Cambridge Language Surveys

The Afroasiatic Languages

Edited by Zygmunt Frajzyngier and Erin Shay

CAMBRIDGE

CAMBRIDGE

more information - www.cambridge.org/9780521865333

THE AFROASIATIC LANGUAGES

Afroasiatic languages are spoken by some 300 million people in Northern, Central, and Eastern Africa and the Middle East. This book is the first typological study of these languages, which are comprised of around 375 living and extinct varieties. They are an important object of study because of their typological diversity in the areas of phonology (some have tone, others do not), morphology (some have extensive inflectional systems, others do not), position of the verb in the clause (some are verb-initial, some are verb-medial, and some are verb-final), and in the semantic functions they encode. This book documents this typological diversity and the typological similarities across the languages and includes information on endangered and little-known languages. Requiring no previous knowledge of the specific language families, it will be welcomed by linguists interested in linguistic theory, typology, historical linguistics, and endangered languages, as well as scholars of Africa and the Middle East.

ZYGMUNT FRAJZYNGIER is Professor and former Chair of Linguistics at the University of Colorado, Boulder.

ERIN SHAY is Adjunct Assistant Professor of Linguistics at the University of Colorado, Boulder.

CAMBRIDGE LANGUAGE SURVEYS

General editors
P. Austin (*University of Melbourne*)
J. Bresnan (*Stanford University*)
B. Comrie (*Max Planck Institute for Evolutionary Anthropology, Leipzig*)
S. Crain (*University of Maryland*)
W. Dressler (*University of Vienna*)
C. J. Ewen (*University of Leiden*)
R. Lass (*University of Cape Town*)
D. Lightfoot (*University of Maryland*)
K. Rice (*University of Toronto*)
I. Roberts (*University of Cambridge*)
S. Romaine (*University of Oxford*)
N. V. Smith (*University College, London*)

This series offers general accounts of the major language families of the world, with volumes organized either on a purely genetic basis or on a geographical basis, whichever yields the most convenient and intelligible grouping in each case. Each volume compares and contrasts the typological features of the languages it deals with. It also treats the relevant genetic relationships, historical development, and sociolinguistic issues arising from their role and use in the world today. The books are intended for linguists from undergraduate level upwards, but no special knowledge of the languages under consideration is assumed. Volumes such as those on Australia and the Amazon Basin are also of wider relevance, as the future of the languages and their speakers raises important social and political issues.

Volumes already published include
Chinese *Jerry Norman*
The Languages of Japan *Masayoshi Shibatani*
Pidgins and Creoles (Volume I: Theory and Structure; Volume II: Reference Survey) *John A. Holm*
The Indo-Aryan Languages *Colin Masica*
The Celtic Languages *edited by Donald MacAulay*
The Romance Languages *Rebecca Posner*
The Amazonian Languages *edited by R. M. W. Dixon and Alexandra Y. Aikhenvald*
The Languages of Native North America *Marianne Mithun*
The Korean Language *Ho-Him Sohn*
Australian Languages *R. M. W. Dixon*
The Dravidian Languages *Bhadriraju Krishnamurti*
The Languages of the Andes *Willem Adelaar with Pieter Muysken*
The Slavic Languages *Roland Sussex and Paul Cubberley*
The Germanic Languages *Wayne Harbert*
Sign Languages *edited by Diane Brentari*
The Afroasiatic Languages *edited by Zygmunt Frajzyngier and Erin Shay*

THE AFROASIATIC LANGUAGES

Edited by

ZYGMUNT FRAJZYNGIER

ERIN SHAY

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521865333

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
The Afroasiatic languages / edited by Zygmunt Frajzyngier, Erin Shay.
p. cm. – (Cambridge language surveys)
Includes bibliographical references and index.
ISBN 978-0-521-86533-3
1. Afroasiatic languages. I. Frajzyngier, Zygmunt. II. Shay, Erin.
PJ992.A37 2012
492 – dc23 2012002683

ISBN 978-0-521-86533-3 Hardback

CONTENTS

FIGURES

MAPS

TABLES

NOTES ON CONTRIBUTORS

AZEB AMHA is a researcher at the Leiden University Centre for Linguistics, in the Netherlands. Her publications include a grammar of Maale and papers on various aspects of Wolaitta and Zargulla. Currently she is investigating Oyda, a little-studied Omotic language spoken in south Ethiopia.

ZYGMUNT FRAJZYNGIER is Professor and former Chair of Linguistics at the University of Colorado, Boulder. His main interests are: foundations of syntax and semantics in crosslinguistic perspective, grammaticalization, and Chadic and Afroasiatic linguistics. He is the author or co-author of seven descriptive grammars of Chadic languages and several theoretical volumes.

GENE GRAGG is Emeritus Professor at the University of Chicago, where he has been since 1969, with joint appointments in the Departments of Linguistics and of Near Eastern Languages and Civilizations. His earliest work centred on Sumerian and other non-affiliated languages of the Ancient Near East. He did linguistic fieldwork in Ethiopia, and his more recent teaching and research have centred on historical linguistics, comparative Semitic, and Cushitic.

ROBERT HOBERMAN is Professor of Linguistics at Stony Brook University. He works on the morphologies and phonologies of Arabic, Hebrew, and Aramaic in both their ancient and modern colloquial varieties. He has studied several modern Aramaic languages through fieldwork with speakers from Iraqi Kurdistan, and has worked on Maltese, which, like Aramaic, is a Semitic language that has undergone radical structural change in contact with Indo-European languages. Other interests include (in various overlapping circles) writing systems, comparative Semitic linguistics, the phonological history of Yiddish, Jewish interlinguistics, and ethnic, linguistic, and religious minorities in the Middle East.

MAARTEN KOSSMANN lectures at Leiden University (Netherlands). He works mostly on languages of Northern Africa and the Sahel, with a particular focus on Berber languages. He has published on language description, historical linguistics, oral literature, and language contact.

ANTONIO LOPRIENO is Professor of Egyptology at the University of Basel (since 2000), Rector of the University (since 2006), former professor of Egyptology at UCLA, and Director of the Basel project 'Diachronic grammar of Egyptian and Coptic'. His major research interests are Egyptological linguistics, morphology and syntax of the Egyptian language, and Egyptian–Semitic interrelations.

MAARTEN MOUS is Professor of African Linguistics at Leiden University, where he studied mathematics and (African) linguistics and did his Ph.D. on a grammar of Iraqw (1993). His interests are language description, in particular morphology, and language contact. He did fieldwork on several Cushitic and Bantu languages and did an extensive study of the mixed language Ma'á-Mbugu.

MATTHIAS MÜLLER is a research assistant at the University of Basel, graduate of the University of Göttingen/Germany, and project leader of the Basel project 'Diachronic grammar of Egyptian and Coptic'. His major research interests are interlanguage in Akkadian, morphology and syntax of the Coptic language, and Egyptian–Coptic phonology.

ERIN SHAY is Adjunct Assistant Professor of Linguistics at the University of Colorado, Boulder. She has written descriptive grammars of two Chadic languages: East Dangla (East branch) and Giziga (Central branch). Her research interests are Chadic languages, syntax/morphology interface, and historical linguistics.

ACKNOWLEDGEMENTS

The work on this volume was supported by grants from the Committee on Research and Creative Work and the Graduate Committee on the Arts and Humanities, both at the University of Colorado at Boulder. The work of the editors was also supported by National Science Foundation Grant No. 0439940.

Maps for this volume were drawn by Monika Feinen of the University of Cologne. It has been a pleasure working with her. We are most grateful to Marian Safran for her editorial work on this volume. We have also benefited from the advice and comments of contributors to this volume and of a number of other colleagues with whom we have discussed various issues. In particular, thanks are due to Guy Deutscher, Geoffrey Khan, Amina Mettouchi, Jonathan Owens, and Mauro Tosco. We are also most grateful to anonymous readers of the volume for pointing out mistakes and oversights and for other insightful comments. All errors of fact or interpretation remain our own.

This volume was inspired by Bob Dixon, who first broached the possibility of preparing the volume for the Cambridge Language Surveys. Our sincere thanks go to Helen Barton of the Cambridge University Press, who has provided wise and very friendly guidance throughout our work on this project. We are also very grateful to Leigh Mueller for her insightful and meticulous copy-editing.

ABBREVIATIONS

The list below contains abbreviations used in various chapters of the volume. In some cases, the same symbol may refer to different categories, or the same category may be indicated by more than one symbol, in different chapters. Also, abbreviations may use either upper- or lower-case letters, depending on which chapter they appear in.

ˋ	low tone
ˊ	high tone
Ø	zero marked (unmarked)
#	word boundary (in chapter 3)
=	clitic boundary
$	syllable boundary (in chapter 3)
1	first person
2	second person
3	third person
3MSG	third-person masculine singular
3FSG	third-person feminine singular
A	Akhnimic (Coptic dialect)
A	aorist
A	class a of verbs (Ts'amakko, Dhaasanac)
ABL	ablative
ABS	'absolutive' = unmarked case
ACC	accusative
AD	preverbal particle *ad* 'NON-REALIZED'
ADJ	adjective
ADJP	adjectival phrase
ADVST	adversative
AFF	affected (Chadic)
AFF	affirmative (Omotic)
Akk	Akkadian
ALL	allative
ANAPH	anaphora
ANN	annex
ANNEX	annexed state
AOR	aorist
APPL	applicative
Ar.	Arabic
ASSC	associative
ASSOC	associative
AT	'at'
ATR	Advanced Tongue Root
ATT	attributive
AUX	auxiliary
AWAY	motion away (verbal extension)
B	Bohairic (Coptic dialect)
B	class b of verbs (Ts'amakko, Dhaasanac)
BCKG	background

BGND	background
C	consonant
C_I	first consonant
CAUS	causative
C_F	final consonant
CN	connector
CNTQW	content question word
CNV	converb
COLL	collective
COM	comitative
COM	comment clause marker
com	common
COMP	complementizer
COMPL	complementizer
COMPL	completive
COMT	comitative
CON	construct case
CON	converb
CONJ	conjunction
CONS	consecutive
CONST	construct-state
COORD	coordinative
COP	copula
D	demonstrative
D	dependent
D.PROG	dependent progressive
D.PVG	distal point of view of goal (Chadic)
DAT	dative
DEC	declarative
DECL	declarative
DED	deduced referent
DEF	definite
DEF.ART	definite article
DEM	demonstrative
DEM1	first degree of distance (proximal) demonstrative
DEP	dependent
DEST	destination
DEST	destinative extension
DET	determiner
DF	definite
DIR	directional
DIST	distal deixis
DO	direct object
DP	determiner phrase
DS	different subject
DST	distal (demonstrative)
dtrm	determinate
DU	dual
DUR	durational
EI	habitual past (Dahalo)
EA	annexed state (*état d'annexion*)
EE	end-of-event marker
EL	free state (*état libre*)
EMPH	emphasis
EP	epenthetic
ES	Ethiopian Semitic
EXCL	exclusive
EXPEC	expectational
EXT	existential
EXT	extension
F	Faiyumic (Coptic dialect)
F	feminine
FEM	feminine
fgenc	foregrounding-enclitic
FIN	final
FM	focus marker
FOC	focus
FOR	'for'
FREQ	frequentative
FROM	'from'
FUT	future
GEN	genitive
GER	gerund
GO	goal
gr	grade (Hausa)
H	high tone
H	Huehnergard (1997) (in chapter 4)
HAB	habitual
Hebr.	Hebrew

HEC	Highland East Cushitic
HON	honorific
HUM	unspecified human subject
HYP	hypothetical
I	imperfective
ICAUS	indirect causative
ICP	Intransitive Copy Pronoun
IDEO	ideophone
IMP	imperative
IMPER	imperative
imperf	imperfect
IMPERS	impersonal
IMPF	Imperfective
IMPFV	imperfective
IMPS	impersonal
IMPT	imperative
IMPTV	imperative
IN	'in'
IN	negative imperfective
INC	inceptive
INCEPT	inceptive
INCL	inclusive
INCLU	inclusive
IND	indicative
INDEF	indefinite
INDEF.ART	indefinite article
INDEP	independent
INF	infinitive
INN	inner space (verbal extension)
INST	instrumental
INT	intensive
INTENS	intensifier
INTER	interrogative
INTERJ	interjection
INTJ	interjection
INV	inverse
IO	indirect object
IPF	imperfective
IPS	impersonal
IRR	irrealis
ITER	iterative extension
JUSS	jussive
L	low tone
L	Lycopocitan (Coptic dialect)
Late Eg.	Late Egyptian
LOC	locative
LOG	logophoric pronoun
M	masculine
MAN	Mood-Aspect-Negation
MASC	masculine
MED	medial (function similar to that of 'converb')
MED	middle voice
MID	middle
MOD	modifier
MR	multiple reference
MSA	Modern South Arabian
N	neuter = plural
N	nominal form
NEG	negative
negenc	negative-enclitic
NFS	non-factual stem
NMZ	nominalizer
NOM	nominalizing marker
NOM	nominative
NONDEF	non-definite pronoun
NORM	normative
NP	noun phrase
NUM	number
O	object
OBJ	object
obl	oblique
old Eg.	Old Egyptian
OPT	optative
OSA	Old South Arabian
OUT	movement out extension
P	Perfective (Berber)
P	person
P, (P)	plural (of gender)

P	predicate
P2	secondary perfective (Berber)
PART	participle
PART	partitive
PAS	passive
PASS	passive
PAST	past
PCPL	participle
PER	perfect
PERF	perfect
PF	perfective
PL	plural
PLUR	plural
PN	Negative Perfective (Berber)
PN	personal name
PNCT	punctual
PNG	person/number/gender
PO	potential
POL	polite
POS	point of view of the subject
POS	possessive
POSS	possessive
PRE.PRO	prepronominal marker
PRED	(locative) predicator (Chadic)
PRED	predicative
PRED	predicative particle *d* (Berber)
PREP	preposition
PRES	present
PRET	preterite
PRF	perfective
PRO	(independent)-pronoun
PROG	progressive
PROH	prohibitive
PRON	pronoun
PROSP	prospective
PROX	proximal
PST	past
PTC	participle
PURP	purpose
PVG	point of view of goal
Q	question
QUANT	quantifier
QUES	question
R	remote deixis
R	verbal root
RDP	reduplication
REAS	reason
REC.PAST	recent past
RECIP	reciprocal
REF	referential
REFL	reflexive
REL	relative
REM	remote deixis
S	Sahidic
S	singular
S	subject
S3	third-person subject
SBJ	subject
SBJN	subjunctive
SCL	subject clitic
Sem.	Semitic
SENT	sentence marker (Ts'amakko)
SEQ	sequential
SEQUEN	sequential
SET	setting
SFX	suffix
SG	singular
SGLTV	singulative
SH	short (Omotic)
SING	singular
SO	source/subject orientation
SPEC	specified
SS	same subject
STAT	stative
STR	strong
SUBJ	subject

SUBJ	subjunctive (Chadic)	UH	unspecified human
T	target		subject
T	tense	UNM	unmarked tense
TAM	tense, aspect, or mood		(Ts'amakko)
	marker	UP	movement upward
TEMP	temporal		extension
TENT	tentative extension	V	verb
TO	'to'	V	vowel
TOG	'together'	V_1	first vowel
TOP	topic	VEN	ventive
TOT	totality extension	VN	verbal noun
TQ	question about truth	VOC	vocative

1

Introduction

Zygmunt Frajzyngier and Erin Shay

1.1 The goal of the present work

Afroasiatic is the fourth largest linguistic phylum in the world, with about 375 living languages spoken by some 300 million speakers (www.ethnologue.org). In the view of the contributors to this volume, this number may well be an overestimation. For the Chadic family alone Ethnologue estimates over 190 languages, while most linguists working in the area estimate the number to be between 140 and 160 languages. The differences boil down to decisions regarding what is a language and what is a dialect. Given the absence of clear-cut criteria for this distinction we leave the question of the actual number of languages open.

This book provides the first-ever typological survey of each of the language families belonging to the Afroasiatic phylum as well as a typological outline of the entire phylum.

The book is addressed to a general linguistic audience, some of whom may be unfamiliar with Afroasiatic linguistics, as well as to linguists who have worked on Afroasiatic languages and would like information about languages from other branches and about the characteristics of the whole phylum.

The approach taken in this book is typological rather than historical, taking for granted the existence of the Afroasiatic family, as confirmed by comparative historical studies. This is because a number of phenomena in contemporary languages can best be explained by the internal structure of the grammatical systems rather than by their origin in an ancestral proto-language. There exist a fair number of studies, chief among them Diakonoff (1988), that take a historical rather than a typological approach to selected issues.

We adhere to the common usage in referring to Afroasiatic as a phylum, rather than a family, on the grounds of the remoteness of the relationships among its various branches (Diakonoff 1988). We also take it as given that the phylum is composed of six language families (from west to east), namely Berber, Chadic, Egyptian, Cushitic, Omotic, and Semitic, even though the internal structure of the phylum remains somewhat

controversial and further internal subgroupings within the phylum cannot be ruled out (for recent proposals based on different criteria, see Diakonoff (1998), Zaborski (2005a), and Ehret (2005)).

The Afroasiatic languages are an important object of study, not only because of their widespread usage but also because of their great typological diversity. It has been claimed that the Chadic family alone is more typologically diverse than the entire Indo-European language family (Diakonoff 1988). The Afroasiatic phylum exhibits great variation with respect to traditional typological criteria such as the position of the predicate in the clause, the structure of the noun phrase, and the structure of the verb. As a result, the question 'What is a typical Afroasiatic language?' cannot, at this stage, be answered. The material in this book is intended, among other things, to document this typological diversity.

The Afroasiatic languages are also an important object of study because the languages and the cultures they embody have been instrumental in shaping Judeo-Christian culture, Islamic culture, and much of what has come to be referred to as Western civilization. Speakers of Semitic languages, which belong to the Afroasiatic phylum, developed the alphabetic writing system which, with numerous modifications, is now used in thousands of languages throughout the world. The development of the alphabetical writing system may have been facilitated by the underlying structure of verbal roots and derived nominal forms in Semitic languages, where the consonantal structure alone conveyed a great deal of semantic information.

While some Afroasiatic languages are widely spoken and robust, many languages of the phylum are endangered and may disappear within a few generations. Although a decline in the use of a given language is usually a result of various social forces, sometimes including speakers' choice, such a loss also means the loss of the most complex intellectual product of those who speak the language. We hope that this book will be a stimulus and a useful tool for scholars to undertake the task of working on hitherto-undescribed or under-described languages.

Although the focus of the book is typological, individual chapters also provide information on the history of the language family; geographical distribution; historical writing systems, if any; and, in some instances, diachronic changes within the family. While the information included in a single volume cannot be exhaustive, we hope that it will serve as a starting point for a more extensive and intensive typological, and eventually historical, study of the families composing the Afroasiatic phylum. The book includes a bibliography of sources and materials for further reading. Since the scholarship on many languages is new, and since there is no agreed-upon standard for many language names, the spelling of language names on the maps and in various chapters may differ.

1.2 History of the recognition of the phylum

The term 'Afroasiatic' was coined by Delafosse in 1914 (cited in Newman 1980) and was reintroduced by Greenberg in 1960. The term captures the fact that this is the only phylum whose member families include languages spoken in Africa and languages spoken in Asia. The phylum has also been called Hamito-Semitic (since F. Müller 1876), Semito-Hamitic (chiefly in the older Russian sources), Afrasian (Diakonoff 1988), Erythraean (Tucker and Bryan 1966; Tucker 1975), and Lisramic (Hodge 1972). The term 'Afrasian' is an Anglicization of the Russian *afrazijskije*, a variation on the term 'Afroasiatic'. The term 'Erythraean' refers to a core geographical area of the family. The term 'Lisramic' is based on the Proto-Asiatic **lis* 'language' and the Proto-Egyptian **rāməč* 'people'. Appellations for the phylum have been the object of vigorous discussion among linguists, and a special session of the Hamito-Semitic conference held in London (Bynon and Bynon 1975) was devoted to naming the phylum. In contemporary writing by various scholars, the most frequently used terms are 'Afroasiatic', 'Hamito-Semitic', and 'Semito-Hamitic' (see later sections concerning the history of the phylum).

1.3 Evidence for genetic relationships within the phylum

The typical evidence for genetic relationships within the phylum includes numerous comparative word lists showing etymologies across the Afroasiatic families. Some of these lists propose sound correspondences, while others simply provide the presumed cognates. The first of such comparisons was Marcel Cohen (1947), followed by Greenberg (1963), Hodge (1966, 1967), and a number of more recent studies dealing with the whole phylum, two or more families of the phylum, or a single family within the phylum. These include Skinner (1984); Belova *et al*. (1994–7), representing the work of Diakonoff's team; Orel and Stolbova (1995); Naït-Zerrad (1998); Takács (2005 and other works); Ehret (1995); Dolgopolsky (1999); Militarev and Kogan (2000); and Rössler's and Vycichl's numerous studies on Egyptian–Semitic relations. Militarev (2000) uses glottochronology as a means of calculating when the phylum split into various families. The largest of the comparative studies are Belova *et al*. (1994–7), Orel and Stolbova (1995), and Ehret (1995). The larger etymological studies have been criticized for the choice of items taken for comparison and often for the validity of postulated cognates. The cumulative effect of these studies, that of reconfirming the genetic unity of the phylum, is not in doubt.

Another piece of evidence for the genetic relationship of Afroasiatic languages comes from morphology. Across the phylum there are morphemes with similar phonological structures and similar functions. Many of these have long been known for their

occurrence in Egyptian, Semitic, Berber, and Cushitic languages, but it was Greenberg (1963) who demonstrated that the same morphemes also occur in various Chadic languages. Hodge (1969a, first presented in 1965) describes the evolution (what we would call today the 'grammaticalization') of determiners in Afroasiatic languages in a way that leaves no doubt as to the genetic relationship of the phylum. David Cohen (2005: 17ff.) provides extensive evidence for the alternation between *a* and non-*a* vowels in the verbal systems of Afroasiatic languages. This is another piece of evidence for the genetic relationship of the six families.

Sasse (1984a) and Blažek (2006) are devoted to the study of case and (mainly locative) prepositions. These studies are driven by the aims of historical linguistics and do not deal with functions of the reconstructed elements within the grammatical systems of the various languages.

Some linguists have claimed that there is a genetic relationship between Afroasiatic and Indo-European languages. Hodge, who called the proposed super-phylum 'Lislakh', argued for this relationship in a number of publications (Hodge 1978, 1979, 1981). Proponents of Nostratic theory (Dolgopolsky 1998) include Afroasiatic as a member of the Nostratic family. The Nostratic hypothesis is highly controversial and has very few supporters among specialists in Afroasiatic languages.

Debate as to the internal division within the phylum involves the status of Omotic as a separate family and the question of whether there may be further subdivision within the phylum. With respect to Omotic, the question is whether it is a separate family or whether it should be incorporated within the Cushitic family. The history of Omotic as a family within the Afroasiatic phylum is described in detail in chapter 7 of this volume.

1.4 A snapshot of the history of scholarship

The awareness of relationships among languages within Afroasiatic goes back at least to the ninth century, when Judah ben Quraysh of Morocco, a physician to the emir of Fez, wrote of lexical and phonological similarities between Berber and the Semitic languages Biblical Hebrew, Aramaic, and Arabic (Becker 1984, reviewed by Wansbrough 1986). Hayward (2000) reports that the French orientalist Postel (1538) also pointed out resemblances among Hebrew, Arabic, and Aramaic. To these languages, Ludolf (1702) added Amharic and Ge'ez (Hayward 2000). In 1781, von Schlözer gave the grouping the name 'Semitic', based on the biblical Sem, son of Noah (Genesis 5:32). Müller (1876) followed the pattern in naming the Hamitic branch, assumed at that time to consist of Egyptian and Berber. Müller also created the term 'Hamito-Semitic' for the larger language family, reflecting the assumption that the phylum could be split into two branches, the Hamitic languages and the Semitic languages. The selection of languages in Meinhof's 1912 *Die Sprachen der Hamiten* was based on a mixture of

anthropological and linguistic typological criteria and included languages that are now not part of the Afroasiatic family.

Marcel Cohen (1924) was the first to reject a division of the phylum into Hamitic and Semitic branches. He stated emphatically that there is no trait shared by the Libico-Berber (now Berber), Cushitic, and Egyptian languages that would group them together and set them apart from the Semitic languages. However, he retained the term 'Hamito-Semitic' as a purely conventional label. A few linguists still interpret the term in the sense in which it was originally coined, as implying two branches within the phylum. Diakonoff (1998) points out that the use of the term 'Hamito-Semitic' by Orel and Stolbova (1995) wrongly implies a division into two branches. The term 'Hamito-Semitic' is still used in French, Italian, Russian, German, and English writings.

According to Sasse (1981a: 132), some of the languages now classified as Cushitic, such as Beja, Somali, Galla, and Harari, were considered, as of the mid nineteenth century, to belong to the same family as Egyptian, Semitic, and Berber. Sasse cites Lepsius (1844), Beke (1845), d'Abbadie (1845), and Lottner (1860–1) as among those who speculated about the existence of the larger linguistic family. We may add to this list Burton (1856), who stated that 'the Harari appears, like the Galla, the Dankali, and the Somali, its sisters, to be a Semitic graft inserted into an indigenous stock' (Burton 1987 (1856): 153).

The Chadic family was the last to be added to the phylum. Marcel Cohen (1924) did not include Chadic languages in his study of the Hamito-Semitic languages. In Cohen (1947), a comparative study of 500 lexical items, he does include Hausa along with Egyptian, Semitic, Berber, and Cushitic. Although the inclusion of Hausa in the comparative study may be construed as a tentative inclusion of Chadic within the phylum, Cohen does not mention Chadic as a family. As late as 1970 he was reluctant to recognize Chadic as a member of the Afroasiatic phylum: 'Si le tchadien doit réellement nous être adjoint (je crois qu'il doit nous être adjoint mais non pas incorporé) la question se pose aussi: comment le tchadien s'est-il formé?' (Cohen 1970 (1937): 24). The unequivocal inclusion of Chadic in the Afroasiatic phylum is due to Greenberg (1950b).

Fleming (1969, 1974) proposed placing a group of languages previously classified as West Cushitic in a separate branch of Afroasiatic, which he called the Omotic branch. This separation has been accepted by some but rejected by others.

Inclusion of Omotic in the Cushitic family, and by implication in the Afroasiatic phylum, is strongly supported by Cerulli's (1951) study of Kafa (once considered part of West Cushitic, now classified as Omotic). Cerulli provides numerous regular phonological and morphological correspondences between Kafa and the Cushitic languages of the western Sidamo province as well as the Central Cushitic Agau. As

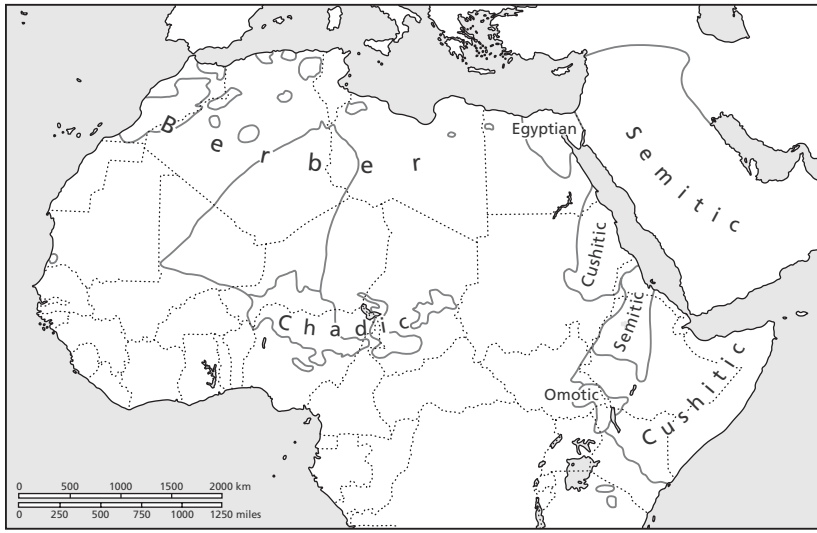
described in chapter 7 in the present volume, the position of these languages within the Afroasiatic phylum remains one of the most controversial issues in Afroasiatic classification. The discipline still lacks a systematic study of regular sound correspondences and of common retentions and innovations that would allow subclassification.

For succinct histories of the concept of the Afroasiatic phylum, the reader is referred to Newman (1980), Diakonoff (1988), and Hayward (2000).

1.5 Geographical range of the Afroasiatic phylum

Afroasiatic languages are spoken in Northern Africa, Central Africa, the Horn of Africa, the Arabian Peninsula, and even in Central Asia (Arabic). Berber languages are spoken in isolated pockets in Mauritania and in Morocco, Algeria, Mali, Niger, Libya, Tunisia, Burkina Faso, as well as in the Siwa oasis in Egypt. The Chadic family, the largest of the phylum, comprises between 140 and 160 languages (estimates vary) spoken in northern Nigeria, southern Niger, northern Cameroon, and southern Chad Republic. Hausa, the Chadic language with the greatest number of speakers, is a vehicular language in West Africa and the official language of Nigeria, and there are pockets of Hausa speakers to be found over large areas of West and Central Africa. Egyptian was the language of ancient Egypt, and its descendant, Coptic, remains the liturgical language of the Coptic Church in Egypt. Cushitic languages are spoken in Eritrea, Ethiopia, Somalia, and northern Kenya, Sudan (Beja) and in isolated pockets in Tanzania. Some Cushitic languages are official in different federal regions of Ethiopia. Somali, also a member of the Cushitic family, is the official language of the Somali Republic. Omotic languages are spoken in southwest Ethiopia. Among the Semitic languages, Hebrew is one of the official languages of Israel, Amharic is one of the official languages of Ethiopia, and Tigre and Tigrinya are the official languages of Eritrea. Arabic, spoken throughout North Africa, the Arabian Peninsula, and in areas outside of Africa, is the official language, or one of the official languages, of Mauritania, Morocco, Algeria, Tunisia, Libya, Egypt, and Chad in Africa, and of Israel, Lebanon, Jordan, Syria, Iraq, Saudi Arabia, United Arab Emirates, Oman, Bahrain, Qatar, and Yemen in the Arabian Peninsula, and in other countries with significant Arabic diaspora. For a map of dialects of Arabic, see Kaye and Rosenhouse (1997: 264). Maltese (Semitic) is one of the official languages of Malta. Despite the wide geographic range of the families of the phylum, and the fact that a number of those languages are the official languages in various countries, many Afroasiatic languages are threatened with extinction because they are spoken by a small number of people in economically and politically unstable environments.

All languages of the Afroasiatic phylum have had extended contact with other Afroasiatic languages and with languages belonging to other families. The Semitic languages



Map 1.1 *Afroasiatic phylum*

of Ethiopia (Ethiosemitic) have been in contact with Cushitic and Omotic languages, and some of the Ethiosemitic languages share a number of typological features with Cushitic and Omotic languages that they do not share with other Afroasiatic languages. Akkadian, an East Semitic language, has been in contact with Sumerian, and North-Eastern Neo-Aramaic has been in extensive contact with Kurdish. Berber languages have been in contact with Arabic and Chadic languages and with Nilo-Saharan languages. Chadic languages have been in contact with Niger-Congo, Nilo-Saharan, Semitic, and Berber languages. Such contacts have no doubt induced changes in Afroasiatic languages. The clause-final position of the verb in Amharic and Tigrinya (Ethiosemitic) languages, for example, is attributed to contact with Cushitic and Omotic languages, and the clause-final position of the verb in Akkadian is attributed to contact with Sumerian. Given the absence of systematic phonological, morphological, and syntactic reconstructions, in most instances we are unable to state categorically which typological features are due to language contact.

1.6 State of the art in Afroasiatic scholarship

A few Afroasiatic languages, including Egyptian, Hebrew, Arabic, and Aramaic, have been the objects of study for more than 200 years, and the amount of literature on these languages is very large indeed. As might be expected, the longer the history of scholarship on a given language or family, the more publications are available. Yet even

for language families with the longest scholarly tradition there are fundamental gaps in the scholarship. Izre'el (2002) and Khan (2002) acknowledge the fact that traditional Semitic scholarship seldom dealt with syntax. Even such fundamental components of grammar as the aspectual and tense systems in Semitic languages remain poorly studied (Izre'el 2002). For the state of the art in Semitic studies, see Izre'el (2002) and chapter 4 of the present volume.

Tosco (1994a) and (1994b) represent typological studies of the syntax of East Cushitic languages. Dolgopolsky (1973) is a massive reconstruction of lexical roots in Cushitic languages. The languages classified now as Omotic are included in Dolgopolsky's study as West Cushitic. Zaborski (1975) is a study of the verbal forms in Cushitic languages. Bender (2000) is a comparative study of Omotic morphology. For a compendium of literature on individual families, the reader is referred to the bibliography at the end of the volume.

Very little scholarship has been devoted to the typology of the Afroasiatic phylum as a whole. The most recent and most complete surveys of Afroasiatic languages are Diakonoff (1988) and Petráček (1989). The latter, a two-volume publication in Czech, is a textbook for students of Semitic, Egyptian, and African languages. The work is conceived as an exhaustive review of the literature on the Afroasiatic phylum and individual families; the history of classification; theoretical issues in genetic classification; sociolinguistic issues; and areal linguistics. The list of references, which ends with 1987, takes up eighty pages. Only ten pages of the two-volume work are devoted to typology, the focus of the present volume.

With the exception of Egyptian and Semitic, there have been no typological studies of any single branch of the Afroasiatic phylum. For most of the branches, and for the phylum itself, there have been no comparative studies of phonological processes, of the syntax of simple or complex sentences, of semantic categories encoded, of reference systems, or of any of the other functional domains that comprise a complete grammar of a language.

Hodge (1971, a reprint of Hodge 1970) is a collection of essays on various Afroasiatic families. These essays, some written by the most eminent scholars of the time, are very brief. Most list the important references for the family surveyed and provide some information about the phonological system and bits and pieces of morphology. None of the essays offers a picture of the grammatical system of any of the families surveyed. Since Hodge (1971), there have been many collections of papers published on Afroasiatic linguistics, e.g. Perrot *et al*. (1981), Lecarme *et al*. (2000), Zaborski (2001), Bender *et al*. (2003), Lecarme (2003), Fronzaroli and Marassini (2005), Lonnet and Mettouchi (2005), and Mettouchi and Lonnet (2006). Studies in these volumes are devoted to single topics in individual languages and do not pretend to offer a survey of any of the families, much less of the entire phylum. The only

work that attempts to look at a large number of issues in Afroasiatic languages remains Diakonoff (1988).

1.6.1 Phonological reconstruction

Most of the work conducted so far on Proto-Afroasiatic reconstruction has dealt with the sound inventory and the lexicon. Diakonoff and his collaborators have worked for many years on reconstructing the Afroasiatic consonants, vowels, tones, and vocabulary. Their results have been published in instalments, first in Russian and then in English (Diakonoff *et al*. 1992; Belova *et al*. 1993; Belova *et al*. 1994–97). Diakonoff (1988) reconstructed a consonantal system involving four manners of articulation for obstruents: voiceless, emphatic, and voiced stops and voiceless continuants. He postulated labial, dental, and palatalized fricatives and affricates, and labial, velar, labialized-velar, post-velar, labialized post-velar, pharyngeal, and laryngeal places of articulation. Diakonoff also postulates two nasals, *m* and *n*; two liquids, *r* and *l*; and palatal and labial glides. He does not include the prenasalized stop *mb*, posited in Greenberg (1965) as a Proto-Afroasiatic phoneme.

Orel and Stolbova (1995) is an attempt to reconstruct 2,672 lexical items. The book is subtitled 'Materials for a Reconstruction'. The authors postulate a Proto-Afroasiatic consonantal system consisting of voiced, voiceless, and emphatic stops and fricatives. They also postulate seven places of articulation: labial, dental, lateral, velar, post-velar, pharyngeal, and laryngeal. Unlike Diakonoff, Orel and Stolbova do not postulate the labialized velar consonants k^w, g^w, and q^w. They postulate a six-vowel system consisting of *i*, *ü*, *e*, *a*, *o*, and *u*. The Orel and Stolbova reconstruction focuses on lexical items rather than on the phonological system. It does not deal with constraints on syllable structure or with phonological processes such as vowel or consonant harmony.

Ehret (1995) reconstructs about forty consonantal phonemes, with three manners of articulation (voiceless, voiced, and emphatic). As places of articulation he posits labial, dental, velar and labiovelar; alveolar and palatal; nasal; and laryngeal. He also postulates glides, an *r*, and four laterals: *l*, *dl* (corresponding to the ɮ of Jungraithmayr and Shimizu's Proto-Chadic reconstruction (1981)), *tl'*, and ɬ. He retains the labial consonants postulated by Greenberg (1958) but explicitly rejects the notion, advocated in Greenberg (1965), of prenasalized stops in Proto-Afroasiatic. Ehret further postulates, albeit tentatively, the existence of tone in Proto-Afroasiatic, basing his conclusion on the analysis of tones in Ngizim (West Chadic) and Mocha (Omotic) and on the fact that tones are attested in Cushitic, Omotic, and Chadic languages. Unlike Orel and Stolbova, Ehret takes phonological constraints into consideration. He takes the emergence of phonological constraints as evidence for innovations that are the basis of his subclassification.

With respect to the consonantal system, most reconstructions agree that Proto-Afroasiatic had three series of obstruents and that the only continuants were voiceless. Note, however, Ehret's reconstruction of the voiced lateral continuant. Diakonoff and his associates postulate a series of labiovelar consonants, while Orel and Stolbova (1995) claim that labial velars derive from 'velar consonants followed by the sequence *au*'. While such sequences often result in a labiovelar and eventually a labial stop (Frajzyngier 1989a), there are also labial velar consonants that cannot be explained as deriving from the sequence velar-*a*-*u*. The evidence that some labial velar stops may, in fact, be underlying is provided by languages where such stops occur in word-final position and where there is no rule of final *a* deletion. This is the case in Hdi (Central Chadic), e.g., the noun *màrkw* 'wife' (Frajzyngier with Shay 2002).

The posited reconstructions show much greater variation with respect to the number of vowels in Proto-Afroasiatic. Diakonoff (1980) considers the possibility that there were only two vowels, the low vowel *a* and a high vowel realized as *i*, *u* or a central vowel. Ehret reconstructs five short and five long vowels for Proto-Afroasiatic: *a*, *aa*, *e*, *ee*, *i*, *ii*, *o*, *oo*, *u*, and *uu*. For a critique of reconstructions in Orel and Stolbova (1995), see Diakonoff (1998), and for a critique of methodologies in Orel and Stolbova (1995) and Ehret (1995), see Ratcliffe (n.d.).

1.6.2 Reconstruction of morphology and syntax

Little work has been done on reconstructing the Proto-Afroasiatic morphological system, and most of this work has been devoted to the morphology of the verb (Diakonoff 1988). Greenberg (1955) described the use of *a*, in the place of a different vowel, as the marker of nominal plurality, and also described the suffix *-en* as a marker of plurality. Greenberg (1952) and (1953) dealt with verbal forms involving gemination, which he analysed as coding present tense. Greenberg (1960) dealt with gender and number agreement, citing *t* as the Afroasiatic marker of feminine gender and citing a contrast between *k* as the marker of masculine gender and *t* as the feminine marker. Greenberg also stresses similarities among pronominal systems, in particular possessive pronouns, among various families of the phylum. Greenberg (1963), which presented a revised classification of African languages, posited a number of Proto-Afroasiatic grammatical morphemes, including pronouns; the causative *-s* (although subsequent studies have shown that this suffix does not occur in Chadic (Frajzyngier 1985a)); the prefix *m-* as a marker of place names, instrument, and agent; and the internal *a*-plural. The internal *a*-plurals occurring in all Afroasiatic languages constitute a strong morphological argument in favour of a genetic relationship among Afroasiatic languages. Greenberg also postulates a variety of other morphemes, all of which have been shown by subsequent studies to be the product of grammaticalization processes whose original source was the same as that posited

by Greenberg (1963) for pronouns. Among these morphemes are the gender markers *k* 'masculine' and *t* 'feminine'; *n* used as a genitive linker; and plurals involving the marker *n*.

Diakonoff (1965; revised 1988) was the first attempt to reconstruct not only the phonological system of the Afroasiatic phylum but also elements of the grammatical system, including morphology and syntax. Given that this reconstruction has more elements of the grammatical system in its scope than other works, we devote more space to it than to other reconstructions that are much narrower in scope. Diakonoff explicitly based his reconstruction of the root and word structure of Afroasiatic on the Semitic languages 'of the Ancient and Middle stages' (Diakonoff 1988: 43). The premise is that the old Semitic languages such as Akkadian, Eblait, and Old Hebrew represent older stages of Afroasiatic and thus preserve much of the inflectional morphology that has been lost in many modern languages. In this approach, the ways in which other languages differ from the old Semitic languages must represent innovations. There is no theoretical basis for this claim or empirical evidence that languages with the oldest written records necessarily represent the oldest forms and functions. The only way information regarding the oldest forms can be obtained is through a reconstruction of the grammatical system. No such reconstruction is available for any language family in the Afroasiatic phylum, where scores of languages remain undescribed, and such a reconstruction is a very remote prospect. Given the biases of Diakonoff's approach, the reliability of his reconstructions for the whole phylum is very much in doubt. However, there are some reconstructions for which Diakonoff provides evidence from at least three branches of Afroasiatic, and we list some of these here. We also mention some hypotheses for which there is no support, or very weak support, but that touch on the totality of the grammatical system or are otherwise interesting.

Diakonoff postulates that no syllable in Proto-Afroasiatic (which he calls 'Common Afrasian') may begin with a vowel and that no syllable may have more than one consonant in the coda ('Auslaut' in his terminology). The first of these generalizations appears to be based on Semitic word structures. In fact, vowel-initial syllables are allowed in many Chadic languages, although they appear mainly in morphemes belonging to closed sets. Some Chadic languages also exhibit syllable- and word-final consonantal clusters (e.g. Hdi), while others do not allow consonant-final words in isolation (e.g. Wandala).

Diakonoff also postulates a sequence of lexicalizations whereby nominal roots were lexicalized first and verbal roots came into being later. This generalization is not supported with any argumentation. The postulated roots had two consonants or two consonants plus a glide that could be realized as a consonant or a vowel. Diakonoff also posits tone in Proto-Afroasiatic. The primary argument for the reconstruction of tone is the postulated existence of numerous identical reconstructed roots. Tone is thus postulated as a phonological means of differentiating lexical items. Interestingly, in many

contemporary Chadic, Cushitic, and Omotic languages, tone has mainly a grammatical function. Because tonal distinctions among lexical items are not frequent, there is no basis for postulating that tone was originally a means of differentiating among lexical elements. Ehret (1995) postulates tone in Proto-Afroasiatic on the more plausible grounds that tone is preserved in three out of the six branches.

Diakonoff postulates nouns, adjectives, and numerals as the nominal categories of Proto-Afroasiatic, but he also states that adjectives are distinguished from nouns through syntactic means. Diakonoff also postulates a two-gender system that was originally marked by a palatal or labial glide for the masculine gender and a glottal continuant or palatal glide, with different stress, for the feminine gender. In his analysis, this marking system evolved further, with *t* replacing the glottal continuant as the marker of feminine gender. Diakonoff proposes that Proto-Afroasiatic had three nominal numbers: singular, dual, and plural. These are attested in Semitic and, he claims, through traces in Cushitic and Chadic.

With respect to the verb, Diakonoff proposes the interesting hypothesis that the primary distinction with respect to verbal categories in Proto-Afroasiatic was between 'action' and 'state', with the distinction between transitive and intransitive verbs emerging later. David Cohen's (2005) comments on apophony in Afroasiatic also postulate the distinction between action and state.

Among the morphological means that were recognized early on as widespread characteristics of Afroasiatic are gemination and reduplication. Diakonoff (1988) rightly points out that these means code a variety of functions, including intensive, iterative, causative, and, in his view, habitual aspect. (Recall that Greenberg proposed that gemination codes present tense, a category that overlaps with habitual aspect.) As shown in chapter 8 of the present volume, gemination and reduplication have an even wider range of functions. Although both of these processes are widespread in Afroasiatic, we are reluctant to claim that they are defining characteristics of the Afroasiatic phylum, since both gemination and reduplication also occur in other language families of Africa, where these means often carry the same functions as they do in Afroasiatic.

In all Afroasiatic families, albeit not in all languages, verbal forms carry inflectional markers coding aspectual and mood distinctions. The alternation between *a* and non-*a* vowels is exploited for a variety of related functions coding aspectual distinctions, stative versus dynamic predications, and intransitive versus transitive predications. This alternation, called apophony or ablaut in Semitic, Cushitic, and Berber languages (D. Cohen 2005), results in Chadic languages from vowel suffixation (Frajzyngier 2004, 1981). In several families, verbal forms also code plurality of the verb, spatial orientation, voice, manner, and point of view. Reconstructions so far have focused on categories attested in Semitic languages, such as aspect, voice, and person. Ehret (1995) proposes that, in addition to these categories, Proto-Afroasiatic also had verbal extensions coding a variety of functions. In his analysis, these extensions later became a third consonant

of the verbal root. This hypothesis has been criticized on the grounds that the thirty-seven extensions postulated by Ehret are too many (Zaborski (1991), as quoted in Kaye (1996)). In view of the fact that some contemporary Chadic languages of the Central branch have a dozen or more extensions each, the test of the validity of the hypothesis should be whether the postulated forms are morphologically distinct and whether they indeed carry the functions ascribed to them.

Diakonoff (1988) postulates VSO as the word order of Proto-Afroasiatic. Verb-initial word order is indeed found in most Semitic languages, in Egyptian, and in some Berber languages. It has also been reconstructed for Proto-Chadic (Frajzyngier 1983). In the Omotic and Cushitic languages, however, the verb occurs in clause-final position. The question of the Proto-Afroasiatic default word order should be revisited. If an assumption is made that Proto-Afroasiatic had the verb in clause-initial position, an explanation should be found for why Cushitic and Omotic languages acquired verb-final default order.

A reconstruction of any lexical entry or grammatical form is a hypothesis that needs to be either confirmed or refuted. All comparative studies of Afroasiatic, including attempts at reconstructing phonological systems, morphology and parts of the grammatical system, have engendered criticism regarding their choice of sources, their methodology, and, not surprisingly, their conclusions. While many criticisms are valid, they should not overshadow the pioneering nature of the work done, the important questions raised by comparative studies, and, most important, the posing of numerous hypotheses regarding the reconstruction of lexical items.

1.7 Where did the Afroasiatic languages come from?

The issue of the origin of Afroasiatic languages is a subject of continuing debate. Two places of origin have been proposed: the Middle East (see Diamond and Bellwood 2003) and Africa (for a recent discussion, see Ehret *et al*. (2004)). The two hypotheses are based on different sets of criteria. Diamond and Bellwood (2003) use mainly archaeological and botanical criteria, with only marginal reference to linguistic criteria. Most linguists, using linguistic data only, would postulate Africa as the home of Proto-Afroasiatic (Ehret *et al*. 2004). Diakonoff (1975, 1988) places Proto-Afrasian in the southeastern Sahara, between the Tibesti mountains in northern Chad and Darfur. He postulates that the first group to break away from the place of origin were speakers of Egyptian, who moved north in about the eighth millennium BC. These were followed soon after by speakers of Proto-Chadic, who moved south and merged with the 'Negroid' (i.e. non-Afroasiatic) substratum. Later, speakers of Omotic moved southeast. Around the seventh millennium BC, in Diakonoff's account, the speakers of various Cushitic dialects moved to the east. He further argues that the speakers of Proto-Semitic separated from the speakers of Proto-Berbero-Libyan around the sixth to the fifth millennium BC, taking a path into

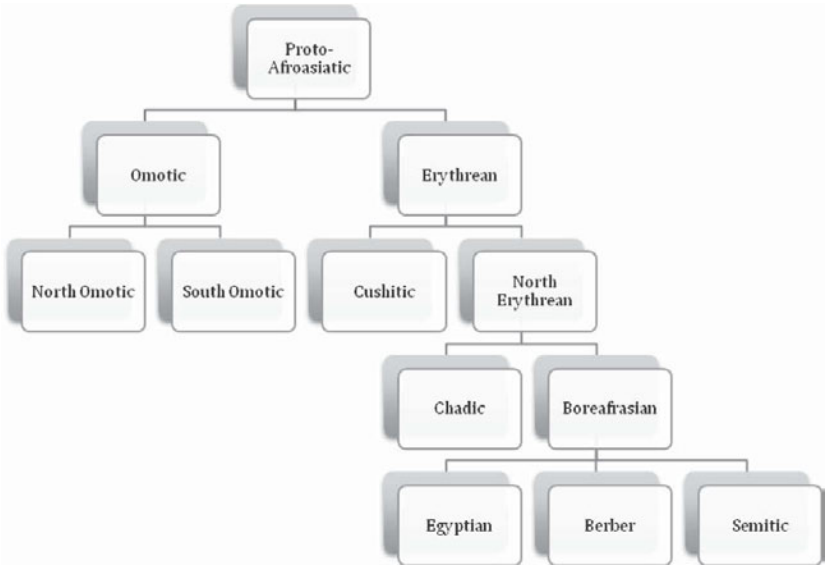


Figure 1.1 *Afroasiatic classification, based on Ehret (1995).*

Asia through the Nile Valley and over the Suez isthmus. The argument for this separation, states Diakonoff, is provided by the fact that the Semitic language in Asia had already divided into various dialects by the fourth millennium BC.

Diakonoff (1998) departs from the notion of a single homeland for Proto-Afroasiatic, positing a different place of origin for each of his major linguistic subgroups. In Diakonoff's view, the distinguishing feature of one of these groups, 'East-West Afrasian', composed of Semitic, Cushitic, and Omotic, is the presence of verbs conjugated by subject prefixes. Verbal prefixes, however, can emerge at a relatively shallow time depth, given the proper phonological conditions. This is the case in Gidar (Central Chadic), where subject pronouns have recently become prefixes in some dialects (Frajzyngier 2008).

Militarev is one of the few proponents of the notion that Afroasiatic originated in the Middle East (see Militarev and Shnirelman 1984). The argument in favour of this hypothesis is that the names of animals and plants reconstructed for Proto-Afroasiatic are of Middle Eastern rather than African origin. Vycichl (1987) also argues for the Asian origin of Afroasiatic.

In 1995, Ehret published a reconstruction of the Afroasiatic vowels, tone, and consonants, a vocabulary of 1,000 roots, and a reconstruction of derivational morphology. Appropriately taking common phonological innovations as a determining criterion for linguistic subclassification, Ehret proposes the classification in figure 1.1.

1.8 **A preview of the book**

The present work takes for granted the genetic relationships within the Afroasiatic phylum that have been established by the most widely accepted historical studies. It also takes for granted the internal classification of the phylum into six families and the status of Omotic as a separate family, although this classification remains controversial. For a list of features that argue for treating Omotic as a separate branch of Afroasiatic, see Hayward (2000). The goal of the book is to provide a survey of the phonetics, phonology, morphology, and syntax of Afroasiatic languages and to provide up-to-date information for each language family, along with a synthesis of similarities and differences among language families. Each chapter highlights those characteristics that are particularly important from a typological point of view and characteristics that make the family interesting and worthy of further study.

The organization of each chapter follows the same general outline, but the chapters differ significantly because the grammatical structures of the languages involved are quite different and because the state of knowledge for different language families varies widely. The Semitic scholarly tradition is more than 1,000 years old, and most of the Semitic languages have been described. In contrast, barely one-third of the Chadic languages have been described, and the scholarly tradition is not yet 100 years old. Moreover, with the exception of Hausa, Munjuk, and Ouldeme, each Chadic language description so far has been the work of just one scholar, so the critical discussion of hypotheses and argumentation that has characterized the Semitic scholarly tradition has not been available for Chadic languages. The same is true, to a lesser degree, of the Cushitic and Omotic languages.

Each of the next six chapters includes information on the following categories and topics, to the extent that such information is available. For some languages, there is no information available on some of the categories mentioned below, and in some languages the categories listed may not even exist. In some cases, the categories themselves are controversial in linguistic theory, but the fact that they are attested in Afroasiatic languages is part of what makes this phylum worthy of further study.

- Phonology
  - Segmental and suprasegmental units (phonemic and phonetic)
  - Major phonological processes
- Lexical categories attested
  - Presence and characteristics of the major lexical categories noun, verb, adjective, adverb
  - Presence and characteristics of adpositions, ideophones, or other lexical categories

Derivational morphology
  Means and direction of derivation
Inflectional morphology
  Inflectional morphology of verbs
  Functional domains coded by verbal morphology
  Inflectional morphology of nouns and pronouns (including case, number, and gender)
  Inflectional morphology of other lexical categories, including adpositions
Structure of noun phrases
  Possession
  Modification
  Other types of relationships between nouns
Position of the predicate
  In the verbal clause (clause-initial, clause-final, clause-internal)
  In verbless clauses
Major functional domains
  Relations between predicate and noun phrases
  Voice distinctions
  Epistemic modality
  Deontic modality
  Negation
  Polar questions
  Content questions
  Pragmatic functions, including topicalization, focus, and backgrounding
Categories of the complex sentence
  Paratactic constructions
  Clausal complementation and clausal order
  Number and functions of complementizers
  Subordinating constructions

The final chapter of the present book is a typological outline whose aim is to establish which coding means and functions are common to the six Afroasiatic families and which characterize a single family or a selection of families. In some cases, the chapter proposes reasons for these typological similarities and differences. For each of the topics and categories listed above, this chapter attempts to answer the following questions:

- What coding means and functions are shared by all language families in the phylum?

- What coding means and functions are unique to one language family or to a subset of language families?
- What coding means and functions, if any, are attested only within one or more subgroups of a single family?
- What are the cross-phylum correlations, if any, between the presence of a given coding means (e.g., position of the verb in the clause) and other coding means?
- What are the cross-phylum correlations, if any, between the presence of a given grammatical function and other grammatical functions?
- Are there correlations between certain form–function relationships (e.g. the coding of mood on auxiliary verbs) and the presence or nature of other form–function relationships?

When a characteristic is present in only one family or subset of families within the Afroasiatic phylum, the question may legitimately be asked whether this characteristic represents an innovation or a retention from Proto-Afroasiatic. Though references are made to historical reconstructions in some of the chapters on individual language families, we do not attempt to answer such questions for the phylum as a whole. The present volume is only one of the prerequisites for a historical reconstruction of the complete grammatical system of Proto-Afroasiatic.

The chapters in this volume were written between 2006 and 2011, with some chapters being submitted earlier than others. Understandably, the authors were not able to take into account all of the literature that has been published after their chapters were submitted.

2

Berber

Maarten Kossmann

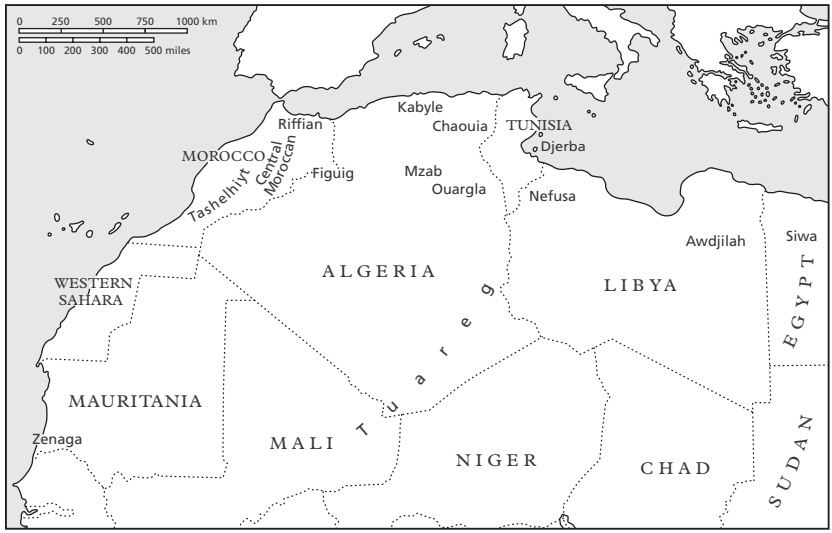
2.1 Introduction

Berber languages are spoken in North Africa, in a discontinuous region stretching from the Atlantic Ocean in the west to the Siwa oasis in Egypt in the east, and from the Mediterranean in the north to Burkina Faso in the south. The languages are quite similar to one another, and many Berberologists prefer to consider Berber one language with many dialects (e.g. Chaker 1995: 7ff.). This seems to be exaggerated, as linguistic variation inside Berber is roughly comparable to that found inside the Germanic or the Romance language families. The following list contains the best-known Berber languages and variants:

MAURITANIA: southwest: Zenaga;
MOROCCO: southwest: Tashelhiyt (also known as Chleuh, Shilha); central and southeast: Central Moroccan Berber (also called Middle Atlas Berber, Tamazight); north: Riffian (also Tarifiyt); northeast: Eastern Riffian (Beni Iznasen); northern Sahara: Figuig;
ALGERIA: northwest: Beni Snous, Chenoua; northeast: Kabyle, Chaouia; northern Sahara: Ouargla, Mzab, Gourara, Touat (now extinct);
TUNISIA: Djerba;
LIBYA: northwest: Djebel Nefusa; Libyan Sahara: Ghadames, Awdjilah, Elfoqaha (now extinct), Sokna (now extinct);
EGYPT: western Egyptian Sahara: Siwa;
TUAREG: Algeria, Libya: Ahaggar; Niger: Ayer, Iwellemmeden; Mali: Adagh des Ifoghas; Burkina Faso: Oudalan.

Subclassification inside the family is extremely difficult, as much of Berber is a kind of discontinuous dialect continuum – i.e., it seems to form a continuum, even though there

I wish to thank the editors of the volume for their valuable comments. Moreover, I wish to thank Maarten Mous, Thilo Schadeberg, and Christian Rapold for reading through and commenting on earlier versions of the manuscript. My special thanks are due to Harry Stroomer, who gave me access to his unpublished Tashelhiyt lexical materials.



Map 2.1 *Berber family*

are sometimes large stretches of Arabic-speaking regions in between different dialects. A few smaller groups, however, are clearly discernible, such as the Tuareg group and the continuum between Tashelhiyt and Central Moroccan Berber. More difficult to prove, but still a reasonable hypothesis, is the existence of the so-called Zenatic subgroup, which would comprise a large number of dialects in the north-central part of Berber, such as the Eastern dialects of Central Moroccan Berber, Riffian, sedentary Northern Sahara Berber (Figuig, Mzab, Ouargla), Chaouia and probably Djerba. On the other hand, the languages of Libya and Egypt do not seem to form one group. Zenaga (Mauritania) is in many points very different from the other Berber languages. While this is partly due to a large number of unique innovations in this language, Zenaga also preserves some very archaic features. Nobody would be surprised if, in a subclassification, Zenaga would surface as the first branching off the Berber family.

Moroccan and Algerian Berber (excluding Tuareg) share a number of typological features, which makes it practical to subsume them under the heading 'Northern Berber'. For the same reason, Libyan and Egyptian Berber will be called 'Eastern Berber'. It must be emphasized that these labels by no means represent a historical subclassification.

As noted above, Berber linguistic variation is not very great, and reminds one most of the variation inside Germanic or Romance. Therefore, it is a reasonable guess to place Proto-Berber somewhere in between 1000 BC and the beginning of the first millennium AD. This implies that it has by no means the time depth of other branches of Afroasiatic. If Berber constitutes a primary branch of Afroasiatic, as is generally assumed, it is highly probable that there existed sister languages to Proto-Berber, which have become extinct

in the course of time. One such branch could be Guanche, the language of the Canary Islands. Unfortunately, our limited data on this language, which became extinct in the seventeenth century AD, are so difficult to interpret that any linguistic classification must remain preliminary.

Many Berber groups refer to themselves as *imaziɣən* (singular *amaziɣ* or *maziɣ*). In other groups, this term may refer to a special class inside society. The language is referred to as *t(a)maziɣt*. By means of regular sound changes and assimilation rules, the well-known Tuareg denominations *tămahăq* (Ahaggar Tuareg), *tămažəq* (Niger Tuareg), and *tămašăq* (Mali Tuareg) go back to the same term. Some Berber activists object to the name 'Berber', which they associate with 'Barbarian', and prefer terms such as 'the Amazigh language'. As most Berberologists, including some who are very active in Berber culture and politics, continue to use the term 'Berber' in scientific publications, this will be done here too.

2.1.1 Historical records and writing systems

The earliest attestations of Berber (or a sister branch) are found in the so-called 'Libyan' (or Libyco-Berber; the term has no direct relation to the present republic of Libya) inscriptions, which are written in an alphabetic script of their own and date from around 800 BC to the first centuries AD (Pichler 2007). Inscriptions in this script are found all over Northern Africa, including the Canary Islands, but the largest number comes from present-day Tunisia. In spite of the enormous quantity of inscriptions, the linguistic materials they contain are quite few, as most of them are very short and consist predominantly of names and titles. Therefore, the exact linguistic classification of the language(s) written in this script is not certain, although some link with Berber is obvious. The Libyco-Berber script has been lost in most of Northern Africa, but an offshoot of it is still used by the Tuaregs, who call it *tifinaɣ* (see below).

From the eleventh century AD on (possibly even earlier), southern Moroccan Berber has been written in Arabic script (van den Boogert 2000). Early attestations include a large Arabic-Berber vocabulary from the twelfth century and a number of smaller textual fragments. Southern Morocco has had a continuous tradition of writing Berber in Arabic script up to modern times. Long texts from the sixteenth century AD onwards provide much information about pre-modern stages of Tashelhiyt (cf. van den Boogert 1997).

Modern Berber languages are written in one of three scripts: Latin script, Arabic script, or *tifinaɣ*. Latin script is preponderantly used in scientific publications and constitutes the standard choice in Kabylia (Algeria), one of the great centres of Berber linguistic planning and activity. A new Arabic orthography, which is independent from earlier scriptural traditions, has met with some success in Morocco, especially in the Tashelhiyt-speaking southern part of the country. Tuaregs traditionally use a script of

their own, based on the ancient Libyco-Berber script, which they call *tifinaɣ*. In spite of official language policy, which has long favoured Latin orthographies in the Sahel countries, *tifinaɣ* remains the most-used script among the Tuaregs. In recent years, a new version of *tifinaɣ* has been created by Berber activists in Algeria and Morocco. Although *tifinaɣ* has not been used by Berber groups other than the Tuareg since Roman times, and despite the differences from the traditional Tuareg script, this neo-*tifinaɣ* is presented as the age-old real traditional script of the Berbers. In 2003, it was chosen as the official script for writing Berber in Morocco, and is now being introduced in experimental teaching at the primary-school level.

2.1.2 Sociolinguistics

Until recently, Tuareg was the only officially recognized Berber language (so recognized in Mali, Niger, and Burkina Faso). In Morocco and Algeria, the existence of Berber has traditionally been denied officially. Nowadays, Berber has the status of official language, next to Arabic, in Algeria and in Morocco. Great efforts are made in Morocco to provide the country with a modernized, standardized, and unified version based on the three main Berber languages of Morocco: Tashelhiyt, Central Moroccan Berber, and Riffian. As of 2011, Berber is still vehemently oppressed by the Libyan government; it plays no role in the official discourse of Tunisia, Egypt, and Mauritania.

There exist no reliable figures for the number of Berber speakers, as in Morocco and Algeria language is either not part of the national census questionnaire, or its results remain unpublished. Therefore one can only make an educated guess. About 30 per cent of the Moroccan population speak a Berber language. The percentage in Algeria is much smaller, and may amount to about 20 per cent. Using 2004 population figures, this would mean about 8 million Berber speakers in Morocco and about 7 million Berber speakers in Algeria. One can add about 1 million Tuareg speakers in the Sahel countries. In the other countries, the number of speakers is much smaller, it seems. Thus there may be about 16 million speakers of Berber languages in the world nowadays.

Most Berber languages are alive and well, and only a few of them can be considered strongly endangered. Among these is Mauritanian Zenaga, which has only a few thousand, mostly elderly, speakers. The language situation in the small Berber pockets in Tunisia seems to be difficult too. In the Libyan oases, Berber has been diminishing for a long time, and a number of dialects have died out during the twentieth century (Sokna and Elfoqaha). The present situation of Berber in Ghadames and Awdjilah is unknown.

2.1.3 History of Berber studies

Berber languages have been the object of a large number of studies during the colonial period and since (see Bougchiche 1997). As the subject falls somewhere between

African studies on the one side (which mainly concentrates on sub-Saharan Africa) and oriental studies on the other side (which tends to focus on written languages), Berberology has developed into a separate discipline with its own dynamics and peculiarities. Thus, Berber historical linguistics has been the subject of very few studies (note, however, the great work by Karl-G. Prasse 1972–4), but under the influence of the most important Berberologist of the post-war period, Lionel Galand, Berber studies has developed great interest in syntactic analysis.

Among the more recent descriptive grammars of Berber languages, one may cite Penchoen (Central Moroccan Berber, 1973), Willms (Central Moroccan Berber, 1972), Bentolila (Central Moroccan Berber, 1981), Chaker (Kabyle, 1983), Kossmann (Figuig, 1997; Eastern Riffian, 2000a), Sadiqi (Central Moroccan Berber, 1997), Naït-Zerrad (Kabyle, 2001b), Sudlow (Tuareg, 2001), and Heath (Tuareg, 2005). There exist only a few general overviews of the Berber languages. The most important and complete is Basset (1952). Galand (1988), although focusing on Tashelhiyt, gives so much information on other variants that it amounts to a general introduction to the field of Berber studies.

In the following, an overview of Berber grammar will be given without focus on one particular variant; where necessary, the differences between the languages will be highlighted. It is of course no coincidence that many of the examples have been taken from Berber languages of northern and eastern Morocco, with which the author is most familiar. This overview is not a bibliographical article, and no effort has been made to describe the exact scientific history of certain analyses, nor should references be interpreted as exhaustive. Unless marked otherwise, examples from Riffian are based on personal notes by the author, or taken from the texts published in Kossmann (2003b). Transcriptions have been uniformized throughout the chapter.

Not many linguists have worked on the historical reconstruction of Proto-Berber. For this reason, few explicit references are made to Proto-Berber. Where historical developments are alluded to, the reader should interpret this as a reference to an earlier, reconstructible state of the variant, which may, but not necessarily must, represent Proto-Berber. Historical statements never refer to states of the language earlier than Proto-Berber.

2.2 **Phonology**

2.2.1 Consonants

Berber consonant systems typically have the constrasting features of voice, pharyngealization, and length. In addition, several languages have labialized consonants. In the northern half of the Berber-speaking territory, there took place a great consonant

Table 2.1 *Consonant system of Ahaggar Tuareg.*

	Labial	Dental	Palatal	Velar	Uvular	Laryngeal
Voiceless plosive		t		k		
		tt		kk	qq	
Voiceless fricative	f	s	(š)		(x)	
	ff	ss	(šš)		(xx)	
Voiced plosive	b	d	g^y			
	bb	dd	gg^y			
Voiced fricative		(z)	(ž)		ɣ	
		zz	(žž)			
Pharyngealized voiceless plosive		ṭṭ				
Pharyngealized voiced plosive		ḍ				
Pharyngealized voiceless fricative		ẓ				
		ẓẓ				
Nasal	m	n	(ɲ)			
	mm	nn		(ŋŋ)		
Glide	w		y			h
	(ww)		(yy)			(hh)
Rhotic		r				
		rr				
Liquid		l				
		ll				

Rare phonemes are put in parentheses.
Source: Prasse (1972–4; modified).

shift called spirantization, which has profoundly changed the pronunciation of the languages.

Berber consonant systems show great differences. These differences are mainly due to three processes: absence of labialization, spirantization, and the insertion of loan phonemes from Arabic (e.g. Kabyle *ṣ*, *ṭ*, *š*, *ž*, *q*, *x*, *ḥ*, *ʕ*). In tables 2.1 and 2.2, the consonant systems of two of the most widely varying dialects, Ahaggar Tuareg and Kabyle, are given. Double writings such as <tt> refer to long consonants.

2.2.1.1 *Pharyngealized consonants*

Pharyngealization, or 'emphasis', as it is sometimes called under influence of Semitic studies, is a central feature in the consonant system of all Berber languages. It has important lowering effects on the pronunciation of the vowels.

Pharyngealization is a prosodic feature that spreads over larger domains than one consonant only. This has led some scholars to consider it a feature on the word level

Table 2.2 *Consonant system of Kabyle*.

	Labial	Interdental	Dental	Pre-palatal	Palatal	Velar	Uvular	Pharyngeal, laryngeal
Voiceless plosive			(t) tt^{s}	(čč)		(k) kk	q qq	
Voiceless fricative	f ff	θ	s ss	š šš	ç		x xx	ħ ħħ
Voiced plosive	(b) bb		(d) dd	(ǧǧ)		(g) gg		
Voiced fricative	β	ð	z zz	ž žž	j̱		ɣ (ɣɣ)	ʕ ʕʕ
Pharyngealized/ labialized voiceless plosive			(ṭ) ṭṭ			(k^{w}) kk^{w}	q^{w} qq^{w}	
Pharyngealized/ labialized voiceless fricative			ṣ ṣṣ	(ṣ̌) (ṣ̌ṣ̌)	ç^{w}		x^{w} xx^{w}	
Pharyngealized/ labialized voiced plosive	bb^{w}		(ḍ)			(g^{w}) gg^{w}		
Pharyngealized/ labialized voiceless fricative		ð̣	ẓ ẓẓ	(ẓ̌) (ẓ̌ẓ̌)	j̱^{w}		ɣ^{w} (ɣɣ^{w})	
Nasal	m mm		n nn					
Glide	w							h hh
Rhotic			r rr	y				
Pharyngealized rhotic			ṛ ṛṛ					
Liquid			l ll					
Pharyngealized liquid			(l̩) (l̩l̩)					

Source: Chaker (1983: 64; modified).

rather than a feature spreading from one consonant. Most Berberologists, however, prefer a segmental analysis, in which pharyngealization is a distinctive feature of some consonants, which spreads to vowels and other consonants. The most obvious reason for doing so is the fact that pharyngealization occurs only in words that include specific consonants. Thus, a word with only the consonants *n* and *g* cannot have pharyngealization, although these consonants may (phonetically) be pharyngealized under the influence of another consonant. In words of Berber origin, only two pharyngealized phonemes are found, *ḍ* (long: *ṭṭ*) and *ẓ* (long: *ẓẓ*). They can constitute the sole consonant of a word. Through borrowing from Arabic and through secondary developments, other pharyngealized consonants have become part of the Berber phoneme inventory.

The domain of pharyngealization spread is an intricate question. At least in some variants of the language, word boundaries are of secondary importance to the spreading of pharyngealization, the make-up of the syllable being much more important (e.g. Tashelhiyt: Boukous 1990). Depending on speech tempo, the domain can be larger or smaller, although it is hardly ever restricted to one segment only.

2.2.1.2 *Long consonants*

In Berber almost every consonant has a long counterpart. Northern Berber languages are unusual typologically in that long consonants are allowed to occur utterance-initially as well as utterance-finally, and contrast in these positions with short consonants. As there exists an important discussion pertaining to the question of whether these long consonants are mono-phonemic tense consonants (cf. Galand 2002a [1997]: 147–61) or bi-phonemic geminates (e.g. Saïb 1977), the neutral term 'long' consonant will be used here. Although length is not necessarily the only or the main feature which differentiates 'short' consonants from 'long' consonants, it always plays a role in the opposition (see Chaker 1984; Ouakrim 1995: 56–8; Louali and Maddieson 1999) and can therefore be taken as definitory for the class.

The short–long contrast plays an important role in morphology, where both consonant lengthening and consonant shortening processes occur. Compare Figuig *ini* 'say!' vs *inna* 'he said'; *ifhəm* 'he understands' vs *ifəhhəm* 'he always understands'; *iɣəlwas* 'jars' vs *aɣəllus* 'jar'. The existence of such pairs allows us to establish two series of consonants, which interact in the morphophonology of the language. With all consonants, phonetic length and phonetic tenseness (Louali and Puech 1994; Ouakrim 1995: 68–75) are among the distinguishing features between the two series. In many consonant pairs these two features are supplemented by other phenomena:

(1) A voiced short plosive may have a voiceless long consonant as its counterpart. In addition to Pan-Berber *ḍ* - *ṭṭ*, dialectally other plosives are

devoiced. This may affect all plosives, as is the case in Figuig and in the female pronunciation of Kabyle.

(2) A fricative short consonant may have a plosive long consonant as its counterpart, as in Pan-Berber *ɣ - qq*. In spirantizing dialects (see below), spirantized consonants correspond to plosive long consonants, thus giving correlations such as Kabyle: *β - bb*, *ð - dd*, *ḍ - ṭṭ*, *θ - tt*^{s}.

(3) In most Berber variants, short *w* has long *gg*^{w} as its counterpart.

(4) A short sibilant may have an affricate long consonant as its synchronic counterpart. This is dialectally attested in a number of Northern languages (e.g. Figuig and Kabyle) in a small number of morphological contexts. In other contexts in these languages the fricative pronunciation is also found in the long consonant.

In most cases there is no difference between lexical long consonants, or those resulting from morphophonological alternations, and long consonants that are products of the conflation of two identical short consonants (e.g. in sandhi). However, when the phonetic differences between long consonants and their short counterparts are very large, there may be a difference between the products of conflation and those found in morphophonology.

A long consonant followed by a short consonant, or the other way round, is not necessarily assimilated to become one long consonant, e.g. Figuig *šətt'tən* ([šətˢ:tˢən]) 'they dispersed', Tashelhiyt *tt'tahaln* [t:tahalṇ] 'they always marry'. A long consonant may function at the same time as the coda of a syllable and the onset of the following syllable. Therefore, in languages which do not permit schwa in open syllables, it is possible to have schwa followed by a long consonant, e.g. Figuig *išəttət* 'he dispersed'.

2.2.1.3 *Labialization*

Labialized velars and uvulars are a common feature of Northern Berber phonology (see Chaker 1984: 90–3; Leguil 1981). In Libyan Berber and in Tuareg, labialization is not attested. The most extensive systems of labialization are found in Kabyle, Central Moroccan Berber, and Tashelhiyt. In these languages, any velar or uvular can be labialized. Although minimal pairs are relatively rare, there is no doubt as to the phonological nature of the opposition labialized–non-labialized. Compare Tashelhiyt *ig*^{w}*ra* 'frogs' vs *igra* 'he threw' (van den Boogert 1997: 243).

While the labialization in *gg*^{w}, the tense counterpart of *w*, is attested in many Berber variants and may very well be reconstructible to Proto-Berber, the labialization of short consonants is probably a secondary development. It has been suggested that it is a remnant of ancient short **ŭ*, which became schwa in these dialects but whose labial feature was taken over by neighbouring velars and uvulars (see Kossmann 1999: 42–59).

Table 2.3 *Absence and presence of spirantization in selected languages.*

Ouargla	Kabyle	Temsamane (Rif)	
abrid	*aβrið*	*aβrið*	'road'
idammən	*iðammən*	*iðammən*	'blood'
tala	*θala*	*θaře, haře*	'well'
aḍil	*aðil*	*aðiř*	'grape'
yugəm	*yuʝəm*	*yuyəm*	'he pulled (water)'
kməz	*çməz*	*šməz*	'scratch!'

2.2.1.4 *Spirantization*

In the Berber variants spoken in the northern part of the Berber territory, reaching from Tunisia in the east to the Moroccan Rif in the west, one finds a pervasive phonetic development, namely spirantization. In this historical process, short stops have become fricatives. Moreover (where possible), their place of articulation is fronted with respect to the original stops. Thus, dental stops have become interdentals, while velar stops have become pre-palatals. By subsequent changes, some spirantized consonants have merged. Very common is the merger of *ʝ* with *y* and of *ç* with *š*. Table 2.3 lists some examples in Ouargla (no spirantization), Kabyle (spirantization without subsequent mergers), and Temsamane Riffian (spirantization with some subsequent mergers).

In a cross-dialectal perspective, the following hierarchy obtains as to which plosives are spirantized and which are not:

velar > dental > bilabial

This means that in those dialects where *b* is spirantized, the dentals and the velars are also spirantized. On the other hand, spirantization of velars does not necessarily imply spirantization of dentals and bilabials.

In those dialects where all three rows of plosives are spirantized, the effects are enormous. Most words contain at least one spirantized consonant, and in many words all consonants have been spirantized, as for example in Riffian *θçəðβəð* < *tkədbəd* 'you lied'.

Spirantization is blocked by a preceding liquid or a preceding homorganic nasal, e.g. Riffian *θandint* < *tandint* 'city' (not *θanðinθ*). In some variants, other preceding consonants have the same effect. Mauritanian Zenaga and the language of the Tunisian island of Djerba have no spirantization in the initial position.

Although originally the result of a regular phonetic process, most spirantized consonants have acquired phonemic status in the affected dialects. This is due to the development of new non-spirantized short consonants, either by the introduction of non-spirantized forms in borrowings or by the irregular shortening of a long consonant. In the

Table 2.4 *Vowel systems in two Berber languages.*

Tashelhiyt	i		u
		a	
Ayer Tuareg	i		u
		ə	
	e	ă[1]	o
		a	

following Riffian example, the latter process is attested twice: *itawi-d* 'he brings' < **yəttawəy-dd*, where both the Imperfective marker **tt*- and the deictic clitic *dd* have been shortened without being spirantized.

2.2.2 Vowels

Berber languages differ considerably as to their vowel systems. Languages such as Tashelhiyt have only three phonemic vowels, while Tuareg and Ghadames Berber have a seven-vowel system (see table 2.4).

In normal speech, central vowels are pronounced shorter than peripheral vowels (which, in this discussion, include *a*). There exists discussion as to the analysis of the central–peripheral contrast. While phonetic data suggest that the main difference lies in vowel quality (see Louali 2000), Tuareg metrification, which is quantificational, shows that in this language central vowels are treated as shorter segments than peripheral vowels (e.g. Mohamed and Prasse 1989: 11ff.). Most scholars call *ə* and *ă* short vowels, whereas the peripheral vowels are referred to as plain vowels. Systems with three plain vowels and one central vowel, *ə*, of questionable phonemic status, are standard in Northern Berber. Eastern Berber and Tuareg systems have more vowel phonemes. Vowel-length oppositions other than the one concerning central vs plain vowels are rare in Berber and are always the result of consonant loss, e.g. in Riffian due to the loss of **r* and in Zenaga of Mauritania due to the loss of **β*. The so-called 'over-long' vowels of Tuareg have now been re-analysed as plain vowels under stress (see especially Louali and Philippson (2005) and Heath (2005)).

2.2.2.1 *Central vowels*

One of the major problems in Berber phonology is the analysis of the central (or short) vowel system. All Berber languages, with the exception of Mauritanian Zenaga, have one or two short central vowels, which cannot be interpreted as simple allophones of plain vowels. Some languages allow central vowels both in closed and in open syllables; in these languages the phonemic status of the central vowels is beyond doubt. In other languages, schwa is only allowed in closed syllables. In these languages the presence

and absence of schwa can be predicted in most cases, and therefore its phonemic status is questionable.

Central vowels in non-final open syllables are attested in Tuareg and in Libyan Berber. In Tashelhiyt, they occur in medieval texts (see van den Boogert 2000) but not from the sixteenth century on. In these dialects the position of the central vowels cannot be predicted from the structure of the word. They can contrast with the absence of a vowel. Compare, for example, Ghadames *i-βðăr* 'he mentioned' vs *abədăr* 'part of a door', and Iwellemmeden Tuareg *talăkat* 'iron ring' and *eləki* 'mattress put behind a saddle' vs *talkit* 'regression'.

A number of varieties which permit central vowels in open syllables present an opposition between two central vowels, *ə* and *ă*. This opposition is well documented in Ghadames and in Tuareg. The contrast may be shown by the following Ghadames forms: *βərǧ* 'dream!' vs *aβărǧ* 'beam' (cf. also *ăβrəǧ* 'pulverize!').

Modern Northern Berber languages do not allow central vowels in open syllables. None of these variants has more than one central vowel. In fact, schwa is a problematic segment here, phonetically as well as phonologically. The pronunciation of schwa differs greatly according to context, speech tempo, and dialect. In many dialects, it is basically a short central vowel that undergoes various assimilations to adjacent consonants, especially to semivowels. In some contexts and speech tempos, it is hardly audible, although in many varieties native speakers, when asked, show a clear notion of its presence, even when it seems to be phonetically absent. On the other hand, schwa may be lengthened in the same way as plain vowels in certain intonation types. An alternative formulation states that what is called schwa here is in fact the inherent syllabicity of a consonant, which may have different pronunciations depending on the consonant and on other factors, such as speech tempo. In the following, this 'consonantal syllabicity' will be written by the sign <ə> put before the consonant.

The main issue at stake in the analysis of Northern Berber schwa (or consonantal syllabicity) is the fact that it seems to function as a syllable-building device, i.e., schwa is inserted in order to create well-formed syllables. Following Dell and Elmedlaoui (1985), I will refer to this process as 'syllabification'.

In Northern Berber two types of syllabification are found. In the first type the place of schwa-insertion (or, to put it differently, the choice of the syllabified consonant) is governed by the inherent sonority of the segments in a string. In the second type of syllable building, schwa is inserted from right to left between two consonants. Because in this type only the formal structure of the string plays a role and the nature of the consonant is of no importance, it will be called structure-based insertion or syllabification. In both types, schwa in an open syllable (or, in other terms, two adjacent syllabic segments) is forbidden. In structure-based syllabification, this leads to the preponderance of insertions such as |ccc| >|cc**ə**c|.

Table 2.5 *Sonority-based syllabification, Tashelhiyt examples.*

Underlying	Syllabification	
iɣml	*iɣ.məl*	'it (M) went mouldy'
islm	*i.səlm*	'it (M) went numb'
tzmt	*tzəmt* ~ *t.zəmt*	'it (F) is stifling'
utx=k	*u.təxk*	'I struck you'
ra tkti	*ra.tək.ti*	'she will remember'
tftkt	*təf.təkt*	'you suffered a sprain'

Source: Dell and Elmedlaoui (1985: 112–13).

In a number of dialects (especially Central Moroccan Berber), the two types are combined. The influence of sonority is restricted to certain segments (vowels, semivowels, liquids and nasals), while the other segments are syllabified according to the second type. Syllabification that is exclusively based on sonority is attested only in Tashelhiyt. Outside of central and southern Morocco, only structure-based syllabification is found.

Sonority-based syllabification In a classic article, Dell and Elmedlaoui (1985) present an analysis of Tashelhiyt syllable-building devices. Their model explains the syllable structure of words and phrases in isolation, but also predicts the situation in larger parts of discourse. The analysis is based on three principles.

(1) Two syllabic segments cannot be adjacent.
(2) Segments have an inherent sonority. The inherent sonority depends on the type of segment, and can be represented as a hierarchy. According to Dell and Elmedlaoui, this sonority scale, which is supposed to be universal, is the following: *a* > *i*/*y*, *u*/*w* > liquids > nasals > voiced fricatives > voiceless fricatives > voiced stops > voiceless stops.
(3) Syllabification follows the sonority hierarchy, the first segments to be syllabified in a string being the highest in the hierarchy, the last being the lowest in the hierarchy. Given an appropriate context, any segment may become syllabic.

Compare table 2.5, which illustrates the principles in Tashelhiyt.

According to Dell and Elmedlaoui (1985), there exists no phonemic contrast between the semivowels [w], [y] on the one hand and the high vowels [u], [i] on the other hand. Vocalic realizations are found when the phoneme is in a syllabic position, while the semivowels are found when the phoneme is in a non-syllabic position. This correctly describes most instances of *y*, *w*, *i*, and *u* in Tashelhiyt, but fails to predict what happens

Table 2.6 *Structure-based syllabification, Figuig examples.*

Underlying form	Phonetic realization		
ils	*i.ləs*		‘tongue’
šrz	*(ə)š.rəz*		‘plough!’
tfhm	*təf.həm*		‘she understands’
itdɣdɣ	*it.dəɣ.dəɣ*		‘he breaks to pieces’
bddl	*bəd.dəl*	** *bddəl*	‘change!’
tssḍṛɣl	*təss.ḍəṛ.ɣəl*	** *tsəs.ḍəṛ.ɣəl*	‘she blinded’
iffxs	*iff.xəs*	** *i.fəf.xəs*	‘it started to blossom’

Source: Kossmann (1995, 1997).

when two segments of this type are adjacent. In fact, one finds here two different solutions: either the first segment is syllabic, or the second. These solutions are lexically determined, e.g. *abrzzuy* ‘piece of bread or meat’ but *isswi* ‘irrigation’. It may therefore be preferable to differentiate between high vowels and semivowels (van den Boogert 1997: 247–9 and 253).

In languages with sonority-based syllabification, the position of the syllabic peak is entirely predictable. There is no reason to posit a central vowel phoneme (or phonemic consonantal syllabicity) there.

Structure-based syllabification Most analyses of syllable building in the Northern Berber languages in which consonantal sonority does not play a role consider schwa as an instance of regular vowel-insertion procedures. In Berberology, this analysis is proposed by followers of French structuralism (beginning with André Basset), as well as by advocates of non-concatenative phonology (e.g. Bader 1985; Dell and Tangi 1992). It will be called here ‘structure-based syllabification’ in order to distinguish it from ‘sonority-based syllabification’, as treated above. The main rules and restrictions in structure-based syllabification are the following.

(1) In a string of two consonants, schwa is inserted. Insertion is from right to left.
(2) Schwa cannot occur in open syllables.
(3) A long consonant can be ambisyllabic, and cannot be dislocated.

Some Figuig examples, presented according to the insertion analysis appear in table 2.6.

These insertion rules predict the place of schwa in the great majority of masculine Berber nouns and most verbal forms in the relevant dialects, but there are a number of problematic forms. In the first place, there exist a good number of lexical exceptions

which have |əcc| rimes instead of the expected |cəc| rimes. Most of these are borrowings from Maghribine Arabic, which has different syllabification procedures. A small number of Berber nouns have similar syllabification, e.g. Figuig *adəxs* 'colostrum', *amərd* 'small grasshopper', *inərz* 'heel'.

The behaviour of nominal and verbal suffixes poses important problems for a strictly phonetic analysis (see Kossmann 1995). In the languages under investigation, there exist some consonantal suffixes which are syllabified (e.g. imperative masculine plural *-ət*, as in Figuig *šərz-ət* 'plough!'), while other consonantal suffixes are not syllabified (e.g. nominal feminine singular *-t*, as in Figuig *ta-məsləm-t* 'Muslim woman'). It is not useful to explain this by positing a difference in the type of boundary between the lexical basis and the suffix. In the first place, there is no reason to assume that verbal inflection would be different in principle from noun inflection. In the second place, the masculine plural suffix of Berber nouns *-ən*, which belongs to the same paradigm as the feminine singular suffix *-t*, does allow for schwa insertion.

As noted by Dell and Tangi (1992: 134), most counter-examples to syllabification rules have a coronal obstruent as their last segment. One should note however, that this is not always the case (e.g. Figuig *ššərħ* 'explication') and that in the majority of cases final coronal obstruents are not exempted from syllabification (e.g. Figuig *imrəz* 'he has a head wound' as opposed to *inərz* 'heel'). Proposals to treat these segments as extra-prosodic are therefore not helpful.

In order to deal with this situation, Kossmann (1995 and 1997) has proposed the positing of phonemic ('inherent') /ə/ in those cases where schwa is not predictable by rule. In addition to this, syllabic structures are obtained by a three-step derivation.

> Rule 1. Base-level insertion: schwa is inserted from right to left into the basis of the word – i.e. the word without affixes – unless this would lead to schwa in an open syllable.
>
> Rule 2. Word-level resyllabification: when an open syllable with schwa is created due to suffixation, the basis is resyllabified from right to left. If not, the structure remains the same.
>
> Rule 3. Splitting of three-consonant groups. When, as a result of the two preceding rules, a group of three consonants (or of two consonants, one of which is long) appears, schwa is inserted, unless this would lead to schwa in an open syllable. See table 2.7.

An alternative solution would be to posit underlying schwa everywhere, in addition to vowel metathesis in all bases, a solution favoured by Jilali Saïb (Saïb 1994). According to this analysis, in those cases where the addition of suffixes leads to schwa in an open syllable, schwa is metathesized.

Which analysis is preferred depends of course on the theoretical premises one has. A phonetic interpretation of schwa insertion cannot account for all forms, and some kind

Table 2.7 *The derivation of some Figuig forms according to Kossmann (1995).*

Underlying form	Insertion rule 1	Insertion rule 2	Insertion rule 3	
a-mərd	*a-mərd*	*amərd*	*amərd*	'small grasshopper'
a-slm	*a-sləm*	*asləm*	*asləm*	'fish (M)'
ta-slm-t	*ta-sləm-t*	*tasləmt*	*tasləmt*	'fish (F)'
fhm	*fhəm*	*fhəm*	*fhəm*	'understand (SG)!'
fhm-ət	*fhəm-ət*	*fəhmət*	*fəhmət*	'understand (PL:M)!'
i-slm-an	*i-sləm-an*	*isəlman*	*isəlman*	'fishes (M)'
t-fhm	*t-fhəm*	*tfhəm*	*təfhəm*	'she understands'
t-fhm-əmt	*t-fhəm-əmt*	*tfəhməmt*	*tfəhməmt*	'you (PL:F)

Source: Kossmann (1995).

Table 2.8 *Examples of distinction between Aorist and Perfective by stress.*

	Aorist 3SG:M	Perfective 3SG:M	
Ghadames	*y-ǎknəf*	*i-knǎf*	'roast'
Djebel Nefusa	*(a) y-ə́fhəm*	*y-əfhə́m*	'understand'
Elfoqaha	*(a) y-ə́gzəm*	*y-əgzə́m*	'cut'

of ad hoc solution is needed in order to explain the exceptions. The intermediate analysis of Kossmann (1995 and 1997) accounts for all forms but demands a rather complicated set of derivational rules. An analysis according to which schwa is always phonemic is much simpler in rule apparatus but misses important generalizations.

2.2.3 Suprasegmentals

Berber languages have no lexical tones. In addition, it is generally assumed that Northern Berber languages have no lexical stress either. Although it is possible to define some rules for the accentuation of words in isolation, the rules may refer to sentence intonation rather than to lexical stress (see Chaker 1995: 97–116). In the Eastern dialects and in Tuareg, lexical stress plays a more important role, but even in these languages, stress is only rarely used to distinguish lexical items.

The Tuareg system is quite complicated (Louali and Philippson 2005; Heath 2005). Some nouns and verb forms have lexically determined stress on a certain syllable, while in other forms stress falls by default on the ante-penultimate syllable.

Stress plays an important, although not always well understood, role in the grammatical structure of those Berber languages which have it. In the verbal system of Eastern and Central Libyan dialects, for example, stress is used to distinguish between Aorist forms and Perfective forms (see table 2.8).

Table 2.9 *Syntactically conditioned stress fronting in Eastern Berber.*

	Form in isolation		Stress fronting	
Siwa	*agbə́n*	'house'	*əlbab n ágbən*	'the door of the house'
	tasárt	'mill'	*fus n tásart*	'the handle of the mill'
	amán	'water'	*y áman*	'into the water'
Nefusa	*uššə́n*	'jackal'	*əssag n úššən*	'the leg of the jackal'
			iẓun úššən	'the jackal divided'
	ɣasrú	'castle'	*s ɣásru*	'from the castle'

Sources: Vycichl (1981 – Siwa) and Brugnatelli (1986 – Djebel Nefusa).

A number of Eastern dialects manifest stress fronting in certain syntactic contexts (Vycichl 1981; Brugnatelli 1986; Louali and Philippson 2005); see table 2.9.

2.3 **Lexical categories**

2.3.1 Overview

All Berber languages attest to the following lexical categories: nouns, verbs, pronouns, adverbs, prepositions. Ideophones occur, but are much less prominent than in many other African languages: both in number and in frequency in actual discourse, they are reminiscent of the restricted standard European usage (English *boom!*).

The most problematic lexical category is the adjective. Tuareg lacks this category altogether and uses relative clauses with stative verb forms in contexts involving modification of a noun. A similar situation is found in Tashelhiyt, which has only one adjective, *lždid* 'new', which is a loanword from Arabic. Other Northern and Eastern Berber languages have nominal forms corresponding to the adjective. These forms have all the morphological characteristics of normal nouns. They can be both heads of an NP and attributes to the NP head. The fact that they can modify a nominal head is the only characteristic which makes them different from normal referential nouns. Therefore, in these languages, adjectives are best considered a sub-category of the noun.

Recently, Catherine Taine-Cheikh (2003) proposed considering the Berber stative conjugation as an adjectival formation. Following this argumentation, one could consider the adjective as a category different from, but close to, the verb.

2.3.2 Derivation of lexical categories

The Berber lexicon is defined by means of abstract 'roots', which are unspecified for the lexical category they belong to. The lexical category is determined by the means of derivation. This general statement must be qualified, as there exist many nouns which

have no verbal counterpart and whose roots can therefore be considered inherently nominal. With very few exceptions, verbs always have at least one nominal counterpart, the Verbal Noun (see Galand 2002b).

The abstract roots are mainly consonantal. In the classical analysis of Berber (and Semitic) word formation, consonantal roots receive vowel schemes (as well as consonant lengthening), which contain grammatical information, e.g. aspect. This is illustrated by the following Ayer Tuareg forms: *t-əkrəḍ* (Aorist 3SG:F) and *t-əkrăḍ*, (Perfective 3SG:F) 'scrape'. In this example, the vowel scheme of the Aorist is *-ə-ə-*, while that of the Perfective is *-ə-ă-*.

In order to express an opposition, several different sets of vowel schemes may be used. Thus, in addition to the scheme found in the Tuareg forms *t-əkrəḍ – t-əkrăḍ* above, Tuareg has a different set of vowel schemes for these aspects with verbs of a slightly different structure, e.g. *t-ăkkărăḍ* (Aorist 3SG:F) and *t-əkkərăḍ*, (Perfective 3SG:F) 'force'. In this verb, the vowel scheme is Aorist *ă-ă-ă* vs Perfective *ə-ə-ă*.

The choice of the set of aspectual schemes is to a large extent ruled by the structure of the verb root, including the number of consonants involved.

A strict analysis with consonantal lexical roots and vocalic grammatical schemes is difficult to maintain for all verbs, however. In a synchronic analysis, assuming that all abstract roots are exclusively consonantal is highly problematic, and one finds many lexical verbs, for instance, which are only differentiated by means of their vowel. This is illustrated by the following verbs in Iwellemmeden Tuareg (Aorist imperative singular forms), which share the consonants *f* and *l*: *əfəl* 'leave; be covered (by a roof)'; *afəl* 'be tanned'; *ăflu* 'appear suddenly'; *ful-ət* 'reinforce (a well) with wooden beams'; *făllu* 'count fully on'. Historically, at least some of these verb forms contain consonants no longer present in modern Tuareg (cf. Vycichl 2005: 68–73; Prasse 1972–4; Kossmann 2001b).

The vowels in these different forms may change in other aspects. When the basic structure of the verb, including vowel positions, is known, one can make predictions about the nature of these changes. Due to phonetic developments in Northern Berber (especially the loss of the opposition between *ə* and *ă*), vowel schemes have collided in many cases, and an analysis featuring grammatical vowel alternations on a more local level (which one might call 'ablaut' or 'apophony') may be more economical than a vowel scheme analysis. In most cases it is arbitrary to consider one of the vowel schemes basic, and the others derived from this vowel scheme. They constitute instead a pattern of vowel scheme sets without obvious directionality.

There is no regular compounding in Berber. There do exist, however, compound words. These are mainly found in the semantic domains of plant names and, to a lesser extent, body parts, e.g. Eastern Riffian: *a-ɣəsmir* 'lower jaw' (i.e. 'beard-bone', compare *iɣəss* 'bone' and *θ-mar-θ* 'beard') and *a-ɣəzðis* 'rib' (i.e. 'belly-bone', compare *iɣəss*

Table 2.10 *Derivation on the basis of the Ayer Tuareg verb root KRZ.*

Basis (without derivation)	KRZ
MAN stem (with the Short Perfective vowel scheme *ə-ă-*)	*əkrăz*
Inflectional prefix (1PL *n-*)	*nəkrăz*
(Extended) basis (with the sibilant causative derivation)	S-KRZ
MAN stem (with the Short Perfective vowel scheme *ə-ə-ă-*)	*əzzəkrăz*
Inflectional circumfix (2PL:M *t- . . . -ăm*)	*təzzəkrăzăm*
(Extended) basis (with the middle derivation *mm-*)	M-KRZ
MAN stem (with the Imperfective vowel scheme *a-ă-a* + the Imperfective prefix *t-*)	*tamăkraz*
Inflectional suffix (3PL:M *-ăn*)	*tamăkrazăn*

'bone' and *a-ʕəddis* 'belly'). It is uncertain whether these words represent remnants of an earlier regular compounding.

2.4 **Verb morphology**

2.4.1 General structure of the verb

The structure of the verb can be described by the following derivational account. The lexical part of the verb is constituted by an abstract root (which is not always exclusively consonantal in Berber). This can be extended by one or more derivational prefixes, which add modifications in voice. The ensuing form will be called the (extended) basis. To the basis (whether extended or not), a vowel scheme is applied, which adds information about mood, aspect, and negation (MAN). In addition to the vocalic scheme, some Imperfective forms have a MAN prefix *t(t)-*. The ensuing form will be called the MAN stem. To this MAN stem, inflectional prefixes and/or suffixes are added, which convey information about person, gender, and number. This is a word. This derivation is illustrated in table 2.10 by the following Ayer Tuareg verb forms, based on the abstract root KRZ: *nəkrăz* 'we obtained', *təzzəkrăzăm* 'you (PL:M) made obtain' and *tamăkrazăn* 'they (always) heal (fractured bones)'.

Although it is impossible to decide which aspectual vowel pattern is basic, for the sake of convenience the Aorist form is taken as the citation form.

2.4.2 Derivational prefixes

Berber has several derivational prefixes which may be attached to the verb root. Derived verbs are inflected for all MAN formations, with the possible exception of the passive, which seems to be excluded in the Imperfective in some Berber languages (e.g. Riffian, Cadi 1989–90: 265).

2.4.2.1 *The 'factitive' prefix* ss-

The prefix *ss-* is a valency-increasing morpheme. It has the following functions.

(1) It is sometimes added to nouns to derive verbs. The result may be intransitive or transitive. Some languages also use the prefix *ss-* in order to make verbs from onomatopoeia. E.g. Ouargla: *awal* 'word' > *ss-iwəl* 'speak'.

(2) It is added to an intransitive or a labile verb (i.e. a verb which may be used as transitive or intransitive without additional marking) to make a transitive verb. E.g. Ouargla: *kkəfkəf* 'foam (verb)' > *ss-kəfkəf* 'make (something) foam'; *kkərdəd* 'be rough' > *ss-kərdəd* 'render (something) rough'; *msəl* 'stop up, be stopped up' > *ss-əmsəl* 'stop up (something)'.

(3) It is added to transitive or labile verbs to make them causative, e.g. Ouargla: *əmḍi* 'taste (something)' > *ss-əmḍi* 'make taste (something)'; *qqən* 'attach (something) / be attached' > *ss-əqqən* 'make (somebody) attach (something)'.

The function of transitivizer of an intransitive verb (2) is productive in many (possibly all) Berber languages. Valency increasing with transitive verbs (3) is in many languages restricted to a small number of verbs; in some languages (Tuareg, Ouargla), it is productive.

2.4.2.2 *The 'middle' prefix* mm- ~ nn-

This prefix has a number of functions, including the reciprocal, e.g. Central Moroccan Berber: *rdəl* 'fall, make fall' > *m-ərdal* 'make each other fall'; *rḍəl* 'lend to somebody' > *m-ərḍal* 'lend to each other'. There exists important dialectal variation as to the semantics of this prefix, which can have several other 'middle' connotations and is not everywhere exclusively reciprocal, as it is in Central Moroccan Berber. It is important to note that the *mm-* derivation is only rarely used as a marker of the reflexive, for which a circumlocution is used. In those Eastern Berber variants which lack the passive morpheme *tt-* (etc.), the nasal prefix functions as a passive, e.g. Ghadames *ănn* 'kill' > *əmm-ənn* 'be killed'.

2.4.2.3 *The 'passive' prefix* tt-*,* ttwa-*, etc.*

There exists much dialectal variation in the form of the passive prefix. The passive prefix is used in order to make transitive verbs intransitive. In most languages, we are dealing with an 'agent-less' passive: it is not possible to express the agent of the passive by means of an agent-phrase. Figuig Berber makes a distinction between two different passive morphemes. The first morpheme, *ttwa-* is a true passive where the speaker has the agent in mind. It can be used with an agent phrase. The second morpheme, *tt(u)-*, has been called a medio-passive, as the speaker does not have the agent in mind.

2.4.2.4 *Combinations of derivational prefixes*

Derivational prefixes can be combined. In Tuareg, there seem to be no morphological restrictions on such combinations, as is illustrated by the following verb forms from Iwellemmeden Tuareg (data from Aghali-Zakara (2001), supplemented by forms from Prasse *et al*. (2003)): *əgru* (understand:A) 'discern, understand'; *sə-gru* (CAUS-understand:A) 'make discern, make understand'; *təwə-gru* (PASS-understand:A) 'be discerned, be understood'; *nă-gru* (MED-understand:A) 'be wise, be tranquil'; *sə-nnə-gru* (CAUS-MED-understand:A) 'have discernment, be intelligent, examine attentively'; *təwə-sə-nə-gru* (PASS-CAUS-MED-understand:A) 'be examined attentively'; *nə-mə-gru* (MED-MED-understand:A) 'understand each other'; *mə-sə-gru* (MED-CAUS-understand:A) = *mə-sə-sə-gru* (MED-CAUS-CAUS-understand:A) 'make oneself understood to each other'; *sə-nnə-mə-gru* (CAUS-MED-MED-understand:A) 'make each other understand, create mutual understanding'; *təwə-sə-ssə-gru* (PASS-CAUS-CAUS-understand:A) 'be made to understand'; *nə-mə-təwə-sə-ssə-gru* (MED-MED-PASS-CAUS-CAUS-understand:A) 'make understand each other'. Such freedom of combination is exclusively found in Tuareg. Outside Tuareg, combinations of more than two prefixes are rare. In some languages, certain combinations have been given specific nuances of meaning that do not constitute a simple sum of the semantics of their constituents. Thus, in Figuig the combination *m-s-* is used to express the reciprocal, while Tuareg uses for this purpose the combination *nə-mə-* (< *mə-mə-*).

2.4.3 Reduplication

In addition to verbal stems that have a more or less simple structure, Berber languages have a large number of alternative root formations, which attest to different forms of reduplication. The functions of some of these groups are clear; in other cases, no specific semantics seem to correlate with the use of a certain reduplication pattern. In all languages except Tuareg, these reduplicating-stem types are normally not derived from attested non-reduplicated roots. Tuareg is different from the other languages, as full reduplication is used as a derivational device for expressing pluractional meaning, as shown in the following forms from Iwellemmeden Tuareg (W) and Ahaggar Tuareg (H): *əbəḍ* 'make a hole' > *(b)băḍbăḍ* 'make here and there a hole' (W); *əfrəs* 'cut' > *fərəsfərəs* 'cut in several pieces' (W); *əbdəɣ* 'beat violently' > *bədəɣbədəɣ* 'beat here and there violently' (H); *ənḍər* 'jump repeatedly up and down' > *nəḍərnəḍər* 'go up and down (buttocks on camel)' (H). This type of formation is not found elsewhere in Berber. It is probably due to sub-Saharan influence, either from Hausa or from Songhay, which both have pluractionals derived by reduplication.

Table 2.11 *Aspectual stems of some classes of verbs in selected Berber languages.*

	Figuig	Ghadames	Ayer Tuareg	Mali Tuareg
Aorist	*ələməd* *atəf*	*ălməd* *atəf*	*ələməd* *atəf*	*ələməd* *aləm*
Imperfective	*ləmmməd* *ttatəf*	*lămmăd* *ttatăf*		*lămmăd* *tiləm*
Secondary Imperfective			*lámmăd* *tátăf*	*lámmăd* *tíləm*
Negative Imperfective	*ləmmməd* *ttitəf*	*ləmmməd* *ttitəf*	*ləmmməd* *titəf*	*ləmmməd* *tiləm*
Perfective	*ələməd* *utəf*	*ələmăd* *utăf*	*ələmăd* *otăf*	*ələmăd* *olăm*
Secondary Perfective			*ələmád* *otáf*	*ələmád* *olám*
Negative Perfective	*ələmid* *utif*	*ələmed* *utef*	*ələmed* *otef*	*ələmed* *olem*
Future		*ələmăd* *utăf*		

Sources: Kossmann (1997); Lanfry (1968); Prasse *et al*. (2003); Prasse and ăgg Albosṭan (1985).

2.4.4 Mood-Aspect-Negation stem formations (MAN stems)

2.4.4.1 *Inventory*

As shown above, MAN stem formation consists of the application of certain vowel schemes to the root. There are different sets of vowel schemes available, which roughly correspond to certain root types; thus most three-consonant verbs without a plain vowel share the same MAN vowel schemes, while four-consonant verbs without a plain vowel have different MAN vowel schemes. Therefore, Berber verbs can be classified according to several formal classes, which are linked to different sets of vowel schemes. Numerically most important is the aforementioned class of three-consonant verbs without a plain vowel. As shown by Prasse (1972–4), it is possible to attach to this class a number of other verbal classes that contain plain vowels and have fewer root consonants. Table 2.11 illustrates two members of these classes in a number of Berber languages, one with three consonants and no plain vowel (LMD ‘learn’), the other with two consonants and an initial plain vowel (vTF ‘enter’, Mali Tuareg vLM ‘open’). Blank boxes refer to aspectual stems that are not attested in the language in question.

The negative MAN stems are exclusively used after a negative particle. It should be emphasized that the choice of the MAN stem is not the only marker of MAN distinctions.

Table 2.12 *Uses of MAN stems in Figuig Berber.*

MAN stem	Example	Main contexts in which the MAN stem is used
Aorist	*atəf!* *y-atəf*	imperative consecutive
ad + Aorist	*ad y-atəf*	non-realized (e.g. irrealis, future)
Perfective	*y-utəf*	past action state (including resultant state)
ul + Negative Perfective	*ul y-utif*	negated past action negated state
Imperfective	*ttatəf!* *i-ttatəf*	habitual/iterative imperative simultaneous action (e.g. progressive) habitual, iterative, durative
ad + Imperfective	*ad i-ttatəf*	non-realized habitual/iterative
ul + Imperfective	*ul ttatəf!*	negated imperative
ul + Negative Imperfective	*ul i-ttitəf*	negated simultaneous action negated habitual, iterative, durative negated non-realized

Source: Kossmann (1997: 347ff.).

There exist a number of pre-verbal particles that convey MAN distinctions in addition to the expression of MAN in the verbal basis.

The main marker of MAN distinctions is the choice of vocalic pattern. In addition, there is one MAN prefix to the verb, the marker *tt-* for the Imperfective, which is used with certain formally defined verbal classes (e.g. verbs with an initial plain vowel, such as Figuig *atəf* 'enter' in table 2.10). This prefix occurs before the derivational prefixes but after prefixal Subject markers. Interestingly, it does not combine with the causative derivation, which may be a hint to the original derivational nature of this prefix (Kossmann 2002).

In table 2.12, the main uses of the MAN stems are given for one language, Figuig Berber, which constitutes a typical Northern Berber system, and arguably constitutes a system similar to that of Proto-Berber. In addition to the MAN stem distinctions, two pre-verbal particles are taken into account, *ad*, which marks a non-realized event (e.g. irrealis and future), and the pre-verbal particle *ul*, which marks negation. The Imperatives have special inflections, while the other forms all share the same inflectional paradigm. In the Imperatives, the examples give the singular form of the verb *atəf* 'go in'. In the other examples third-person masculine singular forms of this verb are given, which have the 3SG:M prefix *y-* ∼ *i-*.[2]

More about the uses of the MAN stems is found below. Different dialects distinguish between different sets of verb stems, as shown in table 2.13.

Table 2.13 *MAN stem inventories in various dialects.*

	Proto-Berber	Kabyle	Figuig	Ghadames	Mali Tuareg
Aorist	+	+	+	+	+
Imperfective	+	+	+	+	+
Secondary Imperfective	−	−	−	−	+
Negative Imperfective	+	−	+	+	+
Perfective	+	+	+	+	+
Secondary Perfective	−	−	−	−	+
Negative Perfective	+	+	+	+	+
Future	?	−	−	+	−

The differences between the dialects are due to several historical developments:

- Loss of the negative stems. The negative imperfective was lost in many varieties, including Tashelhiyt, Central Moroccan Berber, Kabyle, and Siwa. The negative perfective is more stable. It was lost in a few varieties of Tashelhiyt and in Siwa.
- Aspect stem split. In Tuareg, the Perfective and the Imperfective have undergone a split. In addition to the original Perfective and Imperfective MAN stems, which certainly go back to Proto-Berber, Tuareg has introduced two (historically) secondary stems. These stems are both morphologically derived from the original stems by means of vowel lengthening and the insertion of a fixed accent. In an early stage of Tuareg, this morphological device was probably used for a differentiation between non-simultaneous events (for which the original form was used) and simultaneous events (for which the secondary forms were used) (see Leguil 2000). This division is still present in the Perfective in all Tuareg dialects. The opposition of simultaneity has been blurred in the Imperfective. In Malian Tuareg, traces of it persist, while all other Tuareg dialects have generalized the Secondary Imperfective in all contexts, and lost the original Imperfective formation altogether.
- The existence of a future stem. Ghadames Berber has a unique verbal stem, termed 'Future' (Lanfry 1968), which is used only with the particle *d*, the local variant of the pre-verbal particle *ad* 'NON-REALIZED'. The future stem has unique person, gender, and number marking. In most verbal classes, the vocalization of the Future is homophonous either with the Perfective (e.g. triradical verbs), or with the Aorist (e.g. quadriradical verbs). In some verbal classes it has a unique vocalization, which proves it constitutes a MAN stem on its own. Its history is not fully clear (see Kossmann 2000b).

Table 2.14 *Imperfective stems in Ayt Waryaghel Riffian (verb* řməð *'learn').*

Morphological device	Meaning	Example
Gemination	Simultaneous/progressive	*ř**ə**mmə**ð***
Gemination + prefix *t-*	Habitual	*tř**ə**mmə**ð***
Gemination + prefix *t-* + vowel insertion	Iterative	*tř**ə**mma**ð***

Source: Lafkioui and Kossmann (2009).

Table 2.15 *The Perfective and the Secondary Perfective in Siwa.*

Perfective	Secondary Perfective	
i-ɣəḍb-ən 3PL-be.angry:P-3PL	*i-ɣəḍb-**in**a* 3PL-be.angry:P-3PL:P2	'they are angry'
n-əča 1PL-eat:P	*n-əča**ya*** 1PL-eat:P:P2	'we ate'
uɣ-əm=tən buy:P-2PL=3PL:DO	*uɣ-əm=t**in**a* buy:P-2PL=3PL:DO:P2	'you (PL) bought them'

- In some Riffian dialects, an extensive split-up of the original Imperfective stem has taken place. The exact semantics of the oppositions are not entirely clear, but they seem to consist of differentiations inside the broader semantics expressed by the Imperfective stem in other Berber languages. In some verb types, three different stems have come out of this split, as shown in table 2.14.
- The easternmost Berber languages, Siwa and Awdjilah Berber, oppose a simple Perfective form to a Perfective extended by a partly suffixed, partly inserted element (glossed 'P2'). This morpheme is applied to the inflected verb *plus* its verbal clitics e.g., in Siwa (the marking of the Secondary Perfective is in bold type) (see table 2.15). On the functions of this form, see Leguil (1986).

As a result of the loss of phonological opposition in the central (short) vowel system of Northern Berber, some MAN stems have become homophonous in many verbs. In Riffian Berber, for example, the formal opposition between Aorist and Perfective is seen in only 6 per cent of the verbs (Cadi 1987: 55). This homophony follows formal lines and has no semantic correlates. Thus, in Riffian, if a verb consists of three consonants without a full vowel, Aorist and Perfective are homophonous, while a biradical verb without a full vowel in the Aorist has a different form in the Perfective. In spite of the infrequent incidence of the opposition, there is no reason to assume the system as such is breaking down: with those verbs where there has been no vowel coalescence, because

Table 2.16a *Vowel schemes in Tuareg.*

Type	Aorist	Perfective	Imperfective	
a	ə - ə[3]	ə - ă	(ə) - ă - :ă (also subtypes with *tt-*)	
	əlməd	*əlmăd*	*ələmmăd*	‘learn’
b	← ə	← ă	(with *tt-*)	
	kələstəf	*ăklăstăf*		‘speak nonsense’
c	← ă	← ə - ă	(with *tt-*)	
	ăqqătăs	*əqqətăs*[4]		‘snap (thread)’

Table 2.16b *Vowel schemes in Ghadames.*

Type	Aorist	Perfective	Imperfective	
a	ă - ə	ə - ă	(ə) - ă - :ă (also subtypes with *tt-*)	
	ărtək	*ərtăk*	*ərăttăk*	‘be mixed’
b	← ə	ă - ə - ă	(with *tt-*)	
	əbrənšəl	*ăbrənšăl*		‘become adult’

Table 2.16c *Vowel schemes in Zenaga.*

Type	Aorist	Perfective	
a	a - ə	ə - a	
	azgər	*əzgar*	‘go out’
b	← ə	a - ə - a	
	əddərɣəy	*addərɣay*	‘become blind’

the stem has a non-central vowel, the opposition is maintained and Aorist and Perfective are never confused.

2.4.4.2 *MAN stem formation*

In the structure of the system of vowel schemes, there exists a basic distinction between the vowel schemes of the Aorist, the Perfective, and (sometimes) the Imperfective. The vowel schemes of the other stems can be considered as derived from these basic vowel schemes. In Tuareg, the pattern of vowel schemes features three major types (see table 2.16a; the arrows indicate that any vowel preceding has the same quality).

In Ghadames, two main types of vowel scheme patterns are found (table 2.16b).

A similar system is found in Zenaga (see D. Cohen and Taine-Cheikh 2000), as seen in table 2.16c.

Table 2.16d *Vowel schemes in Tashelhiyt corresponding to Tuareg type (a).*

Aorist *ă - ə	Perfective *ə - ă	
ak^{w}r	*ukr*	'steal'
ddu	*dda*	'go'

Table 2.16e *Vowel schemes in Figuig, corresponding to Tuareg type (c).*

Aorist * ← ă	Perfective * ← ə - ă	
llaẓ	*lluẓ*	'be hungry'
ffad	*ffud*	'be thirsty'

Similar patterns are found when the verb contains plain vowels (*a* corresponding to *ă*, all other vowels corresponding to *ə*). This is often the result of compensatory lengthening because of consonant loss.

In Northern Berber the central vowel (contrast) has been lost, and only forms with plain vowels show traces of the old vowel scheme patterns. Pattern (a) is well attested, e.g. Tashelhiyt (table 2.16d).

The other patterns are less readily recognized in Northern Berber. However, forms such as the Figuig Berber verbs in table 2.16e could represent traces of Tuareg type (c).

Originally, these verbs may have had a four-radical structure *LXZX, in which the last consonant was lost without a trace (in Berber languages other than Figuig, such traces can be found in these verbs), while the loss of the second consonant led to compensatory lengthening, thus preserving the vocalic scheme.

2.4.5 Person-number-gender marking

Marking of person, number, and gender (PNG) of the subject is obligatory in all finite verbs. It takes place by means of prefixes, suffixes, or circumfixes, according to the PNG. An interesting feature of the verbal inflection is the difference it displays in its gender-marking system in comparison with the pronominal system. Although pronouns always distinguish between masculine and feminine in the second person singular, this distinction is not marked on the verb. On the other hand, the differentiation between masculine and feminine in the third person singular, which is absent in parts of the pronominal system, constitutes an integral part of verbal inflections.

In all Berber languages, the imperative has unique person, gender, and number marking. In many languages, one finds in addition a special PNG-marking system in the Perfective of a lexically determined set of verbs of permanent state (stative PNG). One

Table 2.17a *The four PNG-marking sets in Ghadames.*

	Imperative PNG	Normal PNG[5]	Stative PNG	Future PNG
1SG		___-ăʕ (< *-ăɣ)	___-ăʕ (< *-ăɣ)	___
2SG	___	t-___-ət (< *-əd)	___-ət (< *-əd)	t-___
3SG:M		y-___	___	y-___
3SG:F		t-___	___-ăt	t-___
1PL		n-___	___-it	n-___
2PL:M	___-ăt	t-___-ăm	___-it	t-___-ăm
2PL:F	___-măt	t-___-măt	___-it	t-___-măt
3PL:M		___-ăn	___-it	___-ăn
3PL:F		___-năt	___-it	___-năt

Source: Lanfry (1968: 224).

Table 2.17b *Examples of PNG-marking in Ghadames.*

	Imperative PNG	Normal PNG (Perfective)	Stative PNG	Future PNG
1SG		*utăf-ăʕ*	*măttit-ăʕ*	*utăf*
2SG	*atəf*	*t-utăf-ət*	*măttit-ət*	*t-utăf*
3SG:M		*y-utăf*	*măttit*	*y-utăf*
3SG:F		*t-utăf*	*măttit-ăt*	*t-utăf*
1PL		*n-utăf*	*măttit-it*	*n-utăf*
2PL:M	*atəf-ăt*	*t-utăf-ăm*	*măttit-it*	*t-utăf-ăm*
2PL:F	*atəf-măt*	*t-utăf-măt*	*măttit-it*	*t-utăf-măt*
3PL:M		*utăf-ăn*	*măttit-it*	*utăf-ăn*
3PL:F		*utăf-năt*	*măttit-it*	*utăf-năt*

Source: Lanfry (1968: 224).

language, Ghadames Berber, has a fourth set of PNG affixes, used exclusively with the future stem. As an example, the maximal system of four different PNG-marking sets, as attested in Ghadames, is reproduced in tables 2.17a and 2.17b. The verb stem is marked by ___.

Compare the examples (vTF ‘enter’, MTT ‘be small’) in table 2.17b.

Here are some remarks about morphological variation within Berber:

- The first-person suffix in Ghadames, *-ăʕ*, is historically derived from *-*ăɣ* (or *-*əɣ*), which is the form found in most other dialects.
- The second-person suffix in Ghadames, *-ət*, goes back to *-*əd* as a regularized form of an assimilation with following consonant-initial clitics. The most commonly attested variant is *-əd*, but a number of languages have a pharyngealized dental instead: Kabyle *-əḍ*, Siwa *-əṭ*, etc. In several Southern Tuareg dialects the second-person singular suffix has been

Table 2.17c *Dialect variations in the stative PNG.*

	Type 1 (Ghadames, Kabyle, Zenaga)	Type 2 (Gourara, Nefusa, etc.)	Type 3 (medieval Tashelhiyt)
3SG:F	-ət	-yət	-ət
PL	-it	-ət	-ət

Source: Kossmann (2009).

Table 2.17d *PNG forms in Siwa.*

	Imperative PNG	Normal PNG	Example of normal PNG (verb 'open')
1SG		___-ax	*fətk-ax*
2SG	___	___-aṭ	*fətk-aṭ*
3SG:M		y-___	*y-əftək*
3SG:F		t-___	*t-əftək*
1PL		n-___	*n-əftək*
2PL	___-wət	___-əm	*fətk-əm*
3PL		y-___-ən	*yə-ftk-ən*

Source: Vycichl (2005).

replaced by a suffix *-ăɣ* or *-ă*, introduced from the first-person singular. As a result, in these dialects, the main difference between the two persons lies in the presence or absence of the prefix *t-*, e.g. Ayer Tuareg: *otăf-ă* 'I entered', *t-otăf-ă* 'you entered'.

- In Zenaga the feminine forms of the plural (normal PNG) have the suffixes 2PL:F *-amɲad* (< **-ămɲăt*) and 3PL:F *-iɲ* ∼ *-iɲɲad* (< **-ənɲăt*).
- There is some dialectal variation in the third-person singular feminine and plural forms of the stative (see table 2.17c).

In several languages which still have the stative PNG, the affixes of the normal PNG have been extended to all non-third-person plural forms. This is the case in Tuareg, among others.

If one allows for the already mentioned variations (and some others) and for the fact that the stative and the future PNG sets are not found in all Berber languages, one may say that the great majority of Berber languages have PNG forms similar to those of Ghadames. There is one major exception to this: the Berber language of Siwa, an oasis in western Egypt (see table 2.17d).

Berber languages show interesting variants in the PNG affixes of the first-person non-singular. Most important is the existence in some languages, e.g. Tashelhiyt, of special

Table 2.17e *Dual and plural PNG in cohortative contexts in Tashelhiyt.*

		example with 'do'
1 DUAL (M/F)	___-aɣ	*skr-aɣ*
1 PL:M	___-ataɣ	*skr-ataɣ*
1 PL:F	___-amtaɣ	*skr-amtaɣ*

Source: Aspinion (1953: 116).

Table 2.17f *First-person non-singular marking in Ghadames.*

		example with 'sow' (Aorist)
1 DUAL / 1 PL:EXCL (M/F)	n-___	*n-ăkrəz*
1 PL:INCL:M	n-___-ăt	*n-ăkrəz-ăt*
1 PL:INCL:F	n-___-măt	*n-ăkrəz-măt*

Source: Lanfry (1968: 327–9).

dual forms in cohortative contexts, which show a category not expressed otherwise in the language (see table 2.17e).

According to Lanfry (1968: 327–9) Ghadames Berber distinguishes between dual/plural and inclusive/exclusive in the first-person plural in all verbal forms, except the imperative (which has no first plural form), and the stative (see table 2.17f).

A distinction between inclusive and exclusive has not been reported for any other Berber language.

2.4.6 The participle

The so-called 'participle' is a verb form, exclusively used in relative clauses, whose subject is the head of clause, e.g.:

(1) *t-wašun-t* *y-iwy-ən* *a-rgaz*
EL:F-girl-SG:F PTC-bring:P-PTC EL:M-man
'the girl that married the man'
(Figuig, Kossmann 1997: 160)

The participle never occurs in other contexts. It is found in almost all Berber languages, the only reported exception being Siwa. Most Berber languages are able to make participial forms of all MAN stems. There exists considerable dialectal variation as to the formal distinctions found in participles (see Drouin 1996, Kossmann 2003a), as shown in table 2.18.

The most elaborate system is illustrated in table 2.19a with forms from Mali Tuareg (Adagh).

Table 2.18 *Formal distinctions in participles in selected languages.*

	Adagh Tuareg	Tashelhiyt	Figuig	Riffian
SG:M <> SG:F	+	−	−	−
SG <> PL	+	+	−	−
Stative <> non-stative	+	+	−	−
Affirmative <> negative	+	−	+	−

Table 2.19a *Non-stative and stative participle system in Adagh Tuareg.*

	Non-stative	Stative		
	Affirmative	Negative (with the negative particle *wər*)	Affirmative	Negative (with the negative particle *wər*)
SG:M	y-___-ăn	(wər) ăn y-___	___-ăn	(wər) ăn ___
SG:F	t-___-ăt	(wər) ăt t-___	___-ăt	(wər) ăt ___
PL	___-nen	(wər) ăn ___	___-nen	(wər) ăn ___

Source: Prasse and ăgg Albosṭan (1985).

Table 2.19b *Participial forms of a non-stative verb in Adagh Tuareg ('learn', Perfective examples).*

	Affirmative	Negative
SG:M	*i-lmăd-ăn* PTC:SG:M-learn:P-PTC:SG:M	*wər ăn i-lmed* NEG PTC:SG:M PTC:SG:M-learn:PN
SG:F	*t-əlmăd-ăt* PTC:SG:F-learn:P-PTC:SG:F	*wər ăt t-əlmed* NEG PTC:SG:F PTC:SG:F-learn:PN
PL	*əlmăd-nen* learn:P-PTC:PL	*wər ăn əlmed* NEG PTC:PL learn:PN

Source: Prasse and ăgg Albosṭan (1985).

Compare the forms in the non-stative verb LMD 'learn' and in the stative verb SDD 'be thin' in Adagh Tuareg (tables 2.19b and 2.19c).

On the other end of the spectrum stands Riffian Berber, which has only one form, used in all contexts: *y-___-ən*.

In languages such as Adagh Tuareg, the participial inflection seems to consist of two parts: the first part, which is always prefixed to the verb, is identical to the prefixal part of third-person verbs. The second part is a suffix in most affirmative contexts. In negative contexts, and in combination with the particle *ad* 'NON-REALIZED', the second part of the participle is fronted to the position before the first part. Note that in several

Table 2.19c *Participial forms of a stative verb in Adagh Tuareg ('be thin').*

	Affirmative	Negative
SG:M	*sădid-ăn* be.thin:P-PTC:SG:M	*wər ăn sădid* NEG PTC:SG:M be.thin:PN
SG:F	*sădid-ăt* be.thin:P-PTC:SG:F	*wər ăt sădid* NEG PTC:SG:F be.thin:PN
PL	*sădid-nen* be.thin:P-PTC:PL	*wər ăn sădid* NEG PTC:PL be.thin:PN

Source: Prasse and ăgg Albosṭan (1985).

languages the first part of the participle is absent when the second part is fronted, as shown in the following participial forms from Figuig (verb *atəf* 'come in'): *y-utf-ən* (affirmative participial form); *un n-utif* (negative participial form with the assimilated negative particle *ul* > *un*). Compare the following examples:

(2a) *argaz dd y-utf-ən*
man HITHER PTC-enter:P-PTC
'the man that came in' (Figuig)

(2b) *argaz dd un n-utif*
man HITHER NEG PTC-enter:PN
'the man that did not come in' (Figuig)

The pre-verbal position of the second part of the participial marker in negative contexts is superficially similar to cases of clitic fronting. However, it cannot be analysed as such, as clitic fronting is obligatory in relative clauses; if the second marker of the participle were sensitive to this specific fronting process, one would expect it to be always preposed, as the participle is a form marking subject relatives, and therefore always occurs in relative clauses.

Most Berber languages do not allow for any elements to be put between the pre-verbal second part of the participle and the verb. Thus it is reasonable to assume that the pre-verbal element constitutes a prefix to the verb. In some Berber languages, however, clitics can come between the pre-verbal second part of the participial marker and the verb form. In such variants, this element cannot be analysed as a prefix, e.g. the following Ghadames phrase:

(3) *w-e wăl ăn ak=t=id=y-ăbbe*
DEF:M-PROX:SG NEG PTC:SG:M 2SG:M:IO=3SG:M:DO=PTC:SG:M-bring:PN
'the one that did not bring to you'
(Ghadames, Lanfry 1968: 336)