unknown. Enough is known, however, to allow for a reasonably certain classification of their position within the Semitic family.

The Sayhadic monumental texts are written in a distinctive consonantal alphabet that is distantly related to, but very different looking from, the Hebrew and Aramaic alphabets. This alphabet was borrowed by the Ethiopians, and survives in the modern Ethiopian script (see above, §1.4).

Since the 1970s, several thousand texts of a different type have been discovered, namely texts written in a cursive script on small wooden sticks and palm-leaf stalks. Unlike the monumental inscriptions, these texts seem to represent a more “everyday” type of writing, including personal letters and economic documents. Until the publication of Stein (2010), only a few dozen of these texts had been published. The unclear script and sometimes unknown vocabulary of these texts, along with their often poor state of preservation, can make them very difficult to read and interpret. Still, they have added to our understanding of the Sayhadic languages.

There is evidence from the comments of Arabic authors that Sabaic continued to be spoken into medieval times. Some have suggested the possibility that a variety of Sayhadic (presumably Sabaic) is still spoken in a small pocket of northwestern Yemen, but this seems very unlikely (Watson et al. 2006). It is

surely the case, however, that Sayhadic has exerted influence on many of the Yemeni dialects of Arabic.

### 1.7 UGARITIC

Ugaritic is the language that was spoken around ancient Ugarit (now called Ras Shamra), a city located just a few miles north of modern Latakia, on the Syrian coast of the Mediterranean. Texts are attested only for a short period in history, from about 1380 to 1180 BCE. Ugaritic was discovered by modern scholars only in 1928. About 1500 texts are known, most of which are very short. These are written in a unique cuneiform script that looks superficially like logo-syllabic cuneiform systems (e.g., Akkadian), but in fact is an alphabet of about thirty characters. Some Ugaritic is also attested in Akkadian syllabic cuneiform, which, since the Ugaritic alphabet rarely indicates vowels, provides important evidence for the vocalization of the language.

The Ugaritic texts cover a wide variety of genres, including epic poems, religious texts, letters, divination texts, and school texts, among others. It is most widely studied by biblical scholars, who have noted many similarities between Ugaritic literature and biblical literature, and who have used this language to shed light on difficult Hebrew words and forms. As one of the earliest attested forms of Northwest Semitic, it is of importance not only to the study of the Hebrew Bible, but also to the study of Semitic in general.

### 1.8 CANAANITE

The most prominent member of the Canaanite branch of the Semitic family is Hebrew. In the ancient period, our main source of Hebrew comes from the Hebrew Bible (or Old Testament), which includes material that was written during the period of (roughly) 1150–150 BCE. However, the earliest biblical manuscripts that are extant today date from only about 100 BCE (from among the Dead Sea Scrolls); the oldest complete Hebrew Bible dates from only around 1000 CE. From the biblical period, a number of Hebrew inscriptions are known, some from as early as the 10th century BCE, though few are of any significant length. Given that the Hebrew Bible contains material written

over the span of a millennium, it is not surprising that one finds differences in the language of the biblical books. Thus, for example, scholars distinguish between Standard Biblical Hebrew and Late Biblical Hebrew, with the dividing point at around 550 BCE. There is also evidence of synchronic dialectal variation, with Standard Biblical Hebrew reflecting the dialect of Judah in general, and with some portions of the Bible reflecting a northern dialect, termed Israelian Hebrew by scholars (Rendsburg 2003a). Still, Biblical Hebrew is overall relatively uniform.

From the post-biblical period comes Qumran Hebrew, which is the language of the non-biblical texts among the Dead Sea Scrolls, dated mainly to the 2nd and 1st centuries BCE. Qumran Hebrew continues Late Biblical Hebrew, though it also exhibits a few linguistic peculiarities unknown from any other variety of Hebrew.

From the 2nd and 3rd centuries CE, we find a dialect of Hebrew that is usually called Rabbinic or Mishnaic Hebrew. In this dialect, we have several important Jewish texts, the most prominent of which is the Mishnah. Rabbinic Hebrew is distinct in many ways from both Biblical Hebrew and Qumran Hebrew, and seems to stem from a northern (Galilean) dialect (Rendsburg 2003b).

Hebrew died out as spoken language by about the 3rd century CE, but remained in use as a literary and liturgical language among Jews. In this capacity, we find writers imitating both the Biblical and Rabbinic types of Hebrew. Though the language was not anyone's native tongue, it continued to develop throughout the Middle Ages, as new vocabulary was invented or borrowed as needed. There is an enormous corpus of medieval Hebrew works, representing a wide variety of genres, much of it still unpublished.

In the late 18th and early 19th centuries, Hebrew began the road to modernization, and from this period we find many original secular works, such as plays, novels, and newspapers. In the late 19th century there began a movement to revive Hebrew as a spoken language, coinciding with the Zionist movement to reclaim Israel as the Jewish homeland. This unprecedented revival led to the creation of Modern Hebrew (or Israeli Hebrew), which is today the official language of Israel, and is spoken by

about six million Israelis. The grammar of the modern language is based heavily on Biblical Hebrew, with many elements from Hebrew of later periods, as well as a large number of newly created and borrowed words.

Hebrew is the only Canaanite language still in use, but there are several others known from the ancient period, the most notable of which is Phoenician. Phoenician is the name that was used by the ancient Greeks to describe the Canaanite peoples who inhabited the coastal plain of what is now Lebanon and northern Israel. The Phoenician language is attested in inscriptions beginning in about 1000 BCE. Because the Phoenicians were seafarers who traveled throughout the Mediterranean, Phoenician inscriptions have been found not only in Lebanon and the vicinity, but also in Cyprus, Greece, Malta, Sicily, Sardinia, Spain, and elsewhere. The dialect of the Phoenician colony that was established at Carthage (near modern Tunis, in Tunisia) is referred to as Punic, which is attested from the 6th century BCE until about the 4th century CE. Inscriptions from after the fall of Carthage (146 BCE) are usually referred to as Neo-Punic or Late Punic. Most Phoenician, including Punic, is written in an alphabet very close to that of ancient Hebrew, but in the latest period we also find inscriptions in Latin or Greek characters.

In addition to Hebrew and Phoenician, there were several other ancient Canaanite dialects, including Moabite, Edomite, and Ammonite. Our knowledge of these dialects comes from a relatively small number of inscriptions from the first millennium BCE, found in what is now western Jordan and Israel. The longest inscription, by far, is a Moabite text of about thirty-five lines from the 9th century BCE, known as the Mesha Stela, since it concerns King Mesha of Moab. The remaining inscriptions are all short and fragmentary, and sometimes it is impossible to determine which dialect is attested in a given inscription. Knowledge of these few Canaanite dialects remains rather poor.

In the late 19th century, an archive of several hundred cuneiform tablets was discovered at Tell El-Amarna in Egypt, about 200 miles south of Cairo. This archive dates from about 1350 BCE, and represents the diplomatic correspondence between Egypt and its vassal states in the Levant, as well as with other powers, like the Babylonians and the Hittites. The corre-

spondence was mainly in Akkadian, since, as noted above (§1.1), Akkadian was the *lingua franca* of the Near East at this time. However, the letters that originated in the various Canaanite cities (in what is now Israel and Lebanon) were often written in very poor Akkadian, with numerous Canaanite words, and even Canaanite grammatical forms and word order. So while the language of the texts is technically Akkadian, many of the Amarna letters provide evidence of Canaanite grammar and vocabulary. And since this corpus pre-dates the attestation of Hebrew or Phoenician by several hundred years, it is an important source for the study of early Canaanite.

### 1.9 ARAMAIC

Aramaic is first attested from about 900 BCE, around the same time as Hebrew. This makes Aramaic and Hebrew the Semitic languages with the longest attested histories (about three thousand years). Yet unlike Hebrew, Aramaic has never ceased to be a living, spoken language. During the nearly three millennia of its attestation, Aramaic can be divided into a large number of dialects, both chronologically and geographically. There are various schemes for dividing Aramaic into chronological periods, but a widely accepted one is the following:

Old Aramaic	(ca. 900–700 BCE)
Imperial Aramaic	(ca. 700–200 BCE)
Middle Aramaic	(ca. 200 BCE–200 CE)
Late Aramaic	(ca. 200–700 CE)
Neo-Aramaic (or Modern Aramaic)	(until the present)

As an emendation to this popular scheme, it is useful and more accurate to divide the Imperial Aramaic period into two: Early Imperial Aramaic (ca. 700–550 BCE) and Imperial Aramaic (ca. 550–200 BCE). Imperial Aramaic (also called Official, Classical, Standard, or Achaemenid Aramaic) became the *lingua franca* of the Near East (promoted by the Assyrian, Babylonian, and Persian empires), and remained widespread even during the Hellenistic and Roman periods. Use of Aramaic began to decline only with the spread of Islam in the 7th century CE.

Discussion of Aramaic dialects can often be confused by the fact that there are several terms that refer to corpora containing more than one Aramaic dialect (e.g., Biblical Aramaic, Targumic Aramaic, Talmudic Aramaic, and Qumran Aramaic), as well as terms which refer to a corpus within a dialect (e.g., Egyptian Aramaic within Imperial Aramaic). For example, Biblical Aramaic refers to the Aramaic of the books of Ezra and Daniel (as well as a handful of words elsewhere in the Bible). Yet the Aramaic of Ezra is a type of Imperial Aramaic (dating from the 5th century BCE), while that of Daniel is a type of Middle Aramaic (dating from the 2nd century BCE).

Already in the Old Aramaic period there is evidence of geographic dialect differences, but it is not until the end of the Middle Aramaic period that such differences fully manifest themselves in the records. At this time, a clear distinction between western (Palestinian and Nabatean) and eastern (Syrian and Mesopotamian) dialects becomes evident.

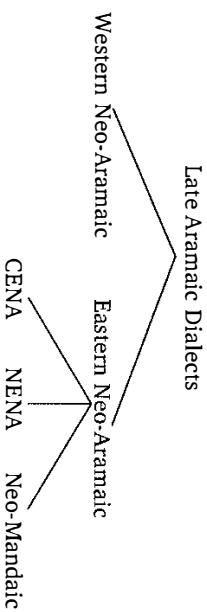
By the Late Aramaic period several very important Aramaic literary traditions developed, and the dialect differences become even more apparent. Syriac, originally the dialect of Edessa (now Sanliurfa [or Urfal], in southeastern Turkey), became the main liturgical language of Christianity in the Fertile Crescent, and is by far the best-attested Aramaic dialect. To the west we find Jewish Palestinian Aramaic (the language of the Palestinian Talmud and Targums), Christian Palestinian Aramaic, and Samaritan Aramaic. To the east of Syriac territory are found the closely related Jewish Babylonian Aramaic (the language of the Babylonian Talmud) and Mandaic. Syriac remains a liturgical language among some eastern churches. The various Jewish dialects of Aramaic continue to be widely read by learned Jews, thanks to its use in the biblical books of Ezra and Daniel, the two Talmuds, and several other compositions important to the Jewish religion.

The Aramaic language has developed into a number of modern dialects, collectively known as Neo-Aramaic. Many of these Neo-Aramaic dialects are unquestionably distinct enough to be called languages. Neo-Aramaic has traditionally been spoken in a noncontiguous area covering parts of Syria, southeastern Turkey, northern Iraq, and northwestern and southwestern Iran, mainly by Jewish and Christian communities. However, as



a result of the great political upheavals of the 20th century (most notably, World War I, the establishment of the State of Israel, and the aggression of Saddam Hussein), many, if not most, Neo-Aramaic speakers have been displaced from these areas. Nearly all Jewish Neo-Aramaic speakers have moved to Israel or the United States since 1948, and many Christians have also emigrated (to the United States, Europe, or Australia), or at least have moved to larger, Arabic-speaking cities.

The split between eastern and western dialects that is seen already in the Middle and Late Aramaic periods has survived into the present, though none of the modern languages is the direct descendent of any recorded ancient dialect. In the modern remnant of the western branch, known as Western Neo-Aramaic, there survive only three closely related dialects, spoken in the Syrian villages of Maflula, Bax'a, and Jubb'adin, located about thirty-five miles northeast of Damascus. The rest of the modern languages, those of the Eastern Neo-Aramaic branch, can be further divided into three subgroups: Central Eastern Neo-Aramaic (CENA), Northeastern Neo-Aramaic (NENA), and Neo-Mandaic.



The Central Eastern Neo-Aramaic branch includes just Turoyo and the closely related, but nearly extinct, dialect of Mlahso. Turoyo is one of the most thriving of the Neo-Aramaic languages, both in its native territory (the Tur Abdin region of southeastern Turkey) and abroad. In Sweden and in Germany, a handful of books in Turoyo have been published in recent years (in both Roman and Syriac scripts). Neo-Mandaic is the highly endangered, modern reflex of Mandaic, the language of the Mandean religion and its followers. Its speakers, living mainly in Iran, number only in the low hundreds.

The Northeastern Neo-Aramaic subgroup includes approximately 150 different dialects, many of which are mutually

unintelligible. They are traditionally spoken in the loosely-defined region known as Kurdistan. A number of the NENA dialects have been well studied in recent years, but many more have yet to be investigated fully. One interesting fact about the NENA languages is that dialect grouping is in many cases based on religious affiliation, rather than geographic location. So, for example, the Jewish Neo-Aramaic dialect of one town may be incomprehensible to the Christian Neo-Aramaic speakers of the same town, but not to Jewish speakers of another village. Such a state of affairs is somewhat analogous to the situation of African American Vernacular English, which defies the traditional geographical dialect boundaries of the United States. The Neo-Aramaic languages, in particular those of the NENA group, have been heavily influenced by neighboring non-Semitic languages (especially Kurdish and Turkish), and therefore are in many ways quite divergent from classical varieties of Aramaic.