A Brief Introduction to the Semitic Languages
A Brief Introduction to the Semitic Languages

Aaron D. Rubin
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 Ṣayḥadīc (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 Suffix 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
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. | circa |
| 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 |
| JUSS | jussive |
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-
languages 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:

[Diagram showing the division of Afroasiatic branches: Proto-Afroasiatic, Semitic, Egyptian, Cushitic, Omotic?, Berber, Chadic]

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 (*yaqtul*) was replaced by the suffixed-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 *yVqattVl*, 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 Ṣayhadic 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 scantly 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

Eblaite 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 Eblaite—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 (Huzechnergard 2006b).

1.3 Modern South Arabian

The Modern South Arabian family includes six languages: Mehri, Jibbali (or Ṣherî), Harsusi, Soqotri, Hobyot, and Baṭḥari. All are spoken in eastern Yemen and western Oman, with the exception of Soqotri, which is spoken on the Yemeni-governed island of Soqotra, located in the Indian Ocean about 150 miles east of the Horn of Africa. Mehri, Jibbali, and Soqotri in turn have a number of dialects; in fact, Harsusi and Baṭḥari 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, Soqotri 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 Bari 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. Lonnet (2006) provides a tentative classification scheme:

![Family Tree of Modern South Arabian Languages](image)

**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:

![Family Tree of Ethiopian Semitic Languages](image)
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 Gafat, 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 Sabeans. 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, ��ci) 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 ŞAYHADIC (OLD SOUTH ARABIAN)

The Şayhadic group—also called Old South Arabian, Ancient South Arabian, or Epigraphic South Arabian—includes four languages (or dialects): Sabaic (or Sabean), Minaic (or Minean), 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 Şayhadic derives from Şayhad, 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. Şayhadic 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 Şayhadic languages are as follows:
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 Israeli 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 1600 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 Stele, 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 Şanlıurfa or Urfa), 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 Mandic. 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 Ma'lula, 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.

Late Aramaic Dialects

- Western Neo-Aramaic
- Eastern Neo-Aramaic
  - CENA
  - NENA
  - Neo-Mandaic

The Central Eastern Neo-Aramaic branch includes just Turoyo and the closely related, but nearly extinct, dialect of Mlašo. 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.
## 2 Phonology

### 2.1 Semitic Consonants

Proto-Semitic had the following consonant phonemes, according to the most recent scholarship (Huehnergard 2004):

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Interc-</th>
<th>Dental/Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops:</td>
<td>voiceless</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td>g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>glottalic</td>
<td></td>
<td>t'</td>
<td>k'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates:</td>
<td>voiceless</td>
<td>'s</td>
<td>'d</td>
<td>'x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>'s</td>
<td>'d</td>
<td>'x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>glottalic</td>
<td></td>
<td>'t'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives:</td>
<td>voiceless</td>
<td>θ</td>
<td>s</td>
<td>x</td>
<td>h</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>ð</td>
<td>s</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>glottalic</td>
<td>θ'</td>
<td></td>
<td>(x')</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral continuants:</td>
<td>voiceless</td>
<td>l</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>l</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>glottalic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasals</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glides</td>
<td>w</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap/Trill</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It is noteworthy that among the series of interdental fricatives, dental/alveolar stops, dental/alveolar affricates, dental/alveolar lateral continuants, and velar stops, there is a three-way contrast: voiceless, voiced, and glottalic. The glottalic consonants are often referred to in Semitic scholarly literature as "emphatic", a rather vague term. In Arabic, these consonants are either velarized or pharyngealized (except *k', which in Classical and Modern Standard Arabic is realized as a voiceless uvular stop q), while in Ethiopian and Modern South Arabian, they remained glottalic (ejective). It is, of course, difficult to know how they were pronounced in the ancient languages (Akkadian, Biblical Hebrew, Sabaic, etc.), but there is good evidence that the glottalic pronunciation can be reconstructed for Proto-Semitic and Akkadian (Aro 1977; Zemánek 1996; Kouwenberg 2003). For this reason, the transcriptions indicating a following glottal stop (e.g., t') are preferable to the more traditional dotted letters (e.g., t) for Proto-Semitic, Akkadian, Ethiopian, and Modern South Arabian. A three-way contrast likely also existed among the velar fricatives, though the reconstructed consonant *x' has not yet been widely accepted. It was first proposed by Huehnergard (2003), on the basis of the standard comparative method. It survives in none of the attested languages; in East Semitic it has merged with *x, while in all of West Semitic it has merged with *h. In Tigrinya, *x is found as a fricativized allophone of k', unconnected with Proto-Semitic *x'. There is no evidence for a Proto-Semitic glottalic labial p', though this phoneme exists in Ethiopian languages. Most Ethiopian words with p' are borrowings from Greek or other languages (e.g., Ge'ez p'ant'ak'ost' e 'Pentecost').

The affricate set *s, *s', *s' has traditionally been reconstructed as *s, *z, *s, but there is compelling evidence to suggest that these phonemes were in fact affricates. It is equally likely that the consonant traditionally reconstructed as palatal *s was really dental/alveolar *s (Faber 1981; 1985). The affricates were deaffricated in much of the Semitic family (e.g., Arabic and Ethiopian), while Proto-Semitic *s shifted to ʃ in a number of the languages (e.g., Biblical Hebrew and Babylonian Akkadian).

The phoneme *t (usually transcribed ʃ by Semitists) is distinct in a number of the ancient languages (e.g., Biblical Hebrew, Old and Imperial Aramaic, Ṣayḥad, and Ge'ez), and there is solid evidence for its pronunciation as a lateral continuant (Steiner 1977). Among the modern languages, ʃ survives only among the Modern South Arabian group. In some languages, like Hebrew and Aramaic, *t came to be pronounced s, while in Arabic it developed into the phoneme ś. The glottalic lateral ʃ likewise survives only among the Modern South Arabian group. In Arabic it became a velarized or pharyngealized d, i.e., [d'ʃ] or [d'ʃ'] (usually transcribed d), while in Hebrew and Akkadian it merged with *s. In Old Aramaic, it became a sound that was written with the same symbol used for k' (though the phonemes did not merge), but in later Aramaic it merged with ʃ.

The phoneme *p became a fricative f in the Arabic, Old South Arabian, Modern South Arabian, and Ethiopian groups, that is, all of the more southerly Semitic languages. This is one of the features that led early Semitists to classify these languages together in a subgroup called "South Semitic". Such a simple phonetic change is, of course, of no value in classification. In fact, the shift *p > f is just an areal phenomenon, one of several that are common to these languages. In the Ethiopian languages, a phoneme p has been re-introduced; like p', it is found mainly in loanwords from European languages (e.g., Amharic polis 'police; policeman').

In most Ethiopian Semitic languages, there also exists a series of labio-velar consonants. For example, in Ge'ez we find the phonemes k̡, k̡', g̡', and x̡. In some cases, it is clear that the labio-velar element derives from the historical presence of the vowel "u, e.g., kall- 'all' < *kull-. In other cases, there is no obvious explanation for the presence of the labio-velar element.

Examples of palatalization within the various Semitic languages are numerous. For example, in most modern Arabic dialects, the Proto-Semitic phoneme *g is realized as some sort of palatal, most often asj, ʃ, g̡', or y (that is, [dʒ], [ʒ̠], [g̡'], or [j]). The phoneme *k is also palatalized to ʒ (that is, [ʃ]) in a few Arabic dialects, including Iraqi dialects. Palatalization of ʃ, k', and g is also known from some Neo-Aramaic dialects. But palatalization is most widespread in the South Ethiopian Semitic languages, where palatalization has also resulted in additional phonemes. For example, Amharic has acquired the palatal phonemes ʒ, ʃ', j, n, ʃ, and ʃ, which also occur as allophones of the
dentals t, t’, d, n, s, and z, respectively. Likewise, the inherited Amharic palatal phoneme y is also an allophone of l.

Some have suggested reconstructing the syllabic sonorants *l, *m, and *n for Proto-Semitic, in order to explain some seemingly irregular correspondences among the descendent languages (Testen 1985; 1993; 1995). How these syllabics would have fit into the phonological system of Proto-Semitic is unclear, as are their implications for Proto-Semitic syllable structure. Despite strong evidence for reconstructing these syllabics, they remain somewhat controversial.

2.2 SEMITIC VOWELS

The vowel system reconstructed for Proto-Semitic is quite simple. There were three vowels, *a, *i, and *u, and each of these could occur either short or long.

<table>
<thead>
<tr>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>û</td>
</tr>
<tr>
<td>u</td>
<td>á</td>
</tr>
</tbody>
</table>

This system is preserved in Classical Arabic, while most other ancient and modern languages exhibit developments in their vowel systems. Of course, for many of the ancient languages whose writing systems are purely consonantal (e.g., Phoenician and Sabaic), our knowledge of their vowel systems is quite limited.

Vowels play an essential part in the case system of Semitic (§3.2.2). For singular nouns and internal plurals (§3.2.1), for example, the case endings in Proto-Semitic were simply the short vowels: -u for nominative, -a for accusative, and -i for genitive. Simple sound change led to the deterioration of case endings in many of the languages. For example, in Proto-Hebrew, final short vowels were lost, leading to the complete loss of the case system. In Ge'ez, the short vowels *u and *i merged to û in all positions, leading to a merger of the nominative and genitive cases.

2.3 ROOTS AND ROOT INTEGRITY

Perhaps the most characteristic feature of the Semitic languages, is the system of consonantal roots, which has had great effect in

the area of phonology. The concept of the consonantal root is fundamental to the morphology and lexicon of Semitic, and is in fact a defining characteristic of the family. All verbs and the majority of nouns possess a set of root consonants, usually three, which are used in conjunction with vocalic templates to form words. These templates, which cause the interlacing of the consonants with vowels and often include the addition of prefixes or suffixes, are used to form all words in the language, save the pronouns and various kinds of particles. The roots carry the lexical meanings, while the templates carry the grammatical functions. For example, the Arabic root ƙƙƙ has a meaning associated with the act of riding, as seen in the words rakib 'he rode', yarkabu 'he rides', râkbûn 'a rider (NOM.)', rakâbûn 'a (professional) horseman (NOM.)', markabatun 'vehicle; carriage (NOM.)', and murakkabun 'mounted (NOM.)'. In each of these words, the root appears with a particular set of vowels, along with a prefix and/or a suffix.

The Semitic roots, as a system, are closely tied in with the study of phonology. This is because there is a set of constraints on which consonants can co-occur within a triliteral root, at least in Proto-Semitic (Greenberg 1950). For example, the first and second root consonants cannot be identical, and homorganic consonants (that is, consonants which share their point of articulation), rarely co-occur within a root. These phonological restrictions thus limit the number of possible Semitic roots.

The root system has also had a great effect on the phonologies of the Semitic languages through the principle of root integrity (Huehnnergard 2004). According to this principle, the triliteral root should remain intact, even at the expense of regular sound changes. For example, in Ge'ez we find the verbs k*arr 'be cold', causative tak*arara 'cool'. It was already noted above that the Ethiopian labio-velars normally derive from the historical presence of the vowel *u. In these verb forms, there is no historical reason to assume the presence of a labio-velar. However, the related nouns k*̄ar and k*̄arrat, both meaning 'cold', derive from *k'arr and *k'urrat, respectively. By regular sound change, we would expect to find the nouns k̄arr and k̄arrat, and the verbs **k'arr 'be cold' and **k'arara, but the pressure to maintain root integrity has led to the spread of the labio-velar (which is a separate phoneme) to all words containing this root.
Another example of root integrity pertains to a more widespread sound change. In Proto-West Semitic, there seems to have been a sound rule by which *s became *h before vowels (Voigt 1987). By such a change, a verb form such as *samīša 'he heard' would have become **hamīša, while a verb form such as *yismafu 'he hears' would have remained unchanged. This alternation of root consonants was deemed unacceptable, as a violation of root integrity, with the result that the sound change of *s > *h was blocked in all verbal roots. We thus see evidence of this sound change only in words not associated with a root: the third person pronouns (e.g., Akkadian šū ← Hebrew hū), the prefix of the causative verbal stem (e.g., Akkadian š- ← Hebrew h-), the conditional particle (e.g., Akkadian summa ← Ugaritic hm), and a few others. Another change of *s > *h took place in the history of the Modern South Arabian languages (cf. Mehri hūma ‘he heard’), but this is separate from, and later than, the sound change that was common to West Semitic.

### 3 Morphology

#### 3.1 Pronouns

##### 3.1.1 Independent Pronouns

All Semitic languages have a set of independent pronouns, that is, pronouns consisting of an independent (non-clitic) morpheme. Proto-Semitic distinguished the following forms (Huehnergard 2004):

<table>
<thead>
<tr>
<th>Gender</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st C</td>
<td>?āna / ?anāku</td>
<td>(none?)</td>
<td>nīnu</td>
</tr>
<tr>
<td>2nd M</td>
<td>?anta</td>
<td>?antumā</td>
<td>?antum</td>
</tr>
<tr>
<td>2nd F</td>
<td>?anti</td>
<td></td>
<td>?antin</td>
</tr>
<tr>
<td>3rd M</td>
<td>suʔa</td>
<td></td>
<td>sum</td>
</tr>
<tr>
<td>3rd F</td>
<td>siʔa</td>
<td></td>
<td>sin</td>
</tr>
</tbody>
</table>

Gender was distinguished for second and third persons only, hence the designation "common" for the first person pronouns. This lack of gender distinction in the first person forms applies to the verbal system as well. The two first person singular pronouns reconstructed for Proto-Semitic seem to have been free variants, as they are in some of the languages (e.g., Hebrew). A first person dual independent pronoun is attested only in the Modern South Arabian languages (e.g., Omani Mehri ʔakay), and so there is little evidence on which to base a Proto-Semitic reconstruction (E. Wagner 1952). All of the Proto-Semitic first and second person forms, with the exception of the first person common plural, seem to be made up of an element ʔan- plus a pronominal marker. The origin of the initial element is un-
known, though it has parallels elsewhere in Afroasiatic (e.g., in Egyptian).

The independent pronouns can serve as subjects of verbless (copular) clauses, as in (1), or they can provide emphasis or topicalization in a verbal clause, as in (2).

Biblical Hebrew
(1) me-layin lattem
from-where 2M.PL
'Where are you from?' (Gen. 29:4)

(2) wa-latem pan1
CONJ-2M.PL be.fruitful:IMPER.M.PL
ü-rabü
CONJ-multiply:IMPER.M.PL
'And you, be fruitful and multiply.' (Gen. 9:7)

In the attested Semitic languages, even in many of the ancient ones, the inherited set of independent pronouns has usually been reduced to include fewer forms. Loss of the dual is very common, and loss of gender distinction is also found in some languages, to varying degrees. For example, in the Northeastern Neo-Aramaic (NENA) Jewish dialect of Arbel (northern Iraq), gender is distinguished only for the second person singular forms, and even then only optionally (Khan 1999). In Maltese, gender is distinguished only in the third person singular forms.

NENA (Arbel)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st C</td>
<td>?āna</td>
<td>?axan</td>
</tr>
<tr>
<td>2nd M</td>
<td>?at / ?āti</td>
<td>?axun</td>
</tr>
<tr>
<td>2nd F</td>
<td>?at</td>
<td></td>
</tr>
<tr>
<td>3rd C</td>
<td>?ō</td>
<td>?ōnī</td>
</tr>
</tbody>
</table>

Maltese

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st C</td>
<td>jien(a)</td>
<td>ahna</td>
</tr>
<tr>
<td>2nd C</td>
<td>int(i)</td>
<td>intom</td>
</tr>
<tr>
<td>3rd M</td>
<td>hu(wa)</td>
<td>hagga</td>
</tr>
<tr>
<td>3rd F</td>
<td>h(a)</td>
<td>hagga</td>
</tr>
</tbody>
</table>

No independent objective pronouns can be reconstructed for Proto-Semitic—at least not for the first and second persons—but a number of languages have developed independent accusative and dative pronouns. Hebrew, for example, has a set of direct object pronouns based on the particle **let** ('t̠y̞yd̠ē), which on its own marks definite direct objects.

<table>
<thead>
<tr>
<th></th>
<th>1st C</th>
<th>2nd M</th>
<th>3rd M</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>?ōt̠i</td>
<td>?ōtkēm</td>
<td>?ōtōn</td>
</tr>
<tr>
<td>PL</td>
<td>?ōtnū</td>
<td>?ōtkem</td>
<td>?ōtān</td>
</tr>
</tbody>
</table>

Similar direct object pronouns based on an original morpheme *t̠y̞yd̠ē are attested in Aramaic, Arabic, and Modern South Arabian (Rubin 2005). Different sources are found in other languages, for example, the Ge'ez independent direct object pronoun is based on a stem *kiyy-, of unclear origin.

The Proto-Semitic third person pronouns, originally anaphoric pronouns, probably had an oblique (objective) case with suffixed -ti. This element is found in the oblique forms of the third person pronouns in Ugaritic (e.g., *huwa 'he', oblique *hūwātī 'him'), while in Akkadian this suffix has spread to all persons, resulting in a new set of independent direct object pronouns based on the suffixed pronouns plus the element -ti (e.g., *yātī 'me' < suffixed pronoun *ya + -ti).

Independent possessive pronouns and possessive adjectives are also not reconstructable for Proto-Semitic, but are found in many Semitic languages. These are most often formed by attaching the suffixed pronouns (§3.1.2) to a genitive exponent, the origin of which can be a relative pronoun or a noun. Examples are Post-Biblical Hebrew *šēl 'of' < relative *še- + l- 'to', and Yemeni Arabic *haqq 'of' < Classical Arabic haqq 'property' (Rubin 2004b; 2005). So from these genitive exponents, we find Modern Hebrew *šēl 'my; mine' (lit. 'of-me'), *šēlx 'your; yours' (lit. 'of-you'), etc., and Yemeni Arabic *hagg 'your; yours' (lit. 'of-you'), etc. In some languages
we find independent possessive pronouns and adjectives based on the genitive exponent (stemming from a relative pronoun) plus the independent subject pronouns, as in Amharic yäñe ‘my; mine’ (< yä + anē, lit. ‘of-I’) and Soqotri dil-hoh ‘my; mine’ (lit. ‘of-I’).

3.1.2 SUFFIXED PRONOUNS

The Semitic languages also possess a set of suffixed pronouns, used to indicate pronominal possession on nouns (3), pronominal objects of verbs (4), and pronominal objects of prepositions (5) and other particles (6). The forms of the pronouns can vary slightly with the different uses. For example, in Iraqi Arabic, the first person singular suffixed pronoun has the form -i after nouns and particles ending in a consonant, -ya after nouns and particles ending in a vowel, and -ni after verbs, while the third feminine singular pronoun has the form -ha in all environments.

- **Iraqi (Muslim Baghdadi) Arabic**
  - (3) bêt, bêt-i, bêt-ak, bêt-ha
    - house, house-1sG, house-2M.SG, house-3F.SG
    - ‘house, my house, your (m.) house, her house’
  - (4) šaf, šaf-ni, šaf-ha
    - see:PAST.3M.SG, see:PAST.3M.SG-1sG, see:PAST.3M.SG-3F.SG
    - ‘he saw, he saw me, he saw her’
  - (5) wara, wara-ya, wara-k, wara-ha
    - behind, behind-1sG, behind-2M.SG, behind-3F.SG
    - ‘behind, behind me, behind you (m.), behind her’
  - (6) wên, wên-ha, wên-hum
    - where, where-3F.SG, where-3M.PL
    - ‘where? where is she? where are they?’ (Erwin 1963)

In the Central Semitic languages, the definite article and a possessive pronominal suffix cannot co-occur. This is probably because nouns with a possessive pronominal suffix are already treated as definite for the purposes of agreement. The opposite is true for Omani Mehri, in which we find the possessive suffixes attached to the definite form of the noun (7).

- **Omani Mehri**
  - (7) a-bayt-i, a-yā-k, ha-brāt-hum
    - the-house-1sG, the-brother-2M.SG, the-daughter-3M.PL
    - ‘my house, your brother, their daughter’ (Rubin 2010b)

3.1.3 REFLEXIVE AND RECIPROCAL PRONOUNS

In the Semitic languages, reflexivity can often be expressed by means of a derived verbal stem (§3.4.1), as in Mehri šak’rū ‘hide oneself’ (vs. k’arū ‘hide, conceal’), Ge’ez taxábbā ‘hide oneself’ (vs. xabbā ‘hide, conceal’), and Arabic ṭaṭafā ‘hide oneself’ (vs. ṭafā ‘hide, conceal’). However, most Semitic languages have developed a means of expressing an independent reflexive pronoun. This is nearly always the result of a grammaticalization based on a word meaning ‘body’ or a part of the body, such as ‘head’, ‘soul’, ‘eye’, or ‘bone’ (Rubin 2005). Some examples are:

- **Amharic**
  - (8) ras-u-n giiddii
    - head-3M.SG-ACC kill:PAST.3M.SG
    - ‘He killed himself.’ (Leslau 1995)

- **Classical Arabic**
  - (9) qāla li-nafs-i-hī
    - say:PAST.3M.SG to-soul-GEN-3M.SG
    - ‘he said to himself’ (Fischer 2002)

Likewise, reciprocity can be expressed by means of a verbal stem (e.g., Modern Hebrew hitkatev ‘correspond’ vs. katav ‘write’; Mehri yarrāb ‘know one another’ vs. yarrūb ‘know’), but many languages have also developed a grammaticalized recipro-
cal pronoun. Examples are Syriac ḫāḏē ‘each other’ (< had ‘one’), Mehri tāt’īday- ‘each other’ (< ṭād ‘one’), Wolane gāg bū-gāg ‘each other’ (< gāg ‘body’), and Akkadian axāmīš or axārīš ‘each other’ (< axu ‘brother’).

3.2 Nominal System

3.2.1 Gender and Number

The Semitic nominal system (including both nouns and adjectives) distinguishes two genders, masculine and feminine. Feminine nouns are normally predictable from the presence of a feminine suffixed morpheme */-VJt. As a result of sound change, this feminine suffix can take a variety of shapes in the languages, for example, Hebrew -ā (< */-at), Mehri -āt (< */-at), and Jewish NENA of Arbel la (< */-a-t < */-a-t) (the feminine marker also has other allomorphs in these languages). Every language includes a set of feminine nouns, both animate and inanimate, which lack an explicit feminine morpheme, such as Mehri ḥāmē ‘mother’, ḫaym ‘eye’, and ḥark’āyib ‘mouse’. The word for ‘mother’ and the words for parts of the body tend to fall in this category. Gender marking on the noun is largely absent from most South Ethiopian languages, though the category of gender is still retained for agreement purposes. In Arabic, which possesses a large number of grammatically singular nouns representing collectives, the feminine form serves to indicate an individual item. For example, in Classical Arabic, we find M. dam‘ ‘tears’, F. dam‘ā ‘a tear’, and M. tamr ‘dates’, F. tamra ‘a date’.

With regard to number, the Proto-Semitic nominal system distinguished singular, dual, and plural. The dual was lost as a productive feature in many of the languages, including Hebrew, Aramaic, Akkadian (except in the oldest periods), and all of Ethiopian, though in most of these languages there are remnants. For example, in Biblical Hebrew, the inherited dual ending -ayim is preserved in the forms of certain numerals, like šayim ‘two’ and mātayim ‘two hundred’; many of the body parts that occur in pairs, like sēnayim ‘eyes’ and ṭanayim ‘ears’; certain time words, like yōmayim ‘two days’ and ṭanayim ‘two years’; and words for tools or devices with two parts, like ṭaḏāyim ‘cymbals’ and melqāḥayim ‘tongs’. Interestingly, in Modern Hebrew, this dual suffix has become semi-productive again, not in spontaneous speech, but in the conscious creation of new words, such as mišqāfayim ‘eyeglasses’ and ṭofanayim ‘bicycle’. In languages without a productive dual, like Hebrew, nouns with vestigial dual forms are treated as plural for purposes of agreement with verbs and adjectives, e.g., Modern Hebrew mišqāfayim yaqarim ‘expensive[pl.] eyeglasses[pl.]’. Languages with a productive dual, like Arabic, have dual forms of verbs and adjectives, as well.

Formation of the Semitic plural is a complex issue. There are two basic ways in which plurality can be indicated. The first is by external means, with the addition of a suffixed morpheme to the noun. The second is by internal means, by replacement of the noun’s vocalic pattern; these are usually termed “internal plurals” or “broken plurals”. In East Semitic (10) and in Northwest Semitic (11), we find only the external means of plural marking, though a few noun types show remnants of internal plural marking (Huehnergard 1987a; 1991). In Arabic (12), Ṣayhadic, Ethiopian, and Modern South Arabian (13), internal plural marking is widespread, though we find external plural marking in these languages as well. In some modern South Ethiopian languages, external plural marking has become more common.

Standard Babylonian Akkadian
(10) mār-ū ‘son’ (nom.), pl. mār-ū (nom.)
    mār-tū ‘daughter’ (nom.), pl. mār-tū (nom.)

Biblical Hebrew
(11) ǧibbōr ‘warrior’, pl. ǧibbōr-im
    šāpā ‘language’, pl. šāp-ōt

Egyptian Arabic
(12) walād ‘child’, pl. awlād
    xādim ‘servant’, pl. xādam

Omani Mehri
(13) ḥāy ‘man’, pl. ḥayyāğ
    ḥēxār ‘old man’, pl. ḥēxār
Sometimes both types of plural marking can be used together in the same word, and so we find plural marking that is indicated by both pattern replacement and suffixation; cf. Ge'ez hagar ‘city’, PL. 7ahgur (internal) or Tahgur-at (internal plus external).

In South Ethiopian, we find an innovative type of internal plural that makes use of reduplication. For example, in Zay, in which external plural marking is the norm, and in which other types of internal plural marking have been lost, some nouns indicate plural through partial reduplication, as in dahá ‘poor, orphan’, PL. dahahu; gála ‘camel’, PL. gálalu; and c’ik’télé ‘elbow’, PL. c’ik’tlálu (Leslau 1999).

3.2.2 NOMINAL CASE

Proto-Semitic distinguished three cases: nominative, accusative, and genitive. In addition to indicating a direct object, the accusative case was used in some adverbial constructions (14). The genitive case was used both for genitive constructions (see below, §3.2.3) and to mark the objects of prepositions.

Classical Arabic

(14) yasta layl-an wa-nahár-an

hurry:NONPAST.3M.SG night-ACC.lNDEF and-day-ACC.lNDEF

‘He hurries night and day.’ (Ryding 2005)

Case marking was done by means of vocalic suffixes on the noun. For singular nouns and internal plurals, the case endings were simply the short vowels: -u for nominative, -a for accusative, -i for genitive. In the dual and plural, there was only a two-way distinction between nominative and oblique cases: -d for nominative dual, -ay for oblique dual, -u for nominative plural, and -i for oblique plural. If a noun contains a feminine marking suffix, the case marker follows this suffix, e.g., Akkadian mārt-u ‘daughter (NOM.)’. A case marker precedes any pronominal suffix (e.g., Arabic li-bayt-i-M to-house-GEN-3F.SG ‘to her house’; Ge’ez hagar-a-ka city-ACC-2M.SG ‘your city (ACC.)’), though in Akkadian, the nominative and accusative (but not genitive) case markers were normally dropped before the possessive suffixes for most singular nouns.

The general tendency towards the loss of word-final vowels led to the disappearance of the case system in most languages. For example, in Proto-Hebrew, final short vowels were lost, leading to the complete loss of the case system in the singular. In the plural, the oblique plural -i, with mimination (see §3.2.3), was generalized, resulting in the external masculine plural -im. In modern Arabic spoken dialects, we find this exact change, with loss of the case system and retention of the oblique external masculine plural. In later Akkadian dialects, loss of final short vowels also led to the decline of the case system; the cases are still used in writing the later dialects, but often indiscriminately or incorrectly. In Ge’ez, the short vowels *u and *i merged to o, in all positions, leading to a merger of the nominative and genitive cases. Word-final a was subsequently lost, though some modern scholars prefer to retain this in transcription. Therefore, Ge’ez distinguishes only an accusative case; cf. hagar (< ‘hagar-’ ‘city’, hagar-a ‘city (ACC.)’).

A number of Semitic languages have developed new means of indicating the accusative. In Canaanite and some early Aramaic dialects, we find a particle derived from a form ?iyyat (e.g., Hebrew 7et, Palestinian Aramaic yelt), which marks definite direct objects (15). In a number of other languages, including a number of Aramaic dialects (73, in §4.5), some Arabic dialects (16), Ge’ez, Tigrinya (17), and Tigre, the dative preposition has developed into a marker of definite direct objects (Rubin 2005). Amharic marks definite direct objects with a special accusative suffix (8, in §3.1.3), which may also be connected with an earlier dative preposition (Appleyard 2004).

Biblical Hebrew

(15) bârâ 7elohim nêt haš-sânayim wá- nét

create:PAST.3M.SG God ACC the-heavens and-ACC haš-sâres

the-earth

‘God created the heavens and the earth.’ (Gen. 1:1)
Iraqi (Muslim Baghdadi) Arabic

(16) 资产重组 來電 the-radio
take.apart.PAST.3M.SG-OBJ.3M.SG DAT-the-radio
‘He took apart the radio.’ (Erwin 1963)

Tigrinya

(17) 重整 a that:F.SG dammu milk
那F.SG DAT-thatM.SG milk
siitayat-o drink:PAST.3F.SG-OBJ.3M.SG
‘The cat drank the milk.’ (Melles 2001)

Besides the three basic Proto-Semitic cases discussed above, there are two other suffixes which some scholars have considered to be case endings, namely the suffixes that appear in Akkadian as -is and -u(m).

The Akkadian suffix -is (the final m was lost by the end of the Old Babylonian period) functions as a locative-adverbial suffix in the standard grammars. It has the basic meaning ‘to, for’, equivalent to the Akkadian preposition ana. It also can be used with adjectives to form adverbs (see §3.5). This suffix likely corresponds to the terminative/directive suffix -(a)h found in Hebrew and Ugaritic, e.g., Hebrew şafor ‘north’, şaforה ‘to the north, northwards’; on the change of s > h, see §2.3.

The Akkadian suffix -u(m) (the final m was lost by the end of the Old Babylonian period) functions as a locative-adverbial suffix in Akkadian, e.g., bit ‘house’, bitum ‘in the house’. This suffix appears identical in shape to the nominative case ending, though the two behave differently when pronominal suffixes are added. It is unclear, therefore, if the two are historically the same. Some have suggested that this suffix provides evidence of ergativity at some pre-Semitic stage, but this is highly speculative (see Waltisberg 2002 for an argument against ergativity). The locative-adverbial suffix also possibly existed in Ugaritic, though evidence is meager, and there may also be vestiges in other languages (Tropper 2000a).

3.2.3 NOMINAL STATE

Proto-Semitic nouns had two forms, known as “states”, the uses of which were connected to a noun’s syntactic function. The

Old Babylonian Akkadian

(18) 重整 master:CONSTR house:GEN
bel bit-im
‘the master of the house’

(19) 重整 master:CONSTR-1PL
bel-ni
‘our master’

(20) 重整 house:CONSTR fall:PAST.3M.SG-SUBORD
bit imqut-u
‘the house that fell’ (von Soden 1995)

When not governing a directly following element, a noun was considered to be in the free (or unbound) state. The free state was marked by a set of nasal suffixes, which appeared after the case endings; the distribution seems to have been *-m after short vowels and *-na after long vowels or diphthongs. For example, in Old Babylonian Akkadian, the nominal base il- ‘god’ appears as ilum in the nominative singular free form but simply as il in the singular construct form (Akkadian singular construct nouns lose case marking); and the nominal base in- ‘eye’ appears as inān in the nominative dual free form but simply as inā in the nominative dual construct form. We find that most languages have levelled one or the other of the nasal consonants, that is, to *-n and *-ma (e.g., Arabic) or to *-m and *-ma (e.g., Hebrew). The shorter allomorph was lost in many languages, including all of Northwest Semitic, Modern South Arabian, Ethiopic Semitic, most modern Arabic dialects, and later phases of Akkadian. It should be mentioned that the addition of the markers *-n/*-m and *-na/*-ma is usually referred to by Semitists as “nunation” and “mimation”, after the names of the letters n (nun) and m (mim) in the Arabic alphabet.
Although the original function of these allomorphs, *-m and *
*na, in Proto-Semitic seems to have been to mark the free state,
this is not necessarily the case in the daughter languages. For
example, in Classical Arabic, these suffixes have become a
marker of indefiniteness in the singular (cf. Classical Arabic
wald-u-n 'a child (NOM.)' vs. al-wald-u 'the child (NOM.)'), though
the suffixes are dropped when the noun is in the construct state,
reflecting their historical function as a marker of the free state
only, e.g., wald-u l-malikat-i 'the child (NOM.) of the queen'. Since
the first member of a genitive chain is syntactically definite, the
lack of the suffix -n seems to have been reinterpreted as a sign of
definiteness in Arabic.

The use of the construct state to express a genitive chain
(as in 18, above) and the use of possessive pronominal suffixes
with the construct form (as in 19, above) are characteristic fea-
tures of the Semitic languages. Yet in a number of the lan-
guages, one or both of these syntactic features have been lost, or
at least have been relegated to less frequent use. As already
noted above (§3.1.1), many of the languages have developed
independent genitive exponents. The use of the construct state
to express a relative clause (as in 20, above) is common only in
Akkadian and Sayhadic (though not unknown in some other
languages, like Biblical Hebrew), and relative clauses are most
often expressed with a relative pronoun (see below, §4.10).

Although in Proto-Semitic the free state was the marked
form and the construct state was the unmarked form, in a num-
ber of languages, as a result of sound change or re-analysis, the
construct has become the marked form. For example, in Hebrew,
many nouns have a marked construct form that results from the
fact that the construct form is regularly unstressed, e.g., bayt
'house' (free) < *bdýt, but construct bêt < *bayt. In Ge'ez, the
construct form of a noun is marked by the suffix -a, the origin of
which is unclear (Tropper 2000b).

3.2.4 ARTICLES

No definite or indefinite articles can be reconstructed for Proto-
Semitic, but articles are found in many of the descendant lan-
guages. In Central Semitic, definite articles have developed from
demonstratives: Hebrew ha- and Arabic al- are prefixed to the
noun, while Aramaic -ä and Sabaic -(h)n are suffixed (on the
development of the Central Semitic articles, see Rubin 2005).
Several modern Semitic languages, including Tigrinya and
Turoyo (Neo-Aramaic), also make use of demonstratives to indi-
cate definiteness, and these are on their way to becoming fully
functioning definite articles. A number of South Ethiopian lan-
guages possess definite articles which derive from third person
possessive suffixes, e.g., Amharic bet-u 'his house/the house'
(Rubin 2010a). This is typologically a more unusual type of
grammaticalization, but one that is known from elsewhere (e.g.,
Indonesian, Yucatec Mayan, and some Turkic languages). Some
other Semitic languages have borrowed definite articles, includ-
ing several NENA dialects, which have borrowed their definite
articles from Kurdish, and Omani Mehri, which has likely bor-
rowed its article from surrounding Arabic dialects.

Indefinite articles, where they have evolved, derive from
the numeral 'one'. Examples can be found in Neo-Aramaic (21),
modern Arabic dialects, and some Ethiopian languages. Occa-
sionally usage of the numeral 'one' as an indefinite article can
also be found in some other languages (e.g., Biblical Hebrew).

NENA (Barwar)
(21) ʔoθ-wa xa-malka
EXIST-PAST one-king
'there was a king' (Khan 2008a)

3.3 NUMERALS

In the classical Semitic languages, numerals from 'one' through
'ten' have masculine and feminine forms. Strangely, the masu-
cline numerals from 'three' through 'ten' (that is, those numerals
that accompany masculine nouns) look morphologically femi-
nine, while the feminine numerals from 'three' through 'ten'
(those numerals that accompany feminine nouns) look morpho-
logically masculine. This peculiar situation has yet to be satis-
factorily explained (Hetzron 1967 is one attempt, using
Afroasiatic evidence). The table below lists the forms of the nu-
merals 'one' through 'ten' in Classical Arabic (without case en-

ings); note the presence of the feminine morpheme -at on the masculine numerals 'three' through 'ten'.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>wāḥīd</td>
<td>wāḥīdat</td>
</tr>
<tr>
<td>2</td>
<td>ʾīḥnām</td>
<td>ʾīḥnāmat</td>
</tr>
<tr>
<td>3</td>
<td>ʾtālāḥāt</td>
<td>ʾtālāḥāt</td>
</tr>
<tr>
<td>4</td>
<td>ʾābarāṣat</td>
<td>ʾābarāṣat</td>
</tr>
<tr>
<td>5</td>
<td>xamsat</td>
<td>xams</td>
</tr>
<tr>
<td>6</td>
<td>sittat</td>
<td>sitt</td>
</tr>
<tr>
<td>7</td>
<td>sabīṣat</td>
<td>sabīṣ</td>
</tr>
<tr>
<td>8</td>
<td>ʾtōmānīyat</td>
<td>ʾtōmānīn</td>
</tr>
<tr>
<td>9</td>
<td>tisṣat</td>
<td>tisṣ</td>
</tr>
<tr>
<td>10</td>
<td>Ṣaṣarāt</td>
<td>Ṣaṣr</td>
</tr>
</tbody>
</table>

In the classical languages, the teens are usually expressed with a digit followed by the numeral 'ten', for example Arabic ʾābarāṣat Ṣaṣarā and Hebrew Ṭarbaṣ Ṣaṣar '14 (M.). Note that in Hebrew and Arabic (among others) the form of 'ten' in the expression of teens exhibits the expected gender, in terms of its morphology, with the result that the digit and 'ten' appear to be of opposite gender. In Ge'ez, the digit follows 'ten' in the expression of teens, and requires a conjunctio, e.g., Ṣaṣarāt we-Ṭaḥadu '11 (M.); both units match in gender, unlike in Hebrew and Arabic. The order 'ten + digit' is common in Ethiopian Semitic, though not always with the conjunctio, e.g., Ṣaṣarāt we-Ṭaḥadu '11 (M.); both units match in gender, unlike in Hebrew and Arabic. The order 'ten + digit' is common in Ethiopian Semitic, though not always with the conjunctio, e.g., Ṣaṣarāt we-Ṭaḥadu '11 (M.). In many of the modern languages, and in fact already in Akkadian, the digit and the numeral 'ten' have fused into a single word, for example Akkadian erbešeret '14 (M.), xamsisseret '15 (M.), and Algerian Arabic xamsaṣ '15 (C.).

In some of the modern languages, gender distinction among the digits and teen numerals has been neutralized, resulting in a single form for both genders. This happens for both the digits and teens in most modern Ethiopian languages and in many NENA dialects, and with the teens in Arabic dialects.

We find two basic means of forming the tens in Semitic. In Akkadian and Ethiopian, the suffix -ā is attached to the numeral 'ten' to form twenty (e.g., Ge'ez Ṣaṣā Ṣaṣā '20') and to the digits 'three' to 'nine' to form 'thirty' through 'ninety' (e.g., Ge'ez xamsā '50'). This suffix -ā is most likely the dual marker, at least historically. In Hebrew, Arabic, and Aramaic, the 'tens' are formed by adding a masculine plural suffix; 'ten' is used to form 'twenty', while the digits are used to form the remaining tens. So we find, for example, Hebrew šešārim '20' and Ṭarbaṣim '40'. It is likely that the dual suffix was originally used only for 'twenty' (the dual of 'ten') and the plural suffix for the others, and that the attested languages all exhibit levelling of the tens. That is to say, within each attested language, the words for 'twenty' through 'ninety' all contain the same suffixed morpheme, which is historically either the dual or masculine plural suffix.

### 3.4 Verbal System

As already mentioned above (§2.3), the most characteristic morphological feature of the Semitic languages is the use of consonantal roots in conjunction with vocalic templates. Nowhere is this more evident than in the verbal system. All Semitic verbs possess a lexical root, which is a set of consonants that contain the general lexical meaning. Each root can be used in conjunction with one or more verbal stems. Within each stem, vocalic templates, prefixes, and suffixes are used in order to construct each verbal form. In short, the roots carry the lexical meanings, while the templates and affixes carry the grammatical functions, indicating person, number, tense, voice, and mood. This is illustrated by the following forms of the Biblical Hebrew root קתוב, which has a basic meaning associated with writing: yiktōb 'he writes', niktōb 'we write', kātōb 'he wrote', kāṭōb 'we wrote', kāṭōb 'write! (M.SG.), kāṭēḇ 'is writing (M.SG.)', nikṭēḇ 'he was written'.

The number of root consonants is normally three, and thus we can speak of the triliteral root. In the classical languages, we occasionally find quadriliteral roots (i.e., roots with four consonants); in some of the modern languages these are more frequent. Even verbs borrowed from non-Semitic languages are normally fitted into the root system, and so we find verbs like Modern Hebrew fiksēs 'he faxed' (root פלקס, from English 'fax') and Ge'ez manḳ*āṣa 'he became a monk' (root ṬMNK*.getProjectRoots(), from Greek monakhos 'monk'), both of which follow the normal patterns for quadriliteral roots. An interesting exception is Maltese,
where borrowings are incorporated into the verbal system as words, rather than as roots, e.g., *iddownload* 'he downloaded' (from English) and *irrakkomanda* 'he recommended' (from Italian *raccomandare*) (Hoberman and Aronoff 2003).

### 3.4.1 Verbal Stems

Every Semitic language has a basic, unmarked verbal stem. This is usually called the G-Stem in Semitic grammars, for German *Grundstamm* 'basic stem'. In addition to this basic stem, there are derived stems, which, among other things, are used to express causative, transitive, reflexive, reciprocal, passive, and iterative meanings. The stems are characterized either by prefixation or infixation of consonantal elements, or by phonological alteration of the root via gemination, vowel lengthening, or reduplication. Each of the major stem types will be treated in turn:

**N-Stem:** This stem is characterized by a prefix *n*- in Akkadian, Hebrew, and Arabic. The N-Stem is usually passive or medio-passive, e.g., Hebrew *nipṭah* 'open (trans.)' vs. G-Stem *pāṭṭaḥ* 'open (trans.)'. In Aramaic, Sayhadic, and most Ethiopian languages, the N-Stem has been lost, while in Mehri and Ge’ez, the N-Stem is only used with some quadriliteral roots. In Akkadian, quadriliteral verbs also appear most often in the N-Stem, and this is likely a Proto-Semitic feature (Gensler 1997); on a possible connection with Egyptian on this point, see Rubin (2004a). An attempt to trace the Afroasiatic origin of the N-Stem was made by Lieberman (1986).

**D-Stem:** This stem is characterized by gemination (*D* = doubling) of the second root consonant. The exact derivational function of this stem is difficult to pinpoint. Grammars of Semitic languages often describe the D-Stem as "intensive", though this vague label is not really warranted. It is true that some D-Stem verbs have a pluralic association (i.e., denoting activity on a plurality of objects), and this can sometimes make it seem as if the verbal action is intensified. For example, from the Arabic G-Stem *kāṣara* 'break (trans.)', we find the D-Stem *kāssara* 'shatter (trans.) into many pieces'. But such verbs are relatively rare. More often, the D-Stem provides the factitive/causative for roots with stative meanings in the G-Stem (e.g., Mehri *māla*? 'fill' vs. G-Stem *mila*? 'be full'), or is used to form denominative verbs (e.g., Mehri *hōnī* 'dye with henna', from *haymē*? 'henna'). In Ethiopian, the D-Stem has lost its derivational value and become simply lexical. A comprehensive study of the Akkadian D-Stem was made by Kouwenberg (1997); see also Joosten (1998).

**C-Stem:** The C-Stem normally adds a causative (= *C*) force to the root, e.g., *NENA* of Qaraqosh *asama* 'to cause to stand', G-Stem *smama* 'to stand'; Zay *ṭabālā* 'feed', G-Stem *bālā* 'eat'; Akkadian *ṣunkūtu* 'cause to fall', G-Stem *makātu* 'fall'. In Akkadian, Ugaritic, and some of the Sayhadic language, this stem is characterized by a prefix *ṣv*- (*< "ṣa-"*), a morpheme which has a cognate in the Egyptian causative prefix *s*. In the remainder of West Semitic, this prefix became *ha-* via a general sound change *s* > *h* (see above, §2.3), and was further weakened to *ʔa*- or *a*- in most languages. In Ethiopian, the C-Stem has become more productive, and we find not one C-Stem, but rather three C-Stems, which are built upon the G-, D-, and L-Stems, respectively.

**T-Stem(s):** T-Stems are characterized by a prefixed or infixed *t*, and are often found in conjunction with another verbal stem. For example, in Aramaic, we find a Gt-Stem, a Dt-Stem, and a Ct-Stem, that is, T-Stems made from the G-, D-, and C-Stems, respectively. The T-Stems have various functions in different languages, but, in general, they designate reflexives, reciprocals, passives, or medio-passives. In Aramaic and the Ethiopian languages, which have lost the N-Stem and internal passives (see below, §3.4.2), the T-Stems have become the primary means of expressing the passive. A lengthy study of the Semitic T-Stems was made by Diem (1982).

**L-Stem:** This is the only stem which is marked purely by vocalic means, namely by the lengthening (= *L*) of the vowel following the first root consonant. The L-Stem is found only in Arabic and the Ethiopian languages, though there are possible vestiges in Hebrew. In those ancient languages whose scripts do not indicate vowels (e.g., Sayhadic and Ugaritic), the L-Stem is impossible to detect. If it ever existed in Modern South Arabian, it has merged completely with the D-Stem. In Ethiopian, the L-Stem, like the D-Stem, has lost all derivational value and become simply lexical. Therefore, the only attested place the L-Stem has any special meaning is in Arabic, where it can have a
variety of meanings, including associative action (e.g., šārōka ‘enter into a partnership with’, G-Stem šārika ‘participate’), attempted action (e.g., qātala ‘attempt to kill’, G-Stem qatała ‘kill’), and behavior (e.g., ḫāṣana ‘treat kindly’, G-Stem ḥāṣana ‘be good’). These categories are not always applicable to modern Arabic dialects.

R-Stem(s): These stems are characterized by the reduplication (= R) of either the second or third root consonant. In Arabic, the two R-Stems, both of which are quite rare, involve reduplication (realized as gemination) of the third root consonant. Such verbs are most often connected with the acquisition of a color or physical trait or defect, e.g., iswadda ‘become black’ and ihwalla ‘squint; be cross-eyed’. This type of R-Stem has remnants in Hebrew, Modern South Arabian, and elsewhere, but it is productive (however rare) only in Arabic. In many modern Ethiopian languages, we find an R-Stem characterized by reduplication and gemination of the penultimate root consonant, and, among other things, it can have pluralic or repetitive associations. This stem is an internal development within the Ethiopian sub-family. Examples are Amharic gānāt‘ṭālā ‘tear into pieces’ (vs. gānāt‘ṭalā ‘tear off; tear out’), làwawwāt‘ā ‘vary (change often); change completely’ (vs. làwaww‘ā ‘change’) (Leslau 1995).

The number of verbal stems differs for each language. For example, in Syriac there are six stems (with vestiges of others), in Classical Arabic there are fifteen, and in NENA of Arbel there are just two. Moreover, the functions of a particular stem in one language do not always correspond with its functions in another language. For example, the Dt-Stem in Hebrew is often a reciprocal or reflexive (e.g., hitnašeq ‘kiss each other’, D-Stem niššeq ‘kiss’; hitqaddēš ‘sanctify oneself’, D-Stem qiḏdēš ‘sanctify’), while in Syriac, it is simply a passive of the D-Stem (e.g., ṭeqqabbal ‘be received’, D-Stem qabbel ‘receive’).

It is important to point out that a derived stem verb need not have a corresponding G-Stem verb. For example, in Arabic, the verb ārsal ‘send’ is a C-Stem verb, but the root RSL does not occur in the G-Stem. There is no other verb from which ārsal can be derived as a causative. As another example, in Mehri, the verb šānōs ‘dare’ (root ŠNS) is a Š1-Stem (a type of C-Stem specific to Modern South Arabian), but this is the only stem in which this root occurs. As these examples illustrate, it can happen that a derived stem verb is simply lexical and not “derived” at all. In Ethiopian, as noted above, the D- and L-Stems have become completely lexicalized. As a corollary to this, a verb that occurs in the G-Stem need not exhibit any derived stem. Also, except in those modern languages that have very few derived stems (e.g., NENA dialects), it is very rare that a verbal root can be used in all stems of a particular language.

3.4.2 Voice

Semitic languages distinguish active and passive voice, though they differ with respect to how these categories are coded. There are two main morphological means of expressing passivity. The first way is by means of a derived verbal stem, usually a T-Stem or an N-Stem, as was already discussed above (§3.4.1). The second way is by means of what can be called an internal passive, meaning that the morphology is by means of the internal vowel pattern only. We find internal passives in Classical and Modern Standard Arabic (22), some modern Arabic dialects (Retso 1983), to a limited extent in Hebrew (though rare in modern spoken Hebrew), in the older Aramaic dialects (all BCE), and to a limited extent in Modern South Arabian (23). For some epigraphic languages (e.g., Sabaic and Ugaritic) there is good syntactic evidence for internal passives, though the lack of written vowels in their respective scripts obscures the data.

Arabic

\[
\begin{align*}
\text{saṭal-nā} & , & \text{suṭil-nā} \\
\text{ask:PAsT.ACT-IPL, ask:PAsT.PASS-IPL} & \quad \text{we asked, we were asked}
\end{align*}
\]

Omani Mehri

\[
\begin{align*}
\text{kōbūr} & , & \text{kōbēr} \\
\text{bury:PAST.ACT.3M.SG, bury:PAST.PASS.3M.SG} & \quad \text{he buried, he was buried}
\end{align*}
\]
In addition to morphological passives, many languages also indicate passivity by syntactic means, either with an impersonal third person plural verb (24), or, less often, by means of an auxiliary verb and a past participle (25).

**NENA (Qaraqosh)**

(24) bet-ah kam-bane-la štqa
house-3M.SG PAsT-build:PRES.3PL-OBJ.3M.SG last.year
‘His house was built last year.’ (lit. ‘His house, they built it last year.’) (Khan 2002)

**Maltese**

(25) ir-ragel kien [or: gie] maqtul
the-man be:PAST.3M.SG[come:PAST.3M.SG] kill:PART.PASS
‘The man was killed.’ (Ambros 1998)

In general, there is a strong tendency among the Semitic languages for the internal passive to be replaced by other means of indicating passivity, whether by derived verbal stems or by syntactic constructions.

### 3.4.3 VERBAL TENSE/ASPECT

The verbal system is the area in which the languages exhibit the greatest variation, and so an overview of the entire family is impossible in a brief format such as this. Here we will limit ourselves to some of the interesting developments in the realm of tense and aspect (Cohen 1984 and Kurylowicz 1973 are important general works on this issue). For Proto-Semitic, suffice it to say that two basic verbal forms can be reconstructed, and it is likely that these mainly distinguished perfective and imperfective aspect. For some of the ancient languages, the issue of tense vs. aspect is a difficult one, and it is very difficult to say that the verbal system of, say, Akkadian or Biblical Hebrew, distinguished only tense or only aspect (see Joosten 2002 and Cook 2006 for recent discussion on this issue in Biblical Hebrew).

**Past Tenses.** In East Semitic (Akkadian), there existed a form known as the stative or verbal adjective, which was characterized by suffixed pronominal forms, as opposed to other verbal forms, which marked person and number by a set of prefixes (Huehnergard 1987b). This stative conjugation is found also in Egyptian, so it can be reconstructed for Afroasiatic. As already noted above (§1), in all of West Semitic, the inherited, prefixed past tense (e.g., Akkadian niššik 'we stole') was replaced by this suffixed conjugation (e.g., Arabic sarqānā ‘we stole’). This is, therefore, a major development in the history of the Semitic verbal system. Subsequent to this, additional innovative past tenses developed in later forms of Aramaic and in Ethiopian. Already in the early first millennium CE, it became possible in some dialects of Aramaic—such as Syriac, Mandaic, and Jewish Babylonian Aramaic—to express a perfect tense by means of a construction which consisted of a past participle plus the preposition l- ‘to, for’ with a pronominal suffix (26).

**Syriac**

(26) gbar lā hkim l-i
man NEG know:PART.PASS.M.SG to-1SG
‘I have not known a man.’ (Nöldeke 1904)

Some believe that the preposition in this construction was originally instrumental in function, and so a phrase that meant something like ‘it was known to/me’ was reinterpreted as ‘I have known it’. However, since the preposition l- does not normally mark the agent of a passive, but does mark possession (see §4.8), others believe that a phrase like Syriac hkim li is a perfect tense, the literal equivalent of English ‘I have known’. It has been suggested that such a development in Aramaic may have been a calque from Persian (Kutscher 1969; though see Bar-Asher 2007 for a counterargument). This new past tense, whatever its origin, has completely replaced the inherited past tense in most Neo-Aramaic languages.

In Ge'ez, there exists a verbal form variously known as the gerund, perfective active participle, converb, or verbal infinitive. This form is always subordinate to the main verb, and can be translated either by an English participial phrase or temporal clause (27).
Ge'ez

(27) wa-s'awim-o ... raxba

CONJ-FAST:GER-3M.SG be.hungry:PAST.3M.SG

'Having fasted ... he was hungry' (Tropper 2002a)

In Tigrinya, this equivalent form has become a simple past tense, replacing the inherited past tense in most contexts, e.g., s'awimu (variant s'oymu) 'he fasted'. Similar developments have taken place in other modern Ethiopian languages.

In most other Semitic languages, secondary past tenses have developed, mostly by means of auxiliary verbs or other particles (Rubin 2005). For example, in many Arabic dialects, the past tense of the verb 'be' is used in combination with other verbal forms to create the past perfect (28), past progressive (29), and past habitual (30).

Iraqi (Muslim Baghdadi) Arabic

(28) čán mākil min

be:PAST.3M.SG eat:PART.M.SG when

xabar-t-a

telephone:PAST-1SG-3M.SG

'He had eaten when I telephoned him.'

(29) čán da-yākul min

be:PAST.3M.SG PROG-eat:NONPAST.3M.SG when

wuṣal-na

arrive:PAST-1PL

'He was eating when we arrived.'

(30) čānat ttrsm zēn

be:PAST.3F.SG draw:NONPAST.3F.SG well

'She used to draw well.' (Erwin 1963)

Present Tenses. In Classical Arabic, Biblical Hebrew, Ge'ez, Akkadian, and other classical Semitic languages, there is a single non-past verbal tense that covers both present and future time. In many languages, however, new forms have developed that are specifically marked as general presents, present progressives, or present indicatives. In some languages, including later forms of Hebrew and Aramaic, the inherited participial form, which is nominal in origin, has become the present tense. In some other languages, a present tense is marked by a particle attached to an existing verbal form. Such particles derive most often from locative verbs or prepositions, i.e., the verbs 'be', 'sit', 'stand', or 'lie', or the preposition 'in, at' (Rubin 2005). For example, the prefix da- in Iraqi Arabic (qa- in Jewish and Christian dialects) is attached to the inherited non-past form, and indicates a present progressive, continuous, or habitual (31; cf. also 29, above). This prefix derives from qāfīd, a participial form of the verb 'sit'.

Christian Baghdadi Arabic

(31) qa-tabbix

PROG-Cook:NONPAST.3F.SG stew

'She is cooking a stew.' (Abu-Haidar 1991)

In some NENA dialects, including the Jewish dialect of Sulemaniyya, a present indicative is marked by a prefix k-. This derives from an earlier prefix qa-, used in older dialects of Aramaic to mark a continuous or habitual present, which ultimately derives from a form qaʔem, a participial form of the verb 'stand'.

In most Ethiopian languages, the inherited non-past form can be combined with a form of the verb 'be'. This compound form indicates the non-past in a main clause, while the inherited, simple non-past is used in subordinate or negative clauses. Thus in many Ethiopian languages, including Amharic, the present and future are not normally distinguished.

Future Tenses. Explicit future tenses have developed in a number of Semitic languages. Often markers of the future derive from grammaticalized forms of verbs meaning 'go' or 'want', though other sources are well attested (Rubin 2005). For example, in Egyptian Arabic (and several other Arabic dialects), the future-marking prefix ha- is derived from the participle rdyih 'going' (32). Several dialects of NENA have a future-marking particle ba- (variants ba-, b-, d-), which is derived from an earlier Aramaic construction bāfe d-'want that' or bīfe d-'it is desired that' (33).
Egyptian Arabic
(32) ha-yirga
bukra
FUT-return:NONPAST.3M.SG tomorrow
‘He will return tomorrow.’ (Abdel-Massih et al. 1979)

NENA (Barwar)
(33) qaṭuθ-i baḍ-mnagra-lux
cat:1SG FUT-gnaw:PREs.3F.SG-OBJ.2M.SG
‘My cat will gnaw you.’ (Khan 2008a)

In Mehri (and the closely related dialects of Ḥarsusi and Baṭhari), the active participle has come to indicate future tense, with the result that the inherited non-past form has become the basic present tense.

3.4.4 Mood

All Semitic languages possess an imperative form (34), though in nearly all of the languages, imperatives occur only in positive commands. Negative commands are normally expressed by negating another verbal form, normally the non-past (35) or the jussive (on this term, see below). Negation of the actual imperative form is found only in some Neo-Aramaic dialects, e.g., in the Jewish NENA dialect of Arbel (36), but not in the Christian NENA dialect of Qaraqosh.

Old Babylonian Akkadian
(34) šupur
send:IMPER.M.SG
‘Send!’

(35) là tašappar
NEG send:NONPAST.2M.SG
‘Don’t send!’ (Huehnergard 2005b)

NENA (Arbel)
(36) la sī
NEG go:IMPER.SG
‘Don’t go!’ (Khan 1999)

There are no first or third person imperatives in Semitic, though there is a jussive mood that fulfills the function of the imperative for these persons. The jussive in Proto-Semitic seems to have been identical in shape to the perfective verbal form. In Akkadian, the jussive (normally termed the “precative” in grammars of Akkadian) is distinguished from the inherited perfective by the addition of an asseverative (or topicalizing) particle l(V)- (Huehnergard 1983; Testen 1993); cf. Akkadian ḫillik ‘he went’ and ḫ-illik ‘let him go, may he go’. In West Semitic, the inherited perfective (of the shape prefix + CCVC) lost its function as the normal indicator of past tense (see above, §1, §3.4.3), but the form itself survived in its jussive function. This is most evident in Modern South Arabian and in Ethiopian Semitic; cf. Ge’ez nabara ‘he sat’ and jussive yanbar ‘let him sit, may he sit’.

The situation in the rest of West Semitic, i.e., Central Semitic, is slightly more complex. The inherited perfective form, in combination with a suffix *-u (originally a marker of subordination) came to indicate non-past, replacing the inherited non-past form (Rubin 2005). The result is that in Central Semitic there was a form of the shape prefix + CCVC that indicated the jussive, and a form prefix + CCVC-u that indicated non-past. This situation is found in Classical and Modern Standard Arabic. As in Akkadian, the Arabic jussive (when used as a true jussive, i.e., as a first or third person imperative) is preceded by an asseverative particle, e.g., indicative non-past yašrāb-u ‘he drinks’ vs. jussive l-yašrāb ‘let him drink, may he drink’. In some Central Semitic languages, including modern Arabic dialects, Hebrew, and Aramaic, the final short -u of this innovative non-past form was lost, with the result that the non-past and jussive merged in most environments. So, in Biblical Hebrew, for example, the jussive is most often not morphologically distinct from the non-past, and a jussive meaning must be gleaned from context. Compare examples (37) and (38), both of which contain the identical verbal form yišpōt.

Biblical Hebrew
(37) hā yišpōt tēbēl ba-sedeq
he judge:NONPAST.3M.SG world with-righteousness
‘He judges the world with righteousness.’ (Psalm 9:9)
54 \( \text{A BRIEF INTRODUCTION TO THE SEMITIC LANGUAGES} \)

(38) \( ȳśp̄āt \) \( \text{YHWH} \) \( bēn-i \)
judge:Juss.3.M.SG DN between-1SG
\( ū-bēn-ekā \)
and-between-2.M.SG

‘May the LORD judge between me and you.’ (Gen. 16:5)

In addition to the indicative and jussive moods, Arabic possesses a subjunctive mood, formed with the same verbal base plus a suffix -\( a \) (e.g., \( yašrāb-a \)). The subjunctive is used in subordinate clauses, most often following subordinating conjunctions. In Akkadian, subordination is marked by attaching the suffix -\( u \) to verbs (20, in §3.2.3); this is the same suffix that in Central Semitic marks the indicative.

The jussive and subjunctive moods of the written language are absent from the modern Arabic dialects, due, as already noted, to the loss of final short vowels. However, there is some distinction of mood by innovative means. In some of the dialects in which a present tense marker has developed, lack of this marker can indicate a subjunctive or jussive. In example (39), we see the Syrian Arabic present tense marker \( b- \) used in an indicative phrase. In example (40), on the other hand, we see how the absence of this (or any other) pre-verbal marker can indicate a jussive meaning.

Syrian Arabic
(39) \( b-\text{i-nāmu} \)
\( \text{Sal-ṣāṣāḥ} \) \( bō-lālēl \)
pres-sleep:nonpast.3.M.PL on-the-roof in-the-night

‘They sleep on the roof at night.’

(40) \( yāšrāb \) \( bēt-o \)
\( \text{be.ruined:nonpast.3.M.SG} \) \( \text{house-3.M.SG} \)

‘May his house be ruined!’ (Cowell 1964)

The situation in many Neo-Aramaic dialects parallels that of Syrian Arabic, which is to say that a subjunctive or jussive can be marked by the absence of an overt present or future tense prefix.

Akkadian exhibits another verbal form, known as the “ventive”, that is sometimes called a mood by Semitists (e.g., Moscati 1964), presumably with the idea that “mood” covers any type of verbal modification. The ventive is really a type of verbal deixis, with no modal functions. In Old Babylonian Akkadian, the ventive suffix has the allomorphs -(a)m and -nim, and may be attached to any finite verbal form. It gives the sense of direction or activity towards the speaker: contrast illik ‘went’ and ventive illik-am ‘he came’; ublii ‘they brought (there)’ and ventive ublii-nim ‘they brought here’; and uṣi ‘I went out’ and ventive uṣi-am ‘I came out here’ (von Soden 1995). Not coincidentally, the first common singular dative pronominal object suffix also has the allomorphs -(a)m and -nim in Akkadian, e.g., taddin-am ‘you gave to me’. It is unclear if the Akkadian ventive has parallels in West Semitic. It has been argued, rather convincingly, that the suffix -\( a \) that can be attached to Biblical Hebrew imperatives has a ventive function (Fassberg 1999). On other connections between the Akkadian ventive and West Semitic morphemes, see the recent study of Hasselbach (2006).

3.4.5 PHRASAL VERBS

Phrasal verbs (also known as composite verbs) are a characteristic feature of the modern Ethiopian Semitic languages. In these constructions, which have developed under Cushitic influence, a verbal concept is expressed by a fixed element and a conjugated form of the verb ‘say’ or, much less often (and not in all the languages), ‘make’, ‘name’, or ‘become’. Examples are Amharic k‘uč‘t‘ \( \text{alā} \) ‘he sat down’, dōss \( \text{alāw} \) ‘he was happy’ (lit. ‘it was pleasing to him’), and \( \text{zəwə} \) \( \text{alā} \) ‘he was quiet’, all based on \( \text{alā} \) ‘he said’ (Leslau 1995); Wolane \( \text{fənčiši} \) \( \text{balā} \) ‘he sneezed’, \( \text{lodd} \) \( \text{balā} \) ‘he was slow’, and \( \text{giff} \) \( \text{balā} \) ‘it smoked (intrans.)’, all based on \( \text{balal} \) ‘he said’ (Meyer 2006); and Tigré \( \text{kaf} \) \( \text{balā} \) ‘he sat’, \( \text{koy} \) \( \text{balā} \) ‘he hurried’, and \( \text{bab} \) \( \text{balā} \) ‘he sweated’, all based on \( \text{balā} \) ‘he said’ (Raz 1983). Only the verb is conjugated, as in Tigré \( \text{koy} \) \( \text{attività} \) ‘I hurried’, \( \text{koy} \) \( \text{attività} \) ‘we hurried’, \( \text{koy} \) \( \text{attività} \) ‘they hurried’, \( \text{koy} \) \( \text{attività} \) ‘they hurry’, and \( \text{koy} \) \( \text{attività} \) ‘hurry!’. Many of the fixed elements that are part of these phrasal verbs do not occur independently, and so it is often not possible to give the fixed elements an independent meaning.

A parallel to the Ethiopian phrasal verbs can be found in some Neo-Aramaic dialects. In the Jewish NENA dialect of Sule-
maniyaa, for example, we find phrasal verbs that have a verbal element (most often 'do', 'become', 'give', 'come', or 'bring') and a fixed element (Khan 2004). The fixed element is normally a loanword, and the phrasal verb itself is in most cases a calque from Kurdish. Examples are ſa раqa wala 'sweat' and rani wala 'hunt' (from wala 'to do'); tiff e dyaya 'spit' and skita dyaya 'stab' (from dyaya 'to give', itself a Kurdish loan); and mirxew ya?a 'snore' and jigrew ya?a 'hate, be angry' (from ya?a 'to come'). In Neo-Mandaic, we find similar phrasal verbs, which are usually calques from Persian (Häberl 2009).

3.5 Adverbs

Semitic languages normally have a limited set of words that function only as adverbs. Each language has numerous demonstrative adverbs (e.g., Yemeni Arabic hāna 'here', hānāk 'there'), adverbs of manner (e.g., Syriac ḥākanā 'thus'), and adverbs of time (e.g., Hebrew ḥētmēl 'yesterday'). However, most languages do not have a productive means of deriving adverbs, and so adverbial expression is often done by means of a prepositional phrase (41), or even an unmarked adjective (42). Alternatively, in a number of the languages, an adverbial sense can be expressed through verb serialization, sometimes termed "hendia dys" in Semitic grammars (43).

Modern Hebrew

(41) ba-simxs, b-often tivši, bli kavana
with-happiness, with-manner natural, without intention 'happily, naturally, unintentionally'

(42) hi mädaberet barur
she speak:FRES.F.SG clear[ADJ].M.SG
'She speaks clearly.' (Coffin and Bolozky 2005)

Ge’ez

(43) k’adam-ku baš’āh-ku
precede:PAST-1SG arrive:PAST-1SG
'I arrived early.' (Tropper 2002a)

Ethiopian Semitic languages, like Amharic, are particularly fond of using verbal phrases to express adverbial connotations (Leslau 1995).

In Akkadian, we do find an identifiable adverbial marker, namely the suffix -iš, which, as already mentioned above (§3.2.2), can be used to form adverbs from adjectives, for example, damk’iš ‘well’, from damik’ ‘good’.

In Syriac and Christian Palestinian Aramaic, we also find a productive adverbial morpheme. The suffix -alit forms adverbs of quality and manner from adjectives and some nouns. Examples are Syriac 謇ppirāl/it 'beautifully' (cf. 謇ppir ‘beautiful’), ṣalāḥāl/it ‘divinely’ (cf. ṣalāḥ ‘God’, ṣalāḥay ‘divine’), and sūrāl/it ‘in Syriac’ (cf. sūrā ‘Syria’). This suffix is known also from Classical Mandaic and Jewish Babylonian Aramaic, though it is not common in either dialect. On the derivation of this Aramaic adverbial suffix, see Gensler (2000) and Butts (2010), the former of which also has some discussion of the Akkadian adverbial -iš.
4 SYNTAX

4.1 WORD ORDER

Proto-Semitic seems to have been a VSO language, that is, the standard word order was Verb-Subject-Object. This is the usual word order for a number of the ancient languages, including Biblical Hebrew, Arabic (44), and Ge'ez. In addition, modifiers such as adjectives (45), genitives (46), and relative clauses (47) followed their head noun.

Classical Arabic
(44) xarajati l-mar7at-u mina l-bayt-i
   go.out:PAST.3F.SG the-woman-NOM from the-house-GEN
   'The woman went out from the house.'

(45) ?al-bayt-u l-bârid-u
    the-house-NOM the-cold-NOM
    'the cold house'

Biblical Hebrew
(46) derek ham-melek
    highway[CONSTR] the-king
    'the king’s highway’ (Num. 20:17)

Ge'ez
(47) ba?si za-mota
    man.NOM REL-die:PAST.3M.SG
    'the man who died'

A number of the languages, both ancient and modern, deviate from the standard Semitic type. Akkadian adopted SOV word order as a result of contact with Sumerian, though all
modifiers still normally follow their head noun (Deutscher 2000). Modern Hebrew, Neo-Aramaic languages, and a number of modern Arabic dialects (e.g., Iraqi, Chadian) exhibit SVO word order (on Arabic, see Brustad 2000). In Modern Hebrew, this is the result of general European influence. In these languages, as in Akkadian, modifiers still follow their head noun.

The most drastic changes in word ordering are found in the modern Ethiopian languages. All modern Ethiopian Semitic languages exhibit SOV word order (48), as a result of Cushitic influence (Leslau 1945a). In addition, nearly all, if not all, the languages place adjectives, genitives (49), and relatives (50) before their head nouns.

Zay

(48) äya t-estate tii-set miit'iihu
   I with-that:F.SG with-woman come:PAST.1SG
   ‘I came with that woman.’

(49) yä:šum-i gär
      of-chief-the house
   ‘the chiefs house’

(50) gaŋi-y yä-gan-e bïyu
      dog-ACC REL-find:PAST.3M.SG-the child
   ‘the child who found the dog’ (Leslau 1999)

4.2 POSITIONAL RELATIONS

All Semitic languages have prepositions, though some languages also have postpositions and circumpositions. Prepositions are the norm in most Semitic languages, including all of the ancient languages. In languages that have retained case marking, prepositions always govern the genitive case. Semitic prepositions are of two types. There are those that are cliticized to their head noun (51) and those that are treated as separate units.

The former type usually number only three to four in a given language, with the great majority of prepositions treated as separate units.

Zay

(51) bi-l-bayt-i
       in-the-house-GEN
   ‘in the house’

(52) mina l-bayt-i
       from-the-house-GEN
   ‘from the house’

Some modern Ethiopian languages, on the other hand, use the independent forms of the pronouns with prepositions (as well as with postpositions and circumpositions); cf. Amharic wäändä ‘to me’ (¼ wäändä ‘to’ + ane ‘I’) and käsäš ‘from him’ (¼ käš ‘from’ + ass ‘he’).

In modern Ethiopian Semitic languages, there is a greater tendency towards circumpositions and postpositions. The postpositions are often the result of grammaticalized nouns, for which there was a preceding genitive phrase used as a modifier (53–54). The prepositional elements of the circumpositions in Amharic can be omitted in both speech and writing, resulting in a postposition (Leslau 1995).

Amharic

(53) ba·bet wast (< *ba·yi·bet wast)
       PREP-house in (< in-of-house interior)
   ‘in the house’ (< ‘in the interior of the house’)

In nearly all of the languages, the expression of the pronominal object of a preposition is by means of a pronominal suffix attached to the preposition (see §3.1.2). For example, following are the forms of the Omani Mehri preposition h- ‘to, for’ (‘to me, to you’, etc.).

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st C</td>
<td>háymi</td>
<td>hâki</td>
<td>hin</td>
</tr>
<tr>
<td>2nd M</td>
<td>hâk</td>
<td>hâki</td>
<td>hîkom</td>
</tr>
<tr>
<td>2nd F</td>
<td>hayš</td>
<td>hâki</td>
<td>hîkan</td>
</tr>
<tr>
<td>3rd M</td>
<td>hâh</td>
<td>hâhi</td>
<td>hîham</td>
</tr>
<tr>
<td>3rd F</td>
<td>his</td>
<td>hîsi</td>
<td>hîsan</td>
</tr>
</tbody>
</table>
In some South Ethiopian languages, if a noun is preceded by a modifier, a preposition is repeated for each element (48, in §4.1). This was also the case in Soqotri a hundred years ago (55). However, in modern Soqotri, only the second occurrence of the preposition has survived, with the result that prepositions can be embedded within a noun phrase (56) (Lonnet 1998).

Soqotri (circa 1900)
(55) min di-heh min amboriye
from of-he from children
‘from his children’

Soqotri (circa 1990)
(56) di-hoh k-amboriye
of-I with-children
‘with my children’

4.3 AGREEMENT

There are a number of issues pertaining to agreement in Semitic that are noteworthy. The case of the numerals, which seem to exhibit reverse gender concord, was already discussed above (§3.3). It was also already noted above (§3.2.1) that in languages in which the dual has ceased to become a productive category (e.g., Hebrew), fossilized dual forms are treated as plural for agreement purposes.

Attributive adjectives in Semitic normally agree with their head nouns in gender, number, and case (where applicable, and referring to the inherited Proto-Semitic case markers only); in Central Semitic (with the exception of Neo-Aramaic) and in Mehri, attributive adjectives also agree in definiteness (57). A noun with a possessive suffix is considered definite for the purposes of agreement (58). As a general tendency, agreement rules are less strict in Ethiopian Semitic.

Classical Arabic
(57) ?al-mar?at-u l-hasan-at-u
the-woman-NOM the-beautiful-F.SG-NOM
‘the beautiful woman’

Biblical Hebrew
(58) bêt-okā hag gādōl
house-2M.SG the-big
‘your big house’

In many Semitic languages, a distinction is made in the agreement of attributive and predicative adjectives. For those languages in which attributive adjectives agree in definiteness (i.e., Central Semitic languages [except Neo-Aramaic] and Mehri), predicative adjectives do not do so. Compare the following examples (59–60) with the two phrases above (57–58).

Classical Arabic
(59) ?al-mar?at-u hasan-at-un
the-woman-NOM beautiful-F.SG-NOM.INDEF
‘The woman is beautiful.’

Biblical Hebrew
(60) bêt-okā gādōl
house-2M.SG big
‘Your house is big.’

In Akkadian, predicative adjectives are marked with a special set of gender and number markers. Moreover, there is a special set of enclitic personal pronouns that are attached to the adjective—which itself has no case, gender, or number markers—to indicate predication (61). Because suffixes have developed for all persons, this construction is often referred to as a stative “tense”. Though this construction is not verbal in origin, it is often convenient to think of it as verbal in Akkadian. This is because the predicative adjectives have verbal properties, in that they are associated with verbal roots; indicate person, number and gender; can take the marker of subordination; and can take a ventive morpheme. Indeed, in West Semitic, this developed into a true past tense (as noted in §1).
A Brief Introduction to the Semitic Languages

Old Babylonian Akkadian
(61) awil-um marus'-O, awil-um mars'-at, mars'-āku
man-NOM sick-3M.SG, woman-NOM sick-3F.SG, sick-1SG
'the man is sick, the woman is sick, I am sick'
(Huelnergard 2005b)

Agreement in Classical and Modern Standard Arabic is slightly more complex than in the other languages. The above rules given for attributive and predicative adjectives apply when the noun is singular or dual. When the noun is plural (and occasionally also when it is a dual), agreement is dependent on animacy. For nouns referring to humans, agreement follows that of singular nouns, so masculine plural nouns referring to humans take masculine plural adjectives and feminine plural nouns take feminine plural adjectives (62). For all other nouns, plurals are treated as feminine singular (63).

Modern Standard Arabic
(62) raz-zuwwār-u
the-visitors[M.PL]-NOM
'the official visitors'
(63) taqālīd-un māshīyyat-un
traditions[M.PL]-NOM.INDEF Christian-F.SG-NOM.INDEF
'Christian traditions' (Ryding 2005)

Agreement between nouns and verbs in Classical and Modern Standard Arabic exhibits similar idiosyncrasies. When a verb precedes its subject (as is normally the case), its form is always singular, even when the subject is dual or plural. The gender of the verb is determined by whether or not the subject refers to humans. Masculine dual or plural subjects referring to humans are preceded by a masculine singular verb (64), while non-human masculine dual or plural subjects (usually) and all feminine subjects are preceded by feminine singular verbs (65). For singular nouns, gender concord between nouns and verbs is strict.

Modern Standard Arabic
(64) dahlī-
laugh:PAST-3M.SG the-students[M.PL]-NOM
'The students laughed.' (Ryding 2005)
(65) ḫitaqāt-ati
agree:PAST-3F.SG the-delegations[M.PL]-NOM
'The delegations agreed.' (Mace 1998)

4.4 Comparison

In the Semitic languages, we find both analytical (syntactic) and inflectional (morphological) means of comparison, though the former is more common. In Arabic, which has an inflectional comparative, the comparative form of the adjective has the basic pattern ḥaSSāC, as in tawāl 'long, tall', ṭatwal 'longer, taller'; pattern ṣaacāC, as in ṭawāl 'long, tall', ṭatwal 'longer, taller'; ḥāsan 'beautiful', ṭāḥsān 'more beautiful'. The comparative form (normally called in Arabic grammars the "relative" form) can also denote an intensive, i.e., 'very tall', 'very beautiful', as required by context. When it occurs with the definite article (66) or as the first member of a genitive chain (67), the comparative form indicates a superlative.

Classical Arabic
(66) ṭaR-rājul-ū l-ṭattwal-ū
the-man-NOM the-tall:COMPAR-NOM
'the tallest man'
(67) ṭattwal-ū r-rijiili
tall:COMPAR-NOM[CONSTR] the-men-GEN
'the tallest of the men'

A very few adjectives of the pattern ḥaSSāC are found lexicalized in Biblical Hebrew (e.g., ḥāsār 'cruel'), but it is unclear if these are archaisms or borrowings. Comparison in Hebrew, like most other Semitic languages, is purely analytical. That is, there is no morphological form indicating comparison (68).

Biblical Hebrew
(68) ūm gādōl wā-rām miinmēn-nū
people great and-tall from-1PL
'a people greater and taller than us' (Deut. 1:28)
The Arabic comparative pattern \( ?aCCaC \) is likely related to the Akkadian pattern \( šuCCuC \), which can give the sense of 'very' or 'most' to adjectives, e.g., \( šurḫu ‘very great’ from \( rôbû ‘great’ \) (Speiser 1952). The extension of this type of intensive pattern into a true inflectional comparative is likely an internal development within Arabic. Some languages—Neo-West Aramaic and Mehri, for example—have borrowed the inflectional comparative pattern \( ?aCCaC \) from Arabic (though in Mehri it is rather rare).

4.5 Coordination

Coordination of elements or phrases is normally indicated by means of a proclitic or enclitic particle. The most common of these is the prefixed particle '\( wa- \) (in some languages realized as \( wo- \) or \( u- \)), which is found in nearly all Semitic languages, with the notable exception of some modern Ethiopian languages (69).

In a few South Ethiopian languages, \(-wa\) has become suffixed (70).

Chadian Arabic

\( al\text{-}gitt \ wa\text{-}l\text{-}jidâd \)

'the cat and the chickens' (Jullien de Pommerol 1999)

Zay

\( bok’olu-wa\text{-}farâz lâm bûrâ \)

'a mule, a horse, a cow, and an ox' (Leslau 1999)

The enclitic coordinating particle \( *-ma \) is known from Akkadian and many Ethiopian languages (71). In some Central Semitic languages, most notably in Ugaritic (Tropper 2000a), this particle seems to be a sort of topicalizer, as it also often can be in Akkadian.

Amharic

\( ē\text{-}aw bârbârre\text{-}mm \)

'salt and pepper' (Leslau 1995)

In Akkadian, the coordinators \( u \) ('\( *wa\)') and \( -ma \) have slightly different functions. For example, two clauses coordinated with \( u \) can be reversed without altering any meaning; not so for clauses coordinated with \( -ma \) (Huehnergard 2005b). Slightly different semantic distinctions are made in Ethiopian languages that have both of these particles (e.g., Zay; see Meyer 2005).

The proclitic coordinating particle '\( pa- \)', indicating sequential coordination ('and then'), is found in some Central Semitic languages, namely Ugaritic, Arabic, Śayhadic, Sam’ālian, and a few Aramaic dialects. In a number of Ethiopian languages, we find an enclitic coordinator with an enclitic element \( n \), including Amharic \(-nâ\) and Tigrinya \(-n\).

Although every Semitic language possesses at least one coordinating conjunction, asyndetic coordination is also well attested in the family (72–73).

Tigrinya

\( 72 \)

\( râ\text{-}ay-ā \ sâmî\text{-}e \)

\( \text{see:PAST\text{-}1SG hear:PAST\text{-}1SG} \)

'I saw and I heard' (Leslau 1941)

Egyptian (Imperial) Aramaic (unvocalized)

\( 73 \)

\( ?elt\text{-}t hûkh-t l-?hyqr \)

\( \text{go:PAST\text{-}1SG find:PAST\text{-}1SG ACC-Ahiqar} \)

'I went and found Ahiqar' (Muraoka and Porten 2003)

4.6 Copular Clauses

In the classical Semitic languages, and in many modern Semitic languages, there is no expressed copula, at least not in the present tense (74–75). This situation can be presumed for Proto-Semitic.

Biblical Hebrew

\( 74 \)

\( mi \ ?attâ \)

\( \text{who YOU:MSG} \)

'Who are you?' (Gen. 27:18)
68 A BRIEF INTRODUCTION TO THE SEMITIC LANGUAGES

Omani Mehri
(75)  ❣māh ḥayr-i u tēθ tēθ-i
thiSM.SG donkey-lsG and woman woman-1sG
'This is my donkey, and the woman is my wife.'  
(Rubin 2010b)

Most languages make use of a verb 'to be' in the past and future
tenses, but the individual verbs vary (i.e., the roots used) (76--77).

Biblical Hebrew
(76)  sebed hāyitā  bo-ʔeres  miśrayim
slave be:PAST.2M.SG in-land[CONSTR] Egypt
'You were a slave in the land of Egypt.'  
(Deut. 15:15)

Omani Mehri
(77)  hēt wak'ūna  a-mbk-h4m
you:SG be:FUT.M.SG the-king-3M.PL
'You will be their king.'  
(Rubin 2010b)

In many of the languages that normally do not express a
copula in the present tense, a demonstrative or personal pro­
noun can fill the spot of a copula, in order to make clear that a
separate subject and predicate are being expressed (78–79).

Ge'eez
(78)  zantu waʔatal nugū
thiSM.SG he  king
'This is the king.'

Modern Hebrew
(79)  našim hen mi-naga  u-gvarim hem mi-madim
women they:F from-Venus and-men they:M from-Mars
'Women are from Venus and men are from Mars.'  
(Coffin and Bolozky 2005)

A number of modern languages, including modern Ethio­
pian languages, Neo-Aramaic languages, and some Arabic dia­
lcts, have developed present-tense copulas, most often from
grammaticalized demonstratives or presentative particles (Rubin
2005).

4.7 EXISTENTIAL CLAUSES

Many Semitic languages have developed existential particles,
the sources of which are varied (Rubin 2005). Some of these
particles derive from locative prepositions. For example, a num­
ber of Arabic dialects (e.g., Syrian, Egyptian, Chadian) use fi
'there is/are' (<fi-hi 'in it'), and Ge'eez uses bo 'there is/are' (lit.
'in it'). A few languages have particles that are fossilized forms
of the verb 'to be', including Moroccan Arabic kāna and, most
likely, Iraqi Arabic ʔaku (both derived from forms of the Arabic
verb kāna 'to be'). The origin of the existential particles in the
Northwest Semitic languages—Hebrew yēl, Aramaic ʔit, Ugaritic
ʔlē, etc.—is not totally clear, though they are almost certainly
related to the Akkadian verb hāmu 'have'. It is quite plausible
that 'have' is the original meaning of this root, and that the exis­
tential meaning is secondary; cf. the French construction il y a.

In some languages, there is a special negative existential parti­
cle, for example, Biblical Hebrew len 'there is/are not' (vs. yēl
'there is/are').

Even languages that have a special existential particle often
make use of the simple verb 'to be' to express existence in the
past and future (80–81).

Biblical Hebrew
(80)  rimmā  lō  hāyitā  b-ō
maggot:COLL NEG be:PAST.3F.SG in-3M.SG
'There were no maggots in it.'  
(Exod. 16:24)

(81)  lō  yihye  maṭār
NEG be:FUT.3M.SG rain
'There will be no rain.'  
(Deut. 11:17)

A few Arabic dialects form existential constructions with the
adverb 'there', just as in English (82).

Maltese
(82)  hemm ilpup f'Malta
there wolves in-Malta
'Are there wolves in Malta?'  
(Ambros 1998)
Sometimes, as in Mehri, an existential is simply understood, and no verb, adverb, or particle is used (83; see also 93, in §4.9)

Omani Mehri
(83) xaurat hokam b-arhabet
once ruler in-town
'once there was a ruler in a town' (Rubin 2010b)

4.8 POSSESSION

It has already been noted above (§3.1.1, §3.1.2) that pronominal possession in Proto-Semitic was indicated by means of a pronominal suffix, and that some languages have developed an independent possessive adjective. It was also noted that a genitive relationship between two nouns was indicated by means of a construct phrase (genitive chain), but that some languages have developed an independent genitive exponent. It is unclear whether Proto-Semitic had a verb corresponding to English 'have'. Akkadian has the verb *iššu* 'have', which is cognate with an existential particle in West Semitic (e.g., Hebrew יִהְיֶה, Ugaritic ḫθ). As mentioned above (§4.7), we cannot know for sure whether an original existential developed into a verb of possession in Akkadian, or whether a verb of possession developed into an existential in some West Semitic languages, since both scenarios are easily conceivable. Regardless, the West Semitic languages make use of various phrasal constructions to express the equivalent of 'have'. These constructions normally contain a locative (84), comitative (85), or dative preposition; the last of these is sometimes found in conjunction with an existential particle (86) or a verb of existence. In Ethiopian languages, we find a construction consisting of an existential particle or a verb of existence plus a suffixed object pronoun (87).

Maltese
(84) omm-i ghand-ha qattus
mother-1sg at-3sg.SG cat
'My mother has a cat.' (Ambros 1998)

4.9 INTERROGATIVES

All Semitic languages exhibit interrogative pronouns and adverbs, which are, on the whole, typologically unremarkable. Occasionally, interrogatives show some syntactic peculiarities. For example, in Egyptian Arabic, interrogatives normally occupy the position in the sentence occupied by the questioned constituent. The result is that interrogatives in Egyptian Arabic—usually standing for direct objects, indirect objects, or adverbial clauses—are often phrase-final (88–89). In most other Arabic dialects, as in other Semitic languages, interrogatives are normally fronted (90).

Egyptian Arabic
(88) tarē-t il-kitāb lē read:past:2sg the-book why
'Why did you read the book?' (Abdel-Massih et al. 1979)

(89) ūwēs ē want:pres:msg what
'What do you want?' (Woidich 2006)

Syrian Arabic
(90) lē šīf-t why go.out:past:2sg
'Why did you go out?' (Cowell 1964)
A number of languages also possess a special interrogative particle, used to turn a statement into a question. Often such particles are proclitic or enclitic elements. Examples are Classical and Modern Standard Arabic fa-, Biblical Hebrew hā-, Ge'ez -nu and -hu (91), Tigrinya -do (92), and Amharic -na. The Arabic, Hebrew, and Ge'ez particles are all attached to the first constituent in the sentence, while the Tigrinya and Amharic particles are sentence-final. Examples of non-clitic interrogative particles are Classical and Modern Standard Arabic hal, Hebrew (mainly post-Biblical) ha?im, and Mehri wālē (93). Most of these particles, whether clitic or non-clitic, are optional. Some are restricted to certain types of questions (e.g., yes-or-no questions).

Ge'ez
(91) santo-nu gobar-ki
this:ACC-INTERROG DO:PAST.2F.SG
'Did you do this?' (Dillmann 1907)

Tigrinya
(92) lomi mbrarti ?allo-do
today school EXIST-INTERROG
'Is there school today?' (Melles 2001)

Omani Mehri
(93) wālē šī rāyd
INTERROG any sardine.COLL
'Are there any sardines?' (Rubin 2010b)

4.10 RELATIVE CLAUSES

In Proto-Semitic, relative clauses could be asyndetic, in which case the antecedent was in the construct state, or syndetic, in which case the antecedent was followed by a determinative-relative pronoun that declined for gender, number, and case. The former type is found in Akkadian (20, in §3.2.3) and in Ṣayhadic, with vestiges in Ge'ez, Hebrew, and elsewhere. The latter type is widespread in Semitic, though the inherited relative has in many languages become indeclinable (i.e., a single frozen form survives). For example, while a fully declinable relative pronoun existed in Old Akkadian (and Eblaite), in Old Babylonian Akkadian, the masculine singular accusative form fa became an indeclinable relative pronoun; in Aramaic, the masculine singular genitive form di (later dī) underwent the same development (94).

Biblical Aramaic
(94) hēn-ā dī hāzēt
dream-the REL see:PAST.1SG
'the dream that I saw' (Daniel 2:26)

In the Canaanite dialects—most importantly, Hebrew and Phoenician—a noun "qābar 'place' has been grammaticalized as a relative pronoun (95), replacing the inherited forms, though the Proto-Semitic relative is attested as such in some archaic Biblical Hebrew passages and in archaic Byblian Phoenician. Huehnergard (2006a) presents a very detailed discussion of this development.

Biblical Hebrew
(95) elle had-sēbirin qāser dibber mōse
these the-words REL speak.PAST.3M.SG Moses
'These are the words that Moses spoke.' (Deut. 1:1)

Relative clauses in Arabic exhibit some interesting syntactic restrictions. The relative pronoun itself, which in the Classical and Modern Standard varieties declines for gender and number (and in the dual, also for case), is an Arabic innovation; the inherited Proto-Semitic forms survive in Arabic only as a determinative pronoun ("the one of"). But the relative pronoun is only used when the antecedent is definite (96). When the antecedent is indefinite, the relative clause is asyndetic (97). This type of asyndetic clause, with an indefinite antecedent in the free (unbound) state, is also found occasionally in Ge'ez, Mehri, and elsewhere, though only in Arabic is it the rule.

Modern Standard Arabic
(96) ʔas-siyyāh-u llaOina yaeiliina
the-tourists-NOM REL:M.PL arrive:NPAST.3M.PL
'the tourists who arrive'
A BRIEF INTRODUCTION TO THE SEMITIC LANGUAGES

(97) rajul-un yamaliku š-ṣajāfat-a
man-NOM.INDEF possess:NONPAST.3M.SG the-courage-ACC
'a man who possesses courage' (Ryding 2005)

In modern Arabic dialects, the relative pronoun is usually a single, indeclinable form, but relative clauses exhibit the same distinction based on whether or not the antecedent is definite or indefinite (98–99).

Iraqi (Muslim Baghdadi) Arabic
(98) l-ḥsān ʿlli ẓylab haš-ṣat
the-horse REL win:PAST.3M.SG this-race
'the horse that won this race'

(99) ibnayya tihēṭi xamis luyāt
girl speak:PRES-3F.SG five languages
'a girl who speaks five languages' (Erwin 1963)

When a Semitic relative pronoun stands in a prepositional relationship, a resumptive pronoun is normally employed. In some languages this is optional, but most often it is obligatory (100). Resumptive pronouns are also found even when the relative pronoun stands for the direct object, and in fact, such use of a resumptive pronoun is obligatory in some languages, including many Arabic dialects (101). Only very rarely do we find a preposition preceding a relative pronoun (like English 'in which'); Ge’ez is a language in which this is possible, though still very infrequent (102).

Biblical Hebrew
(100) hā-ḥāres ʿasher gar-tā b-Āh
the-land[F] REL live:PAST-2M.SG in-3F.SG
'the land in which you lived' (Gen. 21:23)

Egyptian Arabic
(101) ir-rāgīl ʿlli šuf-t-u imbārī ḫ
the-man REL see:PAST-2M.SG-3M.SG yesterday
'the man that you saw yesterday' (Abdel-Massih et al. 1979)

As discussed in §4.1, Proto-Semitic relative clauses followed their antecedents, but in modern Ethiopian Semitic languages, relative clauses precede (50, in §4.1).

4.11 SUBORDINATE CLAUSES

Subordinate clauses in the Semitic languages can be marked either by an explicit conjunction, or, less often, by special syntactic constructions using nominalized verbal forms. We find temporal (103), conditional (104), and causal (105) subordinating conjunctions, among other types.

Omani Mehri
(103) wəzmōna t-ikam fandēl mat
give:FUT.M.SG ACC-2M.PL sweet.potato.COLL when
nakakam
come:PAST.2M.PL
'I will give you sweet potatoes when you come back.' (Rubin 2010b)

Ge’ez
(104) tamma nabar-ka zayya yarakkapu-ka
if stay:PAST-2M.SG here find:FUT.3M.PL-2M.SG
'If you stay here, they will find you.' (Tropper 2002a)

Amharic
(105) sklē-tammāmā al-māt’tā-mm
because-be.sick:PAST-3M.SG NEG-come:PAST.3M.SG-NEG
'Because he was sick, he didn’t come.' (Leslau 1995)

We also find subordinate clauses indicated by means of a preposition plus an infinitive or a verbal noun (106–107).
Biblical Hebrew
(106) lō  yēraf  lōbāb-škā  bā-titt-škā
NEG be.bad:FUT.3M.SG heart-2M.SG PREP-give:INF-2M.SG
lō
to-3M.SG
'Your heart will not be grieved when you give to him (lit.: at your giving).' (Deut. 15:10)

Amharic
(107) mist-u-n  bā-māgdal-u
wife-3M.SG-ACC PREP-kill:INF-3M.SG
tākāssāsā
indict:PASS.PAST.3M.SG
'He was indicted because he killed his wife (lit.: at his killing).' (Leslau 1995)

5 LEXICON

The Semitic languages—at least the classical languages—show a very high proportion of native Semitic roots, with the consequence that the lexicons of the Semitic languages have much in common (Bennett 1998 provides some convenient word-lists). How much the Proto-Semitic lexicon has derived from Afroasiatic is very difficult to determine, since while some of the morphological correspondences between the Afroasiatic branches are reasonably clear, finding regular phonological correspondences has proven very difficult. Therefore, proposed cognates between Semitic and Afroasiatic should always be regarded with a high degree of caution.

As discussed above (§2.3, §3.4), consonantal roots, usually triliteral, form the basis of the Semitic lexicon, since all verbs and nearly all nouns and adjectives are associated with a lexical root. There are, however, a significant number of primary (or isolated) Semitic nouns, which are not associated with a root. Some examples are the Proto-Semitic nouns *tākā 'brother', *dam 'blood', *šapā 'lip', and *tarnab 'hare' (Fox 2003).

Given the history of the regions in which the Semitic languages are spoken, it should be no surprise that borrowing of vocabulary has been frequent. In many cases, the source of borrowing has been another Semitic language. Akkadian was widely used throughout the Near East for much of the second millennium BCE, and because of this, we find a number of Akkadian loans in Biblical Hebrew (Mankowski 2000) and various dialects of Aramaic (Kaufman 1974). Aramaic itself was also long a dominant language in the region, and numerous Aramaic borrowings or calques are found in Arabic (Retsö 2006), all periods of Hebrew (M. Wagner 1966; Kutscher 1982), and elsewhere. Over the last millennium, Arabic has had heavy lexical influence on the Modern South Arabian group (Lonnet 2009).
and on most Neo-Aramaic dialects, and has provided a sizable number of loans to the Ethiopian languages (Leslau 1990; Zaborski 2003).

Influence from non-Semitic languages is found already in Akkadian, which borrowed a large number of nouns (but no verbs) from Sumerian (Lieberman 1977). All other ancient languages possess some non-Semitic vocabulary items, but non-Semitic influence really becomes pervasive in the Hellenistic and Roman periods; Hebrew, Aramaic (especially Syriac), and Ge’ez all have sizable amounts of Greek loanwords (see, e.g., Krauss 1898–99 and Sperber 1984 on Hebrew, and Brock 2007 on Syriac); Arabic has a much smaller, but not insignificant, number of Greek loanwords (Gutas 2007). Large numbers of Persian (usually Middle Persian) loanwords are found in Arabic (Yarshater 1998; Asbaghi 1988) and Aramaic (Ciancaglini 2008). In the modern period, non-Semitic lexical influence is very substantial for a number of languages. All modern Ethiopian Semitic languages have large numbers of Cushitic loans (Leslau 1988), as does Ge’ez, to some extent; many Neo-Aramaic dialects have an abundance of Kurdish loans (Khan 2004, 2008a); Maltese has a copious amount of Italian loans (Aquilina 1976); Moroccan Arabic has scores of French loans (Heath 1989); and Modern Hebrew has countless loans from English. Turkish has provided a fair number of loans into both literary and spoken Arabic, though perhaps less than one would expect, given the centuries of Ottoman dominion over most of the Arab world (Procházka 2004). Studying the lexical inventories of the Semitic languages is really a lesson in the political and cultural history of the Near East.

6 Guide to Further Reading

This section will provide some direction for further research on the Semitic languages. At present, there exists no comprehensive bibliographical work on Semitic. The volumes by Hospers (1973–74) are useful, but well out of date. An extensive online bibliography can be found on the website of the journal Aula Orientalis (http://www.telefonica.net/web2/aulaorientalis/).

It will be observed below that several of the standard grammars of Semitic languages date from the late 19th and early 20th centuries. In these cases, I have also provided more recent sources, in addition to these classic, still valuable works.

General Linguistics: As a basic reference for linguistic terminology, Crystal (2008) and Matthews (2007) are very helpful. For an introduction to historical and comparative linguistics, there are several good books available, including Campbell (2004). Akmajian et al. (2010) and O’Grady et al. (2010) are general introductions to the many subfields of linguistics.

Afroasiatic: Very brief sketches of the non-Semitic branches of Afroasiatic can be found in Hayward (2000) and Huehnergard (2004). The classic survey of Afroasiatic is that of Diakonoff (1988). While this work provides useful data, its reconstructions should not be relied upon. Any comparative grammar or dictionary comparing two or more branches of Afroasiatic must be used with great caution.

General Semitic: The standard comparative grammar of Semitic is probably still that of Brockelmann (1908–13), though it is seriously outdated. A number of works have appeared in recent years, though all of these must be used with some caution. Lipiński (2000) and Kienast (2001) are more comprehensive treatments, while Bennett (1998), Stempel (1999), and Haelewycyck (2007) are much briefer. Moscati (1964), upon which linguists so often rely, and which focuses mainly on the
ancient languages, is outdated. Hetzron (1997), Woodard (2004), and Kaye (1997; 2007) each contain excellent sketches of some of the individual Semitic languages. There is no comparative etymological dictionary of Semitic, though Cohen, Bron, and Lonnet (1994–), with nine fascicles as of 2010 (through the letter $J$, according to the Hebrew ordering of the alphabet) is quite useful. Leslau (1991) is an etymological dictionary of Ge'ez, but it has an index of Semitic roots that makes it a partial substitute for a Semitic etymological dictionary. Miltarev and Kogan (2000; 2005) treat only limited semantic fields.

Akkadian and Eblaite: The standard grammar of Akkadian is that of von Soden (2005). The textbook of Huehnergard (2005b), which focuses mainly on Old Babylonian, is comprehensive enough to be used as a basic reference grammar for most purposes, and is superior to von Soden in its clarity and in its treatment of certain morphological and syntactic features. Hasselbach (2005) is a treatment of Old Akkadian, and Hämeen-Anttila (2000) is a useful work devoted to Neo-Assyrian. A lengthy overview of both Akkadian and Eblaite can be found in Huehnergard and Woods (2004).

Modern South Arabian: Relatively little has been published on the Modern South Arabian languages. A number of grammatical studies were published in the first two decades of the 20th century (see Leslau 1946 and Rubin 2010b for details), but from the rest of the 20th century we have only the syntactical study of E. Wagner (1953) and a few brief grammatical sketches (e.g., Johnstone 1975; Simeone-Senelle 1997; Lonnet 2006). A comprehensive grammar of Mehri, the most widely spoken language of this group, was published in 2010 (Rubin 2010b).

Ethiopian Semitic: For Ge'ez, the classical language of Ethiopia, the standard and most comprehensive reference grammar is that of Dillmann (1907). A more concise, but still very useful grammar was recently published by Tropper (2002a). Lambdin (1978) is a textbook, but is very valuable; in fact, a good amount of the information in Tropper's reference grammar seems to have come from Lambdin. There are grammars published for many, though not all, of the modern Ethiopian Semitic languages. Leslau (1995) is the standard for Amharic, and is an outstanding reference. Leslau has also published grammars, or at least grammatical sketches, for many other Ethiopian languages (e.g., Tigrinya, 1941; Tigré, 1945b; Argobba, 1997; Zay, 1999). Detailed grammars of two Gurage languages have recently been published by Meyer (Zay, 2003; Wolane, 2006). For Tigrinya, the second most widely spoken modern Ethiopian language, there is no recent, comprehensive grammar, though Melles (2001) is a useful short treatment. Raz (1983) is a brief grammar of Tigré. The encyclopedia of Uhlig et al. (2003–), with three of the planned six volumes published (as of 2010), is an excellent reference for all things Ethiopian, including the languages.

Arabic: For a general overview of the various periods of Arabic, see Versteegh (1997) and Ferrando (2001). The standard grammar of Classical Arabic is still that of Wright (1896–98), though the smaller grammar of Fischer (2002) is also useful. On Middle Arabic, Lentin (2008) is a good place to start. Blau (1966–67; 1988; 1999) are also excellent sources. For Modern Standard Arabic, there are a number of very good recent reference grammars, including Ryding (2005) and Mace (1998). The study of modern Arabic dialects is a vast field. Three good handbooks, which provide some overview of the dialects, are those of Fischer and Jastrow (1980), Behnstedt and Woidich (2005), and Corriente and Vicente (2008). Excellent sketches of the dialects, as well as a wide variety of other articles on other Arabic topics, can be found in the recent encyclopedia edited by Versteegh et al. (2006–09). A number of fine Arabic dialect grammars are also published in the series Semitica Viva, published by Harrassowitz. On Arabic-based creoles, see Owens (1997; 2006) and Wellens (2005). On the Ancient North Arabian inscriptional dialects, see Macdonald (2000; 2004).

Sayhadic (Old South Arabian): The only Sayhadic language described at length is Sabaic (also called Sabaean), the phonology and morphology of which are treated in the grammar of Stein (2003). On the stick and palm-leaf texts (also discussed in Stein 2003), see Ryckmans, Müller, and Abdallah (1994) and Stein (2005; 2010). Brief sketches of the other languages of the Sayhadic group can be found in Beeston (1984) and in Nebes and Stein (2004).
Ugaritic, Sam'alian, and Deir 'Allā: Tropper (2000a) is a comprehensive grammar of Ugaritic that has become the standard reference work. The same author also published a much abridged version (2002b), which is also very useful. Watson and Wyatt (1999) is a very useful volume for Ugaritic Studies in general. On the Sam'alian texts, see Tropper (1993), and on the text from Deir 'Allā, see the collection of articles in Hoftijzer and van der Kooij (1991).

Hebrew and Canaanite: The most widely referenced grammar of Biblical Hebrew in English is still that of Kautzsch (1910), though it is a century old. Another widely used grammar, which is more recent, is that of Jotion and Muraoka (2008). For Rabbinic Hebrew, Pérez-Fernández (1999) is a decent introduction; Segal (1927) is useful, but outdated. For Qumran Hebrew, see Qimron (1986). For Modern Hebrew, the most comprehensive grammar is that of Glinert (1989). The more recent works by Coffin and Bolozky (2005) and Glinert (2005), though intended more for students, are also very useful. For a general overview of the history of Hebrew, Kutscher (1982) and Sáenz-Badillos (1993) are very good sources. A multi-volume encyclopedia on the Hebrew language is in production (Khan et al. forthcoming), and this is sure to become a major reference work. For Phoenician, the standard reference grammar is that of Friedrich and Röllig (1999), though Hackett (2004) provides a nice sketch of the language. For the Canaanite of the Tell El-Amarna tablets, see Rainey (1996).

Aramaic: As discussed above (§1.9), Aramaic is the cover term for a large number of languages and dialects. Among the many reference grammars focusing on the ancient dialects, we can cite just a few: Degen (1969) on Old Aramaic; Folmer (1995) and Muraoka and Porten (2003) on Imperial Aramaic; Dalman (1905) on various Jewish dialects; Nöldeke (1904) on Syria; Müller-Kessler (1991) on Christian Palestinian Aramaic; and Macuch (1965) on Mandaic. A wonderful historical overview of the Aramaic family can be found in Brock (2001). Another handy overview of Aramaic, with grammatical sketches and many bibliographical references, is Ferrer (2004). Brock (2006) is an excellent guide to Syriac studies. The study of Modern Aramaic (Neo-Aramaic) has exploded in the last two decades, and there are now excellent grammars of numerous Neo-Aramaic languages. For NENA, noteworthy are those of Khan (1999, 2002, 2004, 2008a, 2008b, 2009); for Turoyo, see Jas-trow (1985); for Neo-West Aramaic, see Arnold (1989–91); and for Neo-Mandaic, see Haberl (2009). There are also more than a dozen volumes (by various authors) devoted to Neo-Aramaic in the series Semitica Viva, published by Harrassowitz. A very nice overview of the many NENA dialects can be found in Khan (2007).
BIBLIOGRAPHY


86  A BRIEF INTRODUCTION TO THE SEMITIC LANGUAGES


BIBLIOGRAPHY


