SOMALI

LONDON ORIENTAL AND
AFRICAN LANGUAGE LIBRARY

Editors
Theodora Bynon
David C. Bennett
School of Oriental and African Studies
London

Masayoshi Shibatani
Kobe University

Advisory Board
James Bynon, Bernard Comrie, Judith Jacob, Gilbert Lazard,
Christian Lehmann, James A. Matisoff, Vladimir P. Nedjalkov,
Robert H. Robins, Christopher Shackle

The LONDON ORIENTAL AND AFRICAN LANGUAGE LIBRARY aims to make available a series of reliable and up-to-date descriptions of the grammatical structure of a wide range of Oriental and African languages, in a form readily accessible to the non-specialist. With this in mind, the language material in each volume will be in roman script, fully glossed and translated. The Library is based at the School of Oriental and African Studies of the University of London, Europe's largest institution specializing in the study of languages and cultures of Africa and Asia. Each volume is written by an acknowledged expert in the field who has carried out original research on the language and has first-hand knowledge of the area in which it is spoken.

Volume 10

John Saeed

Somali

SOMALI

JOHN SAEED
Trinity College, Dublin

Contents

Preface

In writing this book my aim has been to present an up-to-date and accurate picture of the grammar of Somali which will be of interest to non-specialists and linguists interested in language comparison and typology. Thus this volume is not intended as a manual for those wanting to teach themselves Somali nor a reference grammar for such a reader. Such books already exist; see Zorc and Issa (1990), Orwin (1995) and Saeed (1993a), for example. My aim is to highlight those areas that will give the reader a clear view of the grammatical characteristics of Somali and through Somali of what a Cushitic language is like. I have avoided pursuing the concerns of any particular grammatical theory but nonetheless the book is firmly based on the assumptions and methodology of modern descriptive linguistics, and I have felt free to refer to the theoretical and typological literature when the discussion warrants it. The main focus of the book is on the morpho-phonology and syntax of the language, which fits naturally both the size of the book and the state of our knowledge about Somali. Thus the phonetic description serves as background to the discussion of phonological and morphological processes. Similarly, chapter 9 on discourse concentrates on those processes that are most intimately related to grammatical rules.

This work is based on study of the language over some twenty years, in Somalia and Kenya, and with Somali speakers in Europe. I am indebted to a number of academic institutions which have supported my work, particularly my own college, Trinity College Dublin, and my department, the Centre for Language and Communication Studies. I am also grateful for the support and academic hospitality shown to me by the School of Oriental and African Studies of the University of London, the Royal Irish Academy, the University of Nairobi, Lafoole College of the National University of Somalia, and the Somali National Academy of Culture. Many of the examples used in this book are quoted from publications of Somali scholars and teachers associated with the now sadly disbanded Somali National Academy of Culture, the Curriculum Department of the Ministry of Education and Training, and the Somali

List of Abbreviations and Symbols

ABS	absolutive case
ADJ	adjective
ADP	adposition
ADV	adverb
AGR	verbal agreement affixes
ALL	allative adverbial
AP	accentual pattern
AUX	auxiliary verb
C1,C2 etc.	first conjugation, second conjugation, etc.
CAUS	causative (affix or derived verb)
CM	conditional marker
CN	common noun
CNJ	conjunction
COLL	collective noun
COND	conditional verb form
CPRO	clitic verbal pronoun
D1, D2 etc.	first declension, second declension, etc.
DET	determiner
DM	declarative marker
EXC	exclusive
EXP	experiencer (affix or derived noun)
FACT	factitive (affix or derived verb)
FG	falling tone
FEM,F	feminine gender
FOC	focus word
FUT	future
GEN	genitive case
H	high tone
HAB	habitual
IM	imperative marker
IMP	imperative verb form

INC	inclusive
INCH	inchoative (affix or derived verb)
INF	infinitive
INFL	verbal affixes marking tense, aspect and mood
INTR	intransitive
IPRO	independent pronoun
L	low tone
LEX	verbal derivational affixes
MASC, M	masculine
MID	middle (affix or derived verb)
N	noun
NIM	negative imperative marker
NEG	negative word or verb form
NAM	name (proper noun)
NOM	nominative case
NUM	numeral
O	object
O PRO	object clitic pronoun
OM	optative marker
OPT	optative verb form
PAS	passive
PAST	past tense. Used alone: past simple
PHR	phrase
PL	plural
PM	potential marker
POT	potential verb form
PRES	present tense. Used alone: present general
PROG	progressive
QM	interrogative marker
RE	reflexive/reciprocal pronoun
REM	remote definite article
ROOT	base form of a verb
S	sentence
SG	singular
S PRO	subject clitic pronoun
STAT	stativising affix
SUBORD	subordinate clause verb form
SUPR	superlative marker
TR	transitive
V	verb
VEN	venitive adverbial
VGP	verbal group
VOC	vocative case or suffix

1	first person
2	second person
3	third person
*	ungrammatical; or reconstructed form
?	semantically anomalous
#	contextually inappropriate

Glosses: Because of the morphological complexity of Somali and the frequency of coalescence rules, it is often difficult to segment the surface forms into their constituent morphemes. For clarity, in chapters 2-4 each example is accompanied by a morpheme by morpheme breakdown which makes the underlying morphological structure transparent. A gloss and translation accompany this. A simple example is:

> *Wày iigá yeeshay*
> *waa+ay   i+ú+ká   yeel+t-ay*
> DM+she   me+for+from   accept+3SG F-PAST
> 'She accepted it from me for him.'

The following symbols are used: the hyphen - marks surface morpheme boundaries, e.g. *nín-ka* 'man-the'; the colon : marks unsegmentable grammatical features, e.g. *súg* wait:IMP:SG, or when one word in Somali corresponds to several in the gloss, e.g. *xanàaq* 'grow:furious'; the plus sign + marks the underlying morphological structure before rules of cliticisation and coalescence apply, as in the above example.
The glosses are designed to aid the reader and thus clarity has been a greater priority than exhaustiveness or consistency. Thus translations are often used in place of list of features, so that *ahay* is glossed as 'am' rather 'be:PRES:1SG'. Similarly, when the context makes it obvious, information is omitted, so when its subject is overtly a masculine singular noun, a verb like *tegay* will simply be glossed as 'went'. From chapter 5 onwards, the morpheme by morpheme breakdown is omitted for reasons of space. Since the main coalescence and sandhi rules have been described by then, it is hoped that readers will be able to 'read' the morphological structure more easily for themselves.

Data sources: The sources of examples taken from publications are cited in the text in square brackets containing the initials of the publication (listed in Appendix I), followed by page and line numbers, for example [SHS 25.17-19], meaning lines 17-19 on page 25 of *Sheekoxariirooyin Soomaaliyeed, Folktales from Somalia.*

Chapter 1
Introduction

This chapter provides a brief sketch of where Somali is spoken and by whom. It contains a discussion of the cultural context in which the language is used, its relationship to other East Cushitic languages, and the dialect variation within the language. The chapter also highlights some important previous linguistic studies and previews the structure of this book.

1.1   The speakers and the land

Somali is spoken by about nine million people who occupy the north-eastern corner of Africa, a region known as the Horn of Africa. This huge area of mainly arid plateau is bounded to the north and east by sea, the Gulf of Aden and the Indian Ocean, while in the east the desert plateau of the Ogaden runs towards the Ethiopian highlands. Much of the land is arid with light seasonal rainfall and the dominant lifestyle is pastoral nomadism, with the Somalis traditionally moving great distances with their herds of camel, sheep and goats on a seasonal cycle. There is also an ancient tradition of farming in the more fertile areas in the south near the rivers Shabelle and Juba. The coastal towns of Berbera, Mogadishu, Merka, Brava and Kismayo have been trade centres in the Gulf and Indian Ocean for centuries.

Somali society is traditionally organised along patrilineal clan lines with every Somali being taught from an early age of the importance of their clan lineage. Traditional society was characterised by complicated patterns of inter-clan livestock raiding and warfare, as well as intermarriage and political alliance. There is evidence of major migrations both by Somali clans and neighbouring peoples in recent history but the establishment of European protectorates and colonies from the late nineteenth century onwards halted this movement of peoples. The Somalis are in contact with speakers of other Cushitic languages

around almost their entire periphery: 'Afar (Qafar) speakers in the north-west, various dialects of the Oromo language in the west and south-west, as well as Rendille and Boni to the south. There are small groups of non-Somali speakers within the Somali lands. These include Oromo-speakers in the west; speakers of the Bantu language Mushungulu on the river Juba near Jamaame; and speakers of the Swahili dialects, chi-Mwiini, spoken on the lower Shabelle river near Brava, and ki-Bajuni, spoken on the southern coast around Kismayo.

The trading relations between the Horn and the Arabs of Yemen and Oman date from ancient times and the Somalis were converted to Islam peacefully very early in the religion's expansion. The influence of Islam, and of Arabic, on Somali culture has been very great and the population is almost totally Moslem with the majority being Sunni. Many of the Somali clans trace their lineage back to the Arab religious leaders who brought Islam to the region and the majority of personal names, especially of men, are Arabic. One modern reflection of this close relationship was the admittance of the Somali Republic to the association of Arab states, the Arab League. Despite this close religious and commercial attachment to the Arab world, the Somalis maintained their own language, their traditional laws alongside Islamic law, and in many areas, such as the social role of women, preserved their distinct identity. One of the most important expressions of their culture is the rich indigenous tradition of oral poetry consisting of many genres characterised by complex metrical structures.[1] Many foreign writers have commented on the Somalis' fierce independence and pride in their language and culture.[2] Thus, although there are many Arabic loans in Somali[3], especially in the fields of religion, commerce, and politics, many Somalis do not speak Arabic, apart from the verses of the Koran. A person's knowledge of Arabic depends on a number of factors including type and level of education, and proximity to the northern coasts.

The partition of the Horn between Abyssinia and the European colonial powers left the Somalis divided among five states: Abyssinia (later Ethiopia), Kenya, French Somaliland (later Djibouti), Italian Somalia, and the British Somaliland Protectorate. The last two, on their independence in 1960, joined to form the Somali Republic. The remaining partition has been a constant source of instability in the region: the activities of Somali secessionist groups have led to uprisings, and armed clashes and wars between the new republic and neighbouring Kenya, Ethiopia and Djibouti. This local instability was increasingly drawn into the more global conflict of the cold war and the 1978-79 war between Somalia and Ethiopia, involving Soviet Bloc forces, was one of the fiercest wars in post-colonial Africa. The defeat of Somali and secessionist forces led to internal strains within Somalia, where an increasingly despotic regime attempted to shore itself up by the exploitation of inter-clan rivalry. The

increasing violence led eventually to the civil wars of the past ten years which have fragmented the Somali Republic and which at the time of writing have left the country with no central administration.

        Somalis have traditionally formed expatriate communities in east Africa and in Arabia, especially in southern Yemen. More recently small communities were established in Great Britain, Italy and France. In the past decade of troubled times, these communities have greatly increased in size and been joined by new ones in Canada, the United States, and other European countries.

        For a concise history of the Somali people and their role in the modern states of the Horn the reader is referred to Lewis (1980); more recent history is discussed in Laitin and Samatar (1987), Ahmed I. Samatar (1988), Abdi I. Samatar (1989), Drysdale (1994), and Simons (1995). For anthropological studies of the traditional ways of life see Cerulli (1957-64) and Lewis (1961).

1.2   The language

Somali is classified as a member of the Cushitic language family, one of the major subdivisions of the great Afroasiatic or Hamito-Semitic language phylum, along with Omotic, Chadic, Berber, Semitic and Ancient Egyptian. The Cushitic family consists of around thirty languages with Somali having the second largest number of speakers, after Oromo. Current opinion on the position of Somali within East Cushitic is shown in Figure 1.
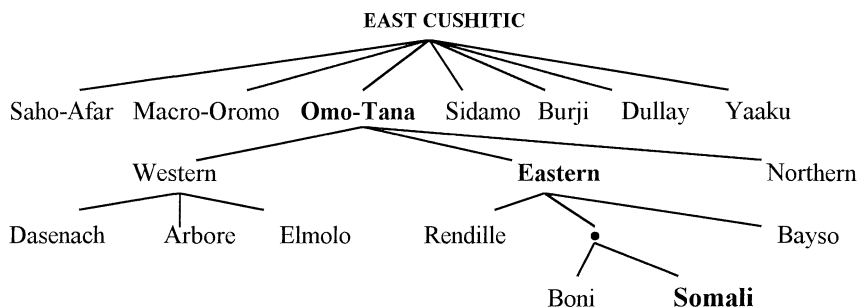
Figure 1.   The classification of Somali in East Cushitic

        After years of heated debate about choosing between Roman, Arabic and indigenous orthographies, a Roman orthography was adopted for Somali in the Somali Republic in 1972 and Somali was adopted as the official language of the

state. Great strides were made in the development of literacy and the introduction of Somali as the medium of government, education and mass media. Arabic was adopted as the second language of the state. At this level of official state languages, Somali thus joined its neighbours Amharic in Ethiopia and Swahili in Kenya, making this region one of the most important in the development of African languages as official media. Though the development of Somali as an official language has been halted by the recent civil wars, it continues to be the written medium in all parts of the Somali-speaking Horn and among and between expatriate communities abroad. Andrzejewski (1974, 1978) and Laitin (1977) give descriptions of the Somali experience of language planning.

Somali shows considerable dialect variation. There are three main dialect groups, Northern, Benadir and May. The Northern dialect group is the most geographically widespread of these; indeed this is a somewhat misleading term since, because of clan migrations, this dialect group is spoken in a broad sweep across the northern, western and southernmost parts of the Somali-speaking area. This dialect seems to have been used as a lingua franca amongst other dialects. It enjoys the most prestige, not least because it has been the dialect used by the greatest Somali poets and by the various political élites since the formation of the Somali state. Not surprising then this dialect formed the basis of the standard used for official purposes in the Somali Republic. Because of this some writers have called this dialect *Common Somali*. It is this dialect that forms the subject of this book, though we will use the term *Standard Somali*.[4] A second large group is formed by the Benadir (Benaadir) dialects spoken along and behind the central Indian Ocean coast, including Mogadishu. There is a fair degree of inter-comprehensibility between Benadir dialects and Standard Somali. The third major dialect group consists of the May (Maay) dialects spoken in a broad band of southern Somalia stretching from almost the Ethiopian border to just short of the coastal strip between Mogadishu and Kismayo. The largest urban centre in this dialect area is Baydhaba (Baidoa). The May dialect group shows considerable differences from Standard and Benadir dialects at all levels and these latter two groups and May are mutually incomprehensible. Because of the role and status given to Standard Somali many May speakers are able to use it as a lingua franca. There has been little use of May dialects in education, official media, or publishing. In addition to these main groups there are a number of smaller, more localised dialects in the southern inter-riverine region and in the southern coastal towns. These include the Jiddu (Jiiddu) dialect (Moreno 1951); the Tunni dialect spoken in the coastal region around Brava (Tosco 1997); the Karre dialect of Qoryooley (Tosco 1989); and the Ashraf dialects of Mogadishu and Merka (Moreno 1953).[5]

The overall picture of the dialects contrasts the relative homogeneity of Standard Somali over its vast range with the much greater dialectal variation in the more sedentary areas of the southern riverine regions and around the southern ports of Mogadishu, Merka, Brava and Kismayo.[6]

1.3   Linguistic studies

Although Somali is possibly the best documented and most studied Cushitic language, in many areas the linguistic description is still very sketchy. Perhaps the most developed area is lexicology. Here there exists a major Somali-Italian dictionary of more than forty thousand entries (Agostini *et al.* eds. 1985), with a related thirty-five thousand-entry Italian-Somali dictionary (Puglielli ed. 1998). There is also a twenty-six thousand-entry Somali-English dictionary (Zorc and Osman eds. 1993), a substantial Somali-French-Somali dictionary (Maxamed 1986) and a monolingual Somali dictionary (Keenadiid 1976). In the area of grammar, there are relatively recent concise reference grammars in English (Saeed 1993a), and in German (Lamberti 1988), as well as older studies in Italian (Moreno 1955, Panza 1974), and Russian (Zholkovskij 1971). A series of grammatical studies, *Studi Somali*, has been published by the Italian Government's Dipartimento per la cooperazione allo svilippo (see for example Cardona and Agostini ed. 1981, Puglielli ed. 1981, Puglielli ed. 1984). Teaching handbooks have been published in English (Zorc and Issa 1990, Orwin 1995), German (El-Solami-Mewis 1987) and French (Abdullahi 1996). Perhaps the weakest areas are in phonetics and phonology where there is still no book-length study published [7], and perhaps less surprisingly, in pragmatics and discourse, though Biber (1995) includes Somali in a study of register variation.

If we move from books to academic articles and papers there has been a steady increase in the number of linguistic studies published. There is notably a three-yearly congress held by the Somali Studies International Association whose *Proceedings* (e.g. Labahn ed. 1984) contain many linguistic papers. Finally any survey, however brief, must acknowledge the groundbreaking series of Somali grammatical studies published by the late B.W. Andrzejewski; for a bibliography see Hayward and Lewis (eds. 1996).

1.4   Design of this book

The book has nine chapters with an appendix of selected texts. Chapters 2 and 3 deal with the sounds of Somali: Chapter 2 outlines the phonological units of the

language, and Chapter 3 describes some of the most important phonological rules. Chapter 4 sets out the basic word classes, or grammatical categories, of the language, while Chapter 5 describes the derivational processes which operate to extend the categories of verb, noun, and adjective. This chapter also discusses compounding. The next three chapters deal with syntax. Chapter 6 is principally concerned with the internal structure of noun phrases and the verbal group. Chapter 7 describes the structure of simple sentences and establishes the major sentence types. Chapter 8 goes on to deal with subordinate clauses, highlighting the central role of relative clauses. Finally, Chapter 9 is concerned with those discourse processes that most fundamentally affect grammatical structure, in particular the roles of focus and topic. Appendix II contains a brief selection of extracts from various types of Somali texts, including both prose and poetry, with morpheme by morpheme glosses and translations.

The Somali examples in this book are taken from a combination of sources. One of my main preoccupations in writing this study has been to use language material that is as authentic as possible. Many of the examples were recorded from native speakers over several years of fieldwork in Somalia and Kenya, and from Somali speakers visiting and resident in the UK and Ireland. Other examples are taken from a variety of published materials: textbooks and other books published by the government of the Somali Republic, newspapers, collections of poems and tales, etc. The principal published sources are listed in Appendix I. As described in chapter 2, tonal accent plays an important role in marking grammatical distinctions in Somali. Since it only makes a restricted number of lexical distinctions, tonal accent is not marked in the official orthography. In examples that were recorded from speech by the author, tonal accent is always marked; in extracts from published sources it is not marked.

Chapter 2
Phonological units

In this chapter we present the inventory of the phonological units in standard Somali, including consonant and vowel sounds, syllables and tonal units. We also introduce the official Somali orthography which will be used to represent examples in the rest of the book. The symbols of the International Phonetic Alphabet (IPA) are used in square brackets [ ] to introduce the sounds and when we need to discuss aspects of their phonetic realisation. We begin with the consonants.

2.1   Consonant inventory

The basic set of twenty-two distinctive consonant sounds is given in IPA symbols in (1) below, followed by a brief description of each sound together with its form in the Somali orthography.

(1)   Consonants

Stops:	b	d	ɖ		g	ɢ	ʔ
		t			k		
Fricatives:	f	s		ʃ		χ ħ h	
							ʕ
Affricate				tʃ			
Nasals	m	n					
Trill		r					
Lateral			l				
Glides	w			j			

2.1.1 Stops

The sound [b] is a voiced bilabial plosive, orthographically *b*. Word initially or finally and before voiceless sounds, there may be little voicing, e.g. [b̥] in *húb*[8] N 'weapon', and *húbka* 'the weapon'. Between vowels *b* often weakens to the corresponding fricative, e.g. [β] in *laba* 'two'.

[d̪] is a voiced dental plosive, orthographically *d*, with similar phonetic qualities to *b*. It is often devoiced in the same positions, e.g. [d̪̥] in *gèed* N 'tree' and *gèedka* 'the tree', and is often weakened to a fricative intervocalically, e.g. [ð] in *hódon* N 'wealthy person'.

[ɖ] is a voiced retroflex plosive, orthographically *dh*. Between vowels it is sometimes pronounced as a flap e.g. [ɽ] in *ádhi* N 'sheep and goats'. The distribution of this sound distinguishes between northern and southern varieties of standard Somali. In southern varieties the distinction between *dh* and the rolled liquid *r* is lost except at the beginning of words or where the northern varieties have a double consonant *dhdh*. This gives the following correspondences:

Northern Standard	Southern Standard		
dháqan	*dháqan*	N	'culture, tradition'
xídh	*xír*	V	'close'
fadhiisó	*fariisó*	V	'sit down'
fádhdhi	*fádhi*[9]	V	'remain sitting'

Some writers, for example von Tiling (1925), Armstrong (1934), and Moreno (1955), have commented on an implosive quality to the sound *dh* in some speakers. This does not seem to be the case for all speakers and the issue remains to be investigated.

[g] is a voiced velar plosive, orthographically *g*, which like *b* and *d* often has little voice initially, finally and before voiceless stops, for example in *góor* N 'time', *lúg* N 'leg', and *lúgta* 'the leg'. This sound seems less likely than *b* and *d* to weaken to a fricative between vowels, e.g. remaining [g] in *xigàal* N 'relative'.

[ɢ] is a voiced uvular plosive, orthographically *q*. This sound occurs with little voice initially or finally and indeed some speakers pronounce it with no voicing at all, i.e. as [q] in these positions. There is of course no contrastive homorganic voiceless stop. Between vowels it is voiced and may weaken to a corresponding fricative, e.g. [ɣ] in *shaqèe* V 'work!'.

[ʔ] is a glottal stop, orthographically represented as '.   It occurs significantly between vowels and finally, e.g. *la'àan* N 'being without', *ló'* N 'cattle'. The sound also occurs initially in words which lexically we would expect to begin with a vowel, e.g. *idan* N 'incense burner', which in isolation is pronounced [ʔídæn]. Here it has no contrastive role and is not marked in the orthography.   We assume that in this position the glottal stop is inserted phonetically as part of a general requirement that word initial syllables in Somali have a consonantal onset, as discussed in 2.3 below.

[t̪] is a voiceless dental plosive, orthographically *t*. This sound has a more limited distribution than the corresponding voiced sound *d*: it only occurs at the beginning of a syllable. As described in section 2.3 below this means that *t* is never doubled to form a geminate. This sound is always pronounced fortis and with aspiration and does not weaken intervocalically, e.g. [t̪ʰ] in *tùug* N 'thief', *hátaq* N 'ditch'.

[k] is a voiceless velar plosive, orthographically *k*. As with the *t/d* pair, it has a more restricted distribution than the corresponding voiced plosive *g*, occurring only at the beginning of syllables.[10] It is always pronounced fortis and with aspiration, and does not weaken intervocalically, e.g. [kʰ] in *kàalay* V 'come!', *túke* N 'crow'.

2.1.2 Fricatives

The sound [f] is a voiceless labio-dental fricative, orthographically *f*, which is always voiceless even intervocalically, e.g. [f] in *fál* N 'verb', *sáfar* N 'journey', *láf* N 'bone'.

[s] is a voiceless alveolar grooved fricative, orthographically *s*. It is always voiceless, even intervocalically, e.g. [s] in *sán* N 'nose', *qósol* N 'laughter', *mìis* N 'table'.

[ʃ] is a voiceless palato-alveolar fricative, orthographically *sh*. It is always voiceless, even intervocalically, e.g. [ʃ] in *shàah* N 'tea', *náshir* N 'broadcast', *kìish* N 'bag'.

[χ] is a voiceless uvular fricative, orthographically *kh*.   It also often pronounced as a velar fricative [x]. This sound is only found in loan words from Arabic. Such loan words may be further 'somalized' by replacing this sound with *q*. Thus one may hear either *khajàar* or *qajàar* 'cucumber'. The choice of pronunciation may say something of the speaker's knowledge of Arabic. The sound *kh* is always voiceless, even intervocalically: e.g. [χ] in *khál* N 'vinegar', *makhaayád* N 'coffee shop', *shìikh* N 'sheikh'.

[ḥ] is a voiceless pharyngeal fricative, orthographically *x*.   It is always voiceless, even intervocalically: e.g. [ḥ] in *xabbád* N 'bullet', *sáxar* N 'dust, speck', *sáx* N 'correctness'.

[ʕ] is a voiced pharyngeal fricative, orthographically *c*.   It is always voiced. e.g. [ʕ] in *cèel* N 'well', *macállin* N 'teacher', *sác* N 'cow'. The sound may occur with creaky voice.

[h] is a voiceless glottal fricative, orthographically *h*.   It is usually voiceless initially and finally but voiced intervocalically, e.g. [h] in *húb* N 'weapon' and *léh* ADJ 'having, owning', and [ɦ] in *dúhur* N 'noon'.

2.1.3 Affricate

The only affricate is the palato-alveolar affricate [tʃ]. Orthographically it is *j*. The sound may be pronounced either voiceless or voiced, i.e. as [tʃ] or [dʒ]. The two variants seem to be in free variation though it is possible that a tendency to use one rather than the other may reflect some regional variation within the Standard dialect. The sound does not occur syllable finally except in Arabic loan words like *xáj* N 'pilgrimage'.   Thus in native Somali words this sound patterns like the voiceless stops *t* and *k*.

2.1.4 Nasals and other consonants

The sound [m] is a bilabial voiced nasal, orthographically *m*. The sound does not occur in word-final position, where it appears instead as *n*, thus collapsing the opposition between the two nasals. We can see an example of this with the root of the word *nín* N 'man' which underlyingly is **ním*. Since the plural is formed by a reduplication process which adds a syllable, the root *m* reappears in *nimán* N 'men'. Compare this with *sán* N 'nose'; *sanán* N 'noses'. As pointed out in Orwin (1994:232-4) this positional contraint may be still stronger, i.e. that this nasal does not occur in syllable final position. Instances where it does occur syllable finally, e.g. *sambáb* N M 'lung', are always before a labial consonant and may be viewed as the result of the assimilation of *n* to *m*.

[n] is a voiced alveolar nasal, written as *n* in the orthography. This sound occurs in all positions in the word. As noted by Armstrong (1934) and Orwin (1994), its pronunciation at the end of a syllable is affected by the previous vowel: after a long vowel the nasal is shorter and may be partially devoiced; after a short, especially high tone, vowel the *n* is longer, and more voiced.

[r] is a voiced alveolar trill, orthographically *r*. It is often pronounced with breathy voice and may be partially devoiced. Between vowels the sound may occur as a single tap. It occurs in all positions in the word.

[l] is a voiced alveolar lateral approximant, orthographically represented as *l*, which occurs in all positions in the word, e.g. *láb* N 'male', *nolól* N 'life'.

[w] is a voiced bilabial approximant, or semivowel, represented as *w* in the orthography.

[j] is a voiced palatal approximant, or semivowel, represented as *y* in the orthography. These two symbols *w* and *y* perform two distinct phonological functions: first, as consonantal onsets of syllables, e.g. *wár* N 'news' and *yár* ADJ 'small'; and secondly as the second vocalic elements in diphthongs, e.g. *háwl* N 'work' and *nàyl* N 'male lamb'. In the latter case the diphthongs count as long vowels in the phonology, as for example in the rules of the scansion system for poetry described by Johnson (1979). At a more purely linguistic level, as we shall see later, the realisation rules for tonal accent also treat diphthongs and long vowels as equivalent units.

2.2   Vowels

Somali has an interesting version of a five vowel system, involving two sets of five vowels, as shown in (2) below where IPA symbols show their principal pronunciation:

(2)	ɪ	u	i	ʉ
	ɛ	ɔ	e	ö
		ɑ		æ

The two sets form five pairs of vowels where in each pair one vowel is pronounced with the tongue more forward in the mouth than the other. Each pair of vowels is represented by a single letter in the standard orthography, as shown in (3):

(3)	Orthography	Back series	Front series
	i	ɪ	i
	e	ɛ	e
	a	ɑ	æ
	o	ɔ	ö
	u	u	ʉ

Each pair of vowels can be differentiated by the phonetic feature **advanced tongue root** (ATR): thus we can label the FRONT series [+ATR] and

the BACK series [-ATR]. An impression of the typical values for these vowels can be shown using a cardinal vowel chart as in Figure 2.
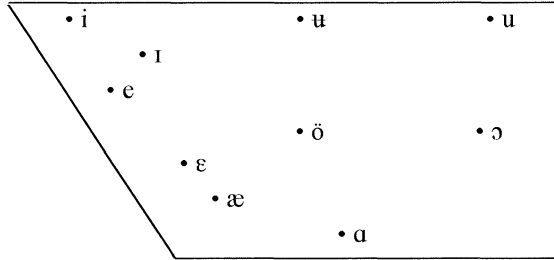
Figure 2.   Cardinal vowel chart for Somali short vowels

Each vowel occurs long with very little change in quality. This is shown by doubling in the orthography, i.e. [ɑː] and [æː] as *aa*, [ɛː] and [eː] as *ee*, etc. Some examples of long and short vowels are given in (4):

(4)　　(a)　Long and short vowels - back series

nín	[nɪn]	N	'man'	*shìiq*	[ʃɪːχ]	N	'sheikh'
Muuse	[muːsɛ]	NAM	'Musa'	*kèen*	[kɛːn]	V	'bring'
dáb	[dɑb]	N	'fire'	*qàad*	[qɑːd]	V	'take'
tól	[tɔl]	N	'kinsfolk'	*góol*	[gɔːl]	N	'lioness'
búste	[bustɛ]	N	'blanket'	*hadhùudh*	[hɑɖuːɖ]	N	'millet'

　　(b)　　Long and short vowels - front series

hílib	[hilib]	N	'meat'	*iibi*	[iːbi]	V	'buy'
dhég	[ɖeg]	N	'ear'	*gées*	[geːs]	N	'side'
cáb	[ʕæb]	V	'drink'	*ràag*	[ræːg]	V	'be late'
tóg	[tög]	N	'river bed'	*gòo*	[göː]	V	'cut'
gúddi	[guddi]	N	'assembly of judges'	*gúul*	[guːl]	N	'victory'

The distinction between short and long vowels is always significant; compare for example *tág* [tɑg] V 'go' and *táag* [tɑːg] N 'strength'.

The relationship between the sets of front and back vowels is interesting. Firstly they are not simply phonetically conditioned variants and thus are not allophones in classical phonemic terms. Individual members of the major lexical categories, for example nouns, verbs and adjectives, must occur with a specific

vowel quality and there are a number of minimal pairs like the nouns in (5)
below and the verbs in (6):

(5) [ɪdɑn]   *ídan*   M 'permission'   [idæn]   *ídan* M 'incense burner'
     [ɔr]   *ór*   M 'chant, chorus'   [ör]   *ór*   M 'cloth'
     [dul]   *dúl*   F 'top'   [dʉl]   *dúl*   M 'nostril'
     [ʕɑːr]   *càar*   M 'spotted coat (livestock)' [ʕæːr]   *càar* M 'camel's penis'
     [gɔːd]   *gòod*   M 'poisonous snake species'   [göːd]   *gòod* M 'piece of cloth'

(6) [sug]   *súg*   'wait for'   [sʉg]   *súg*   'make certain'
     [dɪːd]   *dìid*   'faint'   [diːd]   *dìid*   'refuse'
     [bɑːrɔ] *baaró*   'inspect'   [bæːrö] *baaró*   'writhe, roll about'
     [duːl]   *dìul*   'fly'   [dʉːl]   *dìul*   'attack'

However such minimal pairs are very few and for the most part the back/front
distinction is important for correct pronunciation but not for distinguishing
lexical meaning.

As an exception to this pattern of inherent vowel qualities, there is one
regular grammatical difference signalled by a difference in quality. For many
verbs it distinguishes between the imperative singular and the third person form
of the (rarely used) past independent paradigm, described in 4.3.4:

(7) (a)   [ɖɪs]   *dhís!* 'Build (SG) (it)!'
           [ɖis]   *dhís.* 'He built (it).'
     (b)   [hɛl]   *hél!*   'Find (SG) (it)!'
           [hel]   *hél.*   'He found (it).'
     (c)   [kɛːn] *kèen!* 'Bring (SG) (it)!'
           [keːn] *kèen.* 'He brought (it).'

In this case, since it seems that the back series vowel in the imperative is the
normal root form, the front vowels in the independent past forms may reflect
some historical loss of a suffix.

A second characteristic of this vowel quality distinction is that it is
important to vowel harmony. As mentioned above, roots belonging to the
major lexical categories inherently contain either back or front vowels in their
roots. In polysyllabic roots all the vowels must occur in the same series, for
example:

(8)   (a) Back series
         [mɑn]   *ínan*   N   'boy'
         [dɑmeːr]   *damèer*   N   'male donkey'
         [sɑlɑːn]   *salàan*   V   'to greet'
         [wɑddɔ]   *wáddo*   N   'road'
         [dɪjɑːrsɔ]   *diyaarsó*   V   'prepare for oneself'

         (b) Front series
         [hilib]   *hílib*   N   'meat'
         [wɛbi]   *wébi*   N   'river'
         [wædæːd]   *wadàad*   N   'religious man'
         [sæliːd]   *saliid*   N   'oil'
         [köːfijæd]   *koofiyád*   N   'hat'

The operation of vowel harmony in strings longer than the word has not yet been comprehensively described. As pointed out by Andrzejewski (1955), the stretches over which the harmony applies, what he terms 'harmonic groups', depend on a number of discourse features, in particular the speed and formality of speech and the presence of pauses. The length of harmonic groups increases with speed, and pauses automatically mark the end of a harmonic group. One further feature seems clear: the front series vowels are dominant in that their presence within a harmonic group causes a regressive or anticipatory fronting of any back series vowels in the same group. We can show this with a simple example from Andrzejewski (1955); below are the inherent vowel qualities of three nouns:

(9)   *sáan*   N F 'hide'   [sɑːn]
         fáras   N M 'horse'   [fɑrɑs]
         díbi   N M 'bull'   [dibi]

We can see that *sáan* and *fáras* have back series vowels and *díbi* has front series vowels. When such nouns are combined, we can see that the front series vowels of *díbi* will regressively harmonize all the vowels in the harmonic group, marked < > below:

(10)   (a)   *<Waa sáan   farás>*   [waː sɑːn fɑrɑs]
                   DM hide   horse:GEN
                   'It is a horse's hide.'

(b)   *<Waa   sáan   dibí>*   [wæː sæːn dibi]
        DM   hide   bull:GEN
        'It is a bull's hide.'

Some functional categories, such as the focus word *bàa*, conjunctions, subject pronouns, determiners, and auxiliary verbs are also affected by a progressive vowel harmony: they will agree in vowel series with a preceding word. We can perhaps show this with another example from Andrzejewski (1955):

(11)   (a)   Back series
              <Ma   shabèel   bàa?>   [ma ʃabɛːl baː]
              QM   leopard   FOC
              'Is it A LEOPARD?'
       (b)   Front series
              <Ma   libàax   bàa?>   [mæ libæːħ bæː]
              QM   lion   FOC
              'Is it A LION?'

Here we may assume that the question word *ma* and the focus word *bàa* have back series vowels inherently and are affected in (11b) by regressive and progressive harmony, respectively. We can view the harmony as a spreading of the feature [+ATR] across the harmonic group.

In addition to these short and long vowels, standard Somali also has the following diphthongs, represented in the orthography: *ay*, *aw*, *ey*, *oy*, *ow*, and which also occur long as *aay*, *aaw*, *eey*, *ooy* and *oow*. These diphthongs also occur in back and front series, except for *oy*, which does not seem ever to occur in the back series. Some examples of the back and front series are given for short diphthongs in (12) below:

(12)   Diphthongs: back and front series
              Back   Front
ay   *odáy*   [ɔdɑɪ]   N   'old man'   *dày*   [dæi]   V   'look at'
aw   *àwr*   [ɑur]   N   'male camel'   *càws* [ʕæʉs]   N   'grass'
ey   *sháley*   [ʃɑlɛɪ]   ADV 'yesterday'   *wèyl* [weil]   N   'calf'
oy   -   *qóy*   [qöi]   V   'to wet'
ow   *ká tiirsanòw*   [kɑ tɪːrsɑnɔu] V 'be one of'   *dhòw* [ḍöʉ]   ADJ 'near'

These diphthongs participate in the vowel harmony process described above in the same way as other vowels; for example the vocative suffix -*ow*,

described in chapter 4, harmonises with the vowel quality of the noun to which
it is attached, as shown below:

(13) (a) *Xásanòw!*   [hɑsɑnɔu]   'O Hasan!'
      (b) *Xuséenòw!*   [huseːnöu]   'O Husein!'

2.3   The syllable

Somali has a simple syllable structure: a syllable may only consist of one of the
three following types of combination:

(14)   (a) a vowel, either short vowel, long vowel or diphthong, e.g. *ú* ADP 'to, for',
               oo CONJ 'and', *èy* N 'dog';
          (b) a consonant followed by a vowel, e.g. *kú* ADP 'in, into', *sìi* V 'give to',
               cáy N 'insult';
          (c) a consonant followed by a vowel, followed by a consonant, e.g. *nín* N
               'man'; *qúudh* N 'contempt', *nàyl* N 'male lamb'.

When a word-initial syllable or a monosyllable would consist simply of a vowel,
a glottal stop is inserted to provide a consonantal onset to the syllable, so for
example *èy* N 'dog' is pronounced *'èy*. These 'phonetic' glottal stops are not
marked in the orthography.
        Since the maximum number of consonants at the beginning and end of
syllables is one, the largest consonant cluster that can occur consists of two
consonants across a syllable boundary, as in the following examples where $
represents a syllable boundary:

(15)   *san$dùuq*   N   'box'
       dhag$xán   N   'stones'
       már$kab   N   'ship'
       dam$be   ADJ   'behind'

The consonant clusters are further restricted by the fact, mentioned earlier, that
the sounds *t* and *k* do not occur syllable finally, nor *j* in native Somali words.
        When two instances of the consonants *b d dh g q l m n* and *r* meet at a
syllable boundary they form phonetically long geminate consonants, for
example:

(16)   *rabbèe*   V   'educate'
       áddin   N   'limb'

cádhdho	N	'scabies'
oggól	ADJ	'agreeing'
qallòoc	V	'become crooked'
xammèe	V	'boil'
annága	IPRO	'we (EXC)'
karráani	N	'clerk'

Note that there are no geminate fricatives in Somali. The difference between single and geminate consonants forms minimal pairs like *wáran* N 'spear' and *wárran* V 'tell news'.

2.3   Tone and accent

Impressionistically Somali has a contrast between three tones: High, Low and (High to Low) falling, usually marked as follows: **á** (High), **a** (no mark, Low) and **à** (Falling). Tones on long vowels are marked just on the first vowel symbol. We follow this traditional marking in the body of this book but for the purposes of description, in this section we will mark Falling tone with a circumflex, i.e. **â**. This contrast allows minimal pairs, as in (17) and (18) below:

(17)	*ínan*	N	'boy'	*inán*	N	'girl'
	farîid	N	'wise man'	*faríid*	N	'wise woman'
	wêyl	N	'male calf'	*wéyl*	N	'female calf'
(18)	*êy*	N	'dog'	*éy*	N	'dogs'
	mádax	N	'chief'	*madáx*	N	'chiefs'
	Soomáali	N	'a Somali'	*Soomaalí*	N	'Somalis, the Somali race'

Stress is associated with tone: High tones have strong stress, sometimes described as accent, Falling tones have a diminishing strong to weak stress, while Low tones have no stress, or are unaccented.

The perceived pitch height of these tones is of course relative rather than absolute: a High tone is perceptibly higher than neighbouring Low tones. There are two phenomena which bring this relativity into focus. The first is the phenomenon of downdrift within phrase groups. This is a tendency for successive tones of the same unit to be pronounced with slightly lower pitch. Thus a succession of three High tones will often involve each High tone descending slightly from the previous one, as we show graphically in (19b) below:

(19)　(a)　*Igú*　*sóo*　*céli*!
　　　　　　i-kú　*sóo*　*céli*
　　　　　　me-to　VEN　return:IMP:SG
　　　　　　'Bring him back to me!'

(19b)
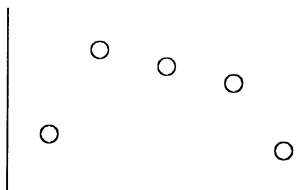
A second, related phenomenon is the lowering of tones in pre-pause position. A tone in Somali is higher in a syllable followed by another syllable than in one followed by a pause. To describe this it is useful to identify three phonetic pitch levels: 3 (highest), 2 (mid) and 1 (lowest); we can use these to show the relative pitch of the syllables of the two words below:

3					kar	
2	as	kar		as		
1						ta

　　　　askár N 'soldiers'　　　　*askárta* N 'the soldiers'

If the two words above, *askár* 'soldiers' and *askárta* 'the soldiers', are uttered in isolation, their High tones are not at the same absolute level. The High tone in *askár* is at level 2 because it is before a pause, while the High tone in *askárta* is at the higher level 3 because it is followed by another syllable. The initial Low tones of the two words are at the same pitch level, level 2, because both are followed by another syllable, while the final Low tone of *askárta* is lower, at level 1 because it is before a pause. We can also see that a pre-pause High tone and a pre-syllable Low tone might both be phonetically mid level in pitch. In terms of the system, though, a two level distinction is maintained in both positions.

　　Both of these tone lowering phenomena can be overruled by speakers: Andrzejewski (1979: 9) for example, describes how speakers sometimes raise the last tone before a pause by one or two pitch levels to show that they are about to continue.

　　Some features of the Somali tonal system have raised the question of where Somali fits into the various typologies of tone languages that have been proposed. We can review these features briefly. The first observation is that

the perceived three tone system can be simplified to two units by treating Falling tone (FG) as a sequence of High (H) and Low (L) tones. This decision makes descriptive sense for two reasons: firstly it explains why Falling tone only occurs on long syllables. Secondly it allows us to recognise that the pattern High-Low on the last two units versus Low-High consistently differentiates between masculine and feminine gender in nouns, whether the nouns are bisyllabic or long monosyllables, as we can see below:

(20)   (a)   *ínan*   H L   N 'boy'
             inán   L H   N 'girl'
       (b)   *gôol*   FG   [←*góol* H L]   N 'male fatted camel'
             góol   H H   [← *goól* L H] N 'female fatted camel'
       (c)   *nâyl*   FG   [←*náyl* H L]   N 'male lamb'
             náyl   H H   [←*naýl* L H]   N 'female lamb'

The conclusion we draw from this is that the tone-bearing unit in Somali is neither a vowel segment nor a syllable, but a distinct unit we can call a mora. A syllable with a short vowel consists of one mora and can bear one tone; a syllable containing a long vowel or a diphthong consists of two moras and can bear two tones. Note that moras attach to vowels: the number of consonants in a syllable does not affect the counting of moras. Thus the syllable structures {V, CV, VC, CVC} all count as a single mora while the structures {VV, CVV, VVC, CVVC} count as two moras, where VV is a long vowel or diphthong.[11] This will become clear from our examples but as noted by Orwin (1994: 18-19) it is also supported by the scansion system of Somali poetry which is based on moras and in which only vowel melodies are relevant for counting.[12]

    As can be seen in (20), we also need to recognise the two phonetic rules for tones given informally in (21) & (22) below:

(21)   On a long vowel or diphthong a sequence of High-Low is realised as a Falling tone.
(22)   On a long vowel or diphthong, a sequence of Low-High is realised as High-High, rather than as a Rising tone.[13]

This analysis allows us to recognise a number of important generalisations about the tone patterns of Somali. One, as we have just seen, is that many tonal minimal pairs of nouns have High tone on the penultimate (or only) mora in the masculine, and High tone on the final (or only) mora in the feminine. Short monosyllables which consist of just one mora have High tone and are not tonally distinguished for gender. Further examples of this pattern are given in (23),

where we can see that not all the noun pairs are semantically related synchronically:

(23)   Masculine   Feminine
        nácas   'foolish man'   *nacás*   'foolish woman'
        xíddig   'star (in astrology)'   *xiddíg*   'star (in astronomy)'
        damêer   'male donkey'   *daméer*   'female donkey'
        islâan   'old man'   *isláan*   'old woman'
        gêes   'horn'   *gées*   'side, direction'
        bêer   'liver'   *béer*   'garden'
        káb   'repairing, mending'   *káb*   'shoe'

The identification of the mora as the tone-bearing unit allows similar generalisations in the verbal system. First conjugation verbs, for example, have in the imperative a tonal pattern High-Low on the last two moras of the root which, depending on the syllable structure, may surface as High-Low on successive short syllables as in (24) below, or as a Falling tone on a long syllable as in (25):

(24)   *dálab*!   'Ask (SG) for (it)!
        hádal!   'Talk (SG)!
        sáfar!   'Travel (SG)!'
(25)   *kèen*!   'Bring (SG) (it)!
        èeg!   'Look (SG)!'
        salàan!   'Greet (SG)!'

We can also see the phonetic rule (22) at work again: in (26) below a High on the second mora of the long syllable produces a High long syllable:

(26)   *áammus*!   (← *aámmus*)   'Be quiet (SG)!'
        dáabac!   (←*daábac*)   'Print (SG)!'
        sáabil!   (← *saábil*)   'Cheat (SG)!'

Once again, a short monosyllable is unable to carry the High-Low pattern and surfaces as High:

(27)   *cún*!   'Eat (SG)!'
        dhís!   'Build (SG)!'
        súg!   'Wait (SG)!'

The imperative plural has a suffix -*a* which changes the syllable structure of the root, allowing verbs which are short monosyllables in the singular to show the characteristic High-Low pattern; compare (28) below with (27):

(28)   *cúna*!   'Eat (PL)!'
        dhisa!   'Build (PL)!'
        súga!   'Wait (PL)!'

These imperative plurals also obey phonetic rule (22): a High tone on the second mora of a long vowel will cause a long High vowel:

(29)   *kéena*!   (← *keéna*) 'Bring (PL)!'
                jóoga!   (← *joóga*) 'Stay (PL)!'

    The second major feature of the Somali tone system is that its function is to signal grammatical rather than lexical contrasts. In our noun examples so far we have seen tone patterns marking distinctions of gender and number. In fact tonal patterns in nouns are predictable from the noun's number, gender and declension class. As we shall see in chapter 4, it is possible to establish declensions of nouns, each with their characteristic tonal pattern, gender, and way of forming plurals. Similarly individual verbal conjugations have associated tonal patterns. Any apparently lexical contrasts, such as between the nouns *ínan* 'boy' and *inán* 'girl' are by-products, so to speak, of the primary grammatical distinction, in this instance of gender. Other functions of the tonal system include case marking on nouns, as we shall see below.

    A third feature of the tonal system is its independence from sentence intonation. Because the tonal system employs pitch for essentially word level grammatical information, it is not free to carry sentence level information such as sentence type distinctions or information structure contrasts such as emphasis, or focus. Thus questions, for example, are not distinguished from statements by pitch tunes as in some other languages like English, but by the use of specific morphemes, as shown in (30):[14]

(30)   (a)   *Má   imáneyaa?*
                        QM   is:coming
                        'Is he coming?'
        (b)   *Wâa   imáneyaa.*
                        DM   is:coming
                        'He is coming.'

Similarly to emphasise or focus on a constituent, Somali speakers use one of the focus particles introduced in 4.7; in example (31) below the focus particle *bâa* marks the noun *búug* 'book' as in focus:

(31)   *Bûug   bûu   keenay*
                 bûug   bâa+uu   keenay
                 book   FOC+he   brought
                 'He brought A BOOK'
                 (1)   Presupposition: He brought something.
                 (2)   Assertion: The thing was a book.

Sentence-level intonation patterns do exist in Somali and carry signals of the speaker's attitude and of emotional colouring, in ways that have not been studied so far. However the use of pitch at word level means that intonation has a smaller communicative load than in languages like English.

A fourth feature of the tonal system is the relatively restricted set of tone patterns. As noted by Andrzejewski (1964, 1968, 1979) and Hyman (1981), all roots of common nouns, verbs and adjectives have one of three basic tone patterns in their citation, or isolation, form, which we can represent as:

(32)   Pattern 1:   High on the last mora, Low on all others
       Pattern 2:   High on the penultimate mora, Low on all others.
       Pattern 3:   Low on all moras.

As we shall see in subsequent chapters, knowing the declension and gender of a noun or the conjugation of a verb, allows one to predict the tone pattern of the basic form. There are of course some complications, especially in other categories: some proper nouns, for example, also allow a fourth pattern:

(33)   Pattern 4:   High on the first and last mora, Low on all others, e.g.
       Bôoramé [←*Bóoramé*] the Somali town of Borama.

Furthermore, as we discuss in chapter 3, some sentence type markers, like the question word *ma*, have variable tone depending on the presence of High tones in their environment. Nevertheless it is very clear that the mapping between tone patterns and words is very restricted and predictable: Somali does not show the full range of tonal patterns we would expect if each syllable or mora were independent of others in the word.

The basic tone patterns of roots are changed by grammatical or derivational processes. This will become very clear in later chapters, but we can select two typical examples here. When nouns are assigned case as part of

sentences, this is marked largely by tone. Example (34) below shows the case forms of the proper noun *Cáli* 'Ali' revealing the associated tonal and suffix changes:

(34)   absolutive case (= citation):   *Cáli*
      nominative case: *Cali*
      genitive case: *Calí*
      vocative: *Cáli* or *Cáliyôw* [← *Cáliyów*]

We discuss accentual case marking further in chapter 3. As a second example we take one of the derivational processes where an affix is used to create verbs from nouns and adjectives, in this case the suffix which has the form *-êe* (Falling tone) in the imperative form used for verb citation forms. The addition of this suffix flattens any High tones in the stem to Low, leaving the derived verb with Pattern 2 in (32) above, High on the penultimate mora. This is phonetically realised as a Falling tone on the final long vowel. We can see examples of this:

(35)   *cásho*   N   'evening meal'   *cashêe*   V   'eat the evening meal'
      dháb   N   'truth'   *dhabêe*   V   'confirm'
      dhêer   ADJ   'long'   *dheerêe*   V   'move fast'

In fact a wide range of morphological and syntactic processes change the basic tone patterns of word roots. These changes will be described in the relevant sections of this book. What is clear is that tone patterns in Somali are word-based and reflect a combination of the word's inherent syntactic class and the tonal markers of grammatical information.

    These basic features allow us to comment briefly on the typology of the Somali tone system. If we adopt the traditional distinction between intonation languages, tone languages and pitch accent languages, it seems clear that Somali is not an intonation language. As we have seen, Somali does not use pitch patterns for the types of sentence level information that we find in English and other European languages. Somali uses tone to make contrasts of meaning at the word level. However, the position of Somali in relation to the distinction between tone languages proper and pitch accent languages is less immediately obvious. We may take the defining characteristics of pitch accent languages like Japanese to be as follows: words may have no more than one accent per word; this accent is realised by pitch prominence; and the accent is relatively unaffected by intonation (Cruttenden 1986). The features we have seen in Somali word roots: their restricted tonal patterns, in particular the fact that common nouns, verbs and adjectives contain only one High tone in their basic

form, makes an accentual analysis seem plausible and such analyses have been proposed by Hyman (1981) and Banti (1988b). The basic idea of this type of approach is that underlyingly, or lexically, Somali roots have no inherent tone. Rules of accent placement determine the position of a single accented mora in each root, which is then phonetically realised as a high pitch segment. These rules of accent placement are governed by grammatical information, for example what syntactic class the root is, which declension or conjugation, etc. Subsequently morphological and syntactic rules may change the placement of this accented mora and thus of the associated High tone. This seems a plausible line of approach for Somali and we discuss it further in chapter 3. What is clear is that Somali uses tone quite differently from Asian tone languages like Chinese and Korean and with significant differences from better known types of African tone languages like Igbo, Hausa, or the Bantu languages. Similar tone accent systems have been reported for other Cushitic languages, notably Oromo (Andrzejewski 1970, Owens 1980, 1982, 1985, Banti 1988b).

Tone is not marked in the official Somali orthography. In the rest of this book we will mark the tone of words and constructions recorded from speech but not attempt to tone mark texts and examples taken from published Somali sources. We will adopt a conservative strategy in the marking of tones, identifying the three 'surface' tones on syllables, rather than the underlying moraic structure, as follows:

(36)   High tone: on short vowels   *á*
                         on long vowels and diphthongs   *áa*   *áy*
        Low tone: on short vowels   *a*
                         on long vowels and diphthongs   *aa*   *ay*
        Falling tone: on long vowels and diphthongs   *àa*   *ày*

We will only mark underlying moraic forms, for example the sequence of High and Low tones underlying the Falling tone, when discussing the details of the tonal system.

Chapter 3
Phonological Rules

In this chapter we outline some of the main types of phonological rules in Somali. We include both phonological rules proper and morphophonological rules, that is, rules which are triggered by or restricted to some grammatical categories. Some rules of allophonic variation, for example the tendency for *b* and *d* to spirantize to [β] and [ð] intervocalically, were described in chapter 2.

3.1   Syllable structure processes

In chapter 2 we described the set of possible syllable structures of Somali, given below:

(1)   {V, CV, VC, CVC}
       where V = short vowel, long vowel or diphthong

In this section we show some compensatory rules to preserve preferred syllable structures against lexical or morphological pressure. In each case we provide just one or two examples for illustration.

3.1.1 Syllable final stops

The voiceless plosives *t* and *k* do not occur in syllable final position. If predicted in this position, their voiced counterparts occur, e.g. *k* → *g*, for example *árag* 'See!' (← **arak*), cf. *arkay* '(I) saw'; *gúnud* 'Knot (it)!' (←**gunut*), cf. *guntay* '(I) knotted it'.

   The nasal *m* does not occur word finally and is there replaced by *n*, for example *n*←*m* in *láan* N branch, cf. *laamo* 'branches'. In fact the only syllable final occurrences of *m* are when followed by *b* or *m*, for example:

(2) *hambábar*   V   'jump up'
     hambályo   N   'congratulations'
     ímmisa   PHR   'how much?'

This restriction can be explained by viewing these cases as underlying *n* assimilating to a following *m* or *b*, in which case the correct observation is that *n* does not occur syllable finally.

3.1.2 Consonant cluster simplification

If a root underlyingly ends in a double consonant, VCC, the cluster will be simplified to VC, unless a following syllable allows a re-syllabification as CVC$CV:

(3) *cáb* 'Drink (it)!' (← **cább*), cf. *wàan cabbay* 'I drank (it)', *wàad cabtay* 'you drank (it)'; *cád* ADJ 'white' (←**cádd*), *caddèe* V 'whiten'.

3.1.3 Restrictions on V syllables

In chapter 2 we noted the rule that syllables beginning words must have consonantal onsets: thus V syllables in this position become CV by the insertion, or epenthesis, of a glottal stop; for example *úl* 'stick is pronounced [ʔul].

There is also a restriction on a V syllable following a CV syllable, i.e. CV$V, wherever this is predicted by the grammar. These sequences are avoided by the insertion of epenthetic consonants, in what we might term a hiatus avoidance rule. Examples below show the consonant *y* inserted after stem final *i* in noun suffixation, and the glottal stop inserted after other vowels in the cliticisation of subject pronouns onto the negative word *má*:

(4) *mindí* N 'knife', *mindiyo* 'knives' (← **mindio*), cf. *náag* N 'woman', *naago* 'women'.
(5) *má* NEG + *aan* 'I' →*má'aan* 'not I', *má* + *uu* → *mú'uu*, 'not he'.

3.1.4 Stem contraction

Whenever the suffixation of elements to a stem creates trisyllabic roots with the form $(C_1) V_1 C_2 V_2 C_3 V_3$, if V_2 is a short vowel it will be deleted, contracting the word to $(C_1) V_1 C_2 C_3 V_3$. Examples from noun plurals are in (6) and from verbal inflection in (7):

(6)	*hílib*	'meat'	**hilibó*	→	*hilbó*	'meats'
	jiríd	'trunk'	**jiridó*	→	*jirdó*	'trunks'
	qódob	'article'	**qodobó*	→	*qodbó*	'articles'
(7)	*dádab*	'block'	**dadabay*	→	*dadbay*	'(I) blocked'
	máqal	'listen'	**ha máqalin*	→	*ha máqlin*	'Don't listen!'
	sáfar	'travel'	**safaray*	→	*safray*	'(I) travelled'

The combination of the stem reduction rule and the restrictions on the distribution of stops described in 3.1.1 gives alternations such as those nouns in (8) and verbs in (9):

(8)	(a)	*ílig*	'tooth'	(from **ílik* by final stop rule)
		ilkó	'teeth'	(from **ilikó* by stem reduction)
	(b)	*gacán*	'hand'	(from **gacám* by final nasal rule)
		gacmó	'hands'	(from **gacamó* by stem reduction)
(9)	(a)	*gúnud*	'knot'	(from **gúnut* by final stop rule)
		guntay	'(I) knotted (it)'	(from **gunutay* by stem reduction)
	(b)	*kában*	'heal up'	(from **kábam* by final nasal rule)
		kabmay	'(it) healed up'	(from **kabamay* by vowel reduction)

The stem reduction rule is blocked if its application would result in a voiceless stop, *t* or *k* occurring at the end of a syllable:

(10)	*fétesh*	'search'	*feteshay*	'(I) searched'	not **fetshay*
	hítiq	'walk slowly'	*hitiqay*	'(I) walked slowly'	not **hitqay*
	cákis	'hinder'	*cakisay*	'(I) hindered'	not **caksay*
	súkul	'pound (grain)'	*sukulay*	'(I) pounded'	not **suklay*

The rule is also blocked if C$_2$ and C$_3$ are identical, that is, if the rule would produce a geminate consonant:

| (11) | *ólol* | 'flame' | *ololay* | '(it) flamed' | not **ollay* |
| | *fududèe* | 'lighten' | *fududeeyey* | '(he) lightened' | not **fuddeeyey* |

In some cases a rule of metathesis operates after stem reduction to produce a preferred consonant cluster, for example:

(12)	(a)	*culús* ADJ 'heavy' + *-aa* 'was' →	
		*culusaa →	
		*culsaa →	(stem contraction)
		cuslaa 'was heavy'	(metathesis)

(b)   *shilís* ADJ 'fat' + *-aan* '-ness' →
         *shilisáan →
         shilsáan →   (stem contraction)
         shisláan   'fatness'   (metathesis)

3.2   Sandhi

When grammatical suffixes attach to lexical roots there are considerable phonological changes. These changes are often specific to particular grammatical morphemes and we will use the Sanskrit-based label *sandhi* as a cover term for such changes. We use as examples in this section the results of suffixation in nouns in 3.2.1-2 and verbs in 3.2.3-5. These rules can be considered a form of word-internal sandhi. We deal with a form of external sandhi, coalescence, in section 3.3.

3.2.1 Determiner suffixes

As described in 6.2.1, the set of determiners attaches as suffixes to nouns. The determiners' initial consonants, *k* in the masculine and *t* in the feminine, undergo a series of changes at the juncture. We list these in (13) and (14), using as an example the non-remote definite article *-ka/-ta*:[15]

(13) (a)   Suffix initial *k* remains *k* after *b*, *d*, *dh*, *f*, *j*, *l*[16], *n*, *r*, *sh*, e.g.
             dáb + ka   →   *dábka*   'the fire'
             siràad + ka   →   *siràadka*   'the lamp'
             xidhìidh + ka   →   *xidhìidhka*   'the connection'
             báraf + ka   →   *bárafka*   'the ice'
             xáj + ka   →   *xájka*   'the pilgrimage'
             cèel + ka   →   *cèelka*   'the well'
             nín + ka   →   *nínka*   'the man'
             fèer + ka   →   *fèerka*   'the punch'
             kìish + ka   →   *kìishka*   'the bag'
      (b)   Suffix initial *k* becomes *g* after *g*, *w*, *y*, *i*, *u*, *aa*, *oo*, *uu*, e.g.
             tùug + ka   →   *tùugga*   'the thief'
             càw + ka   →   *càwga*   'the side'
             cày + ka   →   *càyga*   'the insult'
             qálbi + ka   →   *qálbiga*   'the heart, spirit'
             hú + ka   →   *húga*   'the clothes'
             wàa + ka   →   *wàaga*   'the time'
             adòo + ka   →   *adòoga*   'the father, old man'

	dhùu + ka	→	*dhùuga*	'the leather garment'

(c) Suffix initial *k* becomes *h* after *e* and *o* (usually with accompanying vowel assimilation), e.g.

áabbe + ka	→	*aabbáha*	'the father'
ilkó + ka	→	*ilkáha*	'the teeth'

(d) Suffix initial k is deleted after *q*, *'*, *kh*, *x*, *c*, *h*, e.g.,

dúq + ka	→	*dúqa*	'the elder'
gó' + ka	→	*gó'a*	'the portion'
shèekh + ka	→	*shèekha*	'the learned man'
sáx + ka	→	*sáxa*	'the right way'
sác + ka	→	*sáca*	'the cow'
ráh + ka	→	*ráha*	'the frog'

(14) (a) Suffix initial *t* remains *t* after *b*, *f*, *g*, *n*, *r*, *s*:

xéeb + ta	→	*xéebta*	'the shore'
séef + ta	→	*séefta*	'the sword'
náag + ta	→	*náagta*	'the woman'
sún + ta	→	*súnta*	'the poison'
jéer + ta	→	*jéerta*	'the hippopotamus'
hées + ta	→	*héesta*	'the song'

(b) Suffix initial *t* becomes *d* after all vowels and *'*, *c*, *d*, *h*, *kh*, *q*, *w*, *x*, *y*:

gúddi + ta	→	*gúddida*	'the council of elders'
ayéeyo + ta	→	*ayeeyáda*	'the grandmother'
ló' + ta	→	*ló'da*	'the cattle'
bác + ta	→	*bácda*	'the sound of a slap'
koofiyád + ta	→	*koofiyádda*	'the hat'
báh + ta	→	*báhda*	'the noble person'
taariikh + ta	→	*taariikhda*	'the history'
búq + ta	→	*búqda*	'the spring, source'
qáw + ta	→	*qáwda*	'the thump'
madáx + ta	→	*madáxda*	'the heads, chiefs'
éy + ta	→	*éyda*	'the dogs'

(c) Suffix initial *t* becomes *dh* after *dh*:

gabádh + ta	→	*gabádhdha*	'the girl'[17]

(d) Suffix initial *t* merges with stem final *l* to form *sh*, e.g.

úl + ta	→	*úsha*	'the stick'
méel + ta	→	*méesha*	'the place'

3.2.2 Noun final vowel assimilation

When noun stems end in a vowel, the set is restricted.[18] Masculine nouns may end in *i*, *e* or *o*; and feminine nouns in *i* or *o*. Of these vowels, *e* and *o* are affected by assimilation rules when further suffixes are added to the stem. We

may take as an example singular nouns in declension 7 (4.1.1.3). These are
masculine nouns ending in *i* or *e* which form their plurals by a suffix *-yaal*. This
suffix triggers stem final *e → a*:

(15)   Singular   Meaning   Plural
           árbe   'bull elephant'   *arbayáal*
           bakáyle   'hare'   *bakaylayáal*
           dhámme   'captain (armed forces)' *dhammayáal*
           góle   'enclosure'   *golayáal*
           talíye   'commander'   *taliyayáal*

Final vowel assimilation also occurs when determiner suffixes are added to these
nouns. As shown below, such suffixed forms are also affected by the rules of
determiner sandhi described in 3.2.1 and a tonal rule called Accent Shift in 3.4.2
below. We show an example with the non-remote (*-ka/-ta*) and remote (*-kii/-
tii*) definite determiners:

(16) (a)   *báre* (M) 'teacher' + *ka →*
             *baréka →   (Accent Shift)
             *baréha →   (suffix sandhi)
             baráha 'the teacher'   (vowel assimilation)
     (b)   *báre* (M) 'teacher' + *kii →*
             *barékii →   (Accent Shift)
             *baréhii →   (suffix sandhi)
             baríhii 'the teacher'   (vowel assimilation)

        A similar assimilation rule affects nouns in declension 6 (4.1.1.3).
These are feminine singular nouns ending in *o*. Once again this vowel assimilates
to a following suffix:

(17) (a)   *dáwo* (F) 'medicine + *ta →*
             *dawóta →   (Accent Shift)
             *dawóda →   (suffix sandhi)
             dawáda 'the medicine'   (vowel assimilation)
      (b)   *dáwo* (F) 'medicine' + *tii →*
             *dawótii →   (Accent Shift)
             *dawódii →   (suffix sandhi)
             dawádii 'the teacher'   (vowel assimilation)

As can be seen in (16) and (17) the assimilation rule is more extensive when the
determiner becomes *h* initial. The final vowel becomes *i* before *-hii* but not

before -*dii*. This difference can also be seen in the assimilation rule which affects plural nouns ending in *o*. The feminine plural, *aqalló* (F) 'houses' becomes *aqalláda* 'the houses (non-remote)', and *aqalládii* 'the houses (remote)'. The masculine plural, *maalmó* (M) 'days', on the other hand, becomes *maalmáha* 'the days (non-remote) and *maalmíhii* 'the days (remote)'.

The same pattern of stem final vowel changes with *h*- and *d*- suffixes can be seen when the possessive determiners -*kay*/-*tay* 'my', etc.[19] are added to the nouns *áabbe* M 'father' and *hóoyo* F 'mother':

(18)   Masculine   Feminine

aabbáháy	'my father'	*hooyádáy*	'my mother'
aabbáháa	'your father'	*hooyádáa*	'your mother'
aabbíhiis	'his father'	*hooyádiis*	'his mother'
aabbéhéed	'her father'	*hooyádéed*	'her mother'
aabbáháyo	'our (EXC) father'	*hooyádáyo*	'our (EXC) mother'
aabbéheen	'our (INC) father'	*hooyádéen*	'our (INC) mother'
aabbíhiin	'your (PL) father'	*hooyádiin*	'your (PL) mother'
aabbóhóod	'their father'	*hooyádóod*	'their mother'

As can be seen, the changes in the stem final vowel in the masculine *h* series assimilates to the vowel in the suffix: *e* → *a* before *a* or *aa* , *e* → *i* before *ii*, and *e* → *o* before *oo*. The changes in the feminine *d* series constitute a single rule, stem final *o* → *a* before any suffix. As mentioned above, final *i* vowels are unaffected by these rules:

(19)   *gúri* M 'house'   *gúriga* 'the house'
              hantí F 'property'   *hantída* 'the property'

A similar vowel assimilation takes place in masculine nouns ending in the pharyngeal consonants *c*, *x* and the glottal stop ':

(20)   *súbax* 'morning' + *kii*   →   *súbixíi* 'the morning'
      sác 'cow' + *kii*   →   *sícíi* 'the cow'
      bá' 'suffering' + *kii*   →   *bí'ii* 'the suffering'

This only occurs when the final syllable is short, for example: *subàac* 'wild beasts' + *kii* → *subàacíi* 'the wild beasts'; *muftàax* 'key' + *kii* → *muftàaxíi* 'the key'.

3.2.3 Verbal causative affix

As will be discussed in 4.3.1, the structure of the main suffix type of verb can be described schematically as:

(21)   [ROOT - LEX - AGR - INFL]

In (21) ROOT represents the verb's root; LEX represents one or more of the set of lexical affixes which create derived verbs; AGR are the affixes which mark agreement with the person, number and gender of the verb's subject; and INFL are the markers of tense, mood, and aspect. At each morphological juncture, marked -, phonological changes take place. In this section we provide one or two examples of sound changes at the juncture between ROOT and LEX, using as an example the causative affix. This affix has the form *-i* in most verbal conjugations.[20] When this is added to verbal roots ending in *g* or *q* these consonants palatalise to *j*:

(22)   *dég*   INTR   'alight'   *déji*   TR   'cause to alight, unload'
       jòog   INTR   'stop, stay'   *jóoji* TR   'cause to stop, halt'
       dàaq   INTR   'graze'   *dáaji* TR   'put to graze'
       rúq   INTR   'be uprooted'   *rúji*   TR   'uproot'

This affix is also involved in a vowel assimilation rule similar to that described in 3.2.2: verbs consisting of short monosyllables ending in *x, c,* or the glottal stop ' have their root vowel assimilated to the affix:

(23)   *báx*   INTR   'go out'   *bíxi*   TR   'cause to go out, take out'
       dhác   INTR   'fall, happen' *dhíci* TR   'make something happen, cause'
       bá'   INTR   'be ruined'   *bí'i*   TR   'ruin'

Once again this rule does not apply to long vowels: *ràac* TR 'follow', *ráaci* TR 'cause to follow, send along', but not **riici*.

3.2.4 Verbal middle voice affix

In this section we provide some examples of sound changes associated with the middle voice affix, discussed in 5.1.6. This affix has the underlying form *-at* in most verbal forms[21]. To show the sound changes it is useful to give a typical set of verbal endings: we use the past simple paradigm of the root verb *kèen* 'bring':

(24)	(Clitic PRO)	Verb	Meaning	Ending (AGR + INFL)
	(*aan*)	*keenay*	'(I) brought'	~ + *ay*
	(*aad*)	*keentay*	'(you SG) brought'	*t* + *ay*
	(*uu*)	*keenay*	'(he) brought'	~ + *ay*
	(*ay*)	*keentay*	'(she) brought'	*t* + *ay*
	(*aannu*/*aynu*)	*keennay*	'(we) brought'	*n* + *ay*
	(*aydin*)	*keenteen*	'(you PL) brought'	*t* + *een*
	(*ay*)	*keeneen*	'(they) brought'	~ + *een*

As can be seen, the person markers *t*, *n* and ~ (representing a null marker or gap) intersect with the INFL elements to provide a five way person distinction that is extended further by pronouns.

When the middle voice affix -*at*- is positioned between the verb stem and the verb ending, a number of sound changes result. Firstly, if the affix is followed by a vowel, then its final *t* becomes *d*:

(25)   *joogsaday*   'I stopped myself'   (←*joogsatay*)

Secondly if -*at*- is followed by a person marker *t*, then the sequence *tt* simplifies to *t*:

(26)   *joogsatay*   'she stopped herself'   (←*joogsattay*)

Thirdly this affix often produces the environment for the stem contraction rule described in 3.1.4:

(27)   *dubatay*   →   *dubtay*   '(I) baked (it) for myself'
         guratay   →   *gurtay*   '(I) collected (it) for myself'

As can be seen, stem contraction forms a juncture between the remaining *t* consonant of the affix and the verb stem: in this position the affix *t* undergoes the same sound changes as the *t* of the nominal determiners (3.2.1) with one important difference. The parallels are:

(28)   (a)   *t* → *d* after *x*, *c*, *q*:
         seextay   →   *seexday*   '(I) went to sleep'
         goctay   →   *gocday*   '(I) remembered'
         fiiqtay   →   *fiiqday*   '(I) peeled (it) for myself'
         (b)   *t* may become *dh* after *dh*:
         xidhtay or *xidhdhay*   '(I) closed (it) for myself'

The difference is that when the verb stem ends in *d*, we find regressive assimilation at the juncture, then simplification:

(29)   *qaadatay*   'I took (it) for myself'   →
        qaadtay   (stem contraction)   →
        qaattay   (regressive assimilation)   →
        qaatay   (simplification)

This form contrasts with:

(30)   *qaadattay*   'You took (it) for yourself'   →
        qaadatay   (simplification)

There are two interesting points here: firstly the intervocalic -*t*- produced by simplification does not voice intervocalically like an underlying single *t*, e.g.:

(31)   *waan joogsaday*   (←*joogs-at-ay*)   'I stopped myself'
        waan qaatay   (←*qaattay*)   'I took (it) for myself'

Secondly the regressive assimilation of the sequence *dt* → *tt* with the middle voice affix contrasts with the progressive assimilation of the sequence *dt* →*dd* with the nominal determiners (e.g. *bád* + *ta* → *bádda* 'the ocean'). We shall see further examples of this dependence of phonological rules on grammatical status.

3.2.5 Verbal agreement affixes

In this section we provide one or two examples of sound changes at the juncture in verbs between the verbal stem and the agreement markers, AGR. The agreement marker *t* (2SG and 3SG F) undergoes progressive sound changes at the juncture with the verb stem and thus acts like the *t* of the determiner suffix and not the *t* of the middle voice affix, for example: *qaad* + *tay* → *qaadday* 'you/she took'. Other examples:

(32)   *go'* + *tay*   →   go'day   'you/she cut'
        raac + *tay*   →   raacday   'you/she followed'
        sooh + *tay*   →   soohday   'you/she wove'
        tookh + *tay*   →   tookhday   'you/she boasted'
        faq + *tay*   →   faqday   'you/she consulted secretly'
        bilaw + *tay*   →   bilawday   'you/she began'

bax + tay	→	*baxday*	'you/she went out'
xidh + tay	→	*xidhdhay*	'you/she tied'
hel + tay	→	*heshay*	'you/she found'

The agreement marker *n* (1PL) assimilates to a stem final *r* or *l*:

| (33) | *fur + nay* | → | *furray* | 'we opened' |
| | *hel + nay* | → | *hellay* | 'we found' |

3.3   Coalescence

One of the characteristic features of Somali is the widespread occurrence of coalescence rules which apply when words form part of phrases and sentences. This merging of forms, a kind of external sandhi, is sometimes a feature of the style of speech: whether the speech is fast or slow, formal or informal. However many of the rules are not dependent on the style of speech and are obligatory.   There has been very little study of the variation in coalescence across speech styles but in our examples we try to note those cases which seem obligatory. As will be seen, it is the functional categories like verbal adpositions, focus and negative words, verbal pronouns and conjunctions which are involved in these rules of phonological merger.   We might identify these elements as clitics, entities whose grammatical status lies between the affix and the independent word.   However it is important to note that cliticisation and coalescence are in principle distinct: some clitics like the conjunctions *-na* 'and' and *-se* 'but' (4.9) must attach to a preceding host but do not undergo phonological merger with them, while the elements described below only merge under certain conditions.   Some coalesce only when adjacent to certain grammatical categories, as do the verbal adpositions, for example *ú* 'to, for', while others, like the verbal pronouns, do so only in certain phonological environments.   We select here a few characteristic examples; further examples appear in the texts in Appendix II.

3.3.1 Verbal subject pronouns

The subject pronouns, which as described in 6.1.1 occur in first position in the verbal group, obligatorily coalesce with the sentence type marker *waa* (34 below) and with the focus words *bàa* and *ayàa* (35-6) when they precede the verbal group:

(34)   *Wàan raagay*
        wàa+aan   raagay
        DM+I   was:late
        'I was late.'

Other examples: **waa + aad* 'you' → *waad*, **waa + uu* 'he' → *wuu*, **waa + ay*'
she/they' →*way*, **waa + aannu* 'we (EXC)' → *waannu*, **waa + aynu* 'we (INC)' →
waynu, **waa + aydin* 'you (PL)' → *waydin*.

(35)   *Shàah bàan cabbay*
        shàah bàa+aan   cabbay
        tea   FOC+I   drank
        'I drank SOME TEA.' [22]

Other examples: **bàa + aad* 'you' → *bàad*, **bàa + uu* 'you' → *bùu*, etc.

(36)   *Ninkìi ayàan lá hadlay*
        nín-kìi   ayàa+aan   lá   hadlay
        man-the   FOC +I   with   spoke
        'I spoke with THE MAN.'

Other examples: **ayàa + aad* 'you' → *ayàad*, **ayàa + uu* 'you' → *ayùu*, etc.

There is some dialectal variation in the behaviour of these pronouns with the
negative word *má*. In northern dialects the pronouns cliticise onto the negative
word *má* with an epenthetic glottal stop at the juncture and do not coalesce: *má
+ aan* →*má'aan*, *má + aad* → *má'aad*, *má + uu* →*mú'uu*, etc. In central and
southern dialects of standard Somali, coalescence occurs, together with vowel
assimilation, for example: *má + aan* → *màan*, *má + aad* → *màad*, *má + uu* →
mùu etc. The two forms are:

(37)   *Má'uu/ Mùu helín*
        má+uu   helín
        not+he   found:NEG
        'He did not find it.'

In subordinate clauses, these pronouns are attracted out of the verbal group and
position after the initial constituent of the clause.   This may be a noun phrase
heading a relative clause (38), including one of the noun phrases of time, manner
etc. whose relative clauses act adverbially (39), or the phrase *wáxa* 'what, the
thing' introducing a cleft (40). If these elements end in a final *-a* vowel the
pronouns will always coalesce with them, as shown, with *wáxa* also undergoing
vowel assimilation:

(38)   *baabùurkuu wadó*
      baabùur+ka+uu wadó
      car-the-he   drives
      'the car which he drives'

(39)   *márkuu nimánka arkó*
      már-ka-uu   nimán-ka   arkó
      time-the-he men-the   see
      'when he sees the men' (lit. 'the time (in which) he sees the men')

(40)   *Cali wúxuu cabbay shàah*
      Cali   wáxa+uu cabbay shàah
      Ali   what+he drank   tea
      'What Ali drank was tea.'

The pronouns do not coalesce with words ending in consonants, though they may cliticise, for example *ín* 'that' + *uu* 'he' → *ínuu* 'that he', nor do they usually coalesce with the vowels *-ii*[23], *-ee*, *-o*, *-oo*, *-uu*: thus we find uncoalesced sequences like *dalkíi uu* 'the country which he'; *dugsiyo uu* 'schools which he'; *oo uu* 'and he'; *ee uu* 'and he'.

3.3.2 Focus words

The focus word *bàa*, which as described in chapter 9 focuses a preceding noun phrase, coalesces very freely with both preceding and following words. We have seen in the last section the fusion of *bàa* and subject verbal pronouns; *bàa* will also coalesce with a following object verbal pronoun when the latter begins with a vowel:

(41)   *Nínkíi bày arkáy*
      nín-kíi   bàa+i   arkáy
      man-the FOC+me saw
      'THE MAN saw me.'

Bàa also coalesces with a following negative word *áan* 'not':

(42)   *Nínkíi báan tegín.*[24]
      nín-kíi   bàa+áan   tegín
      man-the   FOC+not   went:NEG
      'THE MAN didn't go.'

Since this negative word itself coalesces with verbal subject pronouns, more complex coalesced forms regularly occur:

(43)   *Nínkii bàanan arkín*
        nín-kíi   bàa+áan+aan   arkín
        man-the   FOC+not+I   saw:NEG
        'I didn't see THE MAN.'
Other examples: **bàa+áan+aad → bàanad, *bàa+áan+uu → bàanu*, etc.

The focus word *ayàa* behaves in the same way as *bàa* as far as rightward coalescence is concerned, thus *ayàa+aan → ayàan, ayàa+áan+aan → ayàanan*, etc. However *ayàa* does not coalesce with preceding elements. Thus their behaviour contrasts after a preceding nominal ending in a vowel:

(44)   (a)   *Lacágtàan ku síinayaa*
               lacág-ta+bàa+aan   ku   síinayaa
               money-the+FOC+I   you   will:give
               'I will give you THE MONEY'
        (b)   *Lacágta ayàan ku síinayaa*
               lacág-ta   ayàa+aan   ku   síinayaa
               money-the   FOC+I   you   will:give
               'I will give you THE MONEY'

This difference in coalescence behaviour may be one of the reasons speakers choose between *bàa* and *ayàa*, which otherwise are identical in function. The use of *ayáa*, preserving as it does more of the uncoalesced structure, is often associated with more careful and slower speech.

These coalesced forms with *bàa* are very common. Question words for example are usually focused and thus many questions involve coalesced forms, for example:

(45)   *Sidùu ú helay?*
        si-dée+bàa+uu   ú   helay
        way-what?+FOC+he   in   found
        'HOW did he find it?' (lit. 'In what way did he find it?')
(46)   *Maxàad sheegtay?*
        maxáy+bàa+aad   sheegtay
        what+FOC+you   reported
        'WHAT did you report?' (a greeting)

3.3.3 *Verbal adpositions*

As described in 6.1.1, the verbal adpositions *ú* 'to, for', *kú* 'in, into', *ká* 'from, away from', and *lá* 'with (comitative)' occur preceding the verb in what we will

will call the verbal group. In this position they obligatorily coalesce with each other and with verbal pronouns in a complex but predictable manner. The pattern of two adpositions merging is in (47) with an example in (48):

(47)

ú + ú	→	*ugú*
ú + kú	→	*ugú*
ú + ká	→	*ugá*
ú + lá	→	*ulá*
kú + kú	→	*kagá*
kú + ká	→	*kagá*
ká + ká	→	*kagá*
kú + lá	→	*kulá*
ká + lá	→	*kalá*

(48)   *Fàarax sídan ugú samèe!*

Fàarax	*sí-tan*	*ú+ú*	*samèe*
Farah	way-this	in+for	do:IMP

'Do it (in) this way for Farah!'

These and all other coalescence rules with adpositions are obligatory. As can be seen, the coalesced forms only have a single HIGH tone, suggesting that the result of the rule is a single phonological word.

In a similar way we can show the pattern of coalescence between a single adposition and a single preceding verbal object pronoun in (49) and an example in (50):

(49) (a) Pronoun *i* 'me' + adposition

i	+	*ú*	→	*íi*
i	+	*kú*	→	*igú*
i	+	*ká*	→	*igá*
i	+	*lá*	→	*ilá*

(b) Pronoun *ku* 'you' + adposition

ku	+	*ú*	→	*kúu*
ku	+	*kú*	→	*kugú*
ku	+	*ká*	→	*káa*
ku	+	*lá*	→	*kulá*

(c) Pronoun *na* 'we (EXC)' + adposition

na	+	*ú*	→	*nóo*
na	+	*kú*	→	*nagú*
na	+	*ká*	→	*nagá*
na	+	*lá*	→	*nalá*

(d)   Pronoun *ina* 'we (INC)' + adposition

ina+	*ú*	→	*inóo*
ina+	*kú*	→	*inagú*
ina+	*ká*	→	*inagá*
ina+	*lá*	→	*inalá*

(e)   Pronoun *idin* 'you (PL)' + adposition

idin	+	*ú*	→	*idíin*
idin	+	*kú*	→	*idinkú*
idin	+	*ká*	→	*idinká*
idin	+	*lá*	→	*idinlá*

(50)   *Way káa qaadday*
       waa+ay   ku+ká   qaadday
       DM+she   you+from   took
       'She took it from you.'

These adpositions also coalesce with the impersonal subject pronoun *la* 'one, someone' which precedes them in the verbal cluster:

(51)   Pronoun *la* 'one' + adposition

la	+	*ú*	→	*lóo*
la	+	*kú*	→	*lagú*
la	+	*ká*	→	*lagá*
la	+	*lá*	→	*lalá*

(52)   *Xaggée bàa lagá keenay?*
       xág-kée   bàa   la+ká   keenay
       place-which?   FOC   one+from   brought
       'Where was it brought from?' (lit. 'From which place did one bring it?')

The examples thus far are of single combinations: more complex clusters regularly occur. We can show some of these in a table (53) below, which shows combinations of single object pronouns with the adposition clusters from (47) above; (54) is an example:

(53)

	ugú	*ugá*	*ulá*	*kagá*	*kulá*	*kalá*
i	*iigú*	*iigá*	*iilá*	*igagá*	*igulá*	*igalá*
ku	*kuugú*	*kaagá*	*kuulá*	*kaagá*	*kugulá*	*kaalá*
na	*noogú*	*noogá*	*noolá*	*nagagá*	*nagulá*	*nagalá*
ina	*inoogú*	*inoogá*	*inoolá*	*inagagá*	*inagulá*	*inagalá*
idin	*idiinkú*	*idiinká*	*idiinlá*	*idinkagá*	*idinkulá*	*idinkalá*

(54)   *Cáli lacágta noogá sóo qàad!*
      Cáli   lacág-ta   na+ú+ká   sóo   qàad
      Ali   money-the   us(EXC)+for+from   VEN   take:IMP
      'Get the money for us (EXC) from Ali!'

With the impersonal subject pronoun *la* 'one', the following coalesced forms appear, with an example in (56):

(55)

	ugú	*ugá*	*ulá*	*kagá*	*kulá*	*kalá*
la	*loogú*	*loogá*	*loolá*	*lagagá*	*lagulá*	*lagalá*

(56)   *Maxàa loogá qaadey?*
      Maxáy+bàa   la+ú+ká   qaaday
      what+FOC   one+for+from   took
      'Why was it taken from him?' (lit. 'What did one take it from him for?')

Once again these coalesced forms have a single High tone, suggesting they constitute a single phonological word.

3.4   Tonal accent rules

In chapter 2 we outlined the tonal system of Somali, outlining the following major features:

(58)   Features of tones
    a) There are two basic tones, High and Low.
    b) The tone-bearing unit is the mora.
    c) At most only one High tone may occur in each word root.
    d) The major function of tone is to mark grammatical rather than lexical information.

We also mentioned some features of the mapping between moras and syllables:

(59)   Moras and syllables
    a) Moras attach to vowels: a (C)V(C) syllable constitutes one mora and a (C)VV(C) syllable constitutes two moras, where VV represents a long vowel or diphthong.
    b) On long vowels and diphthongs, a sequence of High - Low tones is realised as a Falling tone.
    c) On long vowels and diphthongs, a sequence of Low - High is usually realised as High - High (occasionally a Rising tone is heard).

As we mentioned, one response to these characteristics, exemplified by Hyman (1981) and Banti (1988b), is to identify Somali as a tonal accent system and to suggest rules of tonal accent assignment. In this section we turn to the question of what these rules might look like, and also consider some rules of contextual realignment of tonal accent.

To aid our discussion we will identify the three commonest tone accent patterns (hereafter 'accentual patterns' (APs) for short) in word roots:

(60)   AP1: High tone on the last mora, Low elsewhere;
            AP2: High tone on the penultimate mora, Low elsewhere;
            AP3: Low tones on all moras.

3.4.1 Basic assignment of tonal accent

In this approach accentual patterns are assigned to words on the basis of their grammatical status. Thereafter rules of morphology and syntax alter the basic patterns. This means that the details of individual APs and changes on them are best described in the relevant grammatical sections of this book. We can however provide a few examples.

Thus for example the feminine nouns of declension 1 have AP1 in the singular: for example, *warqád* 'letter', and by the rule in (58c) above, *cayáar* 'game'. The masculine nouns of declension 2 on the other hand have AP2 in the singular: for example, *síxir* 'magic, sorcery', by rule (58b) above, *qiràal* 'confession'; and by rule (58c) above, *múudal* 'miser'. It is clear that it is the grammatical distinction between these declensions that is responsible for the different accentual patterns. Some roots, for example, may enter either declension: for example the unaccented underlying form **wiyil* 'rhinoceros' may surface either as a declension 1 noun *wiyíl* 'female rhinoceros' or a declension 2 noun *wíyil* 'male rhinoceros'. Similar pairs are quite common and include: *inán* 'girl', *ínan* 'boy; *matáan* 'female twin', *matàan* 'male twin'; *daméer* 'female donkey', *damèer* 'male donkey'.

In the verbal system individual paradigms are distinguished by a combination of inflection and accentual pattern. The accentual pattern alone does not normally distinguish forms but there are some minimal pairs, for example *kéen* (past independent 3.M.SG) 'He brought (it).'and *kèen* (imperative SG) 'Bring (it)!'. Other grammatical categories have their associated accentual patterns. To use examples we have already seen: adpositions have a characteristic pattern of AP1, while verbal pronouns have the pattern AP3. There are some minimal pairs distinguished by this, for example the object pronoun *ku* 'you' and the adposition *kú* 'in, into', the impersonal subject

pronoun *la* 'one, someone' and the adposition *lá* 'with'. Given the overall pattern of the Somali accentual system, it seems preferable to identify the accentual pattern of these forms with their grammatical category rather than as lexical information. This is supported by the fact that there seem to be no cases where words with the same grammatical features are distinguished by accentual pattern. We give the accentual patterns of individual grammatical categories as they are introduced, for example in chapters 4 and 5.

3.4.2 Accentual patterns and grammatical rules

As will become clear, many grammatical processes influence the basic accentual patterns of word roots. Rules of morphological derivation, for example, often have a tonal as well as segmental element and will produce a single output AP from various input APs. See for example the two following rules of nominal derivation which produce negative antonyms and abstract nouns, respectively, and which both have as their output feminine nouns with AP2:

(61)	*ayàan*	N M	[AP2]	'luck'	→	*ayaandárro*	N F	'bad luck'	
	díin	N F	[AP1]	'religion'	→	*diindárro*	N F	'lack of religion, irreligion'	*rro*
(62)	*carrúur*	N F	[AP2]	'child'	→	*carruurnímo*	N F	'being a child, childhood'	
	dóqon	N M	[AP1]	'fool'	→	*doqonnímo*	N F	'being a fool, foolishness'	*mo*

Similarly the attachment of the interrogative suffix, *-kée/-tée* 'which, what', lowers any High tones in the stem:

(63)	*gúri*	N M	[AP2]	'house'	→	*gurigée?*	'which house?'
	mìis	N M	[AP2]	'table'	→	*miiskée?*	'which table?'
	arrín	N F	[AP1]	'affair'	→	*arrintée ?*	'which affair?'
	náag	N F	[AP1]	'woman'	→	*naagtée?*	'which woman?'

Other suffixes have no effect on the stem AP except to append their own inherent accentual pattern to the new stem, for example the remote definite article *-kíi/-tíi* 'the':

(64)	*gúri*	N M	[AP2]	'house'	→	*gúrigíi*	'the house (remote)'
	mìis	N M	[AP2]	'table'	→	*mìiskíi*	'the table (remote)'
	arrín	N F	[AP1]	'affair'	→	*arríntíi*	'the affair (remote)'

> *náag* N F   [AP1] 'woman'   →   *náagtíi*   'the woman (remote)'

Note that this suffix thus creates a noun phrase with two High tones.

Another important role of AP changes is to mark case on nominals. These accentual changes are also accompanied by suffixation in some noun declensions but the marking is predominantly accentual. There is a basic unmarked case which we term the absolutive (ABS); this is used with nominals in a range of contexts: in isolation, as objects of verbs and adpositions, when focused, etc. The other cases are nominative, genitive and vocative, each of which can be conveniently viewed as modifications to the basic absolutive form. In order to describe case marking we need to add another accentual pattern:

(65)   AP4: High tone on the first mora, Low elsewhere.

Once again word syllable structures will produce various realisations of this AP and some neutralisations: for example, words with two short syllables will not distinguish AP1 and AP4, and of course short monosyllables may collapse AP1, AP2 and AP4.

Although there are exceptions (4.1.1.4), the typical pattern of case marking is for the inherent absolutive case AP of a noun to be changed in the other cases as follows:

(66)   Nominative (NOM)   Genitive (GEN)   Vocative (VOC)
      ABS→   AP3   AP1   AP4

For example:

(67)   ABS   NOM   GEN   VOC[25]
           Cáli   *Cali*   *Calí*   *Cáli*   NAM M   'Ali'
           Faadúmo   *Faadumo*   *Faadumó*   *Fàadumo*   NAM F   'Fatima'
           díbi   *dibi*   *dibí*   -   N M   'bull'
           bisád   *bisadi*   *bisád*   -   N F   'cat'
           carrúur   *carruuri*   *carrúur*   -   N F   'children'
           (←*carruúr*)   (←*carruúr*)

Accentual patterns are also sensitive to, and markers of, other kinds of syntactic information. To take one example, verbs whose subjects are focused are marked by a distinct pattern: they have a HIGH tone on their final mora. Compare for example the forms of *arkay* 'saw' in the following:

(68) (a)   *Áxmed bàan arkay.*
            Áxmed   bàa+aan   arkay
            Ahmed   FOC+I   saw
            'I saw AXMED.'
      (b)   *Áxmed bàa i arkáy.*
            Áxmed   bàa   i   arkáy
            Ahmed   FOC   me   saw
            'AXMED saw me.'

3.4.3 Contextual effects

We claimed in 3.5.1 that the accentual pattern of words is predictable from their grammatical status. This is largely true but there are a small number of words belonging to functional categories whose accentual pattern is dependent upon the surrounding context. The declarative sentence type marker *waa* and the interrogative sentence type marker *ma*, for example, exhibit what we might call a seesaw accentual pattern. If the following word contains a High tone mora, they are Low tone; however if the following word is all Low tone then they have High tone on their first mora, realised as *wàa* and *má*:

(69) (a)   *Wàan diray.*
            Wàa+aan   diray
            DM+I   sent:1SG
            'I sent it.'
      (b)   *Waan dirayaa.*
            Waa+aan   dirayaa
            DM+I   am:sending:1SG
            'I am sending it.'
(70) (a)   *Má dirtay?*
            má   dirtay
            QM   sent:2SG
            'Did you send it?'
      (b)   *Ma diraysaa?*
            ma   diraysaa
            QM   are:sending:2SG
            'Are you sending it?'

Interestingly however, if a following High tone occurs before a pause and is thus lowered to a phonetic Mid tone, then *waa* will have High tone. See for example *waa*'s sensitivity to the presence of a pause after the following High tone:

(71)   (a)   *Wàa   kán.*
             DM   this:one
             'It is this one'
       (b)   *Waa kán ee qaadó.*
             waa   kán   ee   qaadó
             DM   this:one   and   take:IMP:SG
             'It is this one, take it!'

This suggests that the accentual seesaw is sensitive to phonetic rather than phonological tone.

Another example of contextual variation is exhibited by the verb phrase conjunction *oo* 'and'. Before a following word this has High tone while before a pause this becomes Falling tone:

(72)   (a)   *Wùu kacay óo baxay.*
             wáa+uu   kacay óo   baxay
             DM+he   rose   and   left
             'He got up and left.'
       (b)   *Wùu kacay òo, baxay.*
             wáa+uu   kacay òo   baxay
             DM+he   rose   and   left
             'He got up and, [*pause*] left.'

We may treat the accentual pattern on *oo* in (72b) as a form of discourse marker, where the speaker uses a marker of intended continuation to defend against interruption. In other contexts, for example fast speech, *oo* may be de-accented and cliticized onto a preceding element. This often occurs when *oo* marks an appositive relative clause on an independent pronoun (8.2):

(73)   *iságoo   hurdá*
       he+and   sleeps
       'he, who sleeps', 'while he is/was sleeping'

In addition to these variable elements there are a range of contextual accentual changes which remain to be investigated. For example, most plural nouns, and the singular nouns in declension 8, have in their basic absolutive case High tone on their final mora which lowers to a Low tone in certain contexts. This rule, which we can call Accent Lowering, is also associated with a deletion of stress. The contexts form a rather heterogeneous phonetic and syntactic class: lowering does not take place when the noun is before a pause (including isolation), when it is the head of dependent elements like determiners, adjectives

or relative clauses, or when focused by the focus word *ayàa*. The lowering does occur when the bare unfocused noun occurs before the verbal group, or as noted by Banti (1988b), before the focus word *bàa* and the clitic *bá* 'any':

(74)   Non-lowering contexts
      (a)   *nimán*   'men'
      (b)   *nimánka*   'the men (non-remote)'
      (c)   *nimánkíi*   'the men (remote)'
      (d)   *nimán yaryár*   'small men'
      (e)   *Nimán   ayàa   yimíd*
           men   FOC   came
           'SOME MEN came.'

(75)   Lowering contexts
      (a)   *Niman   bàa   yimíd.*
           men   FOC   came
           'SOME MEN came.'
      (b)   *Nimanbá má arag.*
           niman-bá má   arág
           men-any not   saw:NEG
           'I didn't see any men.'
      (c)   *Niman   má   aragtay?*
           men   QM   saw:2SG
           'Did you see some men?'

Banti (1988b), citing evidence for a similar lowering rule in nouns in another Cushitic language Arbore, spoken in Ethiopia, suggests that this rule is a relic of an earlier common Cushitic rule.

    A further accentual rule, which we can term Accent Shift, affects singular nouns of declensions 6 and 7 and shifts their absolutive case pattern of AP1 to AP2 before a pause, for example:

(76)   *hooyó* F 'mother'   →   *hóoyo*
      baré M 'teacher'   →   *báre*

This rule also affects some collective nouns, e.g. *socotó* F 'travellers' → *socóto*. Since isolation is a subset of the pre-pause position, these AP2 forms will occur as elicited forms.

    In our present state of knowledge rules like Accent Lowering and Accent Shift can be described but not given any more general explanation. As we have said, the full range of processes that alter accentual patterns in various phonological and grammatical contexts remains to be investigated.

3.5   Reduplication

Reduplication processes mark a number of grammatical and semantic distinctions, all sharing an iconic meaning of 'increase'. There are two main types: suffixing and prefixing reduplication. Suffixing reduplication applies to mark the plurals of one declension (D4) of masculine nouns; these are quite numerous and we present a few representative examples:

(77)   Singular   Meaning   Plural
             jíd   'road'   *jidád*
             jùr   'rat'   *jiirár*
             wéd   'death'   *wedád*
             bèer   'liver'   *beerár*
             qál   'jaw'   *qalál*
             àas   'burial'   *aasás*
             bóg   'page'   *bogág*
             ròob   'rain'   *roobáb*
             tús   'bookmark'   *tusás*
             sùun   'belt'   *suumán*
             dèyr   'fence'   *deyrár*
             dàwr   'role'   *dawrár*
             càysh   'food'   *cayshásh*

The nouns in (77) form their plurals by appending a final syllable consisting of the short vowel *a* followed by a copy of the stem final consonant, i.e. by a -*aC* reduplicating suffix. Since all nouns in this declension end in a consonant, the only complication in these examples is that the reduplicated suffix allows the underlying nasal *m* in *sùun* (←*sùum*) 'belt' to resurface since it is no longer in word final position. If as discussed in 3.1.1, the correct generalisation is that *m* does not occur syllable finally, this suggests the syllabification of *sùumán* 'belts' is *sùu$mán*.

        There are several prefixing reduplication rules which operate in adjectives and verbs, as shown in (78) - (81) below. Firstly, prefixing reduplication marks the plural forms of an important subset of adjectives, given in (78), with an example in (79):

(78)   Singular   Meaning   Plural
             adág   'hard, difficult'   *adadág*
             ballaarán   'wide, broad'   *balballaarán*
             buurán   'fat'   *buurbuurán*
             cád   'white'   *cadcád*

cás	'red'	*cascás*
culús	'heavy'	*culculús*
cusúb	'new'	*cuscusúb*
dhèer	'tall'	*dhaadhèer* (alt.: *dheerdhèer*)
dhuubán	'thin'	*dhuudhuubán*
fiicán	'good, fine'	*fiicfiicán*
fudúd	'easy, light'	*fudfudúd*
furán	'open'	*furfurán*
gaabán	'short'	*gaagaabán* (alt.: *gaaggaabán*)
gubán	'burnt'	*guggubán*
kulúl	'hot'	*kulkulúl*
macáan	'sweet'	*macmacaan*
madów	'black'	*madmadow*
qaawán	'naked'	*qaaqaawán*
qaybsán	'divided'	*qaybqaybsán*
wèyn	'big'	*waawèyn*
wanaagsán	'good'	*wanwanaagsán*
xardhán	'decorated'	*xarxardhán*
xún	'bad, evil'	*xunxún*
yár	'small'	*yaryár*

(79)   *nínka wanaagsán* 'the good man'   *nimánka wanwanaagsán* 'the good men'

A second function of prefixing reduplication is to mark intensive or iterative forms of verbs and adjectives, as shown below:

(80)   Intensive/iterative reduplicated verbs

bòod	'jump'	*boodbòod*	'jump up and down'
bùus	'dent'	*buusbùus*	'crush'
dalòol	'be pierced'	*daldalòol*	'be full of holes'
dhàaf	'pass'	*dhaafdhàaf*	'pass back and forth near'
dhaqàaq	'move'	*dhaqdhaqàaq*	'move restlessly about, move back and forth'
dír	'send'	*dirdír*	'send repeatedly'
dùub	'fold, roll'	*duuddùub*	'wrap up, coil'
fúr	'open'	*furfúr*	'separate one thing from a bundle; dismantle'
gó'	'cut'	*googgó'*	'cut to pieces'
jáb	'break'	*jajáb*	'break into pieces'
jár	'cut off, slice'	*jarjár*	'mince'
káb	'mend, patch'	*kabkáb*	'piece together, do patchwork'
lìic	'lean, tilt'	*liiclìic*	'swing, sway, rock'
rìix	'push'	*riixrìix*	'push around, move around'

xídh	'tie'	*xidhxídh*	'tie up, pack up, bind'
xóq	'scratch'	*xoqxóq*	'keep on scratching'

(81) Intensive/iterative reduplicated adjectives
jabán	'broken'	*jajabán*	'broken in pieces'
jarán	'chopped'	*jarjarán*	'chopped up'
jeexán	'cracked'	*jeexjeexán*	'cracked all over, split in pieces'
qaroofán	'bent'	*qarqaroofán*	'twisted, bent (around)'
xidhán	'tied'	*xidhxidhán*	'fastened repeatedly, tied in several places'

As noted by Banti (1988a) prefixing reduplication consists of three different types:

(I) *Caa-* prefixation, where the first stem consonant is copied, followed by a long *aa* vowel. This process is the most marked form of prefix reduplication and is restricted to a small number of adjectives, perhaps two; it is exemplified in (78) by *dhaadhèer* and *waawèyn*. Banti (1988a) argues that this is an ancient form of reduplication within Cushitic.

(II) CVC- prefixation, where the copied section of the stem is the first consonant, first vowel and second consonant. The prefix vowel must match the first stem vowel in length. This rule is exemplified by, amongst others, the adjectives *balballaarán*, *buurbuurán*, *qaybqaybsán* in (78) and the verbs *jarjár* and *boodbòod* in (80). The stem of the adjective *adág* begins with a vowel and in this case the reduplication process interestingly blurs the distinction between the phonological and phonetic status of the glottal stop. As noted earlier, word initial vowels normally trigger an epenthetic glottal stop, thus giving the pronunciation *'adág*. In the reduplication of *adág*, the first vowel and consonant are copied but the glottal stop is retained internally by the stem, i.e. *ad'adág*, which of course is normally also pronounced with an epenthetic glottal stop, i.e. *'ad'adág*.

(III) CV- prefixation, where the reduplicated prefix is a copy of the stem's first consonant and vowel. Again the prefix vowel must match the first stem vowel in length. In this rule the initial consonant of the stem is also doubled where this is allowed by the syllable structure constraints. Examples of this process include the adjectives *dhuudhuubán*, and *guggubán* in (78) and the verbs *duuddùub* and *jajáb* in (80).

We can see from our examples that one or two adjectives allow type (II) or (III), for example the more usual *dhaadhèer* or its variant *dheerdhèer*. Moreover there are some forms which need further phonological explanation such as *gó'* → *googgó'* in (80), which is rather like type (III) but in which the prefix vowel does not match the stem vowel in length.

.

Chapter 4
Grammatical Categories

In this chapter we outline the main grammatical categories or word classes of Somali. We concentrate here on the morphology of the basic members of each category, leaving those members derived by productive rules of word formation to chapter 5 and discussion of their phrasal syntax to chapters 6 and 7. Somali is characterised by a high degree of inflectional complexity, especially in the verbal system, and while sometimes it is possible to capture generalisations with rules, often these rules would be so complicated that it is more effective to present the forms in a paradigmatic display.

We identify ten basic categories, with the following abbreviations:

(1)   Noun (N)
        Clitic pronoun (CPRO)
        Verb (V)
        Adjective (ADJ)
        Verbal adposition (ADP)
        Determiner (DET)
        Focus word (FOC)
        Sentence type marker (STM)
        Conjunction (CONJ)
        Adverb (ADV)

Three of these categories are open, that is new members can freely be added: noun, verb and adjective. Six others are closed in that they form systems that cannot easily be extended: clitic pronoun, determiner, focus word, sentence type marker and conjunction. The category of adverb is somewhat marginal, as we shall see, and falls somewhere between an open and closed category.

Most of the categories have subcategories: nouns are divided into such subclasses as common nouns, names, independent pronouns, etc.; verbs into

main and auxiliary verbs, and so on. There are also two groups of rather idiosyncratic elements which we will not list as major categories: ideophones, discussed in 4.11, and the interjections, attention-gainers, etc. described in 4.12.

4.1   Nouns

We can identify four main types of noun: common nouns (CN), names or proper nouns (NAM), independent pronouns (IPRO), and numerals (NUM). The defining grammatical characteristics of nouns are that they are marked for case and that they occur as subjects of verbs, governing agreement features on the verb. They may act as the heads of noun phrases, which share these characteristics, as discussed in chapter 6. These main types or subcategories also have characteristics which distinguish between them, as we shall see; for example, in the modifiers that they allow. Common nouns and numerals, for example, occur with determiners and restrictive relative clauses, while proper nouns and independent pronouns normally do not. We begin our description with common nouns.

4.1.1 Common nouns

Common nouns typically carry information from the grammatical systems of gender, number and case, which are marked by the morphological processes of accentual change, suffixation and reduplication.

4.1.1.1   Gender

There are two genders: all nouns are either masculine or feminine. For the most part gender is not predictable from the meaning of nouns. The exceptions include nouns for people and animals: *nín* 'man' is masculine and *náag* 'woman' is feminine, for example. Even here though there are arbitrary cases: the collective noun *hawéen* 'women' is masculine.

   This arbitrariness of gender is emphasised by the fact that Somali exhibits the gender polarity found in other Cushitic languages. Many nouns reverse their gender in the plural, as shown in the examples below, where the polarity extends over different types of plural formation by suffixes and accentual change:

(2)   Masculine → Feminine

sánnad	M	'year'	*sannaddó*	F	'years'
abtí	M	'maternal uncle'	*abtiyó*	F	'maternal uncles'
túke	M	'crow'	*tukayáal*	F	'crows'
barbàar	M	'bachelor'	*barbáar*	F	'bachelors'

àwr	M	'male camel'	*áwr*	F	'male camels'

(3)   Feminine → Masculine

galáb	F	'afternoon'	*galbó*	M	'afternoons'
káb	F	'shoe'	*kabó*	M	'shoes'
adéeryo	F	'female kudu'	*adeeryóoyin*	M	'female kudus'
hóoyo	F	'mother'	*hooyóoyin*	M	'mothers'

Though this is the dominant pattern, not all plurals involve gender polarity: masculine nouns which mark their plural by reduplication and a subgroup of masculine suffixing nouns remain masculine in the plural:

(4)   Masculine → Masculine

wán	M	'ram'	*wanán*	M	'rams'
bèer	M	'liver'	*beerár*	M	'livers'
dhéri	M	'clay pot'	*dheryó*	M	'clay pots'
wáran	M	'spear'	*warmó*	M	'spears'

Gender is reflected grammatically in three important forms of agreement: agreement with the set of suffixed determiners, subject-verb agreement, and agreement with clitic subject pronouns. We can show these first with singular nouns.

The determiners, discussed in 4.6 below, include definite articles, demonstratives, possessives, etc. and divide into two formal agreement classes: a masculine class which has basic forms with initial *k-* and a feminine class with initial *t-*. We saw in chapter 3 that these forms undergo complex rules of sandhi when attached to the noun but simple examples with the remote definite article *-kíi/-tíi* are:

(5) *baabùur* N M 'truck' + *kíi* → *baabùurkíi* 'the truck (remote)'
     náag N F 'woman' + *tíi* → *náagtíi* 'the woman (remote)'

Subject - verb agreement is marked on the verb by a set of agreement affixes. In the majority class of suffix verbs these are attached after the stem composed of the verb root and any derivational affixes, while in the minority class of prefix or strong verbs, they attach as prefixes to the root. For example in prefix verbs a distinction is made between the affixes *y-* 3SG M and *t-* 3SG F:

(6)   *Baabùurkíi wùu yimi.*
      baabùur-kíi   *wàa+uu*   *yimi*
      truck:M-the   DM+it:M   came:3SG M
      'The truck came.'

(7)   *Náagtíi wày timi.*
         náag-tíi   wàa+ay timi
         woman:F-the   DM+she 3SG F:came
         'The woman came.'

In these examples we can also see agreement with clitic subject pronouns: when
the subject is unfocused these agreeing clitics occur in the verbal group. In (6)
and (7) the agreeing clitics are cliticized onto the sentence type marker *waa* and
are *uu* 3SG M and *ay* 3SG F. They and the agreement affixes identify the gender
of the subject noun.

    In the plural, agreement with verbal affixes and the clitic subject pronouns
does not distinguish gender since both affixes and clitic pronouns have a
common 3PL form, but agreement with determiners still marks the gender
distinction:

(8)   *Baabuurrádii wày yimaadeen.*
         baabuurró-tíi   wàa+ay   yimaadeen
         trucks-the   DM+they   came:3PL
         'The trucks came.'
(9)   *Naagíhii wày yimaadeen.*
         naagó-kíi   wàa+ay   yimaadeen
         women-the   DM+they   came:3PL
         'The women came.'

In (8) the feminine plural noun *baabuurró* 'trucks' takes the feminine determiner
-*tíi*, while in (9) the masculine plural noun *naagó* 'women' takes the masculine
determiner -*kíi*.

4.1.1.2   Number
As with other Cushitic languages, the number system in Somali nouns is quite
complicated. Firstly there are a number of categories of number and secondly
the morphological marking of plural, where applicable, has a number of
exponents. We can begin by dividing common nouns into four number types:
countable, mass, collective and transnumeral.

    Countable nouns form the majority. When used unmodified they refer to
an individual entity, for example: *kòob* 'a cup', *maalín* 'a day'. They can be
counted directly by being used after a numeral, for example *labá kóob* 'two
cups', *áfar maalmóod* 'four days'. As described in 4.1.4., these are genitive
constructions where the numeral is the head and the counted noun occurs in the
genitive case, in some instances in the singular and in others in the plural, as we

shall see. Countable nouns otherwise regularly occur in plural forms, e.g. *koobáb* 'cups', *maalmó* 'days'.

Mass nouns form a smaller group which do not refer to individuals but name a substance or material, for example: *biyó* 'water', *búr* 'flour', *caanó* 'milk', *salíid* 'oil', *shàah* 'tea', *sonkór* 'sugar', *súbag* 'clarified butter', *timó* 'hair'. These nouns can only be counted indirectly by using a relevant unit of measurement, for example:

(10) *labá kóob oo shàah áh*

     labá   *kóob*   *oo*   *shàah*   *áh*

     two-ABS   cup:GEN   and   tea:ABS   be

     'Two cups of tea' (lit. 'Two of cup and which are tea')

Mass nouns do not normally display a singular/plural distinction but the class is divided into nouns which always occur with singular verbal agreement and those with plural verbal agreement; for example, the two masculine nouns *shàah* 'tea' and *caanó* 'milk', respectively:

(11) Mass noun, singular agreement

     Shàahii wùu qubtay.

     shàah-kii   *wàa+uu*   *qubtay*

     tea-the   DM+CPRO:3SGM   spilled:3SG M

     'The tea spilled.'

(12) Mass noun, plural agreement

     Caaníhii wày qubteen.

     caanó-kii   *wàa+ay*   *qubteen*

     milk-the   DM+CPRO:3PL   spilled:3PL

     'The milk spilled.'

Since the mass nouns which take plural agreement all end in final *-ó* (*caanó*, *biyó*, etc.), and since this is, as we shall see, a common plural suffix, we can suggest a possible explanation for this agreement pattern. These represent plural forms historically which no longer have a singular form, i.e. what are sometimes called *plurale tantum* forms.

Collective nouns are another restricted group of nouns, which are used to refer to sets of individuals. They can identify either the usually experienced set, for example *gèel* 'herd of camels' or the universal set, e.g. *rág* 'men, mankind'. If the context allows either interpretation is possible, so that *gèel* can mean ' a herd of camels' or be used generically for 'camels'. Other examples include: *ádhi* '(herd of) goats and sheep, *carrúur* 'children', *dád* 'people', *dúmar* 'women, womankind', *habló* 'girls', *idó* '(flock of) sheep', *ilmó* 'children', *ló'*

'cattle', *xooló* 'livestock'. As with mass nouns, collectives cannot be counted directly, but can be counted indirectly using a relative clause on a head numeral, for example:

(13)   *Sáddex   carrúur   áh*
        three   children:ABS   are
        'three children' (lit.: 'three who are children')

Collectives do not normally show a singular/plural distinction but as with mass nouns, forms ending in -*ó* always take plural agreement:

(14)   Collective noun, plural agreement
        Hablíhii wày yimaadeen.
        habló-kii   wàa+ay   yimaadeen
        girls-the   DM+ CPRO:3PL   came:3PL
        'The girls came'

Speakers do sometimes pluralize collectives when the set referred to is smaller than the universal set. Thus for example, one can hear *geelál* 'herds of camels' from *gèel* '(herd of) camels'.

Transnumeral nouns, to use the term introduced by Serzisko (1992), are a small group of nouns which can be used with the semantic force of any of the categories described so far. Thus *kallùun* 'fish' can be used to refer to an individual fish, or the substance, or a collective, so that the nominal *kallùunkán* (←*kallùun* - *kán* 'this') can be used for 'this (one) fish', 'this fish meat', or 'this group of fish'. Other examples include: *bèed* 'egg', *cámbe* 'mango', *liin* 'orange, lemon', *mùus* 'banana', *sigàar* 'cigarette(s)', *ukún* 'egg'. Transnumeral nouns cannot be counted directly by numerals; they require the use of a relative clause with as its head a numeral or a unit of measure:

(15)   *xabbád liin   áh*
        unit   orange be
        'one orange' (lit.: 'a unit that is orange')
(16)   *labá   liin   áh*
        two   orange be
        'two oranges' (lit.: 'two which are orange(s)')

As we have noted, of these types of common noun, only countable nouns regularly form plurals. Plurals are formed by a range of morphological operations which we describe in the next section when we discuss declensions of

countable nouns. We postpone discussion of case to section 4.1.1.4, since the noun declensions are helpful in discussing case marking.

4.1.1.3   Noun declensions

There have been a number of declension systems proposed for Somali nouns. Some approaches, as in Hyman (1981) and Banti (1988b), are predominantly phonological in orientation and classify nouns according to accentual patterns. Others, like Andrzejewski (1964, 1979), Caney (1984) and Saeed (1993a), have established declensions on combinations of phonological and grammatical criteria. Here we follow Saeed (1993a) with one or two modifications.[26] This system classifies countable nouns according to their gender, how they form plurals and their resulting accentual patterns. In this discussion and our account of case in the next section, we will make use of the major accentual patterns introduced in chapter 3:

(17)   AP1: HIGH tone on the last mora, LOW elsewhere;
             AP2: HIGH tone on the penultimate mora, LOW elsewhere;
             AP3: LOW tones on all moras;
             AP4: HIGH on the first mora, LOW elsewhere.

The seven major declensions are identified as follows.

   Declension 1 nouns
1. Exhibit polarity: are feminine in the singular, masculine in the plural.
2. Form the plural by suffixing -*o*, or -*yo* if the singular ends in *i*.
3. Have AP1 in both the singular and plural.
4. Undergo Accent Lowering (3.4.3) in the plural.
5. May provide in the plural the environment for Stem Contraction (3.1.4). Some example are:

aayád	'miracle'	*aayadó*	'miracles'
cabsí	'fear'	*cabsiyó*	'fears'
gacán	'hand'	*gacmó*	'hands'
xudúud	'border'	*xuduudó*	'borders'

   Declension 2 nouns
1. Exhibit polarity: are masculine in the singular, feminine in the plural.
2. Form the plural by suffixing -*o*; if the singular ends in *i, x, c, j, q* (and sometimes *s* and *g*), the suffix is -*yo* ; if it ends in *b, d, dh, l, r, n* the final consonant will double.

3. Have AP1 in the plural.
4. Undergo Accent Lowering (3.4.3) in the plural.
5. In the singular are polysyllables not ending in *e*.

This declension can be divided into two sub-declensions on the basis of the accentual pattern of the singular. The vast majority belong to declension 2A and have AP2 in the singular. Some examples are:

áabbur	'plug'	*aabburró*	'plugs'
abwàan	'scholar'	*abwaannó*	'scholars'
bábbis	'fan'	*babbisyó*	'fans'
kúray	'lad'	*kurayó*	'lads'
ólol	'flame'	*ololló*	'flames'
salìib	'crucifix'	*saliibyó*	'crucifixes'

Declension 2B nouns form a small group, probably not more than fifty words in the language. They have AP1 in the singular and unusually, undergo Accent Lowering (3.4.3) in the singular as well as the plural. Some examples are:

abtí	'maternal uncle'	*abtiyó*	'maternal uncles'
durbáan	'drum'	*durbaannó*	'drums'
gorgór	'vulture'	*gorgorró*	'vultures'
sambáb	'lung'	*sambabbó*	'lungs'
suldáan	'sultan'	*suldaannó*	'sultans'

Declension 3 nouns

1. Do not exhibit polarity: are masculine in both the singular and plural.
2. Form the plural by suffixing -*o*; if the singular ends in *i*, the suffix is –*yo*.
3. Have AP2 in the singular and AP1 in the plural.
4. Undergo Accent Lowering (3.4.3) in the plural.
5. Since many roots consist of two short syllables, the plural often involves Stem Contraction (3.1.4).
    Some examples are:

béri	'day'	*beryó*	'days'
édeg	'lamb pen'	*edgó*	'lamb pens'
ílig	'tooth'	*ilkó*	'teeth'
nàas	'breast'	*naasó*	'breasts'
máraq	'udder cord'	*marqó*	'udder cords'

Declension 4 nouns

1. Do not exhibit polarity: are masculine in both the singular and plural.
2. Form the plural by a reduplicating *-a*C suffix described in 3.5.
3. Have AP2 in the singular and AP1 in the plural.
4. Undergo Accent Lowering (3.4.3) in the plural.
5. In the singular are monosyllables.
    Some examples are:

áf	'mouth, language'	*afáf*	'mouths, languages'
dáb	'fire'	*dabáb*	'fires'
nín	'man'	*nimán*	'men'
sán	'nose'	*sanán*	'noses'

Declension 5 nouns

1. Exhibit polarity: are masculine in the singular, feminine in the plural.
2. Form the plural by a change of accentual pattern from AP2 to AP1.
    Some examples are:

àwr	'male camel'	*áwr*	'male camels'
Cárab	'an Arab'	*Caráb*	'Arabs'
díbi	'bull'	*dibí*	'bulls'
Soomáali	'Somali'	*Soomaalí*	'Somalis'
yèy	'wolf'	*yéy*	'wolves'

These nouns have an idiosyncratic agreement pattern. In the plural they may take plural or feminine singular agreement; thus both (18a) and (b) below are possible:

(18)   (a)   *Carábtii wày yimaadeen.*
               Caráb-tii   *wàa+ay*   *yimaadeen*
               Arabs-the   DM+CPRO: 3PL   came:3PL
               'The Arabs came.'
        (b)   *Carábtii wày timi.*
               Caráb-tii   *wàa+ay*   *timi*
               Arabs-the   DM+CPRO:3SG F   came:2SG F
               'The Arabs came.'

Puglielli and Ciise (1984: 81-5) suggest that this singular agreement allows a form of collective interpretation with nouns in this declension.

Declension 6 nouns

1. Exhibit polarity: are feminine in the singular, masculine in the plural.
2. Form the plural by suffixing *–oyin*.
3. Have variable accentual pattern in the singular. They undergo Accent Shift (3.4.3): AP1 becomes AP2 before a pause. The plural has AP2.
4. In the singular end in *o*.
   Some examples are:

ayeeyó	'grandmother'	*ayeeyóoyin*	'grandmothers'
dawó	'medicine'	*dawóoyin*	'medicines'
hooyó	'mother'	*hooyóoyin*	'mothers'
magaaló	'town'	*magaalóoyin*	'towns'

Declension 7 nouns

1. Exhibit polarity: are masculine in the singular, feminine in the plural.
2. Form the plural by suffixing *–yaal*.
3. Have variable accentual pattern in the singular. They undergo Accent Shift (3.4.3): AP1 becomes AP2 before a pause. The plural has AP1.
4. In the singular end in *e*, which usually assimilates to *a* before the plural suffix.
5. Include many derived nouns. One important class consists of compound nouns formed from N - V or N - ADJ combinations, e.g. *xoghayé* 'secretary' from *xóg* NF 'confidence' and *hay* V 'hold'; *madaxweyné* 'president' from *mádax* NM 'head, leader' and *wèyn* ADJ 'big'. Another consists of nouns formed from simple verbs by an agentive suffix -*e*, e.g. *baré* 'teacher' from *bár* V 'teach'.
   Some examples are:

aabbé	'father'	*aabbayáal*	'fathers'
agaasimé	'director'	*agaasimayáal*	'directors'
danjiré	'ambassador'	*danjirayáal*	'ambassadors'
tuké	'crow'	*tukayáal*	'crows'
xakamé	'bridle'	*xakamayáal*	'bridles'

In addition to these seven major declensions there are a small number of nouns which have idiosyncratic ways of forming plurals. Some loans from Arabic, for example, maintain for some speakers their Arabic form of plurals, e.g. *kúrsi* M 'chair', *kuraasí* F 'chairs'; *márkab* M 'ship', *maraakíib* F 'ships'. It is interesting though that these exhibit gender polarity. Some speakers however 'somalicize' these nouns further by placing them in declension 2A, allowing

kursiyó 'chairs'. Another small group form their plural by suffixing *-aan* or *-an* and changing AP2 to AP1:

(19)	*úgax*	F	'egg'	*ugxáan*		F	'eggs'
	dhágax	M	'stone'	*dhagxáan* (*dhagxán*)	F		'stones'
	qálin	M	'pen'	*qalmáan*		F	'pens'

Again these forms are under pressure from more regular plural formation rules and *qálin* can be treated as a declension 2 noun and pluralized as *qalimmó*. A third group, perhaps only two words but in very common use, are masculine non-polar nouns with AP1 in both singular and plural: *réer* 'family', *reeró* 'families' and *waláal* 'brother', *walaaló* 'brothers'. These nouns undergo Accent Lowering in both singular and plural.

The basic system of seven declensions is affected by regional variation in two ways that we cannot adequately explore here. The first is that for some nouns there is variation within Standard Somali about the declension they belong to. For example the noun *tùug* M 'thief' is in some varieties treated as a declension 4 noun and given a plural *tuugág* M 'thieves', in others as declension 5, giving *túug* F 'thieves', while in still others it is a declension 2 noun, giving *tuuggó* F 'thieves'. The second is that other dialects have quite different distributions of plural forms, including some suffixes not found in Standard Somali; see Banti (1985) and Lamberti (1986) for some details of these.

4.1.1.4   Case

As described briefly in chapter 3, the marking of case on nouns is largely accentual, with some use of suffixes. The basic unmarked case we termed the absolute (ABS). This is the case nouns show when in isolation, as objects of verbs and adpositions, or when focused. It is convenient to treat it as the base form because it is the most basic form morphologically and also because its accentual patterns are less predictable than the other cases. There is a general tendency for feminine singular nouns to have AP1 in the absolute, while masculine singular have AP2 and all plurals, AP1. However there are exceptions: masculine nouns in declensions 2B and 7 have AP1 in the singular while declension 6 nouns have AP2 in the plural. The other cases are more predictable and can best be viewed as modifications to the underlying absolute form. In chapter 3 we gave the following typical accentual case marking pattern:

| (20) | Nominative (NOM) | Genitive (GEN) | Vocative (VOC) |
| ABS → | AP3 | AP1 | AP4 |

All nouns follow this pattern for the genitive case and for vocatives marked by accentual change alone. We discuss other vocatives below. Most nouns follow this pattern for nominative case too except for non-plural feminine nouns ending in *-o* and masculine nouns ending in *-e* which have AP2 in the nominative, for example *hooyó* F ABS 'mother' → *hóoyo* NOM; *tuké* M ABS 'crow' → *túke* NOM.

In addition to these accentual changes the case system involves suffixes. In the nominative and genitive cases suffixes attach to phonologically specified classes of nouns. In the nominative, feminine nouns ending in a consonant have a suffix *-i*, for example *náag* F ABS 'woman' → *naagi* NOM. In the genitive case, feminine nouns not ending in *-o* add a suffix *-eed*, (or *-yeed* if the noun ends in *i*), for example:

(21)   *áf shimbiréed*
         áf   *shimbiréed*
         mouth:M:SG:ABS   bird:F:SG:GEN
         'mouth of bird, a beak'

A few nouns for domestic animals, e.g. *ló'* F 'cattle', *daméer* F 'female donkey', *hál* 'female camel' take a suffix *-aad*, e.g.

(22)   *caanó lo'áad*
         caanó   *lo'áad*
         milk:M:COLL:ABS   cattle:F:SG:GEN
         'cow's milk'

Polar masculine plurals ending in *o*, formed from feminine singular nouns, have a genitive suffix *-od*, e.g.

(23)   *hádal naagóod*
         hádal   *naagóod*
         talk:M:SG:ABS   women:M:PL:GEN
         'women's talk'

Names and nouns used as forms of direct address occur in the vocative case. There are two basic forms: vocatives by accentual change alone, as we have already seen, and those formed with suffixes. Cross-cutting this formal distinction there are two functional classes: the first is used for people's names, nouns of kinship and friendship, for the names of occupations when used respectfully, and for the names of God. This class reflects degrees of solidarity and/or respect. The second class is more neutral and tends to be used with common nouns. We can call the first class name vocatives and the second noun

vocatives, and show the relationship between the formal and functional classes as follows:

(24)	AP change	Suffix
Name	✓	✓
Noun	-	✓

Turning to suffixed forms first, both name and noun vocatives have two forms. One is used with feminine singular nouns, including feminine collectives and mass nouns which take singular concord, and the other with masculine singular nouns and plurals:

(25)   Vocative suffixes

	Name	Noun
F SG	*-èey/àay/òoy*	*-yahay*
M SG & PL	*-òw*	*-yohow*

The choice between F SG *-èey/àay/òoy* is phonologically conditioned: nouns ending in *a* take *-àay*; those ending in *o* take *-òoy*; and all others take *-èey*. These suffixes also occur in short vowel variants: *-èy/ày/òy*. When name vocative suffixes are added to stems ending in the short vowels *e*, *a* or *o* the vowel is deleted, losing any associated tone. Stems ending in *i* will usually have an epenthetic consonant *y* inserted. All vocative suffixes append their own AP without altering the stem's original AP.[27] See below for some examples of these phonological effects:

(26)   Suffixed name vocatives

Cánabèey	F SG	'Anab!', 'O Anab!'	(cf. *Cánab*, a woman's name)
hooyòoy	F SG	'Mother!', 'O Mother!'	(cf. *hooyó* 'mother')
Cáliyòw	M SG	'Ali!', 'O Ali!'	(cf. *Cáli* 'Ali')
Xásanòw	M SG	'Hasan!' 'O Hasan!'	

(27)   Suffixed noun vocatives

carrúuryahay	F COLL	'children!', 'O children!'
hályahay	F SG	'female camel!', 'O female camel!'
ninyohow	M SG	'man!', 'O man!'
gabdháyohow	M PL	'girls!', 'O girls!'

Name vocatives can also be formed by accentual change alone as shown in (20) above, ie. ABS → AP4:

(28)   Vocatives by accentual change
        Absolutive   Vocative
        Cáli   *Cáli*   'Ali'   (no change)
        Faadúmo   *Fàadumo*   'Fatima'
        Múuse   *Mùuse*   'Musa'   (a man's name)
        Warsáme   *Wársame*   'Warsame'   (a man's name)
        Guddoomiyé   *Gúddoomiye*   'Chairman'
        Madaxweyné   *Mádaxweyne*   'President'

There is some variation in the use of these vocatives because of their sensitivity
to pragmatic factors: their use reflects the speaker's estimation of the
relationship with the addressee, including such factors as solidarity, respect,
affection, etc. The distinction between names and nouns is, of course, itself
subject to speaker control, for example in animal fables where we find animals
addressing each other, using nouns for species with name vocatives, e.g.
Shimbírèy 'Bird!' and *Shinníyèy* 'Bee!'.

When nouns form noun phrases, case is marked phrasally, i.e. just once
on the final element of the noun phrase. Since noun phrases are head initial in
Somali, this means that the exponents of case may be marked not on the noun
but on modifying non-nominal elements such as determiners, adjectives, and
subordinate clause verbs. In subsequent sections we shall see the effect of this
phrasal case marking on a range of non-nominal categories. When the case
marking occurs on this phrase final position, the noun occurs in the default
absolutive case, as we can see in the subject marked noun phrases in the (b)
sentences below:

(29)   (a)   *nin*   'a man'   [ABS]
        (b)   *nin*   'a man   [NOM]
(30)   (a)   *ninka*   'the man (non-remote)' [ABS]
        (b)   *ninku*   'the man (non-remote)' [NOM]
(31)   (a)   *ninka dhèer*   'the tall man'   [ABS]
        (b)   *ninka dheeri* 'the tall man'   [NOM]

Only when the noun *nín* 'man' stands alone, as in (29), is case marked on it; in
(30) the marking shifts to the determiner *ka* 'the', and in (31) to the adjective
dhèer 'tall'. In (30) and (31) the head noun *nín* shows absolutive case, though
the whole NP is subject.

4.1.2 Names

Names are nouns: they have inherent gender and are case marked, as we have seen. Names do not normally occur in plural forms, with determiners, or as heads of restrictive relative clauses, except in special circumstances, for example *Caligée* 'which Ali?' and its possible reply:

(32)   *Cáligii Berbera jóogi jíray*
          Cáli-kii   Berbera   jóogi   jíray
          Ali-the   Berbera   stay:in:INF   used
          'the Ali who used to stay in Berbera'

Many personal names are of Arabic origin, for example the male names *Áxmed, Cábdi, Cáli, Maxámed, Saciid* and female names *Faadúmo* ('Fatima'), *Cáasho* ('Ayesha'), and *Maryámo* ('Miriam'). Somali names are also regularly used, for example, the male names *Róoble* 'Rain Bearer' and *Warsáme* 'Bearer of Good News' and female names like *Káaha* 'Distant Light, Dawn', *Déeqa* 'Gift' and *Hódan* 'Wealth'. The naming system for both men and women consists of a personal name followed by the father's personal name and then the grandfather's name, forming a compound genitive list 'X (child) of Y (child) of Z', etc. The three names in normal use can be extended to a complete lineage. Women do not change their names on marriage. Since case marking is phrasal only the last element in a complex NP is case marked and thus only the last name carries genitive case, the preceding names being in the absolute:

(33)   (a)   *Maxámed*   *Áxmed*   *Xasán*
                 Mohamed:ABS   Ahmed:ABS   Hasan:GEN
                 'Mohamed Ahmed Hasan'
          (b)   *Cáasha*   *Xásan*   *Axméd*
                 Ayesha:ABS   Hasan:ABS   Ahmed:GEN
                 'Ayesha Hasan Ahmed'

Nicknames are commonly used and are useful since many people share a relatively small stock of the commonest personal names. The nicknames may reflect an attribute of the person's physique or character or may be tied to specific incidents, perhaps in childhood. The nicknames are usually also nouns, e.g. *madóobe* N M 'the black one' in *Aadan Madóobe* 'Adam the Dark' or *musmàar* N M 'nail' in *Cali Musmàar* 'Ali "Nail"'. Such nicknames are frequently in public use. Name and nickname combinations act accentually like compound nouns (5.4.1), thus these examples have just one high tone on both

names. This allows contrasts like the following, pointed out to me by B.W.
Andrzejewski:

(34)   (a)   *Cali Khàyr*
               'Ali Khayr' (whose nickname is *khàyr* N M 'goodness')
        (b)   *Cáli Kháyr*
               'Ali (son) of Khayr'

4.1.3 Independent pronouns

The independent pronouns, which are also sometimes called substantive or
emphatic pronouns, contrast in function with the clitic or verbal pronouns
described in 4.2. The independent pronouns are grammatically nouns. They are
case marked, have inherent gender and show the syntax of nouns, including
functioning as subject and object, occurring as heads of relative clauses, being
focused by the nominal focus words *bàa* and *ayàa*, and occurring in isolation as
one word utterances. In Standard Somali they occur with the suffixed non-
remote definite article *-ka/-ta* but may occur in an unsuffixed short form when
followed by a conjunction or focus word. The suffixed and short forms are
shown below, in (35) followed by the verb phrase conjunction *-na* 'and' and in
(36) by the focus word *bàa*:

(35)   (a)   *adígana*   (b)   *adna*
               adi-ka-na   adi-na
               you:SG-the-and   you:SG-and
               'and you...'   'and you...'
(36)   (a)   *adígàa*   (b)   *adàa*
               adi-ka-bàa   adi-bàa
               you:SG-the-FOC   you:SG -FOC
               'YOU' (focused)   'YOU' (focused)

The set of independent pronouns is given below, with the determiner
separated off by a hyphen to show the short form. They provide an eight-person
system, with EXC 'exclusive' meaning 'not including the addressee(s)' and INC
'inclusive' meaning 'including the addressee(s)'. As the suffixed determiners
show, all of the pronouns are masculine in gender, except *iyáda* 'she, her'.

(37)   1 SG   *aní-ga*
        2 SG   *adí-ga*
        3 SG M   *isá-ga*
        3 SG F   *iyá-da*
        1 PL EXC   *anná-ga*

1 PL INC	*inná-ga*
2 PL	*idín-ka*
3 PL	*iyá-ga*

The independent pronouns can occur with other determiners like demonstratives and even the interrogative determiner, for example *anigée*? 'which me?', an expression that may be used to protest when unfairly accused.

As described in chapter 9, independent pronouns like other nominals can occur with coreferential clitic pronouns, for example:

(38) *Adíga, maxàa ku keenáy?*
      adi-ka   maxáy+bàa   ku   keenáy
      you:IPRO-the   what+FOC   you:CPRO   brought
      'You, what brought you?'

Since independent pronouns normally occur with the suffixed determiner, we deal with subject marking in the discussion of case marking on determiners in 4.6.

4.1.4 Numerals

The cardinal numerals are nouns: they are case-marked, have inherent gender, allow the full range of suffixed determiners and show nominal syntax, including functioning as subjects, objects, heads of relative clauses, etc. The set of numerals 1-10 and examples of higher numerals are given below in the absolutive case; they are case marked regularly as described in 4.1.1.

(39)
ków	F	'one'		*líx*	F	'six'
lába	F	'two'		*toddobá*	F	'seven'
sáddex	F	'three'		*siddéed*	F	'eight'
áfar	F	'four'		*sagáal*	M	'nine'
shán	F	'five'		*tobán*	M	'ten'

(40)
kóob iyo tobán	M	'eleven'
labá iyo tobán	M	'twelve'
sáddex iyo tobán	M	'thirteen'
labaatán	M	'twenty'
sóddon	M	'thirty'
afártan	M	'forty'
kónton	M	'fifty'
líxdan	M	'sixty'
toddobáatan	M	'seventy'
siddéetan	M	'eighty

sagaáashan	M	'ninety'
boqól	M	'hundred'
labá boqól	F	'two hundred'
labá boqól iyo kónton	F	'two hundred and fifty'
kún	M	'one thousand'
malyùun/milyùun	M	'million'

The numeral *lába* undergoes the rule of Accent Shift (3.4.3): it is given here in its prepause or isolation form with AP2; before other elements the absolutive has AP1, i.e. *labá*. The numerals *sagáal* 'nine', *tobán* 'ten', *labaatán* 'twenty' and *boqól* 'hundred' undergo the rule of Accent Lowering (3.4.3).

The numeral 'one' has several suppletive forms. When counting, the numeral is *ków*, which becomes *kóob* when followed by *iyo* 'and' as shown in (40). When used with nouns this word is replaced by *hál* 'one', e.g. *hál baabúur* 'one truck'. When the numeral occurs alone as the argument of a verb, it is replaced by *míd* 'one', e.g. *míd kèen!'* 'Bring one!' This word *míd* is also used in isolation when a verb is understood:

(41)   Question: *Ímmisàad dóonaysaa?*
                     ímmisa+bàa+aad   dóonaysaa
                     how:many+FOC+you   are:wanting
                     'How many do you want?'
        Answer: *Míd*
                     'One'.

Approximate numerals are formed by adding to numerals either the suffix *-éeyo* to form feminine nouns, or the suffix *-éeye* to form masculine nouns:

(42)   *tobanéeyo*   F   'about ten, ten or so'
        labaatanéeyo   F   'about twenty'
        afartaméeye   M   'about forty'
        kontoméeye   M   'about fifty'

When used with nouns, the cardinal numerals act as the head of the construction and the counted noun occurs in the genitive case, for example:

(43)   *áfar*   *naagóod*
        four:ABS   women:GEN
        'four women' (lit. 'four of women')

This construction also occurs with composite numerals like *labá boqól* 'two hundred' where *labá* 'two' is a feminine noun in absolutive case and *boqól* 'hundred' is a masculine noun in genitive case. Since *labá* is head, the whole nominal is feminine.

In such constructions nouns of declension 1 occur in plural genitive forms after numerals greater than one, as in (43) above and (44) below, while all other nouns occur in the singular genitive form after all numerals (45):

(44)   *sáddex gabdhóod*   'three girls'   cf. *gabdhó* M PL   'girls'
        toddobá maalmóod   'seven days'   cf. *maalmó* M PL   'days'
(45)   *sáddex wíil*   'three boys'   cf. *wìil* M SG   'boy'
        toddobá baabúur   'seven trucks'   cf. *baabùur* M SG   'truck'

Thus *sáddex wíil* is literally 'three of boy'.

When such constructions are made definite, it is the head numeral which attracts the determiner; in the following examples the determiners are feminine to match the head of the NP, despite the counted noun being masculine. The numeral NPs below will of course take feminine agreement with a main sentence verb.

(46)   (a)   *labádíi waraabé*   (cf. *waráabe* M SG 'hyena')
               labá-tii   *waraabé*
               two:F:ABS-the:REM   hyena:SG:GEN
               'the two hyenas (remote)'
        (b)   *sáddexdáas nín*   (cf. *nín* M SG 'man')
               sáddex-táas   *nín*
               three:F:ABS-those   man:SG:GEN
               'those three men'
        (c)   *sáddexdàydíi adéer*   (cf. *adèer* M SG 'paternal uncle')
               sáddex-tày-tii   *adéer*
               three:F:ABS-the-my   uncle:SG:GEN
               'my three paternal uncles'

Nouns occur in these dependent genitive forms after several other quantifying nouns like *dhàwr* 'several', *ímmisa* 'how many?, *méeqa* 'how many?

The ordinal numerals are formed from cardinal numerals by the addition of a suffix *-áad*, which may produce the environment for Stem Contraction (3.1.4), for example:

(47)   *kowáad*   'first'
              afráad   'fourth'

> *boqláad*   'hundredth'
> *kumáad*   'thousandth'

These ordinal numerals are not nouns: they modify and follow the counted noun, which acts as head of the construction and thus attracts determiners, percolates its gender to the whole NP, etc:

(48)   *nín saddexáad*   'a third man'
             nínkíi saddexáad   'the third man'

4.2   Clitic Pronouns

Clitic pronouns form a separate category, which unlike independent pronouns do not exhibit the morpho-syntax of nouns. Clitic pronouns can only occur with a verb and must occur in the strict sequence of clitic elements in the verbal group described in chapter 6. They form two main sets, subject and object pronouns, and also include the impersonal subject pronoun *la* 'one' and the reflexive and reciprocal pronoun *is* 'self, each other'.

4.2.1 Subject clitic pronouns

The set of these pronouns is:

(49)   1 SG   *aan*
             2 SG   *aad*
             3 SG M   *uu*
             3 SG F   *ay*
             1 PL EXC   *aannu* (*aan*)
             1 PL INC   *aynu*
             2PL   *aydin*
             3PL   *ay*

Once again EXC 'exclusive' means 'not including the addressee(s)' and INC 'inclusive' means 'including the addressee(s)'. The forms in parentheses are optional short forms that may be used when the context permits. Note that 3 SG F and 3 PL are both *ay*; however the ambiguity is resolved by the accompanying verb forms. The 1 SG and 2 SG pronouns may be used for their corresponding plurals, especially in informal conversation.

As described in chapter 3, these pronouns cliticize onto and coalesce with preceding elements, for example, the declarative sentence type marker *waa* and the focus word *bàa*:

(50) *Wùu yimi.*
      waa+uu   yimi
      DM+he   came
      'He came'

(51) *Dáar bày dhísayaan.*
      dáar   bàa+ay   dhísayaan
      house   FOC+they   are:building:3 PL
      'They are building A HOUSE.'

4.2.2 Object clitic pronouns

There are two sets of object clitic pronouns: a primary and a secondary set. The primary set is:

(52)   1 SG   *i*
              2 SG   *ku*
              1 PL EXC   *na*
              1 PL INC   *ina*
              2PL   *idin*
              3 SG/PL   -

As shown, there is a general third person object pronoun marked by a gap in the paradigm. This means that the absence of any other object clitic pronoun will result in a third person interpretation. See for example:

(53) *Wùu keenay.*
      wàa+uu   keenay
      DM+he   brought
      'He brought it/her/him/them.'

The specific type of third person has to be recovered from context. Note that for economy we have not given an English gloss for all the third person options in each example in this book, simply giving whichever pronoun seems likeliest in context, or in some cases writing 'him etc.', 'her etc.' and so on. One effect of this use of a gap is that the argument structure of a verb is always filled: it is not possible in Somali to remove an argument from a verb simply by omitting it as in the English pair *He ate lunch/He ate*. In Somali omitting an overt NP will trigger a third person object pronoun interpretation as in (54a); to express the

English *He ate*, Somalis must use an indefinite object *wáx* N M 'thing, something' as in (54b):

(54)   (a)   *Wùu cunay.*
                wàa+uu   cunay
                DM+he   ate
                'He ate it/her/him/them.'

        (b)   *Wáx bùu cunay.*
                wáx   bàa+uu   cunay
                thing   FOC+he   ate
                'He ate something'

The second series object pronouns are:

(55)   1 SG   *káy*
                    2 SG   *káa*
                    1 PL EXC   *kayó*
                    1 PL INC   *kéen*
                    2 PL   *kíin*
                    3 SG/PL   -

Again, these pronouns also have a gap for third person. The second series pronouns are used when two overt object clitic pronouns occur in the same verbal group. This most commonly occurs when one is the object of a verb and the other the object of an adposition.[28] Thus, if only one object pronoun occurs it will be a primary series form, whether governed by a verb or adposition:

(56)   *Wáy idin arkeen.*
        wàa+ay   idin   arkeen
        DM+they   you:PL   saw:3 PL
        'They saw you (PL).'
(57)   *Wày idinká qaadeen.*
        wàa+ay   idin-ká   qaadeen
        DM+they   you:PL-from   took:3 PL
        'They took it /etc. away from you (PL)'

In (56) *idin* 'you (PL)' is the object of the verb *árag* 'see' and in (57) it is the object of the adposition *ká* 'from'. Note that in (57) the object of *qàad* 'take' is marked by a gap.

    The second series forms only occur when two overt pronouns are required, that is when two non-third person object pronouns occur in the same

verbal group. This limits their occurrence somewhat in conversation. An example is:

(58)   *Wày igá kiin qaadeen.*
       wàa+ay   i-ká   kiin   qaadeen
       DM+they   me-from   you:PL   took:3 PL
       'They took me away from you (PL).'

As this example shows, the primary and second series clitic pronouns occupy different positions in the verbal group. As described in chapter 6, the primary series occur before adpositions while the second series occurs immediately before the verb. Note that in this position, the second series pronouns do not participate in the coalescence between clitic pronouns and adpositions described in 3.3. One result of this sequence is that *i* 'me' in (58) coalesces with *ká* 'from' even though there is no syntactic or semantic government between these two elements, *i* being the object of the verb and *ká* governing its object *kiin*.

One interesting feature of the use of the two series is that when both forms are required there is optional variation in the choice of matching between the series and the type of object. For example if a verbal object and an adposition object are both represented by clitic pronouns, the primary and secondary series may represent either. Thus paralleling (58), where the verbal object is primary series and the adpositional object is secondary, we can find (59):

(59)   *Way idinká káy qaadeen.*
       wàa+ay   idin-ká   káy   qaadeen
       DM+they   you:PL-from   me   took:3 PL
       'They took me away from you (PL).'

Here the adpositional object is from the primary series and the verbal object is from the second series.

This optionality also occurs in the rarer examples of purely ditransitive verbs, which control two objects without the use of adpositions. Such verbs are relatively rare in Somali, but include the common verb *sìi* 'give to'. Example (60) shows the possible choices of object pronouns when two non-third person forms are required:

(60)   *Ilàah bàa na káa siiyey.*
       Ilàah   bàa   na   káa   siiyey
       God   FOC   us   you:SG   gave
       'God gave you (SG) to us' or 'God gave us to you (SG).'

It is clear that in form the second series object pronouns are quite distinct from the primary series and resemble the possessive determiners *-kày/-tày* 'mine' etc. described in 4.6, except of course that the pronouns have no overt third person forms.

4.2.3 Non-specific subject pronoun

The non-specific subject pronoun *la*, which we gloss in English as 'one', takes third person masculine singular verbal agreement and is typically used with a transitive verb when the speaker does not know or wishes to suppress the identity of the subject, for example:

(61)   *Wàa la xaday.*
        wàa   la   xaday
        DM   one   stole:3SG M
        'It was stolen.', lit. 'One stole it'
(62)   *Wàa lagu arkay.*
        wàa   la-ku   arkay
        DM   one-you:SG   saw:3SG M
        'You (SG) were seen.', lit. 'One saw you (SG)'

This pronoun is also used when the subject is so general as to be vague, for example in the formulaic beginning to stories *wáxaa la yidhí* 'what one said…' or 'It is told…', or as in the following example:

(63)   *Dúmarku waa beló lóo baahányaháy.*
        dúmar-ku   waa   beló   la+ú   baahan-yahay
        women-the   DM   disaster   one+of   needful:ADJ-is:3SG M
        'Women are a necessary evil',
        (lit. 'Women are a disaster which one is needful of')

These uses mean that constructions with *la* are functionally equivalent to some uses of passive constructions in other languages, and indeed there is no syntactic passive in Somali. There are other grammatical mechanisms which share the functional domain of passives in other languages, as we shall see, for example the verbal derivational affix *-am* described in 5.1.4, which removes an agent from a causative verb and promotes the causee to subject.

The pronoun *la* has often been called an impersonal subject pronoun but we avoid the term here because unlike some impersonal constructions in other languages, such as impersonal passives, *la* cannot be used with intransitive

verbs. Thus one cannot say (64a) below but must instead use a full nominal *qóf* N M 'person' in (64b):

(64) (a) **Wàa   la   dhintay.*
           DM   one   died
           'One died', 'Someone died'
      (b) *Qóf   bàa   dhintaý.*
           person FOC   died
           'A person died', 'Someone died'

Note too that *la* can only be used for human entities and is thus restricted to verbs which could have a human subject.

As described in chapter 6, *la* occurs in the verbal group in the same position as other subject pronouns and thus precedes adpositions and object clitic pronouns, with which it coalesces. Earlier, in 3.3.3 we described its coalescence with adpositions. The following shows *la* combining with object clitic pronouns:

(65) *la*   +   *i*   'me'   →   *lay*
      la   +   *ku*   'you (SG)' →   *lagu*
      la   +   *na*   'us (EXC)   →   *nala*
      la   +   *ina*   'us (INC)   →   *layna*

An example is in (63) above. Note that *la* undergoes metathesis when combining with a monosyllabic morpheme beginning with a nasal, as in *nala* in (65). This also occurs in combinations with the negative word *má*, which normally precedes the elements of the verbal group; compare (66a) and (66b):

(66) (a) *Má'uu   cunó.*
           má-uu   cunó
           not-he   eat:NEG
           'He doesn't eat it.'
      (b) *Lamá cuno.*
           la-má   cunó
           one-not   eat:NEG
           'One does not eat it.', 'It isn't eaten.'

4.2.4 Reflexive/reciprocal object clitic pronoun

The pronoun *is*, which we gloss as 'self' or 'each other' in English, patterns like an object clitic pronoun: it can only occur with a verb and occurs in the verbal

group preceding the verb. It is invariable in form except that *is* becomes *isa*
when the following word begins with *s*, i.e. there is vowel epenthesis as in *isa
sìi!* 'Give it to yourself!'. The pronoun can have a reflexive (67) or reciprocal
interpretation (68), depending on context:

(67)   (a)   *Wày is dishay*
                waa+ay   is   dishay
                DM+she   self   killed
                'She killed herself'
        (b)   *Wuu iská warramay*
                waa+uu   is-ká   warramay
                DM+he   self-about   told:news
                'He told the news about himself'
(68)   (a)   *Wày is arkeen.*
                waa+ay   is   arkeen
                DM+they   self   saw
                'They saw each other'
        (b)   *Sháley bày islá hadleen.*
                sháley   bàa+ay   is-lá   hadleen.
                yesterday   FOC+they   self-with   spoke
                'They talked with each other yesterday.'

Note that (68a) could also mean 'They saw themselves.'
    The combinations of *is* and the adpositions *ká* 'from, away, out of' and *kú*
'in, into, on, onto, with (INSTR)' produce some very common idioms. *Is* and *ká*,
often contracted to *ská*, feature in a number of abrupt commands often used in
informal conversation, given below in singular forms:

(69)   *Iská báx!*   'Get away!'   (cf. *báx* V 'go out, exit')
        Iská dàa!   'Never mind!'   (cf. *dàa* V 'let be')
        Iská èeg!   'Watch out!'   (cf. *èeg* V 'watch')
        Iská jír!   'Take care!'   (cf. *jír* V 'protect')
        Iská tág!   'Go away!'   (cf. *tág* V 'go')

    The combination of *is* and *kú* has an unusual de-accented pattern of LOW -
LOW and produces an idiosyncratic meaning of 'together, joined' or 'likeness',
as in the phrasal verb *isku dár* 'put together, join' (cf. *dár* V 'put'), or in the
verbless sentences:

(70)   *Wàa isku ìib.*   'They are the same price.'   (cf. *ìib* N M price')
        Wàa isku jàad   'They are the same type.'   (cf. *jàad* N M 'type')

Wàa isku míd	'They are the one, the same', 'It's all the same',	
	'It makes no difference'	(cf. *míd* NUM 'one')
Wàa isku qólo	'They are of the same clan.'	(cf. *qólo* N M 'clan')
Wàa isku àfar	'There are four of them together'	(cf. *áfar* NUM 'four')

4.3   Verbs

Verbs as a category are identified by the tense, aspect and mood information they carry, which is signalled largely by affixes but also by vowel alternations and accentual patterns. Syntactically they show agreement with the clause subject, and function as the head of a preceding group of satellite clitics (6.1.1). Verbs can be subdivided in two intersecting ways: by their morphological shape and by their syntactic function with other verbs. Three classes of verb can be identified by their morphology. The largest group is the suffix (or 'weak') verbs, which carry inflection as suffixes. Then there are the four prefix (or 'strong') verbs: *yidhi* 'say', *yimi* 'come', *yiil* 'be (in a place)' and *yiqiin* 'know', which in Somali show the ancient Afroasiatic pattern of carrying inflection as prefixes and vowel alternations. Finally there is the single verb *yahay* 'be', which is a kind of degenerate prefix verb. There are two major syntactic functions: while members of all three morphological classes function as main verbs, a small subset of suffix verbs function as auxiliary verbs. We begin with the morphological classes.

4.3.1 Morphological classes

As we have noted, there are three morphological classes of verbs. The suffix verbs have a basic morphological structure that may be represented schematically as:

(71)   [ROOT - LEX - AGR - INFL]

where ROOT is the base form of the verb, LEX represents the various lexical affixes which form derived verbs, such as the causative affix *-is* (5.1.5), AGR represents the agreement affixes which mark concord with a subject nominal, and INFL represents the inflectional endings which give information about tense, mood and aspect. For this class of verbs the reference form is the imperative singular form since it corresponds to the form of the basic root. Below are examples of this structure with past simple and present general forms of the root verb *súg* 'wait for':

(72)   (a)   *Wàannu sugnay.*
                waa+aannu   [$_{ROOT}$ *sug*]- [$_{AGR}$ *n*] - [$_{INFL}$ *ay*]
                DM+we:EXC   wait:for   1 PL   past simple
                'We (EXC) waited for it etc.'
        (b)   *Wày sugtaa.*
                waa+ay   [$_{ROOT}$ *sug*]- [$_{AGR}$ *t*] - [$_{INFL}$ *aa*]
                DM+she   wait:for   3 F SG   present habitual
                'She waits for it etc.'

The division between the AGR and INFL endings is not as simple as our schema
in (71) suggests, as we can see from the complete past simple and present
general paradigms for this verb *súg* 'wait for':

(73)	(Clitic PRO)	Past simple ('I waited for' etc.)	Present general ('I wait for' etc.)
1 SG	*(aan)*	*sugay*	*sugaa*
2 SG	*(aad)*	*sugtay*	*sugtaa*
3 SG M	*(uu)*	*sugay*	*sugaa*
3 SG F	*(ay)*	*sugtay*	*sugtaa*
1 PL EXC	*(aannu)*	*sugnay*	*sugnaa*
1 PL INC	*(aynu)*	*sugnay*	*sugnaa*
2 PL	*(aydin)*	*sugteen*	*sugtaan*
3 PL	*(ay)*	*sugeen*	*sugaan*

We can see from (73) that the combination of subject clitic pronouns and verb
forms differentiate eight person-number forms, while the verb form alone
differentiates five. The verb form shows a characteristic cross pattern, where
the three AGR affixes, zero (1 SG, 3 SG M and 3 PL), *t* (2 SG, 3 SG F and 2 PL), and
n (1 PL), combine with the two INFL shapes, *-ay/-een* and *-aa/-aan*, to give five
person-number distinctions. This binary split in INFL occurs in all verb
paradigms.

The identification of the AGR affixes is supported by the occurrence of
similar affixes, but with *y* replacing zero for 3 M SG and 3 PL, as prefixes in the
second morphological class, the prefix verbs. For comparison we give the past
simple and present habitual paradigms of the prefix verb *yiqiin* 'know'. Note
that since this is a stative verb the distinction contrasts a past state with a
present state, as discussed in 4.3.3 below.

(74)   (Clitic PRO)   Past simple   Present general
                                    ('I knew' etc.)   ('I know' etc.)

1 SG   *(aan)*   *iqiin*   *aqaan*
2 SG   *(aad)*   *tiqiin*   *taqaan*
3 SG M   *(uu)*   *yiqiin*   *yaqaan*
3 SG F   *(ay)*   *tiqiin*   *taqaan*
1 PL EXC   *(aannu)*   *niqiin*   *naqaan*
1 PL INC   *(aynu)*   *niqiin*   *naqaan*
2 PL   *(aydin)*   *tiqiinneen*   *taqaanniin*
3 PL   *(ay)*   *yiqiinneen*   *yaqaanniin*

Here we can see the morphological contrast with suffix verbs: the AGR affixes are prefixed to the stem and the tense/aspect is marked by contrasts of vowel quality within the root. Unlike suffix verbs, we use as the reference form for this class of verbs the third masculine singular of the past simple, e.g. *yiqiin* 'know'. The imperative forms, as we shall see, are not diagnostic of the root, and moreover, since two of the four verbs are stative, they do not occur in imperative forms.

The third class contains just one stative verb, *yahay* 'be', which combines the morphological structures of the two other classes: in the past simple it is a suffixing verb while in the present habitual it is a prefixing verb:

(75)   (Clitic PRO)   Past simple   Present habitual
                                    ('I am' etc.)   ('I was' etc.)

1 SG   *(aan)*   *ahaa*   *ahay*
2 SG   *(aad)*   *ahayd*   *tahay*
3 SG M   *(uu)*   *ahaa*   *yahay*
3 SG F   *(ay)*   *ahayd*   *tahay*
1 PL EXC   *(aannu)*   *ahayn*   *nahay*
1 PL INC   *(aynu)*   *ahayn*   *nahay*
2 PL   *(aydin)*   *ahaydeen*   *tihiin*
3 PL   *(ay)*   *ahaayeen*   *yihiin*

We use as the reference form for this verb a prefix form, the third masculine singular of the present habitual, i.e. *yahay* 'be'.

As might be expected, the prefix verb class is under pressure from the majority suffix class and there is considerable dialect variation in the forms of prefix verbs. In many dialects there occur mixed forms with, for example, suffix verb inflectional endings added to prefix stems, producing such forms as *aqaannay* or *iqiinnay* for 'I knew'; see Banti (1985) for some examples of this.

We may speculate that such pressure is historically the cause of the mixed forms of the verb *yahay* 'be'.

Of these three classes, only suffix verbs allow the productive attachment of lexical affixes, or root extensions, like the causative and middle voice morphemes described in chapter 5. As described in chapter 3, the morphophonological changes that follow from the attachment of these affixes to stems are complex. For clarity, writers on Somali have tended to set up conjugations of suffix verbs based on the presence of lexical affixes. A typical system, such as in Bell (1953), identifies three conjugations. Conjugation 1 consists of root verbs, for example *jòog* INTR 'stay, remain'. Conjugation 2 consists of verbs with causative affixes, for example *jóoji* TR 'cause to stop, stop' (← *joog+i(s)*). Finally, conjugation 3 consists of verbs with middle voice affixes, for example *joogsó* INTR 'stop oneself, come to a stop' (← *joog+is+o*). The, perhaps logical, step of identifying as many conjugations as there are lexical affixes is not however taken here. Firstly because the morphophonological shape of other affixes is more predictable, and secondly, because causative and middle stems exhibit unpredictable gaps and semantic shift to an extent which suggests they are more lexicalised than other verb-affix combinations. We follow this practice here, identifying three conjugations of suffix verbs in section 4.3.3.

4.3.2 Main and auxiliary verbs

Crosscutting the three-way morphological classification outlined in the last section is a syntactic distinction betwen main and auxiliary verbs. All verbs, whatever their morphological class, employ auxiliary verbs to mark certain distinctions in the tense-aspect-mood (TAM) system. These constructions all have the form:

(76)   [Main verb   -   Auxiliary verb]
                   (infinitive)   (+inflection)

The main verb occurs in an invariable form that we will term the infinitive, while the auxiliary verb is inflected for agreement affixes, etc.; for example:

(77)   (a)   *Wày*   *keeneen.*
                DM+they   brought:3PL
                'They brought it.'

        (b)   *Wày*   *kéeni*   *jireen.*
                DM+they   bring:INF   used:3PL
                'They used to bring it.'

In (77b) the auxiliary verb *jír* signifies the past habitual TAM form and carries the inflection.

As we shall see in sections 4.3.4-6, there are four auxiliary constructions that act as part of the tense-aspect-mood system. The future marked by *dóon* 'will'; the past habitual with *jír* 'be, exist'; the conditional with the fused adjective and verb form *lahaa* 'was possessing', from *léh* ADJ 'possessing' and *yahay* 'be'; and the progressives marked by partial coalescence with *hày* 'have'. In these constructions the auxiliary verbs have lost their independent lexical content and function as markers of the relevant TAM distinction. In addition there are a small number of verbs which act like these auxiliaries while maintaining their lexical content. The most important of these are given below:

(78) Verb   Meaning   Example
     kár   'can, be able'   *Waan   tégi   karaa*
                                       DM+I   go:INF   can-PRES
                                       'I can go'
     wàa   'fail, be unable'   *Waan   tégi   waayey*
                                       DM+I   go:INF   fail-PAST
                                       'I failed to go, I couldn't go'
     gàadh   'just fail, almost do' *Waan   tégi   gaadhay*
                                       DM+I   go:INF   almost:do-PAST
                                       'I almost went, I nearly went'

Since they do not function as part of the TAM system and retain lexical content, we will analyse these verbs as main verbs which, unusually for Somali, take infinitival complements; these are discussed further in 6.1.3 and 8.4.2.

4.3.3 Inflectional categories

Verbal inflection carries information from a number of sentence-level semantic and syntactic systems. We provide a basic list of these below:

(79) *Inflectional categories of the verb*
     Tense
     Aspect
     Mood
     Negation
     Sentence subordination

These systems interact on a relatively restricted number of inflectional endings and auxiliary verbs with the result that verbal inflection in Somali is most

economically and clearly represented in a paradigmatic form, as in the next section. These sentence-level systems also interact with lexically-based semantic distinctions, such as those of *aktionsart*, or situation type, for example between stative and dynamic verbs. Stative verbs occur in all three morphological classes, for example:

(80)   (a)   Suffix class statives
              búk   'be ill'
              haysó   'have, possess'
              húb   'be certain'
              húrud   'be asleep, sleep'
              jòog   'be (in a place), remain, stay'
              muuqó   'appear, seem'
        (b)   Prefix class statives
              yiil   'be (in a place)'
              yiqiin   'know'
        (c)   *yahay*   'be'

As the terms suggest, stative verbs portray situations as static or unchanging for their duration, whereas dynamic verbs, such as inchoatives and causatives, allow speakers to view situations as having internal temporal structure, or subparts. This distinction has an impact on the sentence-level systems of aspect and mood. Stative verbs do not, for example, occur in progressive forms or for the most part in imperative sentences.[29] In addition, in paradigms where both do occur, stative and dynamic verbs may have different aspectual interpretations. With dynamic verbs the paradigm we label present habitual describes a repeated or regular activity, like *kèen* 'bring' in *Wùu keenaa (maalín kastá)* 'He brings it (every day)'. With stative verbs, on the other hand, it signifies a state ongoing at the time of speaking, like *húb* 'be certain' in *Wàan hubaa* 'I am certain'.

In this chapter we simply identify the verbal inflectional forms, leaving all but the sketchiest discussion of the use and meaning of forms to later chapters. The inflectional categories in (79) above exhibit the following formal contrasts:

(81)   *Contrasts in inflectional categories*
       Tense: Past/Present/Future
       Aspect: Simple/Progressive/Habitual
       Mood: Declarative/Imperative/Conditional/Optative/Potential
       Negation: Positive/Negative
       Sentence subordination: Main/Subordinate

Not every possible combination of these categories occurs, as we shall see: for example, tense and aspect are only marked in declarative sentences; there is no negation in potential sentences, etc. We can group the possible combinations into the twelve verbal paradigms below, details of which are given in the next three sections for suffix verbs, prefix verbs and *yahay* 'be':

(82)				
	1.	Imperative	7.	Present progressive
	2.	Infinitive	8.	Future
	3.	Past simple	9.	Conditional
	4.	Past progressive	10.	Optative
	5.	Past habitual	11.	Potential
	6.	Present habitual	12.	Subordinate clause forms

4.3.4 Suffix Verb Paradigms

As mentioned earlier, because of the morphophonological complexity associated with verbal derivational affixes, it is useful to identify three major conjugations of suffix verbs as follows:

(83) Conjugation 1: Root verbs with no lexical affixes, e.g. *súg* TR 'wait for', *kár* INTR 'boil, cook';

Conjugation 2A: Verbs derived from root verbs by the causative affix *-i/-is*, e.g. *kári* TR 'cook' (from conjugation 1 *kár* INTR 'boil, cook');

Conjugation 2B: Verbs derived from nouns and adjectives by the causative/factitive affix *-ee/-ayn,* e.g. *yarèe* 'make small' (from *yár* ADJ 'small');

Conjugation 3A: Verbs derived from verbal stems by the middle voice affix *-o/-at*, e.g. *karsó* 'cook for oneself' (from conjugation 2 *kári* TR 'cook');

Conjugation 3B: As conjugation 3A but verbs whose syllable structure triggers Stem Contraction (3.1.4) and subsequent sandhi rules, e.g. *qaadó* 'take for oneself' (from conjugation 1 *qàad* TR 'take').

The derivational affixes mentioned above are described in chapter 5. In the list of paradigms below, conjugations 2B and 3B are only listed when they contain significant differences from 2A and 3A. Our example verbs are C1 *súg* TR 'wait for', C2A *kári* TR 'cook', C2B *yarèe* TR 'make smaller, reduce', C3A *joogsó* INTR 'stop oneself, come to a halt', C3B *qaadó* TR 'take for oneself'.

The paradigms for suffix verbs list only the five persons distinguished by verb inflection alone, though as we have seen, this may be extended to an eight person distinction by the use of subject pronouns.

4.3.4.1   Imperative

These forms are used for commands, which are more commonly used in Somali than, say, English since they seem to carry little or no implication of power or authority: e.g. *Súg!* 'Wait (SG) for it!', *Kári!* 'Cook (SG) it!', etc.[30]

	C1	C2A	C2B	C3A	C3B
2SG	súg	kári	yarèe	joogsó	qaadó
2PL	súga	kariya	yaréeya	joogsáda	qáata

The negative imperatives occur with a sentence type marker *ha* (4.8):

	C1	C2A	C2B	C3A	C3B
2SG	súgin	kárin	yaréyn	joogsán	qaadán
2PL	sugína	karinína	yaraynína	joogsanína	qaadanína

4.3.4.2   Infinitive

This form never occurs alone but combines with auxiliary verbs to form compound tenses like future and conditional, as we will see.

	C1	C2A	C2B	C3A	C3B
	súgi	kárin	yarèyn	joogsán	qaadán

4.3.4.3   Past simple

This tense/aspect form is used of completed action in the past, e.g. *Wàad sugtay* 'You waited for it.'. It also occurs in negative clauses, where the negative verb shows no person agreement distinctions, in what we will term an invariable form: *Má sugín* 'I/you/he etc. did not wait for it'. As usual the negative word *má* 'not' occurs in main clauses, while *aan* 'not' occurs in subordinate clauses.

	C1	C2A	C2B	C3A	C3B
1SG/3SG M	sugay	kariyey	yareeyey	joogsaday	qaatay
2SG/3SG F	sugtay	karisey	yareysey	joogsatay	qaadatay
1PL	sugnay	karinney	yareyney	joogsannay	qaadannay
2PL	sugteen	kariseen	yareyseen	joogsateen	qaadateen
3PL	sugeen	kariyeen	yareeyeen	joogsadeen	qaateen

The negative forms of this paradigm, which must occur with a negative word *má* or *aan*, are invariable in form:

Invariable:	sugín	karín	yaréyn	joogsán	qaadán

4.3.4.4 Past progressive

This tense/aspect form is used for an action in the process of occurring at a time referred to in the past, e.g. *Márkay timi, wàan shaqéyneyey* 'When she came, I was working', where *timi* '(she) came' is past simple and *shaqéyneyey* '(I) was working' is past progressive. As with other progressive paradigms it is not used with stative verbs. The paradigm has negative forms, which again are invariable, e.g. *Má súgéyn* 'I/you/he etc. was not waiting'. The forms currently in use are clearly contractions of auxiliary constructions, where the *ay* or *ey* segments are the remnant of the auxiliary verb *hày* 'have' now separating the infinitive stem from the agreement affixes. The final *-i* of the declension 1 infinitive has been elided. The three main declensions may serve to illustrate this paradigm:

	C1	C2	C3
1SG/3SG M	*súgayey*	*karínayey*	*joogsánayey*
2SG/3SG F	*súgeysey*	*karíneysey*	*joogsáneysey*
1PL	*súgeyney*	*karíneyney*	*joogsáneyney*
2PL	*súgeyseen*	*karíneyseen*	*joogsáneyseen*
3PL	*súgayeen*	*karínayeen*	*joogsánayeen*

The negative forms of this paradigm, which as usual must occur with a negative word *má* or *aan*, are invariable in agreement and have two optional forms, a longer and shorter version:

Invariable: *súgéyn/súgeynín   karínéyn/karíneynín   joogsánéyn/joogsáneynín*

There is some variation in the operation of vowel harmony in the *e* and *a* vowels before *y* in these forms: we assume that the underlying form is *ay* (from *háy* 'have'), which may become *ey* in the relevant context. In the forms given here *ay* → *ey* when followed by a CVC syllable, where V is long or short *e*. Speakers of other varieties of Standard Somali, however, may not apply this rule so that the endings are *-ayay*, *-aysay*, etc. or generalise the environment to a following VC syllable so that the endings are *-eyey*, *-eysey* etc.

4.3.4.5 Past habitual

This tense/aspect form is used of a repeated or habitual action in the past which no longer occurs. The form consists of the verb *jír* 'be (in a place), exist' used as an auxiliary verb in the past simple. It follows an infinitive main verb, e.g. *Wàan dhegeysán jirey* ' I used to listen to it'; and in the negative, *Má daawán jirín* 'I/you/he etc. didn't use to watch it'.

	C1	C2	C3
1SG/3SG M	*súgi jirey*	*karín jirey*	*joogsán jirey*
2SG/3SG F	*súgi jirtey*	*karín jirtey*	*joogsán jirtey*
1PL	*súgi jirney*	*karín jirney*	*joogsán jirney*
2PL	*súgi jirteen*	*karín jirteen*	*joogsán jirteen*
3PL	*súgi jireen*	*karín jireen*	*joogsán jireen*

The negative forms of this paradigm, which as usual must occur with a negative word *má* or *aan*, are invariable in agreement:

Invariable: *súgi jirín*   *karín jirín*   *joogsán jirín*

4.3.4.6   Present general

This tense/aspect form has two interpretations: with dynamic verbs it signifies a habitual or repeated action which still occurs at the time of speaking; and with stative verbs, it signifies a state holding at the time of speaking, e.g. the dynamic *Daawáda góormàad cabtaa*? 'When do you take the medicine?' or the stative *Wùu bukaa* 'He is ill', or the negatives, dynamic *Sigàar má cabbó* 'I don't/he doesn't smoke cigarettes' and stative *Má bukó* 'I am not/he is not ill'.

	C1	C2A	C2B	C3A	C3B
1SG/3SG M	*sugaa*	*kariyaa*	*yareeyaa*	*joogsadaa*	*qaataa*
2SG/3SG F	*sugtaa*	*karisaa*	*yareysaa*	*joogsataa*	*qaadataa*
1PL	*sugnaa*	*karinnaa*	*yareynaa*	*joogsannaa*	*qaadannaa*
2PL	*sugtaan*	*karisaan*	*yareysaan*	*joogsataan*	*qaadataan*
3PL	*sugaan*	*kariyaan*	*yareeyaan*	*joogsadaan*	*qaataan*

The negative forms, which occur with negative words *má* or *aan*, show full person agreement:

	C1	C2A	C2B	C3A	C3B
1SG/3SG M	*sugó*	*kariyó*	*yareeyó*	*joogsadó*	*qaató*
2SG	*sugtó*	*karisó*	*yareysó*	*joogsató*	*qaadató*
	or	or	or	or	or
	sugtíd	*karisíd*	*yareysíd*	*joogsatíd*	*qaadatíd*
3SG F	*sugtó*	*karisó*	*yareysó*	*joogsató*	*qaadató*
1PL	*sugnó*	*karinnó*	*yareynó*	*joogsannó*	*qaadannó*
2PL	*sugtàan*	*karisàan*	*yareysàan*	*joogsatàan*	*qaadatàan*
3PL	*sugàan*	*kariyàan*	*yareeyàan*	*joogsadàan*	*qaatàan*

4.3.4.7   Present progressive

This tense/aspect form is used for an action in the process of occurring at the time of speaking, e.g. *Wàan shaqéyneyaa* 'I am working'.   It is also very commonly used as a proximate future tense, of action that is about to happen, especially when planned in the present, e.g. *Ma tégaysaa?* 'Are you going?', 'Are you about to go?'.   As with other progressive paradigms it is not used with stative verbs. As noted for the past progressive, the positive forms are clearly contractions of auxiliary constructions, where the *ay* or *ey* segments are the remnant of the auxiliary verb *hày* 'have'. The negative still has optional forms which are overtly auxiliary constructions, e.g. contracted *má súgayo* and uncontracted *súgi maayó* both occur meaning 'I am not/he is not waiting for it'. Note that the negative word *má* precedes the contracted form, while it precedes and fuses with the auxiliary in the uncontracted form. In the tables below we give both forms. The three main declensions may serve to illustrate this paradigm, starting with the positive forms:

	C1	C2	C3
1SG/3SG M	*súgayaa*	*karínayaa*	*joogsánayaa*
2SG/3 SG F	*súgeysaa*	*karíneysaa*	*joogsáneysaa*
1PL	*súgeynaa*	*karíneynaa*	*joogsáneynaa*
2PL	*súgeysaan*	*karíneysaan*	*joogsáneysaan*
3PL	*súgayaan*	*karínayaan*	*joogsánayaan*

The main clause negatives, with coalesced *má* 'not', may occur in the following uncontracted forms:

	C1	C2	C3
1SG/3SG M	*súgi maayó*	*karín maayó*	*joogsán maayó*
2SG	*súgi maysó*	*karín maysó*	*joogsán maysó*
	or	or	or
	súgi maysíd	*karín maysíd*	*joogsán maysíd*
3SG F	*súgi maysó*	*karín maysó*	*joogsán maysó*
1PL	*súgi maynó*	*karín maynó*	*joogsán maynó*
2PL	*súgi maysàan*	*karín maysàan*	*joogsán maysàan*
3PL	*súgi maayàan*	*karín maayàan*	*joogsán maayàan*

The main clause contracted forms, used with *má*, are:

	C1	C2	C3
1SG/3SG M	*súgayó*	*karínayó*	*joogsánayó*
2SG	*súgaysó*	*karínaysó*	*joogsánaysó*
	or	or	or
	súgaysíd	*karínaysíd*	*joogsánaysíd*
3SG F	*súgaysó*	*karínaysó*	*joogsánaysó*
1PL	*súgaynó*	*karínaynó*	*joogsánaynó*
2PL	*súgaysàan*	*karínaysàan*	*joogsánaysàan*
3PL	*súgayàan*	*karínayàan*	*joogsánayàan*

In subordinate clauses, the uncontracted auxiliary forms do not occur. Speakers either use the contracted forms (*súgayó* etc.) with *aan*, or, more commonly, use the same invariable forms as with the past progressive (4.3.4.4) again with *aan*:

Invariable: *súgéyn/súgeynín*   *karínéyn/karíneynín*   *joogsánéyn/joogsáneynín*

This of course neutralises the tense difference between the two paradigms.

4.3.4.8   Future

This form is used for actions or states that are situated in the future of the act of speaking, and carries an implication of speaker certainty. The form is an auxiliary construction comprising the infinitive of the main verb and the present habitual of the C1 verb *dòon*, which as a main verb means 'wish' or 'want', e.g. *Way kéeni doontaa* 'She will bring it' or *Má árki doontíd* 'You won't see it.'. The positive forms are:

	C1	C2	C3
1SG/3SG M	*súgi doonaa*	*karín doonaa*	*joogsán doonaa*
2SG/3SG F	*súgi doontaa*	*karín doontaa*	*joogsán doontaa*
1PL	*súgi doonnaa*	*karín doonnaa*	*joogsán doonnaa*
2PL	*súgi doontaan*	*karín doontaan*	*joogsán doontaan*
3PL	*súgi doonaan*	*karín doonaan*	*joogsán doonaan*

The main clause negative forms occur with *má* 'not' and have the auxiliary *dòon* in the negative present general, e.g. *súgi doonó* 'I/he will not wait', *súgi doontó* 'you/she will not wait', etc. The subordinate clause negative forms occur with *aan* 'not' and consist of an invariable form:

	C1	C2	C3
invariable:	*súgi doonín*	*karín doonín*	*joogsán doonín*

4.3.4.9   Conditional

This is an irrealis mood form and signifies a hypothetical situation in the present or past, e.g. *Waan tégi lahaa* 'I would go', 'I would have gone'. As this example shows, this is an auxiliary construction formed with the main verb infinitive followed by the adjective *léh* 'having, possessing' acting as the complement of past simple forms of *yahay* 'be' (4.3.6). As is usual with adjectives (4.4) the past tense of the copula *yahay* fuses with the adjective stem, e.g. *leh + ahaa → lahaa* '(I/he) was having/had'. Interestingly, the negative forms are not auxiliary constructions, e.g. *Má diidnéen* '(We) would not refuse', '(We) would not have refused'.[31] The positive forms are:

	C1	C2	C3
1SG/3SG M	*súgi lahaa*	*karín lahaa*	*joogsán lahaa*
2SG/3SG F	*súgi lahayd*	*karín lahayd*	*joogsán lahayd*
1PL	*súgi lahayn*	*karín lahayn*	*joogsán lahayn*
2PL	*súgi lahaydeen*	*karín lahaydeen*	*joogsán lahaydeen*
3PL	*súgi lahaayeen*	*karín lahaayeen*	*joogsán lahaayeen*

The negative forms, which occur with *má* or *aan*, are:

	C1	C2	C3
1SG/3SG M	*sugéen*	*kariyéen*	*joogsadéen*
2SG/3SG F	*sugtéen*	*kariséen*	*joogsatéen*
1PL	*sugnéen*	*karinnéen*	*joogsannéen*
2PL	*sugtéen*	*kariséen*	*joogsatéen*
3PL	*sugéen*	*kariyéen*	*joogsadéen*

4.3.4.10 Optative

This is a second type of irrealis mood form and is used mainly to express wishes, hopes and blessings, e.g. *Há tago!* 'May he go!', *Ilaahày há ku barakeeyo!* 'May Allah bless you!', and the negative *Khámri yàanu cábbin!* 'May he not drink alcohol!'. The form is also used for strong suggestions and weakened commands, e.g. *Aan saméeyo* 'Let me do it!', *Aynu tágno* 'Let's (INC) go!'. In the positive, third person forms are preceded by the optative sentence type marker (STM) *há*.[32] Other persons are obligatorily preceded by the appropriate subject clitic pronoun, which may occur in the shortened forms *an* for 1SG *aan*; *ad* for 2SG *aad* (also used here for 2PL *aydin*); and *annu* for 1 PL EXC *aannu*. These short forms are more commonly used in speech and are listed below; in writing and print the longer forms are more usual. The positive forms are:

	C1	C2	C3
1SG	*an súgo*	*an karíyo*	*an joogsádo*
2SG	*ad súgtid*	*ad karísid*	*ad joogsátid*
3SG M	*há sugo*	*há kariyo*	*há joogsado*
3SG F	*há sugto*	*há kariso*	*há joogsato*
1PL EXC	*annu súgno*	*annu karínno*	*annu joogsánno*
1PL INC	*aynu súgno*	*aynu karínno*	*aynu joogsánno*
2PL	*ad súgteen*	*ad karíseen*	*ad joogsáteen*
3PL	*há sugeen*	*há kariyeen*	*há joogsadeen*

The negative forms are invariable and are preceded by the sentence type marker *yàan*, which coalesces with a following short form of the appropriate subject clitic pronoun:

	C1	C2	C3
1SG	*yàanan súgin*	*yàanan karín*	*yàanan joogsán*
2SG	*yàanad súgin*	*yàanad karín*	*yàanad joogsán*
3SG M	*yàanu súgin*	etc.	etc.
3SG F	*yàaney súgin*		
1PL EXC	*yàanan(u) súgin*		
1PL INC	*yàynu súgin*		
2PL	*yàanad súgin*		
3PL	*yàaney súgin*		

4.3.4.11 Potential

This is a further irrealis mood form which signifies the possibility of the predication, e.g. *Shòw keenee* 'Maybe he'll bring it', 'Perhaps he'll bring it'. There are no corresponding negative forms. The inflected verb forms are preceded by the sentence type marker *shòw*:

	C1	C2	C3
1SG/3SG M	*sugee*	*kariyee*	*joogsadee*
2SG/3SG F	*sugtee*	*karisee*	*joogsatee*
1PL	*sugnee*	*karinnee*	*joogsannee*
2PL	*sugteen*	*kariseen*	*joogsateen*
3PL	*sugeen*	*kariyeen*	*joogsadeen*

4.3.4.12 Independent past

This is a special form of the past simple which is used for emphasis, usually occurring without nominals as a one word sentence. It is not very commonly used. Unlike other verbal forms it occurs without subject clitic pronouns, focus words or sentence type markers, e.g. *Arké* 'They saw it'. Its interpretation is

the same as for the past simple. There are no negative forms of this verbal paradigm. The forms are:

	C1	C2	C3
1SG	súgay	karíyey	joogsáday
2SG/3SG F	súgtay	karísey	joogsátay
3SG M	súg	karí	joogsáy
1PL	súgnay	karíney	joogsánay
2PL	sugté	karisé	joogsaté
3PL	sugé	kariyé	joogsadé

4.3.4.13 Subordinate clause forms

As described in chapter 8, for the most part verbs in subordinate clauses are distinguished from corresponding main clause forms by a combination of accentual pattern and, to a lesser degree, segmental inflection. The differences are caused by the intersection of two marking systems. The first marks the syntactic distinction between main and subordinate clauses; the second is case marking. Clauses may perform various sentential functions such as subject, object and adverbial and consequently are case marked. Again, as with nominals, this case marking applies to the final or rightmost element in the constituent. Since subordinate clauses usually are verb final, this case marking falls on and influences verbal inflection.[33] In outlining the system it is useful to distinguish between relative clauses on the one hand, and other subordinate clauses like complement clauses, and adverbial clauses introduced by elements like *márka* 'when', *haddíi* 'if', *ilaa* 'until', etc. All clauses share many of the features we describe but relative clauses have some extra formal distinctions, as we shall see.

The first marking system differentiates main clause verbs from subordinate clause verbs in their basic absolutive case. The differences are both segmental and accentual. The segmental differences are limited to forms where the main verb ends in *-aa*; here the subordinate clause forms end in *-o*, or optionally *-id* for 2SG. The paradigms affected include present general, and present progressive, as shown below in the subordinate clause forms of *súg* 'wait for':

	Present general	Present progressive
1SG/3SG M	sugó	súgayó
2SG	sugtó or sugtid	súgaysó or súgaysid
3SG F	sugtó	súgaysó
1PL	sugnó	súgaynó
2PL	sugtàan	súgaysàan
3PL	sugàan	súgayàan

The accentual differences may be informally characterised as a modification of main clause verb accentual patterns:

(84)   (a) The last syllable becomes high tone if short; falling tone if long.
       (b) If the last syllable is already high, as in some negative paradigms, it remains
           so.

Thus we can compare the main clause and subordinate verbs in the (a) and (b) pair below:

(85)   (a)   *Wùu keenaa.*
                 waa+uu   keenaa
                 DM+he   brings:MAIN:PRES GEN:3SG M
                 'He brings it'
         (b)   *ínuu keenó*
                 in-uu   keenó
                 that-he   brings:SUBORD:PRES GEN:3SG M
                 'that he brings it'

     We can show the second marking system, of case, by comparing subject marked and absolutive subordinate verb forms. Segmentally, subject marked verb forms have the same ending as main clause verbs; in other words the change -*aa* to -*o*/-*id* found in absolutive forms does not occur. Accentually, subject marking follows the main rule for nominals: the verb becomes AP3, i.e. any high tones are lowered. Since most main clause positive verbs have AP3, this means that subject marked positive subordinate verbs have the same form as the corresponding main clause forms. However, as can be seen from the paradigms earlier in this section, many main clause negative verb forms have high tones and for these verbs, subject marked subordinate clause forms (with AP3) are distinct from main clause forms, e.g.:

(86)   (a)   *Inaanán arkini waa yàab*
                 in-aan-aan   arkini   waa   yàab
                 that-not-I   saw:NEG:SUBJ   DM   surprise
                 'That I did not see it is surprising.'
         (b)   *Má'aan arkín.*
                 má-aan   arkín
                 not-I   saw:NEG
                 'I did not see it.'

Since the clause in (86a) is subject of the sentence, the verb *arkini* carries the markers of subject case: AP3, i.e. all low tones, and the subject case suffix *-i* described in 4.1.1.4. The corresponding main clause verb *arkín* in (86b) has neither of these inflectional features.

We can see in this asymmetry between positive and negative verbs the result of an interaction between different uses of the same accentual distinction. high versus low is a marker of absolutive versus subject case in nominals, while in the verbal system low tones are characteristic of positive paradigms and high tones of negative paradigms. When the two systems interact in the case marking of clauses, we find that subordinate absolutive forms and main negative forms pattern together accentually, while subordinate subject forms and main positive forms are accentually similar.

All subordinate clauses verbs are case-marked in this way. Relative clause verbs, however, show some extra agreement features compared with other clause verbs. The structure of relative clauses is described in 8.1; we might note in particular that the relative clause follows the nominal it modifies and that there is no relative pronoun: the position of the relativised element, coreferential with the head nominal, is marked by a gap. Our first observation is that, as phrase final elements in the NP, relative clause verbs are marked for the role of their head nominal in the main clause. Compare the following where we schematically represent the relative clause structure (REL = relative clause):

(87) (a)   *Èeg*   [_{NP} *qálabka*   [_{REL} *aannu kú   shaqèynaynó*]] !
            look:IMP:SG   equipment-the   we   with are:working
            'Look at the equipment that we are working with!'

      (b)   [_{NP} *Qálabka* [_{REL} *aannu kú   shaqèynaynaa*]] *waa   dhàwr   jáad.*
            equipment   we   with are:working   DM   several   kind
            'The equipment that we are working with is (of) several kinds.'

In (87a) the head nominal *qálabka* 'the equipment' is object of the verb *èeg* 'look at' and the absolutive case marking occurs on the relative clause verb *shaqèe* 'work'. In (87b) the same head nominal is subject of a verbless sentence and here the relative verb carries subject marking. We can compare the different inflection on the relative clause verb *shaqèe* 'work'.

A second agreement pattern emerges when the gap in the relative clause is subject of the relative clause verb. When this occurs the relative verb shows a reduced agreement pattern, called 'convergent' by Andrzejewski (1968). These reduced agreement forms show a smaller number of person distinctions and, in present tense forms, a segmental change. The person reduction alone can be

shown in the past simple forms of *súg* 'wait for' below, which can be compared with the full paradigm in 4.3.4.3:

(88)   Reduced past simple in relative clauses
        1/2SG/3SG M/2PL/3PL   *sugáy*
        3SG F   *sugtáy*
        1PL   *sugnáy*

These forms have the accentual pattern of absolutive case. The segmental changes which affect present tense forms like present general can be seen below:

(89)   Reduced present general in relative clauses
        1/2SG/3SG M/2PL/3PL   *sugá*
        3SG F   *sugtá*
        1PL   *sugná*

A comparison with the main clause forms in 4.3.4.6 reveals three differences: the verbs in (88) and (89) have the accentual pattern of subordinate clause verbs in absolutive case; they exhibit reduced person differentiation; and the ending *-aa* is reduced to *-a*.

   As with other relative clause verbs, these reduced agreement forms carry the subject marking for the whole NP, as usual both segmentally and accentually. The marking is similar to other subordinate clause case marking: forms ending in *-a* will lengthen to *-aa*; and the accentual pattern will become AP3. Thus we can compare the following, where again the nominal head and modifying relative clause are schematically represented (and [SUBJ] represents the gap subject of the relative clause):

(90)   *Wáxaan raadínayaa* [NP *nimánka* [REL [SUBJ] *kéenayá*]].
        wáxa+aan   raadínayaa   nimán-ka   kéenayá
        what+I   searching:for men-the   are:bringing
        'What I'm searching for are the men who are bringing it.'
(91)   [NP *Nimánka* [REL [SUBJ] *kéenayaa*]] *waa tuugág.*
        nimán-ka   kéenayaa   waa   tuugág
        men-the   are:bringing DM   thieves
        'The men who are bringing it are thieves.'

In both (90) and (91) the subject of the relative clause verb *kèen* 'bring' is a gap coreferential with the head *nimánka* 'the men', resulting in a reduced verb agreement pattern. In (90) the relative clause head is object in the main

sentence and therefore the verb *kéenayá* carries absolute case for the whole noun phrase; in (91) the relative clause head is subject of its main sentence and *kéenayaa* is marked for nominative case. This example enables us to see the syntactic information carried by the inflectional system in verbs. The present progressive form *kéenayá* in (90) signals three independent pieces of syntactic information: (a) that the verb is in a subordinate clause; (b) that its subject is a gap coreferential with the relative clause head; and (c) that this head nominal is not subject of the main clause. [34]

4.3.5 Prefix Verb Paradigms

As mentioned earlier, in Standard Somali there are four prefix or 'strong' verbs: *yidhi* 'say', *yimi* 'come', *yiil* 'be (in a place)' and *yiqiin* 'know', giving them in their reference forms of 3.SG.M past simple. We mentioned too that under pressure from the suffix verbs, there is considerable regional variation in the forms of these verbs. The following are the forms commonly used in northern varieties of the language, with some common variations shown by the use of parentheses. We list the verbal paradigms in the same order as for suffix verbs, abbreviating the description somewhat to avoid repetition. Note that unlike suffix verbs these verbs distinguish between 1SG and 3SG M in simple forms.

4.3.5.1 Imperative
The main forms are as follows:

	yidhi	**yimi**	**yiil**	**yiqiin**
SG	*dhé(h)*	*káalay*	*òol*	*oqòow*
PL	*dháha*	*kaalaya*	*óolla*	*oqáada* (*oqóoda*) (*aqóoda*)

The imperative forms of these verbs are irregular in a number of ways. Firstly we might note that the positive imperative forms of *yimi* 'come' are suppletive, not reflecting the root shape. Secondly the forms given for *yiil* and *yiqiin* are not commonly used since these are stative verbs. They are usually replaced by dynamic verbs, for example, for 'Know it!' speakers commonly use the suffix verb forms *ogòw*! (SG) and *ogáada* (PL) derived from the adjective *óg* 'aware, cognisant'. The negative imperatives, used with the sentence type marker *ha*, are:

	yidhi	**yimi**	**yiil**	**yiqiin**
SG	*ódhan(in)*	*íman(in)*	*óollin*	*óqoon(in)*/ *áqoon(in)*
PL	*odhanína*	*imannína*	*oollína*	*oqoonína*

4.3.5.2   Infinitive
These forms are used in auxiliary constructions.

yidhi	**yimi**	**yiil**	**yiqiin**
odhán	*imán*	*óolli*	*oqóon*

4.3.5.3   Past simple

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*idhi*	*imi(d)*	*iil*	*iqiin*
2SG/3SG F	*tidhi*	*timi(d)*	*tiil*	*tiqiin*
3SG M	*yidhi*	*yimi(d)*	*yiil*	*yiqiin*
1PL	*nidhi*	*nimi(d)*	*niil*	*niqiin*
2PL	*tidhaahdeen*	*timaaddeen*	*tilleen*	*tiqiinneen*
3PL	*yidhaahdeen*	*yimaaddeen*	*yiilleen*	*yiqiinneen*

The negative forms, used with *má* 'not' in main clauses and *aan* 'not' in subordinate clauses are invariable:

odhán	*imán*	*oqóon*	*oqóon*
(*odhanín*)	(*imanín*)	(*oollín*)	(*aqoonín*)

4.3.5.4   Past progressive
As with suffix verbs this is a contracted auxiliary construction.   The verbs *yiil* and *yiqiin* are stative and thus do not occur in this paradigm.

	yidhi	**yimi**
1SG/3SG M	*odhánayey*	*imánayey*
2SG/3SG F	*odhánaysey*	*imánaysey*
1PL	*odhánayney*	*imáneyney*
2PL	*odhánayseen*	*imáneyseen*
3PL	*odhánayeen*	*imánayeen*

Negative forms are invariable:

odhánéyn (*odháneynín*)   *imánéyn* (*imáneynín*)

4.3.5.5   Present general
As with suffix verbs, with stative verbs like *yiil* and *yiqiin* this tense/aspect form signifies a state which obtains at the time of speaking, while with dynamic verbs like *yidhi* and *yimi* it signifies a repeated or regular activity.

	yidhi	yimi	yiil	yiqiin
1SG	idhaahdàa	imaaddaa	aal	aqaan
2SG/3SG F	tidhaahdaa	timaaddaa	taal	taqaan
3SG M	yidhaahdaa	yimaaddaa	yaal	yaqaan
1PL	nidhaahnaa	nimaadnaa	naal	naqaan
2PL	tidhaahdaan	timaaddaan	taalliin	taqaanniin
3PL	yidhaahdaan	yimaaddaan	yaalliin	yaqaanniin

Negative forms:

	yidhi	yimi	yiil	yiqiin
1SG	idhaahdó	imaaddó	áal	aqáan
2SG/3SG F	tidhaahdó	timaaddó	táal	taqáan
3SG M	yidhaahdó	yimaaddó	yáal	yaqáan
1PL	nidhaahnó	nimaadnó	náal	naqáan
2PL	tidhaahdàan	timaaddàan	taallìin	taqaannìin
3PL	yidhaahdàan	yimaaddàan	yaallìin	yaqaannìin

4.3.5.6 Present progressive

As with suffix verbs, the positive forms in this paradigm are contracted auxiliary constructions, while the negative forms remain uncontracted: the negative word *má* fusing with the auxiliary verb *hày* 'have'. Once again the stative prefix verbs do not occur in this progressive paradigm:

	yidhi	yimi
1SG/3SG M	odháneyaa	imáneyaa
2SG/3SG F	odháneysaa	imáneysaa
1PL	odháneynaa	imáneynaa
2PL	odháneysaan	imáneysaan
3PL	odháneyaan	imáneyaan

Negative forms:

	yidhi	yimi
1SG/3SG M	odhán màayó	imán màayó
2SG	odhán màaysó (màaysíd)	imán màaysó(màaysíd)
3SG F	odhán màaysó	imán màaysó
1PL	odhán màaynó	imán màaynó
2PL	odhán màaysàan	imán màaysàan
3PL	odhán màayàan	imán màayàan

4.3.5.7 Auxiliary constructions

The future, conditional and past habitual paradigms are formed, as in suffix verbs, by auxiliary constructions consisting of the main verb infinitive and the

auxiliaries, *dòon*, *léh* and *jír*, e.g. *Wuu imán doonaa* 'he will come', *Wuu odhán lahaa* 'He would say it', 'He would have said it', etc.

4.3.5.8   *Optative*
The positive forms are:

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*an idháahdo*	*an imáaddo*	*an áallo*	*an aqáanno*
2SG	*ad tidháahdo*	*ad timáaddo*	*ad táallo*	*ad taqáanno*
3SG M	*há yidhaahdo*	*há yimaaddo*	*há yaallo*	*há yaqaanno*
3SG F	*há tidhaahdo*	*há timaaddo*	*há taallo*	*há taqaanno*
1PL EXC	*annu nidháahno*	*annu nimaádno*	*annu náallo*	*annu naqáanno*
1PL INC	*aynu nidháahno*	*aynu nimaádno*	*aynu náallo*	*aynu naqáanno*
2PL	*ad tidháahdeen*	*ad timáaddeen*	*ad táalleen*	*ad taqáanneen*
3PL	*há yidhaahdeen*	*há yimaaddeen*	*há yaalleen*	*há yaqaanneen*

Negative forms:

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*yàanan odhán*	*yàanan imán*	*yàanan oollín*	*yàanan oqóon*
2SG	*yàanad odhán*	*yàanad imán*	*yàanad oollín*	*yàanad oqóon*
3SG M	*yàanu odhán*	etc.	etc. etc.	
3SG F	*yàaney odhán*			
1PL EXC	*yàanan(u) odhán*			
1PL INC	*yàynu odhán*			
2PL	*yàanad odhán*			
3PL	*yàaney odhán*			

4.3.5.9   *Potential*
This paradigm has the same meaning of possibility as with suffix verbs, e.g. *Shòw yimaaddee* ' Maybe/Perhaps he will come.' The forms occur with the sentence type marker *shòw*, again with no corresponding negative forms:

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*idhaahdee*	*imaaddee*	*aallee*	*aqaanee*
2SG/3SG F	*tidhaahdee*	*timaaddee*	*taallee*	*taqaanee*
3SG M	*yidhaahdee*	*yimaaddee*	*yaallee*	*yaqaanee*
1PL	*nidhaahnee*	*nimaaddee*	*naallee*	*naqaanee*
2PL	*tidhaahdeen*	*timaaddeen*	*taalleen*	*taqaanneen*
3PL	*yidhaahdeen*	*yimaaddeen*	*yaalleen*	*yaqaanneen*

4.3.5.10 Independent past

This is used as with suffix verbs; all the forms are uncommon but the stative verbs *yiil* and *yaqaan* are seldom heard in this paradigm. When they are used, their forms show some analogical pressure from the suffix verb structure, as can be seen by comparison with past simple forms (4.3.4.2).

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*ídhi*	*ími(d)*	*áallay*	*iqíinay*
2SG/3SG F	*tídhi*	*tími(d)*	*táallay*	*tiqíinay*
3SG M	*yídhi*	*yími(d)*	*yaalláy*	*yiqiináy*
1PL	*nídhi*	*ními(d)*	*náallay*	*niqíinay*
2PL	*tidhaahdé*	*timaaddé*	*taallé*	*tiqiiné*
3PL	*yidhaahdé*	*yimaaddé*	*yaallé*	*yiqiiné*

4.3.5.11 Subordinate clause forms

The marking of subordinate forms follows basically the same pattern as for suffix verbs: subordinate clause verbs are distinguished from main clause verbs and will carry the case marking for the role of the clause in the main sentence. Relative clause forms show reduced agreement when their subject is coreferential with the head of the clause. The effect of this reduction is shown below for present general and past simple forms, where the forms exhibit the default absolutive case:

(92)   Reduced present general in relative clauses

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*idhaahdá*	*imaaddá*	*aal*	*aqáan*
2SG/3SG M/2PL/3PL	*yidhaahdá*	*yimaaddá*	*yáal*	*yaqáan*
3SG F	*tidhaahdá*	*timaaddá*	*táal*	*taqáan*
1PL	*nidhaahná*	*nimaadná*	*náal*	*naqáan*

(93)   Reduced past simple in relative clauses

	yidhi	**yimi**	**yiil**	**yiqiin**
1SG	*idhí*	*imí(d)*	*iil*	*iqíin*
2SG/3SG M/2PL/3PL	*yidhí*	*yimi(d)*	*yiil*	*yiqíin*
3SG F	*tidhí*	*timí(d)*	*tiil*	*tiqiin*
1PL	*nidhí*	*nimí(d)*	*níil*	*niqiin*

As usual, when these forms are subject marked the high tones will be lowered. In addition any short final *-a* vowel in the present general will lengthen to *-aa*.

4.3.6 The copula verb yahay

The verb *yahay* 'be' was introduced in section 4.3.1 where we gave its present general and past simple forms, noting that the verb mixes morphological features of suffix and prefix type verbs. Below we give the negative forms of these paradigms, along with other important forms. Note that as a stative verb *yahay* does not normally occur in progressive paradigms, though it can be used as an imperative. The compound tenses, future, conditional and past habitual are formed as in other verbs by auxiliary constructions consisting of the infinitive and the auxiliaries, *dòon*, *léh* and *jír*, and are not listed here, except for the negative conditional which is not an auxiliary construction.

4.3.6.1   Imperative
The positive forms are: *ahàw* (SG); *aháada* (PL). The negative forms, used with *ha*, are: *áhaan*(*in*) (SG); *ahaanína* (PL).

4.3.6.2   Infinitive
As with other verbs this is used in compound tenses: *aháan*.

4.3.6.3   Past simple
The negative forms are invariable: *ahàyn*.

4.3.6.4   Present general
The negative forms, used with *má* or *aan*, are:

1SG	*ihí*
2SG	*ihíd*
3SG & PL	*ahá*
1PL	*ihín*
2PL	*ihidín*

4.3.6.5   Conditional
The negative forms, used with *má* or *aan*, are:

1SG/3SG M/3PL	*ahaadeen*
2SG/3SG F/2PL	*ahaateen*
1PL	*ahaanneen*

4.3.6.6   Optative
This is used as with other verbs, e.g. *Gabádh quruxsán há ahaato*! 'May she be a beautiful girl!'. The positive forms are:

1SG	*an aháado*
2SG	*ad aháato*
3SG M	*há ahaado*
3SG F	*há ahaato*
1PL EXC	*annu aháanno*
1PL INC	*aynu aháanno*
2PL	*ad aháateen*
3PL	*há ahaadeen*

The negative form is the invariable *ahaannín*, preceded by the sentence type marker *yàan* fused with the subject clitic pronoun as with other optatives, i.e. *yàanan, yàanad, yàanu,* etc.

4.3.6.7   Potential
This paradigm has the same uses as with other verbs, e.g. *Shòw nácas ahaadee*, 'Perhaps he is a fool'. The forms are:

1SG/3SG M	*ahaadee*
2SG/3SG F	*ahaatee*
1PL	*ahaannee*
2PL	*ahaateen*
3PL	*ahaadeen*

4.3.6.8   Subordinate clause forms
There is no distinction between main and subordinate clause forms of *yahay* except for relative clauses, where the verb shows a reduced agreement pattern or restrictive form, as described in 4.3.4.13, when agreeing with the relative clause subject. The restrictive form of the present general is the invariable form *á(h)* and the forms of the past simple are given below:

1SG/2SG /3SG M/2PL /3PL	*aháa*
3SG F	*aháyd*
1PL	*aháyn*

These are the absolute case forms which occur when the relative clause is not subject of the main clause; when the clause is subject, all high tones will be lowered and the present general form *a(h)* will carry the nominative suffix *-i*, becoming *ihi* by vowel assimilation.

4.4   Adjectives

Adjectives as a class are characterised by occurring attributively as post-nominal modifiers, as in (94) below, and predicatively as complements of the copula *yahay*, as in (95). They are differentiated from nouns in the following ways. They do not occur with suffixed determiners, they do not have inherent number and gender, and when they do occur as complements of *yahay*, unlike nouns, they are positioned between the satellite clitics and the verb (6.1.4), as we shall see. They are unlike verbs in that they occur as complements of *yahay*, which does not function as an auxiliary verb. In addition, as described in 3.5, adjectives may mark plural agreement with a nominal head by reduplication, which does not occur with this function in verbs.

(94)   (a)   *náag   dhèer*
             woman   tall
             'a tall woman'
       (b)   *méel   fóg*
             place   distant
             'a distant place'
(95)   (a)   *Wuu yár yahay*
             waa+uu   yár   yahay
             DM+he   small   is
             'He is small'
       (b)   *Way yaryár yihiin*
             waa+ay   yaryár   yihiin
             DM+they   small   are
             'They are small'

In (95) the adjective *yár* occurs between the subject clitic pronouns *uu* and *ay* and the verb *yahay*; neither nouns nor verbs may occur thus in this construction. In (95b) we see the plural marked by reduplication.

   In predicative uses adjectives coalesce with the copula *yahay*. When the copula stem begins with the prefixed agreement markers, as in present general forms, this is a stylistic optional rule, as for the adjective *wanaagsán* 'good' in (96) below. However, when the copula stem begins with *ah-*, as in the past simple, this coalescence is obligatory, as in (97) and (98):

(96)   (a)   *Wuu wanaagsán yahay*
             waa+uu   wanaagsán   yahay
             DM+he   good   is
             'He is good'

(b)   *Wuu wanaagsányay*
'He is good'

(97)   (a)   **Wuu wanaagsán ahaa*
waa+uu   wanaagsán   ahaa
DM+he   good   was
'He was good'

(b)   *Wùu wanaagsanaa*
'He was good'

(98)   (a)   **nínkíi wanaagsán aháa*
nín-kíi   wanaagsán   ahaa
man-the   good   was
'the man who was good'

(b)   *nínkíi wanaagsanáa*
'the man who was good'

As can be seen, both the optional and obligatory coalescence rules delete the segment *ah* from the copula and cliticize the remnant onto the adjective stem. This has led some writers, for example Andrzejewski (1969) and Banti (1994), to recognize these coalesced predicative adjectives as a form of verb, presumably under the formal criterion that they bear (*yahay*'s) inflection and also perhaps from the understandable reluctance to posit a synchronic obligatory coalescence rule. If recognised as such, they form an unusual type of verb, or a category that flips between adjective and verb depending on context, perhaps reflecting an ongoing historical process, hence Andrzejewski's (1969) term 'hybrid verb'; see Saeed (1988) and Banti (1994) for further discussion. The attributive and predicative uses of adjectives are discussed further in chapter 6.

4.4.1 Basic adjectives

The number of basic adjectives is so small that we can provide a comprehensive list as follows, where we also give the past tense forms produced by coalescence with *yahay*. Note that this coalescence often produces the environment for Stem Contraction (3.1.4).

Adjective	Past tense (fused with *yahay*)	
adág	*adkaa*	'hard, difficult'
bisíl	*bislaa*	'ripe'
cád	*caddaa*	'white, clear'
cás	*casaa*	'red'
ceedhíin	*ceedhíinaa*	'raw'
culús	*cuslaa*	'heavy'
cusúb	*cusbaa*	'new'

dihín	*dihnaa*	'having udders full of milk'
dhán	*dhammaa*	'complete, all'
dhèer	*dheeraa*	'tall, long, deep'
dhów	*dhowaa*	'near'
ú…ég	*ú…ekaa*	'similar…to'
kú…ég	*kú…ekaa*	'appropriate, complete'
feeyíg	*feeyigaa*	'alert, cautious'[35]
fóg	*fogaa*	'far, distant'
fudúd	*fududaa*	'light, easy'
idíl	*idlaa*	'all'
irmáan	*irmaanaa*	'in milk'
jecél	*jeclaa*	'liking, loving'
kulúl	*kululaa*	'hot'
lá'	*la'aa*	'lacking, missing'
le'ég (*léeg*)	*le'ekaa* (*leekaa*)	'equal (in size)'
léh	*lahaa*	'owning, possessing'
mác	*macaa*	'sweet, plump'
macáan	*macaanaa*	'sweet'
madów	*madoobaa*	'black'
mòog	*moogaa*	'ignorant of, unaware'
necéb	*nebcaa* (*necbaa*)	'hating'
nég	*negaa*	'permanently settled, stable'
nòol	*noolaa*	'alive, living'
nugúl	*nuglaa*	'delicate, not hardy'
óg	*ogaa*	'aware, cognisant'
oggól	*oggolaa*	'agreeing'
qabów	*qaboobaa*	'cold'
qadháadh	*qadhaadhaa*	'bitter'
qaydhíin	*qaydhiinaa*	'raw'
ròon	*roonaa*	'very good, better'
sán	*samaa*	'good'
shilís	*shislaa* (*shilisaa*)	'fat'
wèyn	*weynaa*	'big'
xún	*xumaa*	'bad'
yár	*yaraa*	'small'

The derivational affixes which productively create adjectives are discussed in chapter 5.

4.4.2 Comparative and superlative

In comparative and superlative constructions the morphological form of adjectives does not change; instead they form constructions with adpositional

clitics. The adposition *ká* 'from' (4.5) is the marker of the comparative, as in (99) below where we loosely gloss it as 'than'. The word order is: subject first, standard second as in (99a), though of course the standard can be implicit as in (99b):

(99) (a) *Náagtaasu náagtán way ká dhèer tahay.*
         Náag-taas-u   náag-tán   waa+ay   ká   dhèer tahay
         woman-that-NOM woman-this DM+she than tall is
         'That woman is taller than this woman.'

    (b) *Náagtaasu way ká dhèer tahay.*
         Náag-taas-u   waa+ay   ká   dhèer tahay
         woman-that-NOM DM+she than tall is
         'That woman is taller.'

The marker of the superlative is the adpositional cluster *ugú* (←*ú*+*kú* 'to, for'+ 'in, at' or *ú*+*ú* as described in 3.3.3) as in (100). Since this cluster cannot be easily translated into English we gloss it simply as SUPR for 'superlative'.

(100) (a) *Afsoomáaligu waa áfka ugú adág afáfka aan bartáy.*
      Af-soomáali+ku   waa áf-ka   ugú adág   afáf-ka
      language-Somali+the DM   language-the SUPR **difficult languages-the**
      aan   bartáy
      I   studied
      'The Somali language is the most difficult language of the languages I have studied.'

    (b) *Afsoomáaligu wuu ugú adágyahay.*
      af-soomáali+ku   waa+uu   ugú   adág-yahay
      language-Somali+the   DM+it   SUPR **difficult-is**
      'The Somali language is the most difficult.'

4.4.3 Attributive adjectives

While most adjectives can occur both attributively and predicatively, there is a small group that occurs only attributively. Many, but not all, of these are derived from nouns by a suffix -*e*, for example *dhexé* 'between' from *dhéx* N F 'middle, centre'. Some other important examples are:

        dambé   'behind, later, last, next'
        hoosé   'lower' (cf. *hóos* N F 'lower part, bottom')
        horé   'first, former, previous' (cf. *hór* N F 'front')
        kalé   'other'

kastá	'each'
koré	'upper' (cf. *kór* N M 'top, surface')
saré	'topmost, uppermost, high'
shishé	'far, further, beyond'
soké	'near, close'
wál	'each, every'

Nouns modified by the adjective *wál* are usually followed by the adverbial *ba* 'any' (4.10.4), i.e. *nín walba* 'each man, every man'. For some speakers this combination is frozen into a composite attributive adjective. This can be seen from subject marking, where instead of *nín waliba* with the nominative suffix *-i* attaching to the adjective (as the last element in the nominal), the form used is *nín walbi* (←*walba-i*), where the nominative case suffix attaches after the adverbial.

Another element which is intermediate between an attributive adjective and an adverbial is the word *uun* 'some, other, only, just, mere(ly)', discussed in 4.10.5. This word appears adjective-like as a post-nominal modifier, e.g.:

(101)   *már uun*   'sometime or other'   (*már* N M 'time')
           méel uun   'someplace or other' (*méel* N F 'place'
           wáx uun   'something or other' (*wáx* N M 'thing')

4.4.4 Agreement features

Adjectives show no gender differentiation in agreement with nouns but the reduplication process described in 3.5 marks plural agreement, e.g. *nimán yaryár* 'small men', *naagó waawèyn* 'big women'. The use of these reduplicated plural forms is subject to some variation, with the tendency being that the longer the adjective the less likely speakers are to use the reduplicated forms. Because of the phrasal operation of case marking in Somali (described in 4.1.1.4) modifying adjectives do not agree in case with the head noun: as with all noun phrases only the last element carries case marking. If the adjective is the last element then subject marking will apply in the usual way, i.e. to lower any high tones in the adjective and since all regular adjectives end in a consonant, to add the nominative suffix *-i* described for nouns in 4.1.1.4. See for example:

Absolutive case	Nominative case	
dhèer	*dheeri*	'tall'
wanaagsán	*wanaagsani*	'good'
yár	*yari*	'small'

Note that the purely attributive adjectives, except for *wál* 'each', end in vowels, and thus do not show the nominative suffix but still have lowered tones, e.g. *saré* ABS 'topmost' → *sare* NOM (but *wál* ABS 'each' →*wali* NOM).

4.5   Verbal adpositions

In Somali as in other closely related languages like Boni, Rendille, Elmolo, Dasenach (Appleyard 1990) and Arbore (Hayward 1984)[36] there are pre-verbal morphemes which semantically govern oblique elements of the predication much in the way that nominal adpositions do in other languages.   In fact these elements functionally parallel nominal postpositions and oblique case suffixes in less closely related Cushitic languages like Oromo (Owens 1985).   The four morphemes are given below with their main meanings:

ú	'to, for'
kú	'in, into, on, at, with (by means of)'
ká	'from, away from, out of'
lá	'with (in company with)'

As we can see, the English glosses are prepositions and these pre-verbal morphemes serve very similar functions.   These include marking nominals as having the semantic roles of location (102) & (103), source (104), goal (105), beneficiary (106), instrument (107), manner (108) and comitative (109), where the morphemes and relevant governed nominals are underlined:

(102)   *Wúxuu ká shaqeeyaa warshádda.*
      Wáxa+uu   ká   shaqeeyaa   warshád-da
      what+he   at   works   factory-the
      'He works at the factory, Where he works is at the factory.'

(103)   *Magaaládii bàa lagú arkay.*
      magaaló-díi   bàa   la-kú   arkay
      town-the   FOC   one-in   saw
      'One saw him in the town, He was seen in the town.'

(104)   *Nínkán bàa Cadán ká yimí.*
      nín-kán   bàa   Cadán ká   yimí
      man-this   FOC   Aden   from   came
      'This man came from Aden.'

(105)   *Yày ú tagtay?*
      Yàa+ay   ú   tagtay
      who+she   to   went
      'Who did she go to?'

(106)   *Cáli shàah <u>ú</u> samèe!*
         Cáli   shàah   ú   samèe
         Ali   tea   for   make:IMP
         'Make some tea <u>for</u> Ali!'

(107)   *Wúxuu sameeyey <u>dár</u> uu <u>kú</u> waraabiyó géela.*
         wáxa+uu   sameeyey   dár   uu   kú   waraabiyó   géel-a.
         what+he   made   trough he   with   water   camels-the
         'What he made was <u>a trough</u> <u>with</u> which to water the camels.'

(108)   *Dégdeg <u>ú</u> samèe!*
         dégdeg   ú   samèe
         haste   with   do:IMP:SG
         'Do it <u>with</u> <u>haste</u>!, Do it quickly!'

(109)   *Kumàad <u>lá</u> joogtay?*
         kuma+bàa+aad   lá   joogtay
         who+FOC+you   with   stayed:2SG
         '<u>Who</u> did you stay <u>with</u>?'

As described in Saeed (1993b) these constructions have the following characteristics:

(110)   (a)   The morpheme remains in pre-verbal position regardless of the position of the NP it semantically governs;

       (b)   This governed NP occurs in the basic ABSOLUTIVE form (the default case used for objects, etc.), whichever morpheme occurs;

       (c)   When more than one non-subject NP is involved, these structures can give rise to ambiguity, which is resolved by contextual information.

The semantic and functional similarity between these morphemes and prepositions and postpositions in other languages is clear; and indeed there have been attempts, for example Biber (1984a), to derive them historically from stranded postpositions. However, their synchronic syntactic position as members of the pre-verbal cluster leads to difficulties in labelling the category. In Saeed (1993a), for example, written for English-speaking learners I followed Bell (1953) in calling them prepositions, changing in Saeed (1993b) to adpositions. Other terms used for them have included prepositional particles (Andrzejewski 1960), locative particles, and case markers (Biber 1984a). If these morphemes only governed non-arguments of the verb one might term them 'adverbial clitics'. However, as with English prepositions, some verbs require them to introduce a second or third obligatory argument (for example *kú...ríd* 'put in'). In still others they are lexicalised or 'frozen' into verb-particle units which are semanticaly opaque (for example *lá...báx* 'elope', literally 'leave

with', or *ká...céli* 'defend', literally 'send back from'). For present purposes we will continue to call them verbal adpositions, while recognising the shortcomings of the label. The position of adpositions in the verbal group is described in chapter 6, and the coalescence rules which affect them are described in chapter 3.

We might note that there are other ways of marking spatial and temporal relations: firstly, Somali has borrowed one nominal preposition from Arabic, *iláa* 'until', which therefore forms a class of its own. It is used pre-nominally to form prepositional phrases, e.g. *iláa berríto* 'until tomorrow', *iláa Hargeysá* 'as far as Hargeysa'. Secondly, a very productive way of describing location involves nominal possessor-possessed constructions using the suffixed possessive determiners, as described in 4.6.4, for example:

(111)   *gúriga hóostìisa*
        gúri-ga   hóos-tìisa
        house-the   underneath-its
        'underneath the house' lit. 'the house its underneath'
        (cf. *hóos* N F 'underneath')
(112)   *gúriga hórtìisa*
        gúri-ga   hór-tìisa
        house-the   front-its
        'in front of the house' lit. 'the house its front' (cf. *hór* N F 'front')

In these constructions both elements are noun phrases. Such nouns of location, *ág* N F 'near(ness)', *dhéx* N F 'middle, centre', *hóos* N F 'underneath', *hór* N F 'front', etc. may also function as adverbs: occurring immediately before the verb in the verbal group to modify the meaning of the head verb, for example:

(113) *biyaha   wasakhda ah   laga   dhex   helaa*   [BCH 25.11-12]
      water-the dirt-the   is   one-in middle finds
      'one finds (them) within dirty water'

4.6   Determiners

The determiners are a class of elements that are suffixed to nouns. There are four types: definite articles, demonstratives, interrogatives and possessives. Each type has a masculine form beginning with *k-* and a feminine form beginning with *t-* to agree with the head noun's gender. The determiners do not show plural agreement with the head noun. Their initial consonants undergo various sandhi processes when attached to nominal stems, as described in 3.2.1. Many

of the determiners can stand alone as verbal arguments like a kind of pronominal, as we shall see.

4.6.1 Definite articles

There are two definite articles, carrying a distinction of distance in space and time, shown with examples below:

(114)   Non-remote article   Remote article
Masculine   *ka*   *kíi*
Feminine   *ta*   *tíi*

    náagta 'the woman (non-remote)'   *náagtíi* 'the woman (remote)'
    nínka 'the man (non-remote)'   *nínkíi* 'the man (remote)'

If the referent of the noun is distant from the speaker, or is in the past, the remote forms *kíi/tíi* are used; the article *ka/ta* is unmarked spatially, and is used for referents temporally located in the present and the future. This distinction has a discourse analogue: the determiner *kíi/tíi* is used as a form of discourse anaphora, picking out entities introduced earlier in the discourse, as described in chapter 9.

    When the article is the last element in the noun phrase it will carry subject marking, and undergo the changes exemplified below:

(115)   Absolutive   Subject
    nínka   *nínku*
    náagta   *náagtu*
    nínkii   *nínkii*
    náagtíi   *náagtíi*

The article *ka/ta* undergoes a segmental change to *ku/tu*, while the article *kíi/tíi* undergoes a tonal lowering.

    The articles may occur without a head nominal as a kind of pronominal, e.g. (116); in this function they may occur in the plural forms in (117), undifferentiated for gender:

(116) (a)   *kíi*   *yimi*
          the:M came
          'the one (MASC) who came'
    (b)   *tíi*   *kalé*
          the:F other
          'the other one (FEM)'

(117) (a) *kúwa* 'the ones (non-remote)'
       (b) *kúwii* 'the ones (remote)'

There is no indefinite article, nor as exist in some Cushitic languages, paucal or individuating suffixes.[37] There is however one element *ku* 'one' (feminine *tu*; plural *kuwo*) with a very limited distribution, which may be a relic of such an element. This word does not occur suffixed to nominals but as the head of a few constructions, especially with the adjective *kalé* 'other', e.g.:

(118) *ku kalé*   'another one (MASC)'
       kuwo kalé   'some others, some more'

4.6.2 Demonstratives

There are four demonstrative determiners, which are suffixed to the noun and which undergo the same rules of sandhi as other determiners (3.2.1). The demonstratives divide up the space around the speaker into four zones:

(119) | Masculine | Feminine | |
       |-----------|----------|---|
       | *kán* | *tán* | 'this, these' (close to speaker) |
       | *káa(s)* | *táa(s)* | 'that, those' (further away from speaker) |
       | *kéer* | *téer* | 'that, those' (in the middle distance) |
       | *kóo* | *tóo* | 'that, those' (in the far distance) |

The basic use then is to provide four zones of increasing distance from the speaker. The demonstratives *kán/tán* and *káas/táas* are also used for temporal and discourse deixis.

When carrying the phrasal subject marking for the NP, these demonstratives undergo the changes exemplified below, with optional variants in parentheses:

(120) | Absolutive | Subject | |
       |------------|---------|---|
       | *nínkán* | *nínkanu (nínkani)* | 'this man' |
       | *náagtán* | *náagtanu (náagtani)* | 'this woman' |
       | *nínkáa(s)* | *nínkaasu (nínkaasi)* | 'that man' |
       | *náagtáa(s)*| *náagtaasu (náagtaasi)* | 'that woman' |
       | *nínkéer* | *nínkeeri* | 'that man' (middle distance) |
       | *náagtéer* | *náagteeri* | 'that woman' (middle distance) |
       | *nínkóo* | *nínkooyi* | 'that man' (far distance) |
       | *náagtóo* | *náagtooyi* | 'that woman' (far distance) |

The demonstratives can also occur alone as pronominal forms, and in this role have plural forms which are not differentiated for gender, e.g.:

(121)   Absolutive   Subject
          kúwán   *kúwanu (kúwani)*   'these ones'
          kúwáa(s)   *kúwaasu (kúwaasi)*   'those ones'
          kúwéer   *kúweeri*   'those ones' (middle distance)
          kúwóo   *kúwooyi*   'those ones' (far distance)

4.6.3 Interrogatives

The interrogative determiners *kée* 'which (masc)?' and *tée* 'which (fem)?' are used to question nouns, for example:

(122)   *ninkée*   'which man?'   *naagtée* 'which woman?'

These determiners undergo the same rules of sandhi as other determiners but, unusually, lower any high tones in the noun to which they are attached, as in the above examples. Again they may be used alone, when a gender-neutral plural form occurs:

(123)   MASC SG   FEM SG   PL
         'which one(s)?'   *kée*   *tée*   *kuwée*

It is worth mentioning here that nouns may also be questioned by suffixing the question word *ma*, also used to question sentences, which can also be attached to independent pronouns and the word *ku/tu* 'one' described in the last section, for example:

(124)   *nínma?*   'which man?'
         góorma?   'which time?' 'when?'
         idinma?   'which of you (PL)?'
         kúma?   'which one (MASC)?' 'who (MASC)?
         túma?   'which one (FEM)?' 'who (FEM)?'

4.6.4 Possessives

The possessive determiners show once again the class features of being suffixed to a head noun, of agreeing with it in gender and of undergoing the sandhi rules in 3.2.1. The set is:

(125)   Masculine   Feminine
              kàyga   *tàyda*   'my'
              kàaga   *tàada*   'your (SG)'
              kìisa   *tìisa*   'his'
              kèeda   *tèeda*   'her'
              kayága   *tayáda*   'our (EXC)'
              kèenna   *tèenna*   'our (INC)'
              kìinna   *tìinna*   'your (PL)'
              kòoda   *tòoda*   'their'

For example:

              gúrigàyga   'my house'   (cf. *gúri* N M 'house')
              fáraskìisa   'his horse'   (cf. *fáras* N M 'horse')
              gèelayága   'our (EXC) camels'   (cf. *gèel* N M COLL 'camels')

These forms are interesting in their morphology. Firstly we may note that *k-* forms attach to masculine nouns and *t-* forms to feminine nouns as usual with determiners. However these forms consist of two elements: a possessive determiner followed by the definite article *ka/ta* described in 4.6.1, e.g. *kày-ka* → *kàyga* 'my'. In this case the definite article undergoes normal determiner sandhi but there are also some unusual rules. Firstly, wherever the possessive element ends in a consonant other than *n* the first consonant of the article is dropped, e.g. *kìis-ka* → *kìisa* rather than the predicted **kìiska*. Secondly when the possessive ends in *n*, the initial consonant of the article assimilates to *n*, e.g. *kèen-ka* → *kèenna* 'our (INC MASC)', *tèen-ta* → *tèenna* 'our (INC FEM)'.

This combination of possessive and article is the basic form but possessives occur with the full range of determiners, with associated meanings, for example:

(126)   remote article *kíi/tíi*:   *gúrigàagíi*   'your house (remote)'
        demonstrative *káas/táas*:   *gúrigàagáas*   'that house of yours'
        interrogative *kée/tée*:   *gurigaagée?*   'which house of yours?'

The possessive determiners occur without the article *ka/ta* to mark a semantic distinction similar to the distinction between alienable and inalienable possession found in many languages: possessives without suffixed articles are used with nouns for family members, parts of the body, and by extension for close friends:

(127)   *aabbáháy*   'my father'
             hooyádáa   'your mother'
             lúgtéed   'her leg'

     saaxìibkíin   'your (PL) friend'

When possessives occur in this way without a following second determiner, they are noun phrase final elements and must carry case marking, exhibited by accentual changes. The accentual patterns on the pre-article possessive determiners in (125) above on the other hand show no case marking[38]. When possessives are noun phrase final, they show a distinction between absolutive and nominative case as follows (shown on *aabbé* N M 'father'):

(128)   Absolutive   Nominative
          aabbáháy   *aabbàhày*   'my father'
          aabbáháa   *aabbahàa*   'your (SG) father'
          aabbíhìis   *aabbihìis*   'his father'
          aabbáhéed   *aabbahèed*   'her father'
          aabbáháyo   *aabbahayo*   'our (EXC) father'
          aabbáhéen   *aabbahèen*   'our (INC) father'
          aabbíhíin   *aabbihìn*   'your (PL) father'
          aabbáhóod   *aabbahòod*   'their father'

The possessive determiners perform some important semantic functions. As shown below, they are used to mark partitive quantification, including the use of numerals. In this function the possessives are used in their inalienable form.

(129)   *dádka giddigóod*   (cf. *gíddi* N M 'all, entirety')
       dád-ka   gíddi-kóod
       people-the   all-their
       'all of the people' (lit. 'the people their entirety')
(130)   *dádka qàarkóod*   (cf. *qàar* N M 'portion')
       'some of the people' (lit. 'the people their part')
(131)   *dádka bádidóod*   (cf. *bádi* N M 'multitude, majority')
       'most of the people'
(132)   *dádka áfartóod*   (cf. *áfar* NUM M 'four')
       'four of the people'
(133)   *áfartéen*
       'we four, the four of us'

They also mark spatial and temporal locations:

(134)   *gúriga dabádìisa*   (cf. *dabó* N F 'rear, behind')
       gúri-ka   dabó-tìis-a
       house-the   front:F-its-the
       'behind the house' (lit. 'the house its rear')

(135) *gúriga kórkìisa*   (cf. *kór* N M 'top')
     'on top of the house'

(136) *gúriga gúdihìisa*   (cf. *gúde* N M 'interior, iside')
     'inside the house'

(137) *gúriga ágtìisa*   (cf. *ág* N F 'nearness, proximity')
     'near the house'

(138) *guryáha dhéxdòoda*   (cf. *dhéx* N F 'middle, centre')
     'between the houses'

(139) *labá   bilóod   hórtèeda*
     two   months   front-its
     'two months ago'

(140) *labá   bilóod   dabádèeda*
     two   months   behind-its
     'after two months'

Note that in these spatial and temporal uses the possessives occur in their alienable form.

4.7 Focus words

There are three words which function at the pragmatic level of information structure: *bàa*, *ayàa* and *wáxa(a)*.[39] They mark certain sentence constituents, such as NPs, as being in focus. The pragmatic role of focus and the elements to which it can be applied are discussed in 9.1, but briefly, focus is typically applied to elements introduced as new information or assigned contrastive emphasis. The first two words, *bàa* and *ayàa*, are lexically empty morphemes which follow the constituent in focus, e.g. (where as usual the gloss uses small capitals for focus):

(141) *Nimán   bàa   yimí.*
     men   FOC   came
     'SOME MEN came.'

(142) *Nimán   ayàa   yimí.*
     men   FOC   came
     'SOME MEN came.'

As described in 3.3.2 the main difference between these words seems to be that *ayàa* undergoes fewer rules of coalescence and is consequently associated with more formal and measured speaking styles.

    The insertion of *bàa* or *ayàa* after a constituent, such as an NP in (141-2) above, is not a syntactically inert process: these constructions show a range of

similarities to relative clauses, which has led some writers, e.g. Saeed (1984), to formally relate the structures to clefts. The third focus word *wáxa(a)* is clearly related to the noun *wáx* 'thing'. As described in 7.4, the construction in which it occurs has been analysed in various ways. In one view it is related to a relative clause on the noun phrase *wáx+ka* → *wáxa* 'the thing'. Compare for example the normal cleft in (143) below, where the main clause is a verbless sentence with the sentence type marker *waa*, and the focus construction with *wáxa(a)* in (144), without *waa*:

(143)   *Wáxa yimí waa baabùur.*
        wáx+ka   *yimí*   *waa*   *baabùur*
        thing+the   came   DM   truck
        'The thing which came was a truck.'
(144)   *Wáxa yimí nimán.*
        wáxa   *yimí*   *nimán*
        what/who   came   men
        'Who came was SOME MEN.', 'There came SOME MEN.'

These constructions are discussed in chapters 7 and 9, but briefly, the element *wáxa* in (144) has no referential function and is semantically 'bleached' so that there are no selectional restrictions between it and the complement of the cleft, unlike in (143). This complement is informationally equivalent to an item focused by *bàa* or *ayàa*. Along with this semantic bleaching, *wáxa(a)* is reanalysed as a unitary word: it is not possible, as with the lexical NP *wáx* 'thing' in (143) to attach the normal range of deteminers in place of the article *-(k)a*, i.e. *wáxán* 'this thing', *wáxáas* 'that thing', etc.

    Some writers, for example Andrzejewski (1975a) and Antinucci and Puglielli (1980), have identified as a corresponding verbal focus word the morpheme *waa*. However this approach causes various syntactic and semantic problems, described in 9.1.3, and we will view *waa* as one of the set of main clause sentence type markers, described in the next section.

4.8   Sentence type markers

The term sentence type marker is used for the set of morphemes which, together with verbal inflection, identify clauses of different types, i.e. declarative, imperative, conditional, optative, etc. As described in 4.3 the sentence type system interacts with other semantic systems like tense, aspect and negation to produce verbal paradigms of some inflectional complexity. The syntactic distinction between main and subordinate clauses is also important since the

range of tense, aspect and modality distinctions is much reduced in subordinate clauses, as are inflectional patterns of agreement. These features are discussed in chapter 8. For now, we describe the system applying to main clauses without nominal focus.

For clarity, we distinguish between positive and negative clauses; (145) below gives the sentence type markers and associated 3 SG MASC forms of *súg* 'wait for' found in positive clauses. The full paradigms are given in section 4.3 earlier.

(145) Types of positive main clauses

SENTENCE TYPE	MARKER	VERB FORM	MEANING
Declarative	*waa*	*sugaa*	(He) waits
Interrogative	*ma*	*sugaa*	Does (he) wait?
Imperative	-	*súg* (SG) / *súga* (PL)	Wait (for it)!
Conditional	*waa*	*súgi lahaa*	(He) would wait
Optative	*há*	*sugo*	May (he) wait!
Potential	*shòw*	*sugee*	(He) may wait

In negative clauses the range of types and markers is slightly reduced:

(146) Types of negative main clauses

SENTENCE TYPE	MARKER	NEGATIVE WORD	VERB FORM	MEANING
Declarative	-	*má*	*sugó*	(He) doesn't wait
Interrogative Type 1	*ma*	*áan-u*	*sugín*	Didn't he wait?
Interrogative Type 2	*sòw*	*má*	*sugó*	Doesn't (he) wait?[40]
Imperative	*ha*	-	*súgin/suginína*	Don't wait!
Conditional	-	*má*	*sugéen*	(He) would not wait
Optative	*yàan-u*	-	*súgin*	May (he) not wait!

Only interrogative clauses have an independent sentence type marker and negative word: for the most part either morpheme and the verbal form identify the type of sentence in negative clauses.

The class abbreviation used for sentence type markers is STM, with DM representing the declarative marker; QM, the interrogative marker; IM, the imperative marker; CM, the conditional marker; OM, the optative marker; and PM, the potential marker. Abbreviations for negative markers are prefixed by N, for example *ha* NIM, the negative imperative marker. Though it makes sense to

view the negative words *má* (in main clauses) and *aan* (in subordinate clauses) as members of the set of sentence type markers, we will gloss them separately as 'not'.

4.9   Conjunctions

There are a number of different conjunctions which link various types of grammatical category. The basic set is as follows (where a hyphen '-' marks a clitic form):

(147)   Conjunctions
      (a)   Clausal conjunctions
            oo   'and
            ee (e)   'and'
            -na   'and'
            -se   'but'
            laakìin   'but'
            ama   'or'
            misé   'or?'
      (b)   Nominal conjunctions
            iyo   'and'
            ama   'or'
            misé   'or?'

We briefly introduce these conjunctions in the next sections.

4.9.1 Clausal conjunctions

Verbal groups and clauses are conjoined in a number of ways, depending upon the type of clause. The most basic strategy, used when the clauses are of the same type, is to place the conjunction *oo* 'and' between the verbal groups or clauses, for example:

(148)   *Wày kú sóo noqotay óo iibsatay.*
        waa+ay   *kú*   *sóo*   *noqotay*   *óo*   *iibsatay*
        DM+she   to   VEN   returned   and   bought
        'She came back (to it) and bought it.'

(149)   *Káalay óo hálkán fadhiisó!*
        káalay   *óo*   *hál-kán*   *sóo*   *fadhiisó*
        come:IMP:SG   and   place-this   VEN   sit:down:IMP:SG
        'Come and sit down here!'

When the clauses differ in polarity or in sentence type, the conjunction *ee* 'and' is placed between them, e.g. the positive and negative declarative clauses in (150) and the declarative and imperative in (151): [41]

(150)   *Wáxba má cunín ee wùu baxay.*
        wáx-ba   má   cúnin   ee   waa+uu   baxay
        thing-any   not   ate   and   DM+he   left
        'He didn't eat anything and he left.'

(151)   *Waa méel khátar áh ee iská rów!*
        Waa   méel   khátar   áh   ee   iská   ráw!
        DM   place   danger   is   and   self-about take:care
        'It's a dangerous place, take care!'

This conjunction is also used when there is an element of contrast between the clauses, as in the following proverb, where we gloss it as 'but':

(152)   *Áwr kala guurtaa má xumá èe ùur kala guuràa xún.*
        Áwr   kala   guurtaa   má   xumá   èe   ùur   kala   guurá+bàa xún
        camels   apart   move   not   bad-is and   heart apart   moves+FOC bad-is
        'It's not bad for camels to separate but a heart which is divided is an evil thing.'

An alternative construction has a shortened form *e* 'and' following the second clausal conjunct, for example:

(153)   *Warkiisa   ha   qaban,*   *waa   waalanyahay*   *e.*
        news-his   IM   take:IMP:NEG:SG   DM   mad-is   and
        'Don't listen to his story, he is mad.' [SHN 194.27]

A further type of clause linking employs a clitic conjunction *na* 'and' which may link any type of clause and which is placed after the first constituent of the second clause in what is sometimes called the Wackernagel position, or clause second position. These first constituents are most typically noun phrases (however long, together with any following focus particle); and sentence type markers (including negative words). See the following where we show the first constituent in bold:

(154)   *Beeni márka horé waa málab, **márka dambéna** waa méleg.*
        beeni   már-ka   horé waa málab   már-ka   dambé-na waa méleg.
        lies   time-the first DM honey   time-the second-and DM death
        'Lies are honey the first time and the second time, death.' **(Proverb)**

(155)   *Axmed wùu yimi, **wùuna** ilá hadlay.*
        Axmed   waa+uu   yimi   waa+uu+na   i-lá   hadlay
        Ahmed   DM+he   came   DM+he+and   me-with   talked
        'Ahmed came and he talked with me.'

The commonest adversative conjunction *-se* 'but' links clauses in the same way, occurring in clause second position, for example:

(156)   *Awr xámilkí wàa qaadaa, **xádhig qalloocánse** má qaadó.*
        Awr   xámil-kí   wàa   qaadaa   xádhig qalloocán-se   má   qaadó
        camel   load-his   DM   takes   rope   twisted-but   not   take:NEG
        'A camel takes his load but will not take a crooked rope.' (Proverb)

Here *-se* follows the NP *xádhig qalloocán* 'a twisted rope'. Somali has also borrowed from Arabic the conjunction *laakìin* 'but' which is placed between clauses, for example:

(157)   *Wàan raadiyey laakìin má helín.*
        waa+aan   raadiyey   laakìin   má   helín
        DM+I   looked:for   but   not   found:NEG
        'I looked for it but didn't find it.'

A similar role is performed by the adversative phrase *hasé yeeshee* 'but, nevertheless' (literally: 'but let it be'), for example:

(158)   *Wàan ku raadínayay hasé yeeshee ku màan arkín.*
        waa+aan   ku   raadínayay   hasé yeeshee   ku   má+aan   arkín
        DM+I   you   was:looking:for   nevertheless   you   not+I   saw:NEG
        'I was looking for you but I didn't see you.'

    Turning to disjunctives, clauses can be linked by two elements: the first *ama* 'or' is used when there is no interrogative force to the disjunction, and the word can be used alone between clauses or can begin each clause in an 'either...or' construction, for example:

(159)   *Tág   ama jòog!*
        go:IMP:SG   or   stay:IMP:SG
        'Go or stay!'
(160)   *Ama táag   lahàw   ama tamár   lahàw!*
        or   strength have:IMP:SG   or   cunning   have:IMP:SG
        'Either have strength or cunning!'   (Proverb)

The second element *misé* 'or' is used when the linked clauses together have the force of an interrogative, for example:

(161)   *Ma jóogaysaa misé waad báxaysaa?*
       ma   *jóogaysaa*   *misé*   *waa+aad*   *báxaysaa*
       QM   are:staying   or   DM+you   are:leaving
       'Are you staying or leaving?'

Note that the clause introduced by *misé* has the form of a declarative not an interrogative though the whole sentence is interpreted as a question.

4.9.2 Noun phrase conjunction

Noun phrases are linked by the conjunction *iyo* 'and'; in a string of NPs only the last pair need be linked by the conjunction:

(162)   *shabèel iyo wèer*   'a leopard and a hyena'
       maróodi, shabèel iyo wèer   'an elephant, a leopard and a hyena'

Noun phrases containing relative clauses are also linked by *iyo*, as in the following proverb:

(163)   *Sáddex nín bàa rág   ú   liitá: bakhàyl baryó   badán i̠*
       three   men FOC mankind of worse-are   miser   begging   much:is a
       sáani   hinaasé   badán   iyo   fulé   nabaddìid   áh.
       adulterer   jealousy   much:is and   coward   peace-rejector   is
       'Three men are the worst of mankind: a miser who begs a lot, and an adulterer
       who is very jealous, and a coward who rejects peace.'   (Proverb)

Noun phrases may also be linked by the disjunctive elements *ama* 'or' and *misé* 'or?', for example:

(164)   *Wúxuu dóonayaa ínuu tagó Maréykan ama Yúrub.*
       wáxa+uu   dóonayaa   in-uu   tagó   Maréykan   ama   Yúrub
       what+he   is:wanting   that-he goes   America   or   Europe
       'What he wants is to go to America or Europe.'
(165)   *Cáli misé Maxámed?*   'Ali or Mohamed?'

4.10 Adverbs

The category of adverb is somewhat problematic in Somali. There is no doubt that the sentential function of adverbial is an important part of sentence grammar, as discussed in chapter 7, but this role is usually performed by noun phrases and clauses. Even the narrower role of modifying verbs and adjectives is normally performed by verbal adpositions governing NPs, as discussed in chapter 6. There are however a restricted number of words whose function can be described as adverbial. They do not however form as coherent a group as other categories.

4.10.1   Time adverbs

There are a small number of adverbs that are clearly derived from corresponding nouns signifying periods of time but which are themselves not nouns. They may be focused by the focus words *bàa* and *ayàa* but they do not occur with the suffixed nominal determiners (4.6), do not have inherent number, or gender and are not case marked. Their function is clearly adverbial in that they locate in time the situation described in the clause. The main examples, together with corresponding nouns, are:

(166) *berríto*   'tomorrow'   cf. *berri*   N F 'tomorrow'
       sháleyto   'yesterday'   cf. *sháley*   N F 'yesterday'
       xáleyto   'last night'   cf. *xálay*   N F 'last night'
       dórraato   'day before yesterday' cf. *dórraad* N F 'day before yesterday'

There are a few other adverbs formed by this suffix *-to*, e.g. *kaléeto* 'like, similar' from the attributive adjective *kalé* 'other'; *hadhówto* 'later on, soon' from the noun phrase *hád dhów* 'a near time'; *dhowéyto* 'recently' cf. *dhów* ADJ 'near', *dhowèe* V CAUS 'bring near'. The derivation does not seem to be productive.

There are other time adverbials which have no obvious corresponding nouns, for example:

(167)   *cáawa* (*cáawey*)   'tonight'
               ciyòw (*ciyòon*)   'immediately, soon'
               dúrba   'quickly, soon'
               ímminka (*ámminka*)   'now'
               sáa   'then, at that time'
               sáaka   'this morning'

See for example:

(168) *Ciyòw ú shèeg arrínta!*
      ciyòw   *ú*   *shèeg*   *arrín-ta*
      immediately   to   tell   matter-the
      'Tell him the matter immediately!'

(169) *Márkaannu sóo qaadnay sheekádíi, dúrba wùu xanaaqay.*
      már-ka-aannu   *sóo*   *qaadnay*   *sheeká-díi*   *dúrba*   *waa+uu*   *xanaaqay*
      time-the+we:EXC VEN   took   story-the   suddenly   DM+he   grew:furious
      'When we started the story, he suddenly grew furious.'

4.10.2   *Intensifier àad*

The word *àad* is used to intensify adjectives, as in (170b) below, and verbs, as in (171b). It occurs governed by the adposition *ú* in the verbal group:

(170) (a)   *Wàan jeclahay.*
            waa+aan   *jecel-ahay*
            DM+I   liking-am
            'I like it.'

        (b)   *Àad bàan ú jeclahay.*
            àad   *bàa+aan*   *ú*   *jecel-ahay*
            much   FOC+I   ADP   liking-am
            'I like it very much.'

(171) (a)   *Ròob bàa dá'ayá.*
            rain   FOC   is:falling
            'Rain is falling,' 'It's raining.'

        (b)   *Ròobku àad bùu ú dá'ayaa.*
            ròob-ku   *àad*   *bàa+uu*   *ú*   *dá'ayaa*
            rain-the   much   FOC+it   ADP   is:falling
            'The rain is falling heavily,' 'It's raining heavily.'

As is usual, in (170b) and (171b) the intensifier is focused by *bàa*. The intensifier *àad* can be repeated for extra emphasis, when it is conjoined by *iyo* 'and': *Àad iya àad bàan ú jeclahay* 'I like it very very much.'

4.10.3   *Preverbal adverbial clitics*

There are four words which modify the meaning of verbs and which occur as clitics in the preverbal cluster, as described in chapter 6. These are given below with the additional meaning they add to verbs:

(172)   *sóo*   'in the direction of the speaker'
         sii   'away from the speaker'
         wada   'together, towards each other'
         kala   'apart, away from each other, separately'

The first two adverbial clitics, *sóo* and *sii* are very productively used for verbs; *wada* and *kala* are more limited in that they for the most part require verbs with plural subjects. We gloss *sóo* as a venitive adverbial (VEN) and *sii* as an allative adverbial (ALL). Examples of their effects are:

(173)   (a)   *Wuu socónayaa*!
                  Waa+uu   socónayaa
                  DM+he   is:going:on
                  'He is going on!', 'He is travelling on.'
         (b)   *Wuu sóo socónayaa*!
                  Waa+uu   sóo   socónayaa
                  DM+he   VEN   is:going:on
                  'He is coming over here', 'He is approaching.'
         (c)   *Wuu sii socónayaa*!
                  Waa+uu   sii   socónayaa
                  DM+he   ALL   is:going:on
                  'He is going over there', 'He is going away.'
(174)   (a)   *Wày hadleen*
                  waa+ay   hadleen
                  DM+they   spoke
                  'They spoke.'
         (b)   *Wày wada hadleen*
                  waa+ay   wada   hadleen
                  DM+they   together   spoke
                  'They spoke together', 'They conversed with each other.'
(175)   (a)   *Wày tageen*
                  waa+ay   tageen
                  DM+they   went
                  'They went'
         (b)   *Wày kala tageen*
                  waa+ay   kala   tageen
                  DM+they   apart   went
                  'They went separately', 'They separated.'

Since these adverbial clitics occur in the preverbal clitic cluster after verbal adpositions, it is important to note that *kala* 'apart' is distinguished by its accentual pattern from the coalesced adpositional form *kalá* from *ká* 'from,

away' and *lá* 'with'. As an example of the possible difference in meaning, compare (175b) above with (176) below:

(176)   *Wày kalá tageen*
        waa+ay   ká-lá   tageen
        DM+they   from-with   went
        'They stole it etc. from him etc.', 'They abducted her etc. from him etc.'
        (lit.: 'They went away with him etc. from him etc.')

4.10.4   ba

The word *ba* exhibits a continuum of uses and meanings. It can be translated in English by the following words, depending on context: 'each, every, both, too, any, ever, at all'. Its use varies between a more adjectival and a more adverbial function.   Its adjectival function, mentioned in 4.4.3, involves modifying a preceding noun to which it may be cliticized, for example (with the meaning 'any'): *wáx* N M 'thing', *waxbá* 'anything'. When a noun already has a modifying adjective, *ba* is added, not in the usual way that another adjective would be, by conjoining with *oo* 'and' (as in (177b) below) but simply by being cliticised onto the NP, e.g. (177c) below (with the meaning 'each'):

(177)   (a)   *qóf kastá*   'every person'
        (b)   *qóf kastá oo wanaagsán*   'every good person'
        (c)   *qóf kastába*   'each and every person'

Similarly, *ba* can mean 'both, all': *labáda* 'the two', *labádaba* 'both of them'. *Ba* can also be used more adverbially, meaning something like English 'ever':

(178)   *Méeshaad tagtába wáx hubsó!*
        méel+ta+aad   tagtó-ba   wáx   hubsó
        place+the+you   go-ever   thing   be:sure:of:IMP
        'Wherever you go, be sure!'   (Proverb)

Its use is more clearly adverbial when, in its post-NP position, *ba* functions as an intensifier in negative clauses, meaning something like '(not) at all' or '(not) ever', for example:

(179)   *Súudow, dagàalka ba anígu dóoni màayó e*
        Súudow   dagàal-ka   ba   anígu   dóoni màayó   e
        Suude-VOC   war-the   at:all I   want:PRES PROG:NEG   and
        'O Suude, I don't want this war at all'[42]

4.10.5   uun

The word *uun* is an intensifier which shares some characteristics with *ba:* it occurs post-nominally and varies between adjectival and adverbial functions. As discussed in 4.4.2, in its adjectival use we might gloss it as 'some…or other', for example:

(180)   *méel uun*   'someplace or other'   (cf. *méel* N F 'place')

In its more adverbial role, its main function is to allow the speaker to add emphasis: we can parallel it with some English uses of 'just', 'only', or 'right', for example:

(181)   (a)   *Míd   qaadó!*
                  one   take:IMP:SG
                  'Take one!'
        (b)   *Míd   uun   dooró!*
                  one   just   choose: IMP:SG
                  'Just choose one!'
        (c)   *Már uun bùu xusuusatay.*
                  már   uun   bàa+uu   xusuusatay
                  time   just   FOC+he   remembered
                  'Just then he remembered.'

4.11 Ideophones

Somali, like many other African languages, has a stock of ideophones: words which are used to add intensity to descriptions of sounds, images and actions, especially movement. Some ideophones are onomatopoeic in that they mimic sounds; for example, the ideophone *sháw* used to describe water that is showering down, such as rain. Typically though the relationship between the ideophone and what it describes is conventional and arbitrary, for example *sám* used to describe the sudden extinction of light. Ideophones are also more restricted grammatically than the general phenomenon of onomatopoeia: they are all feminine nouns (with the accentual pattern AP1) and they occur as complements of a verb or verbal adposition. The most comprehensive discussion of Somali ideophones is Dhoorre and Tosco (1998), on which this section is based.

　　　Phonologically, ideophones follow the normal phonotactic and syllable structure rules, except that word finally the nasal *m* occurs, for example *hám* used to describe the action of eating in big bites. The majority are monosyllables

but bi- and, less commonly, tri-syllabic words do occur, for example *halalác* used to describe intermittent or sparkling light, for example reflecting off metal. Ideophones are regularly reduplicated to add further intensity or repetition, with the dominant pattern being total reduplication. Grammatically, ideophones are nouns but with some peculiarities. They do not form plurals, they occur in a basic undefined form and they only occur with two verbs, *yidhi* (*yiri*) 'say' and *sìi* 'give' [43], as complements of the verb or an associated verbal adposition, as we shall see. There is a tendency for *sìi* to be used with causative predications, and with subjects high in agency, for example:

(182)   *Albàabkíi bàa qáb yidhí*
        door-the   FOC   *qáb*   said
        'The door banged.' (lit. 'The door said *qab*')

(183)   *Mindidii baa shuq tiri*. (Dhoorre & Tosco 1998:152)
        knife-the   FOC   *shuq*   said
        'The knife slipped into him' (lit. 'The knife said *shuq* to him')

(184)   *Mindidii buu shuq ku siiyay*. (ibid.)
        knife-the   FOC+he   *shuq*   with   gave
        'He thrust the knife into him.' (lit. 'He gave *shuq* to him with the knife')

Dhoorre and Tosco (1998) divide ideophones into four groups, given below where we repeat some of their examples:

(185)   A.   Movement (often with accompanying sound), e.g.
            búl   'rushing off noisily'
            shulúx   'falling down, softly'
            fúd   'coming out'
            wáf   'skimming past quickly'
            Example: *Galayr baa hortaydii waf yiri*.
                        Hawk   FOC   front-my   *waf*   said
            'A hawk swept across in front of me.'

            B.   Hitting and breaking, e.g.
            xáf   'cutting neatly and sharply'
            dhác   'tearing, with a ripping noise'
            dhág   'puncturing, opening a hole'
            gúb   'knocking (wood)'
            Example: *Sakin buu dhag ku siiyay*.
                        Blade   FOC+he   *dhag*   with   gave
                        'He made a little hole with a blade.'

C. Emitting sound and light, e.g.

jíiq   'squeaking, as of wood and birds'
dhíiq   'hissing, as of compressed air'
halalác   'giving off sparkling, shimmering light'
wác   'giving off piercing light'
Example: *Hillaacii   baa   wac   yiri.*
                    Lightning   FOC   *wac*   said
'Lightning flashed.'

D. Miscellanea[44], e.g.

julúq   'gulping down, without chewing'
qaláf   'dying suddenly'
gám   'wiping out, killing off'
dúg   'falling into a deep sleep'
Example: *Dug   dheh*!
                    dug   say:IMP:SG
'Hush!, Go to sleep!' (said to a child)

4.12 Interjections and other items

In addition to these categories discussed so far there are a number of other types of words that do not enter into sentence grammar in the same way. Somali speakers, for example use attention-gathering expressions, when addressing someone at a distance. These include *wàar* or *waaryàa* for addressing men and *nàa* or *naayàa* for addressing women. Hearing these a Somali speaker will stop and turn towards the speaker. They can also be used to prefix names, e.g. *Wàar Áxmed*! 'Hey Ahmed!'. There is also a group of words we could call interjections, which express the reactions or emotions of the speaker, e.g.:

(186)   Word   Expressing
                     àa   sorrow
                     aqás/akhás   disgust
                     álla   surprise
                     bága   joy
                     cáku   distaste
                     cár   warning challenge
                     úf   experiencing a bad smell
                     wày   distress, disappointment
                     yáh   irritation

Sometimes these words can be integrated into phrases and clauses, for example, linked by a conjunction to a NP following the exclamation uttered on hearing of a relative's death:

(187)   *Wày iyo tollàay (tollá'ày)!*
        wày   iyo   tól-lá'-kày
        wày   and   kin-lacking-my
        'Oh, and my lost kin!'

Similarly there is a range of words used to animals for specific purposes, for example:

(188)   Word   Use
           dèeg   stopping a horse
           dòoh   calming a nervous camel
           ís   urging an animal on
           jú'   ordering a camel to kneel
           yúr   frightening off wild animals

Words with special social purposes would also include *hée*, the word used by listeners to show attention and encourage a narrator to continue. The narrator in turn periodically utters prompts such as *hée dhé*! literally 'Say *hee*!'. Similarly a sequence of utterances of the word *hooyáale*, adapted to the relevant metrical pattern, is used to introduce a poem or song. Or as an example of a different social function, the word *hóoddi* is uttered before entering a house.

Finally we might include in this section the words *háa* 'yes', *háye* (*háyye*) 'yes, okay', and *máya* 'no', which operate as single words, though it seems possible that they have a historical source as verbal expressions meaning 'It is (so)' and 'It is not (so)'.

Chapter 5
Word formation

In this chapter we describe some characteristic processes by which new words are formed. This can be done either by shifting a word from one semantic type to another within a category, for example changing one kind of verb into another, or by allowing one category to be transformed into another, for example forming verbs from nouns. Some roots can occur in different categories without modification, for example *éngeg* N M 'dryness', *éngeg* V INTR 'dry, dry out'; and *siràad* N M 'lamp, light', *siràad* V TR 'light (a lamp), illuminate'. For the most part however, new words are formed by the addition of affixes and by compounding. We cannot attempt a comprehensive listing of the derivational affixes within the space of this chapter; instead we list the most important and productive affixes in order to give an idea of how the system works. As we shall see, all derivational affixes are suffixes and we divide our description according to the main types of category produced by them: derived verbs, adjectives and nouns. These suffixes are not themselves independent words; in a fourth section we discuss compounding, where new words are created by the combination of independent words.

5.1   Derived verbs

As described in chapter 4, root verbs can be assigned to various semantic classes, membership of which determines their grammatical behaviour. For example we can make a basic distinction between stative and dynamic verbs where, as we saw, the distinction governs a verb's participation in the set of morphological paradigms. Stative verbs do not occur in some paradigms, for example the progressive forms, while in others, they have a different interpretation from dynamic verbs. Dynamic verbs themselves fall into several classes: for example inchoative verbs, which describe an entity becoming or

achieving a state. Grammatically these form a class of intransitive verbs, of which some examples are below:

(1)   *dacìif*   'weaken, become feeble'
               éngeg   'become dry, dry out, wither'
               fásas   'become infected'
               jílic   'become soft'
               kór   'grow, develop, increase'
               máshmash   'become overripe (fruit), overcooked (food)'
               míridh   'rust, get rusty'
               qóy   'become wet, damp'
               xáaluf   'become grazed out, barren of pasture'

A second class consists of lexical causative verbs, where an Agent causes a Patient entity to achieve a state. Grammatically these form a class of transitive verbs, for example:

(2)   *díl*   'kill'
               fúr   'open'
               gúnud   'knot'
               jáb   'break'
               káb   'mend, repair'
               shíd   'light, ignite'
               sìib   'pull out, uproot'
               xídh   'close'

A further class consists of activity verbs, which can either be intransitive or transitive, for example:

(3)   *cún*   TR   'eat'
               kèen   TR   'bring'
               órod   INTR   'run'
               qósol   INTR   'laugh'

As mentioned in chapter 4, new verbs can be created by the attachment of lexical derivational affixes to root verbs. These affixes allow the shifting of a root into a different semantic class, for example creating causative verbs from inchoative roots. There are also affixes which allow the creation of verbs from non-verbal roots, for example creating inchoative verbs from nouns. In this section we outline the most important of these verbal derivational affixes.

5.1.1 Derived inchoative verbs

In addition to the class of root inchoatives mentioned above, inchoatives can be formed from other grammatical categories by the use of derivational affixes. Two in particular are very productive: the first is the affix *ow/oob* which for a noun with the meaning N creates an intransitive verb meaning 'become N', for example:

(3)	*barafòw*	'turn into ice, freeze'	←*báraf*	N M 'ice'
	biyòw	'become water, melt'	←*biyó*	N M 'water'
	gaalòw	'become an infidel'	←*gàal*	N M 'infidel'
	magaalòw	'become a town'	←*magáalo*	N F 'town'
	nacasòw	'become a fool'	←*nácas*	N M 'fool'
	tuugòw	'become a thief'	←*tùug*	N M 'thief'

The affix has two allomorphs: *ow* before a consonant or at the end of a word, and *oob* before a vowel, compare for example (where the inchoative affix is marked INCH):

(4)	(a)	*Wùu*	*baraf-oob-ay*
		DM+it:M	ice-INCH-3SG M:PAST
		'It (M) turned to ice, it (M) froze'	
	(b)	*Wày*	*baraf-ow-day*
		DM+it:F	ice-INCH-3SG F:PAST
		'It (F) turned to ice, it (F) froze'	

This alternation seems to be phonological since it is not limited to this affix: we saw in 4.1.4 that the numeral *ków* 'one' becomes *kóob* before a vowel, as in *kóobiyo tobán* 'eleven ('one-and ten')'; similarly the adjective *madów* 'black' ~ *madoobaa* 'was black'.

The second affix is *ow/aan* which for an adjective describing a state S creates an intransitive verb meaning 'become S', for example:

(5)	*adkòw*	'become hard, harden'	←*adág*	ADJ 'hard'
	dhowòw	'get near, approach'	←*dhów*	ADJ 'near'
	kululòw	'get hot'	←*kulúl*	ADJ 'hot'
	yaròw	'get small, shrink'	←*yár*	ADJ 'small'

The examples above show this affix attaching to basic adjective roots but it also productively attaches to the derived adjectives, described in 5.2:

(6)   *gaabnòw*   'become short'   ←*gaabán*   ADJ 'short'
        sahlanòw   'become easy'   ←*sahlán*   ADJ 'easy'
        qaalisanòw   'become expensive'   ←*qaalisán*   ADJ 'expensive'
        wanaagsanòw   'become good'   ←*wanaagsán*   ADJ 'good'

This affix has several forms, whose distribution seems to be grammatically conditioned: *aw* occurs only in the imperative, *aan* occurs in the infinitive, and therefore also in auxiliary constructions and progressives, while *aad* occurs everywhere else. Note though that when *aad* occurs before the agreement affix *t* (2SG/3SG F/2PL), rules of regressive assimilation and simplification apply in the same way described for the middle voice affix in 3.2.4, i.e. *aad* + *t* → *aat*, as in *wày ogaatay* 'She became aware'←(*ogaad-t-ay*).[45]

5.1.2 Derived experiencer verbs

A suffix *ood* is used to create derived verbs from nouns: for a noun with the meaning N this affix creates an intransitive verb meaning 'to experience N', for example:

(7)   *cadhòod*   'be angry'   ←*cádho*   N F   'anger'
        dhaxamòod   'feel cold'   ←*dháxan*   N F   'cold'
        gaajòod   'be hungry'   ←*gáajo*   N F   'hunger'
        cidlòod   'feel lonely'   ←*cídla'*   N F   'lonely,uninhabited place'
        riyòod   'have dreams, dream'   ←*ríyo*   N F   'dream

This experiencer affix (EXP) has the form *oon* in the infinitive and associated paradigms like the progressives, e.g. *Waan gaajóonayaa* 'I am hungry'; and *ood* elsewhere.[46] Note that the use of progressives reflects the fact that these derived verbs are not stative.

5.1.3 Factitive

The affix *ays* is a very productive factitive affix which creates verbs from both nouns and adjectives. It has the principal forms: *ee* in the imperative; *ayn* (*eyn*) in the infinitive and derived paradigms, and *ays* (*eys*) elsewhere.[47] The morphophonological effects of adding it to noun and adjectival roots are captured in chapter 4 by listing such verbs as a separate conjugation, C2B (4.3.3). Typically the derived verb is transitive; when this affix is added to a noun root N, the derived verb means either 'make X into N' or 'add N to X', for example:

(8)   *afèe*   'put an edge on, sharpen'   ←*áf*   N M   'edge, blade'
        birèe   'make into iron'   ←*bír*   N F   'iron'
        biyèe   'add water to, water down'   ←*biyó*   N M   'water'
        dharèe   'clothe'   ←*dhár*   N F   'clothes'
        gaalèe   'turn s.o. into an infidel'   ←*gàal*   N M   'infidel'
        irbadèe   'prick, inject'   ←*irbád*   N F   'needle'
        sumèe   'poison'   ←*sún*   N F   'poison'

When added to an adjectival root describing a state S, the derived verb means 'make X achieve state S', e.g.

(9)   *adkèe*   'harden, strengthen'   ←*adág*   ADJ   'hard, tough'
        caddèe   'whiten'   ←*cád*   ADJ   'white'
        fududèe   'lighten, make easy'   ←*fudúd*   ADJ   'light, easy'
        le'ekèe   'make equal to'   ←*le'ég*   ADJ   'equal (size, number)'
        yarèe   'make smaller, reduce'   ←*yár*   ADJ   'small'

Although most verbs derived by *ays* are transitive, there are some derived intransitive verbs, which we can assign to two types. The first are derived from nouns and convey a change of state, for example:

(10)   *ammuurèe*   'become troubled, act worried'   ←*ammúur*   N F   'problem'
        candhèe   'fill with milk (of an udder)'   ←*cándho*   N F   'udder'
        jirjirèe   'turn on/lie on one's back'   ←*jírjir*   N M   'back'
        laamèe   'put out branches, branch out'   ←*láan*   N F   'branch'

The second type are verbs derived from the attributive adjectives described in 4.4.3, most of which describe locations in space or time. Where the attributive adjective describes a location L, the derived intransitive verbs mean 'be in location L', for example:

(11)   *dambèe*   'be behind, last'   ←*dambé*   'behind, last'
        hoosèe   'be under, low'   ←*hoosé*   'lower'
        sarrèe   'be above, on top'   ←*saré*   'topmost, upper'

5.1.4 Passive

We mentioned in 4.2.3 that Somali does not have a syntactic passive and that some of the corresponding functional load is borne by constructions using the non-specific subject pronoun *la* 'one'. A related function is performed by the derivational affix *am* which is added to lexically causative verbs to remove the

Causer argument and allow the Patient argument to become subject. Compare the causative root verb *fúr* TR 'open' and the derived passive *fúran (← *fúram)* INTR 'be/get opened':

(12)   (a)   *Cali   albàab-kíi   bùu   fur-ay.*
               Ali   door-the   FOC+he   open-PAST
               'Ali opened the door.'
        (b)   *Albàab-kii   wàa   fur-m-ay.*[48]
               door-the   DM   open-PAS-PAST
               'The door was opened.', 'The door opened.'

Typological accounts of the passive (e.g. Foley and Van Valin 1984, Givón 1984) often divide its function into two parts: foregrounding a non-Actor argument and backgrounding the Actor argument. This affix *am* performs both since, firstly, the Causer argument of the root verb must be suppressed and secondly, the Patient argument becomes subject. However this effect of this affix is narrower than many morphological passives because the input root verb must be causative. The affix cannot be attached to, for example, experiencer verbs like *árag* 'see' or *máqal* 'hear', or activity verbs like *cún* 'eat' or *síd* 'carry'. The effect of the affix is to remove the Causer from and thus de-transitivise causative root verbs. We can see below that the lexical causative verbs in (2) earlier may all form corresponding derived passives (where word-final *m → n*, as always):

(13)   *dílan*   'get/be killed'
               fúran   'get/be opened'
               gúntan   'get/be knotted'
               jában   'get/be broken'
               kában   'get/be mended, repaired'
               shídan   'get/be lighted, ignited'
               sìiban   'get/be pulled out, uprooted'
               xídhan   'get/be closed'

Two further points are important: firstly this affix has no stativising effect. The derived verbs remain dynamic, as can be seen by that fact that they occur in progressive forms:

(14)   *Albàab-ku   waa   fúr-m-ayaa.*
        door-the   DM   open-PAS-PRES PROG
        'The door is getting opened, is opening.'

Thus the passive verb *fúran* (←**fúram*) 'be/get opened' contrasts with the derived adjective *furán* 'open', which is of course stative, as in:

(15) *Albàab-ku waa fur-án-yahay.*
     door-the DM open-STAT-is
     'The door is open.'

This adjective has of course no corresponding progressive form. The respective affixes are *am*, passive, and *an*, stativiser, where the latter produces derived adjectives as described in 5.2.

As dynamic verbs, these derived passives freely occur in imperative forms, as in the following traditional prayer:[49]

(16) *Belaayá-da hor-àay róor-an,*
     evils+the before-EXCL drive-PAS
     Ta damb-àay réeb-an,
     those behind- EXCL leave-PAS
     Ta baalléed-àay ríih-an,
     those beside- EXCL push-PAS
     Ta saar-àay ráar-an,
     those above- EXCL suspend-PAS,
     Ta hoos-àay ná radéeb-an!
     those below- EXCL and blunt-PAS

     'Oh evils before us, be driven away,
     Oh those behind us, be left behind,
     Oh those on our sides, be pushed aside,
     Oh those above, be suspended,
     Oh those below us, be blunted!'

The second point is that the passive affix seems to preserve the notion of causation in the verb, even though the Causer is suppressed. We can contrast this with one important function of the middle voice affix which, as described in 5.1.6, marks uncontrolled inchoation, i.e. where a process affects but is not under the control of the subject argument. We can thus find contrasts like the following on the causative verb *bilàab* 'cause to begin, start':

(17) (a) *Cali wùu bilaab-ay shír-kíi.*   (Causative)
        Ali DM+he start-PAST meeting-the
        'Ali started the meeting.'

(b)   *Shír-kii*   *wàa*   *bilaab-m-ay.*   (Passive)
      meeting-the   DM   start-PAS-PAST
      'The meeting was/got started'

(c)   *Shír-kii*   *wàa*   *bilaab-t-ay.*   (Middle)
      meeting-the   DM   started-MID-PAST
      'The meeting started.'

The causative verb in (17a) requires a Causer and signifies causation; the passive derived verb in (17b) suppresses the Causer but still identifies causation; the middle derived verb in (17c) suppresses the Causer and communicates no sense of causation. This example gives one glimpse of the semantic interplay between lexical affixes and the semantics of root verbs which is so characteristic of Somali.

5.1.5 Causatives

There are two causative affixes. The first, *is*, creates causative verbs from other verbs. Its prototypical use is to create causative verbs from root inchoative verbs: where the inchoative means 'X achieve state S' the causative verb means 'Y cause X to achieve state S', e.g. (where the affix occurs as *i*): [50]

(18)   *búuxi*   TR   'fill'   ←*bùux*   INTR   'fill up'
       engéji   TR   'dry'   ←*éngeg* INTR   'become dry, dry out, wither'
       jílci   TR   'soften'   ←*jílic* INTR   'become soft'
       kári   TR   'cook'   ←*kár*   INTR   'boil, cook'
       kóri   TR   'rear, bring up' ←*kór*   INTR   'grow, develop, increase'

For example:

(19)   (a)   *Berkéd-díi*   *wày*   *buux-day.*
             trough-the   DM+it:F   fill-3SG F:PAST
             'The trough filled up.'

       (b)   *Naagí-hii*   *berkéd-díi*   *wày*   *buux-iy-een.*
             women-the   trough-the   DM+they   fill-CAUS-3PL:PAST
             'The women filled the trough.'

Comparing the root and derived verbs we can see that the causative affix *is* triggers an argument structure change: it adds an extra argument, a Causer, which functions as subject. The intransitive inchoative root verb thus becomes transitive. The single subject argument of the inchoative, the Patient, becomes

object of the derived causative verb. Less typically this affix can be attached to activity verbs; compare for example:

(20)   (a)   *Ilmí-hii   wùu   ooy-ey (←*ooh-ey)*
              child-the   DM+he   cry-3SG M:PAST
              'The child cried.'

       (b)   *Ilma-ha   ká   óoh-i!*
              child-the   ADP   cry-CAUS:IMP
              'Make the child cry!'

       (c)   *Ilmí-híi   waa   la-gá   ooh-iy-ey.*
              child-the   DM   one-ADP   cry-CAUS-3SG M:PAST
              'Someone made the child cry.'

(21)   (a)   *Ardá-dii   wày   qosl-een.*
              students-the   DM+they   laugh-3PL:PAST
              'The students laughed.'

       (b)   *Barí-hii   ardá-dii   wuu   ká   qosl-iy-ey.*
              teacher-the   students-the DM+he   ADP   laugh-CAUS-3SG M:PAST
              'The teacher made the students laugh.'

We can see from (20b & c) and (21b) that with these verbs the argument structure change is different: the subject of the root verb occurs not as object of the causative verb but as an oblique NP governed by the adposition *ká* 'from'.

This affix does not usually occur with transitive verbs. When it does, the derived verb is subject to semantic shift, i.e. the derivation is neither fully productive nor completely transparent semantically, e.g.: *dhál* 'beget' →*dháli* 'originate'; *tír* 'cancel' → *tíri* 'count'; *màal* 'milk' → *máali* 'lend (a milch animal) to someone'. The affix has a number of forms, including *i*, and *is*; these and the morpho-phonological effects of adding *is* to verb roots is captured in chapter 4 by listing such verbs as a separate conjugation, C2A (4.3.3).

The second causative affix is *sii*. This affix, which is less commonly used than *is*, is typically attached to activity verbs, especially with human Actor subjects. It may be attached to either transitive or intransitive root verbs. With transitive verbs the effect is to add a third argument: the Causer is subject; the Actor-subject and the Patient-object of the root verb both become objects of the derived causative, for example:

(22)   (a)   *Ilmú-hu   wùu   cabb-ayaa   biyá-ha.*
              child-the   DM+he   drink-PRES PROG   water-the
              'The child is drinking the water.'

    (b)   *Ilmá-ha*   *cab-sìi*   *biyá-ha!*
           child-the   drink-CAUS:IMP   water-the
           'Make the child drink the water!'

(23)   (a)   *Carrúur-tii*   *wày*   *cun-een*   *barìis-kii.*
           children-the   DM+they   eat: 3PL:PAST   rice-the
           'The children ate the rice.'

    (b)   *Faadúmo*   *ayàa*   *carrúur-tíi*   *cun-síi-say*   *barìis-kíi.*
           Fatima   FOC   children-the   eat-CAUS-3SG F:PAST   rice-the
           'FATIMA caused the children to eat rice/fed the children rice.'

When this affix is attached to intransitive verbs, the usual pattern is for the root verb's Actor-subject to become an oblique NP governed by the adposition *ká* 'from', e.g.:

(24)   (a)   *Maxamed*   *wùu*   *híndhis-ay.*
           Mohamed   DM+he   sneeze-PAST
           'Mohamed sneezed.'

    (b)   *Siigáda*   *ayàa*   *ká*   *hindhi-síi-sáy*   *Maxámed.*[51]
           dust-the   FOC   ADP   sneeze-CAUS:PAST   Mohamed
           'The dust caused Mohamed to sneeze.'

(25)   (a)   *Cali*   *wùu*   *feker-ayay.*
           Ali   DM+he   think-PAST PROG
           'Ali was thinking.'

    (b)   *Way*   *ká*   *feker-síi-say*   *Cáli.*
           DM+she   ADP   think-CAUS-PAST   Ali
           'She caused Ali to think,' 'She made Ali think.'

There are some cases, however, where an intransitive verb patterns like a transitive verb when *sii* is attached, i.e. where the root verb's subject becomes a second object of the derived causative, e.g.:

(26)   (a)   *Maxamed*   *wùu*   *qadeyn-ayaa.*[52]
           Mohamed   DM+he   lunch-PRES PROG
           'Mohamed is lunching/eating lunch.'

    (b)   *Wày*   *qaday-síin-aysaa*   *Maxámed*
           DM+she   lunch-CAUS-PRES PROG   Mohamed
           'She is giving lunch to Mohamed' (lit. 'She is lunching Mohamed.')

(27)   (a)   *Inamá-dii*   *wày tartam-een.*
           boys-the DM+they   compete: 3PL:PAST
           'The boys competed with each other.'

(b)   *Inamá-dii wàan   tartan-siiy-ey.*
        boys-the DM+I   compete-CAUS-1SG :PAST
        'I caused the boys to compete with each other.'

It is not clear why these intransitive verbs pattern like transitives in this way, unless we speculate that there are semantically understood objects in *qadèe* 'eat lunch' and the reciprocal verb *tártan* 'compete (with each another)' which influence the argument structure of the derived causative.

One interesting point is that this affix *sii* is homophonous with the verb *sìi* 'give', as shown below:

(28)   (a)   *Cáli   cab-sìi!*
              Ali   drink-CAUS:IMP:SG
              'Cause Ali to drink (it)!', 'Make Ali drink (it)!'
       (b)   *Cáli   wúxuu   cabbó   sìi!*
              Ali   what+he drinks   give:IMP:SG
              'Give Ali what he drinks!', 'Give Ali something to drink!'

The sentence (28a) is only suitable if Ali is a child or medical patient, while (28b) would be normal in a café. All of the inflectional forms of the affix *sii* are the same as if the verb *sìi* 'give' was suffixed onto the root verb and it is therefore possible to view these forms as an instance of grammaticalization, where a lexical verb has acquired the status of a derivational affix. If so, the process does not seem to be complete: there are significant differences between this morpheme *sii* and the first causative affix *is*, which seems to exhibit more typically affixal behaviour. Firstly, *is* is much more productive: for example in Zorc and Osman's (1993) 26,00-entry Somali-English dictionary there are about 1300 verbs with *is* compared with around 180 causative verbs with *sii.* Secondly *is* causatives are involved in other derivational processes while *sii* causatives are not: for example, stems in *is* form verbal nouns, e.g. *jóoji* 'cause to stop' → *joojín* N F 'stopping', as described in 5.3.2. Thirdly *is* causatives undergo semantic shift as we mentioned earlier, for example *tíri* 'count' from *tír* 'cancel, eliminate', which *sii* causative seem not to do. Lastly, *sii* does not seem as semantically 'bleached', to use Givón's (1975) term, as *is*: while *is* causatives have no semantic restriction on the animacy of the causee, *sii* causatives are largely restricted to an animate, in fact sentient, causee. This restriction is shared by the second (Recipient) argument of the homophonous verb *sìi* 'give'.

5.1.6 Middle voice

The middle voice affix *at* forms derived middle verbs from other verb stems, for example *laabó* INTR 'turn oneself back, head back' derived from the causative root verb *làab* TR 'bend, fold'. It has several forms: some are morphologically conditioned, like *o* in the imperative and *an* in the infinitive, while others are phonologically conditioned by context, as when *at* → *ad* before vowels. The morphophonological effects of adding this affix to root verb stems are outlined in chapter 4 by listing such verbs as a separate conjugation, C3B (4.3.3). The affix is also very commonly added to derived causatives, for example *hub-s-ó* (sure-CAUS-MID) 'make sure of for oneself, ascertain' derived from *húb-i* (sure-CAUS) 'make sure of, verify', itself derived from the root verb *húb* 'be sure'. The combination of causative and middle affixes on a verb root gives rise to rather complicated morphophonological changes: these are summarised in chapter 4 by listing such verbs as a separate conjugation, C3A (4.3.3). The affix can also be attached to other derived stems, for example *daaweysó* INTR 'treat oneself with medicine' from *daawèe* TR 'treat with medicine', formed by the factitive affix *ays* (5.1.3) from the noun *dáawo* F 'medicine'. In our examples in this section we use middle verbs formed on various stem types.

   This affix has been given a number of names in the literature including *riflessivo* (Moreno 1955), *medio-riflessivo* (Bruno 1984), and *autobenefactive* (Andrzejewski 1968, Puglielli 1984), each concentrating on one aspect of its use. As described in Saeed (1995), this affix has a number of distinct but related meanings, all of which share the characteristic noted for middle voice in other languages (e.g. Givón 1984, Klaiman 1991, Kemmer 1994) of the subject of the verb being affected by the action described in the verb. This 'affectedness' has different expressions in Somali: we can outline six main meaning types, which form a semantic continuum. One important point is that the effect of this affix on the transitivity of the stem is not uniform. Although in many instances the derived middle verb is intransitive, removing an object argument from a transitive stem as in the examples so far, the effect depends on the semantics of the stem verb and the type of middle meaning produced.

   The first type of middle describes an action that is inherently reflexive, i.e. where the subject is both the Agent of the action and the entity affected or Patient, e.g.:

(29)   *dheefsó* TR   'nourish oneself on'   ←*dhéefi* CAUS 'nourish on'
       diirsó INTR   'warm oneself'   ←*díiri* CAUS 'warm'
       jirsó INTR   'shelter oneself (from rain etc.)'←*jíri*   CAUS 'shelter'
       xoqó   TR   'scratch oneself'   ←*xóq*   TR   'scratch'

A distinction is drawn in the grammar between inherently reflexive verbs, expressed by middle verbs, and reflexive uses of action verbs that are not inherently reflexive, where the reflexive/reciprocal pronoun *is* (4.2.4) is used with non-middle verbs, e.g.:

(30)   *Wùu*   *is*   *dil-ay.*
       DM+he   RE   killed
       'He killed himself.'

A second related type of middle is where the action affects the subject's body. There are three main subtypes: the first describes bodily motion and posture, for example in (31a) below; the second are grooming verbs (31b), and the third are actions involving the body and clothes, as in (31c):

(31)   (a)   *arrumó* INTR 'crouch, curl up (of a camel)' (←*arrún* TR 'cause s.t. (e.g. a camel) to crouch')
             dhambacaadsó INTR 'lie down on one's stomach' (cf. *dhambacàad* INTR 'be lying down on one's stomach')
             foororsó INTR 'bend over, stoop, bow' (←*fooróri* CAUS 'cause to bend' ←*fóoror* INTR 'be in a stooped, bent over position')
       (b)   *firó* TR 'comb (one's own hair)' (←*fir* TR 'comb')
             maydhó INTR 'wash oneself' (←*màydh* TR 'wash')
             rifó TR 'pluck, shave (one's own hair, eyebrows)' (← *ríf* TR 'pluck')
             tidcó TR 'braid (one's own hair)' (←*tídic* TR 'braid (hair)')
       (c)   *guntó* INTR 'wrap clothing around one's waist' (←*gúnud* TR 'knot')
             xidhó TR 'dress, put on one's clothes' (←*xídh* TR 'fasten, close')
             xijaabó INTR 'veil oneself' (←*xijàab* TR 'veil, cover')

A third type of middle verb describes actions which also affect the subject but involve emotions or cognitive abilities rather than the body, including experiencing emotions, expressing emotions, talking, thinking and religious experience. Some examples are:

(32)   *ashahaadó*   INTR   'say the profession of the Muslim faith'
       barooró   INTR   'mourn, keen, wail (of women)'
       dhibsó   TR   'feel annoyed at'
       durraamó   TR   'pray to'
       gedmó   INTR   'be confused, mistaken'
       karahsó   TR   'dislike, have an antipathy for'
       lahó   INTR   'be consumed with desire, have a burning desire'
       miyirsó   INTR   'come to, regain consciousness'

| *quusó* | INTR | 'lose heart, despair' |
| *xusuusó* | TR | 'remember, call to mind' |

A fourth type of middle verbs is used for actions that are inherently reciprocal. Typically, transitive reciprocals have both arguments as Agents, participating in a bi-directional activity, e.g. *geddisó* TR 'exchange, barter', while in intransitive reciprocals the subjects participate in a joint activity, e.g. *googgaaleysó* INTR 'play a game of riddles' and *jaraysó* INTR 'play Somali checkers'. As with reflexives, the grammar allows a distinction between inherently reciprocal activities, expressed by middle verbs, and reciprocal occurrences of other verbs, using the reflexive and reciprocal pronoun *is* (4.2.4), e.g.:

(33)   *Wày   is   arkeen.*
        DM+they   RE   saw
        'They saw each other.'

Examples of transitive and intransitive middle reciprocals include:

(34)   (a)   *dhinacsocó*   TR   'walk side by side with'
               hoobsó   TR   'share out the milk when milking, or selling milk'
               miileysó   TR   'do by turns, take turns'
               nabadaysó   TR   'greet'
        (b)   *qorituuró*   INTR   'draw lots using varying lengths of wood'
               sheekeysó   INTR   'tell each other stories, chat, socialize'
               waalaysó   INTR   'play a hand-clapping game'
               yoobsó   INTR   'assemble, gather together'

One important subset of reciprocals describe family forming or family strengthening actions, for example:

(35)   *alkumó*   TR   'get engaged, ask for a girl in marriage'
        casabaysó   TR   'take as one's relative, admit kinship with'
        guursó   TR   'marry, take a wife'

A fifth type of middle consists of verbs where the subjects, which may be animate or inanimate, undergo a change of state over which they have no control, what we earlier termed uncontrolled inchoation. These middle verbs are all intransitive. Middle voice verbs used in a similar function are called 'neuter intransitives' in studies of Ancient Greek (e.g. Bakker 1994) and a

similar middle function is also reported for Sanskrit (Klaiman 1991). Examples include:

(36)   *bukó*   'become sick, ail'
         dhimó   'die'
         fooló   'go into labour, start to give birth'
         garaadsó   'reach the age of reason, reach maturity'
         habsó   'collapse, fall to pieces'
         ilkaysó   'teethe, grow teeth'
         naqaysó   'grow green again, sprout (after rain)'
         waaló   'become insane, go mad'

The sixth and final type of middle to be discussed here is the very common and productive class of autobenefactive middle verbs, where the middle affix signals that the Agent subject performs the action of the verb for his or her own benefit. While typically in this use the middle affix is attached to transitive stems, it does not affect the syntactic transitivity of the stem: it simple adds the meaning 'for self' to both intransitives and transitives[53]. Some examples are:

(37)   *beeró*   TR   'farm for oneself'   (←*bèer* TR 'farm, cultivate')
         ducaysó   INTR   'pray for oneself'   (←*ducèe* INTR 'pray')
         kaydsó   TR   'store for oneself' (←*káydi* CAUS 'store, preserve')
         qaadó   TR   'take for oneself'   ←*qàad* TR 'take')

5.2   Derived adjectives

As described in 4.4 there is a restricted set of basic adjectives: most adjectives are derived from verbs by the stativising suffix *an*, which may also derive adjectives from noun stems. With transitive verbs the derived adjective is what we might call 'object oriented': for example with causative verbs, either root or derived, the adjective describes the state achieved by the Causee, for example (where again *s* is an allomorph of the causative affix *is*):

(38)   *baduugán*   'broken in pieces'   ←*badùug*   TR   'break in pieces'
         buusán   'dented, squashed'   ←*bùus*   TR   'dent, squash'
         carbisán   'tamed, trained'   ←*cárb-i*   CAUS   'tame, train'
         fuqsán   'pulled out, uprooted'   ←*fúj-i* (←**fuq-i*) CAUS 'pull out'
         karsán   'cooked'   ←*kár-i*   CAUS   'cook'
         laabán   'folded'   ←*làab*   TR   'fold'

solán	'roasted'		←*sól*	TR	'roast, grill'

The effect with factitive verbs is similar, for example:

(39)	*afaysán*	'sharpened'	←*afèe*	FACT	'put an edge on, sharpen'
	buleysán	'tasselled'	←*bulèe*	FACT	'put tassels on'
	sumaysán	'poisoned'	←*sumèe*	FACT	'give poison to'

With other transitive verbs the derived adjective describes the effect of the verb on the object argument, e.g.:

(40)	*basaasán*	'spied on'	←*basàas*	TR	'spy on'
	caayán	'insulted'	←*càay*	TR	'insult'
	xulán	'chosen, selected'	←*xúl*	TR	'choose, select'

With intransitive inchoative verbs the derived adjective describes the resulting state, for example:

(41)	*aamusán*	'silent, quiet'	←*áamus*	INTR	'become silent'
	bararán	'swollen'	←*bárar*	INTR	'swell up'
	cayilán	'fat'	←*cáyil*	INTR	'get fat'
	engegán	'dry, dried out'	←*éngeg*	INTR	'become dry'
	qoyán	'wet, damp'	←*qóy*	INTR	'become wet, damp'

The ability of this affix to attach to either root verbs or derived stems sometimes allows contrasts of meanings. The intransitive inchoative *yàab* 'be surprised', for example, forms the adjective *yaabbán* 'in a state of surprise', while the derived causative *yáabi* 'cause to be surprised, surprise' forms the adjective *yaabsán* 'in a state of having been surprised'.

The affix also derives adjectives from nouns, for example *gaabán* 'short' from *gàab* N M 'shortness' or *ballaadhán* 'wide, broad' from *ballàadh* N M 'width, space'. Interestingly there are only three basic colour adjectives, described in 4.4.1: *cád* 'white', *cás* 'red' and *madów* 'black'. Others are derived from nouns, for example:

(42)	*cagaarán*	'green'	←*cagàar*	N M	'greenery, green grass'
	cawlán	'yellow, tan'	←*càwl*	N M	'yellowness'
	guduudán	'red-brown'	←*gudùud*	N M	'reddish-brownness'[54]

5.3 Derived nouns

As with verbs, derivational processes may be used to create nouns from other grammatical categories or from other nouns. We outline some of the main types in this section.

5.3.1 Verb roots as nouns

As mentioned earlier, many verb stems can be used as nouns without any additional suffixes. We may call this process category shift. If the verb stem is a bare root, then the same form may be used as a masculine noun, e.g.:

(43) *dàaq* V 'graze' *dàaq* N M 'grazing, pasture'
      káb V 'repair' *káb* N M 'a repair, repairing'
      qósol V 'laugh' *qósol* N M 'a laugh, laughter'

As shown in the glosses the noun may be either concrete or abstract. If the verb stem includes lexical affixes, these affixes occur in their basic form, e.g.:

(44) *gaajòod* N M 'experience of hunger' (cf. *ood* EXP)
      dúqow N M 'ageing, getting old' (cf. *ow* INCH)
      yaràys N M 'reducing, making smaller' (cf. *ays* FACT)
      káris N M 'cooking, cookery' (cf. *is* CAUS)
      háfad N M 'drowning' (←*háfat*) (cf. *at* MID)

5.3.2 Nominalising affixes

There are several types of nouns formed by adding nominalising suffixes to verb and adjective stems. Feminine abstract verbal nouns may be created from verb stems by the suffixes *id* for root verbs, *n* for factitive and causative verbs, and *asho* for middle verbs, for example:

(45) Verb Verbal noun
  dàaq INTR 'graze' *daaqíd* N F 'grazing, browsing'
  dhammèe FACT 'end' *dhammáyn* N F 'ending, completion'
  dáaji CAUS (←*dáaqi*) 'put to pasture' *daajín* N F 'pasturing'
  beeró MID 'farm for oneself' *beerashó* N F 'farming for oneself'[55]

The root verb nominalising affix *id* is sometimes replaced by an affix *is*; thus, for example, from the verb *qór* TR 'write' we find both *qoríd* N F 'writing' and *qorís*

N F 'writing'. Currently there seems to be no consistent meaning difference between these two forms.

Root verbs may also form masculine verbal nouns with a suffix *itaan*, for example:

(46)   Verb   Verbal noun
    bàadh TR   'search'   *baadhitàan* N M   'research'
    báx   INTR 'depart'   *bixitàan* (←*baxitàan*) N M 'departure'
    fúr   TR   'open'   *furitàan*   N M   'opening'
    jòog TR   'stay in/at' *joogitàan*   N M   'staying'

In addition to these basic verbal nouns, there are a number of affixes which derive nouns from verbs with specific semantic modifications. There is for example an affix *niin* which for a root verb meaning V creates a noun meaning 'the act of V-ing'; some of the derived nouns are masculine, others feminine, e.g.:

(47)   Verb   Derived noun
    dég   V INTR 'descend'   *degnìin*   N M 'act of descending, landing'
    díg   V TR   'warn'   *digníin*   N F 'act of warning'
    dhís V TR   'build'   *dhisnìin*   N M 'act of building, construction'
    fúr   V TR   'open'   *furnìin*   N M 'act of opening, divorce'
    tún   V TR   'hammer'   *tuníin*   N F 'hammering'
    (←*túm*)   (←*tumníin*)

A very productive agentive/instrumental suffix *e* (*ye* after stems ending in *i*) creates from a verb meaning V a noun meaning 'doer of V'. The derived noun is a member of declension 7 and therefore masculine, e.g.:

(48)   Verb   Derived noun
    bár   TR   'teach'   *baré*   N M 'teacher'
    bíxi   CAUS   'pay'   *bixiyé* N M 'payer'
    fúr   TR   'open'   *furé*   N M 'opener, key'
    qór   TR   'write'   *qoré*   N M 'writer'
    qósol   INTR   'laugh'   *qoslé* N M 'laughcr'

This affix may also be used to form nouns from adjectives, where for an adjective A, the derived noun means 'one who is A', for example *weyné* N M 'one who is big' from *wèyn* ADJ 'big' or *yaré* N M 'one who is small' from *yár* ADJ 'small'.

A further affix *aal* creates for a verb meaning V, a derived noun meaning 'product of V'. These nouns are usually masculine, though feminine nouns do occur:

(49)   Verb   Derived noun
      gúd   INTR 'travel by night'   *gudàal*   N M   'night spent travelling'
      húb   TR   'be sure of'   *hubàal*   N M   'certainty'
      qór   TR   'write'   *qoràal*   N M   'writing, text, document'
      túb   TR   'pile, heap'   *tubáal*   N F   'pile, heap'
      tús   TR   'show'   *tusàal*   N M   'indicator, index'

A suffix *aan* creates nouns from adjectives. From an adjective describing a state A, this affix creates a feminine abstract noun meaning 'A-ness', e.g.:

(50)   Adjective   Derived noun
      cás   'red'   *casáan*   'redness'
      dhèer   'long'   *dheeráan*   'being long, length'
      sán (←*sam*)   'good'   *samáan*   'goodness'
      wèyn   'large'   *weynáan*   'being large, largeness'
      yár   'small'   *yaráan*   'being small, smallness'

There are a number of other, more restricted nominalising affixes whose use is either lexically restricted or subject to semantic shift. Indeed the more productive affixes we have described thus far are also subject to these lexical forces, in ways that remain to be documented. Among the more restricted affixes, there are, for example, two agentive affixes which parallel the affix *e* mentioned above: the first is *aa* which seems to be archaic: for example *gabyàa* N M 'poet' from *gábay* V INTR 'recite a poem', *qoràa* N M 'writer' from *qór* V TR 'write'. The second is the agentive/occupational affix *to* which forms collective nouns belonging to declension 6, for example:

(51)   Verb   Derived noun
      gargàar V INTR   'help'   *gargaartó* N F   'nurses, helpers'
      ganacsó V MID   'trade, do business' *ganacsató* N F   'traders'
      kalluumèe V FACT 'fish, catch fish'   *kalluumaysató* N F 'fishermen/women'

Once again the derived nouns contain the basic forms of the lexical affixes *at* MID and *ays* FACT. These collective nouns can also be used to refer to a single member of the collective, e.g. a single trader or fisherman.

There are also a number of more restricted ways of forming verbal nouns, including the affix *tin*, for example *báq* V INTR 'be afraid' →*baqdín* (←*baqtín*) N F 'fear'; *kúf* V INTR 'fall, slip' → *kuftín* N F 'falling over'.

5.3.3 Noun to noun derivation

There are a number of suffixes which alter the meaning of nouns to form derived nouns. Once again such suffixes form a continuum from very productive elements to others which are more restricted and in some cases have become semantically opaque. Here we list some of the more productive and semantically transparent suffixes.

5.3.3.1 Derived abstract nouns

There are several suffixes which when added to concrete nouns create related abstract nouns. The difference of meaning between these affixes is sometimes difficult to characterise. This difficulty is compounded by degrees of variability and semantic shift associated with particular words. We attempt to reflect semantic distinctions in the English glosses but these are of course only approximations.

A suffix *nimo* is added to a concrete noun N to form an abstract noun 'being N' or 'the state of N'. The derived nouns belong to declension 6 and are therefore feminine, for example:

(52)   Noun   Derived noun
  askár N F 'soldiers' *askarnimó* 'being a soldier, military service'
  dóqon N M 'fool'   *doqonnimó* 'being a fool, foolishness'
  míd N M/F 'one'   *midnimó*   'being one, unity'
  náag N F   'woman' *naagnimó* 'being a woman'
  saaxìib N M 'friend' *saaxiibnimó* 'being a friend, friendship'
  wìil N M   'boy'   *wiilnimó*   'being a boy, boyhood'

A suffix *tooyo* is more restricted in use: typically for concrete nouns denoting people it creates abstract nouns meaning variously 'acting like N', 'experience of N' or 'role of N'. Once again the derived nouns belong to declension 6 and are therefore feminine, for example:

(53)   Noun   Derived noun
  bóqor N M 'king'   *bortooyó*   'kingdom'
  mádax N M 'head'   *madaxtooyó* 'presidency'
  saaxìib N M 'friend' *saaxiibtooyó* 'friendliness'
  walàal N M 'brother' *walaaltooyó* 'brotherhood'

One suffix (*i*)*yad* seems to be a borrowed affix from Arabic, -*iyat*, and is used to form feminine abstract nouns from loan words, for example: *dibloomasiyád* 'diplomacy', *teknoolojiyád* 'technology', *Maarkisiyád* 'Marxism'.

5.3.3.2   *Possessive suffix* le

The addition of the suffix *le* to a noun N creates a masculine noun belonging to declension 7, meaning 'owner of N' or 'possessor of N', for example:

(54)   Noun   Derived noun
    biyó   N M   'water'   *biyóole*   'water-owner, one who possesses water'
    géel   N M   'camels'   *géelle*   'camel-owner'
    hílib   N M   'meat   *hílible*   'possessor of meat, butcher'
    wáran   N M   'spear'   *wáranle*   'possessor of spear, warrior'

As shown above, if the base noun ends in *o* the vowel is lengthened before the suffix.

These derived nouns are in fact on the borderline between derivational suffixes and compounds since they are clearly related to headless relative clauses with the adjective *léh* 'possessing, owning'. See for example the relative clause:

(55)   *nín   dukàan   léh*
              man   shop   owning
              'a man who owns a shop'

By deleting the head noun phrase and suffixing the adjective (dropping *h*) we can form the associated derived noun *dukáanle* 'shop-owner'. Thus it is possible to view these derived nouns as the result simply of suffixation on a base noun, like *dukàan* N M 'shop', or as a way of nominalising a clause. Moreover, depending on the analysis of *le*, this can be seen as either morphological derivation or as the kind of noun-verb/adjective compounding we describe in 5.4.3 below.

5.3.3.3   *Antonym suffix* darro

The suffix *darro* may be attached to certain abstract nouns, usually denoting something good or praiseworthy and forms a feminine declension 6 noun denoting an unfavourable antonym, for example:

(56)   *ayàan*   N M   'good luck   *ayaandarró*   'bad luck'
       aqóon   N F   'knowledge'   *aqoondarró*   'ignorance'

áxdi	N M	'oath, treaty'	*axdidarró*	'oath/treaty-breaking'
bállan	N M	'promise'	*ballandarró*	'promise-breaking'
diin	N F	'religion'	*diindarró*	'lack of religion'
gár	N F	'justice'	*gardarró*	'injustice, wrong'

The affix can also be used as a more general privative suffix with concrete nouns:

(57)	*áf*	N M	'mouth, speech'	*afdarró*	'muteness, speechlessness'
	áf	N M	'edge, sharpness'	*afdarró*	'lack of edge, bluntness'
	dhìig	N M	'blood'	*dhiigdarró*	'anaemia'

5.4   Compounds

Despite the fact that Somali has a wide and productive range of compounding processes, this area of the grammar has received little attention.[56] The important role of compounding in the creation of new vocabulary after the adoption of Somali as the national language of Somalia is described in Caney (1984). However there have been very few linguistic studies of compounds. Here we limit ourselves to an overview of the main types. Since, as outlined in chapters 3 and 4, accentual patterns (APs) are associated with words, they provide a formal criterion for identifying compounds. We can take as an example the sentence in (58), in which each independent word has its own accentual pattern:

(58)	*Hór*	*ú*	*mári*!
	front	to	pass-CAUS-IMPER:SG
	'Move it to the front, move it forward!

These words can be combined to form a compound verb *horumári* CAUS 'cause to progress, advance' and this can be nominalised to form the compound noun *horumarín* N F 'progress, development'. These compounds are single words and therefore have the single accentual pattern, AP2 for the verb and AP1 for the noun, associated with their grammatical status. As we can see, the tones which realise the accentual patterns make clear the distinction between the sentence and the compound verb.

Despite this formal criterion, it is still sometimes difficult to draw a line between compounds and phrases, with a corresponding problem distinguishing between lexical and syntactic processes. We can provide one example to illustrate this. The quantifying adjective *badán* 'much, many' is typically used

alone with a subject nominal as in (59) below; as usual the subject and any other nominals must occur outside the verbal group:

(59)   *Dád-ku*   *wày*   *badan-yihiin.*
            people-the   DM+they many-are
            'The people are many.'

However there is a construction in which a noun may occur within the verbal group immediately preceding this adjective:

(60)   *Nín-ku*   *wuu*   *xòog*   *badan-yahay.*
            man-the   DM+he   strength   much-is
            'The man is very strong.'

For any noun N preceding the adjective *badán* the meaning is something like 'much in N'; thus we can find:

(61)   *Gabádh-dhu*   *way*   *qurúx*   *badan-tahay.*
            girl-the   DM+she beauty   much-is
            'The girl is very beautiful.'

The problem here is how to treat combinations like *xòog badán* and *qurúx badán*. The accentual patterns do not identify a single word, a compound adjective like **xoogbadán* 'great-in-strength' or *quruxbadán* 'great-in-beauty'.[57] Accentually, the nouns remain independent, showing the same tonal pattern as when, for example, heading a noun phrase, e.g. *qurúx badán* 'beauty which is much, great beauty'. However the nouns *xòog* and *qurúx* in (60) and (61) are occupying positions inside the verbal group not normally available to nominals (see chapters 6 and 7). Thus it is not possible to analyse *qurúx badán* as a nominal complement of the copula *yahay* in (61) (meaning something like 'She is a great beauty'): such a complement cannot occur between the sentence type marker *waa* and the verb. What we seem to have here is a process somewhere between straightforward compounding and normal phrasal syntax, where this adjective may incorporate nouns to form a phrase. Moreover the process is fairly productive: any semantically plausible nominal may participate.[58] It may seem preferable on balance to view this as a syntactic rather than a lexical process but the decision is not an easy one. With such difficulties in mind we outline some of the major types of compounding.

    Compounds can be classified in a number of ways, for example by the category of the derived word, the relationship to syntactic phrases, whether the

meaning of the compound is predictable from its parts, etc. In this section we discuss compounds by their grammatical category.

5.4.1 Compound nouns

Compound nouns are very common and consist of a number of different types. One class we could call lexicalised noun phrases. These have the form of syntactically well-formed noun phrases, though of course their accentual pattern identifies them as single words. These compounds are therefore head-initial and may include various types of modifier elements, for example adjectives in (62); nouns with genitive case endings in (63); and as described earlier, relative clause constructions containing a head, a complement noun and the adjective *léh* 'possessing' in (64):

(62)	(a)	*badwéyn*	N F	'ocean' (lit. 'big sea')	
		from:	*bád*	N F	'sea'
			wèyn	ADJ	'big'
	(b)	*afwéyn*	N F	'first cervical vertebra' (lit. 'big mouth')	
		from:	*áf*	N M	'mouth'
			wèyn	ADJ	'big'
(63)	(a)	*libaaxbadèed*	N M	'shark' (lit. 'lion of sea')	
		from:	*libàax*	N M	'lion'
			bád	N F	'sea'
			-eed		genitive case suffix (see 4.1.1.4)
	(b)	*abeesabadéed*	N F	'moray eel' (lit. 'snake of sea')	
		from:	*abbéeso*	N F	'venomous snake'
			bád	N F	'sea'
			-eed		genitive case suffix (see 4.1.1.4)
	(c)	*sunindhòod*	N M	'tear gas' (lit. 'poison of eyes')	
		from:	*sún*	N F	'poison'
			indhó	N M	'eyes'
			-ood		genitive case suffix (see 4.1.1.4)
(64)	(a)	*baqabéenle*	N M	'feigned retreat' (lit. 'retreat which has/is a lie')[59]	
		from:	*baqó*	N F	'retreat'
			béen	N F	'lie'
			le		nominalising suffix 'possessor' from *léh* ADJ 'possessing' (see 5.3.3.2)
	(b)	*mansa-cagáley*	N F	'monitor lizard species' (lit. 'snake having feet')	
		from:	*mansó*	N F	'snake'
			cagó	N M	'feet'
			léh	ADJ	'possessing'
			y		nominalising suffix

A second class of compound noun is composed of the sequence noun + noun (N + N). These compounds are also head-initial and parallel independent genitive constructions, though once again they are tonally marked as words, for example:

(65)　(a)　*afgúri*　N M　'dialect'　　　(lit. 'language-of-house')
　　　　from:　*áf*　　N M　'language'
　　　　　　gúri　N M　'house, home'
　　(b)　*birdánab*　N M　'magnet'　　　(lit. 'iron-of-electricity')
　　　　from:　*bír*　　N F　'iron, metal'
　　　　　　dánab　N M　'thunderbolt, electricity'
　　(c)　*dábdhul*　N M　　'volcano'　　(lit. 'fire-of-earth')
　　　　from:　*dáb*　　N M　'fire'
　　　　　　dhúl　N M　'earth, ground, soil'
　　(d)　*xiddígdhul*　N M　'hedgehog'　　(lit. 'star-of-ground')
　　　　from:　*xiddíg*　N F　'star'
　　　　　　dhúl　N M　'earth, ground, soil'

There are few examples of N + N compounds where the head occurs finally: one productive exception is provided by compounds with the noun *aqóon* N F 'knowledge' as the second element, for example:

(66)　*beeraqóon*　　N F　'agronomy'　　cf.　*béer*　　N F 'farm'
　　cayayaanaqóon　N F　'entomology'　　cf.　*cayayàan*　N M 'insects'
　　ciraqóon　　N F　'meteorology'　　cf.　*cír*　　N M 'sky, atmosphere'
　　nololaqóon　　N F　'biology'　　cf.　*nolól*　　N F 'life'[60]

The third and perhaps the commonest class of compound nouns are deverbal, or synthetic, compounds, where the final element is a verbal noun derived from a verb root. In these compounds the preceding noun is semantically related to the verbal noun in the same way as it would be to a verb in a corresponding clause. We can classify these compounds by the particular role the first noun bears in relation to the verbal noun. We begin with probably the commonest type: where the verbal noun is formed from a transitive verb stem and the preceding noun functions semantically like its object. Here there are two sub-types. The first, exemplified in (67) below, consists of compounds where the verbal noun is derived by category shift of the root as described in 5.3.1 (sometimes called 'zero affixation' or 'conversion'), and the subject argument is unspecified.

(67) (a)   *adhíjir*   N M   'shepherd' (lit. 'sheep and goats-watch over')
              from: *ádhi*   N M   'sheep and goats'
                        jír   V TR   'watch out for, take care of '
      (b)   *bádmar*   N M   'seafarer, sailor' (lit. 'sea-pass-over')
              from: *bád*   N F   'sea'
                        már   V TR   'pass over, by, through'
      (c)   *dabdámis* N M   'fireman, fire-fighter' (lit. 'fire-extinguish')
              from: *dáb*   N M   'fire'
                        dámi(s) V TR CAUS 'extinguish, put out'

The second sub-type consists of compounds where the subject argument is
overtly present in the form of an agentive suffix in the verbal noun:[61]

(68) (a)   *afcelíye* N M   'interpreter' (lit. 'language-returner')
              from: *áf*   N M   'language
                        céli   V CAUS   'turn back, send back'
                        -e   agentive suffix (see 5.3.2)
      (b)   *xoghayé* N M   'secretary' (lit. 'confidence-keeper')
              from: *xóg*   N F   'secret, confidence'
                        hày   V TR   'hold, keep'
                        -e   agentive suffix (see 5.3.2)

Deverbal compounds also occur with the first noun in the compound acting as
an oblique argument, i.e. governed by and followed by a verbal adposition, as in
(69) below, or performing an adverbial function, as in (70):

(69) (a)   *cidlajóog* N F 'domestic animal' (lit. 'stay-with-people')
              from: *cíd*   N F   'people, family'
                        lá   ADP   'with'
                        jòog   V TR/INTR 'stay'
      (b)   *dalkucelín* N F 'repatriation'   (lit. 'send-back-to-country')
              from: *dál*   N M   'country'
                        kú   ADP   'to'
                        céli   V CAUS   'turn back, send back'

(70) *shindhálad*   N M   'child born at full term' (lit. 'proper-time-be born')
       from: *shín* N M   'season, proper time'
                dhaló V MID INTR 'be born'

In these compounds, like those in (67), the verbal noun is formed by simple
category shift (or zero affixation) and the subject argument is unspecified.

In some deverbal compounds the first noun acts as subject of the verbal noun, which is derived by category shift from an intransitive verb, for example:

(71) (a) *biyádhac* N M 'waterfall' (lit. 'water-fall')
      from: *biyó* N M 'water'
             dhác V INTR 'fall'
      (b) *caloolxanùun* N M 'stomach ache' (lit. 'stomach-ache')
      from: *calóol* N F 'stomach'
             xanùun V INTR 'be in pain, ache'
      (c) *dabaqallòoc* N M 'scorpion' (lit. 'tail-crook')
      from: *dabó* N F 'tail'
             qallòoc V INTR 'be bent, curved'
      (d) *qorraxsóobax* N M 'dawn, daybreak' (lit. 'sun-come-out')
      from: *qorráx* N F 'sun'
             sóo VEN
             báx V INTR 'go/come out'

In our examples so far of deverbal compounds we have used relatively simple examples. More complicated examples are very common since either of the nominal elements may be more complex. For example, corresponding to the compounds in (67) we find examples where the first nominal is a noun phrase rather than a simple noun:

(72) *warxungèe* N M 'large poisonous spider sp.' (lit 'bad-news-carry')
      from: *wár* N M 'news'
             xún ADJ 'bad'
             gèe V TR 'carry, convey, bear'

Similarly the verbal noun may be derived from a verb which is itself derived from another category, noun or adjective. For example, paralleling the compounds in (68) we find:

(73) *codweyneeyé* N M 'loudspeaker' (lit. 'voice-magnifier')
      from: *cód* N M 'voice'
             wèyn ADJ 'big'
             ays factitive affix (here *ee*; see 5.1.3)
             ye agentive suffix (here *ye*; see 5.3.2)

Here the agentive verbal noun *weyneeyé* 'magnifier' (also 'microscope') is based on the verb *weynèe* 'make bigger, increase', which is derived from the adjective

wèyn 'big'. The processes of derivation and compounding may be repeated in this way to form increasingly complicated compound nouns.

5.4.2 Compound verbs

Compound verbs may be formed by a process paralleling the verbal compound nouns described in the last section, where the verb is joined by one or more nominal arguments to form a compound verb. A transitive verb may, for example be joined with an object argument with the result being either an intransitive compound verb as in (74) or a transitive verb as in (75):

(74) (a)   *biyyabèel*   V INTR   'dry out'   (lit. 'lose-water')
            from: *biyó*   N M   'water'
                       bèel   V TR   'lose'
      (b)   *cagajìid*   V INTR   'walk slowly'   (lit. 'drag-feet')
            from: *cagó*   N M   'feet'
                       jìid   V TR   'drag'
(75) (a)   *afdùub*   V TR   'kidnap'   (lit. 'bind-mouth')
            from: *áf*   N M   'mouth'
                       duub   V TR   'fold, bind (up)'
      (b)   *dhígdhab*   V TR   'slaughter'   (lit. 'strike-jugular')
            from: *dhíg*   N F   'jugular vein'
                       dháb   V TR   'strike'

Once again compounds on intransitive verbs may include a subject nominal:

(76) (a)   *hiyíkac*   V INTR   'become excited, emotional'
            from: *híyi*   N M   'heart, emotions'
                       kác   V INTR 'rise'
      (b)   *néefgur*   V INTR   'pant, gasp'
            from: *néef*   N F   'breath'
                       gúr   V INTR 'run dry, become empty'

Some verbs are very productive in the formation of compounds, for example *báx* TR 'go/come out (of), leave, exit' as shown in examples like (77); and *gál* TR 'enter, go in' as in (78):

(77)   *áarbax*   INTR   'take revenge'   cf. *àar*   N M   'revenge'
       arrádbax INTR   'get new clothes'   cf. *árrad* N M   'lack of clothing'
       dhíigbax INTR   'bleed, haemorrhage'   cf. *dhìig* N M   'blood'
       jéelbax   INTR   'satisfy one's needs'   cf. *jèel* N M   'need, desire'
       rúuxbax   INTR   'die'   cf. *rùux* N M   'soul, spirit'

(78)   *calóolgal*   INTR   'be conceived'   cf. *calóol*   N F   'stomach'
     dabbágal   TR   'follow'   cf. *dabó*   N F   'after, behind'
     dhaqángal   INTR   'be established'   cf. *dháqan*   N M   'custom'
     garábgal   TR   'support'   cf. *gárab*   N M   'shoulder'
     nabádgal   INTR   'be safe'   cf. *nabád*   N F   'peace'
     tabárgal   INTR   'get strong'   cf. *tabár*   N F   'strength'

Although they are not strictly compounds since the two elements do not form a single word, we can also identify a class of phrasal verbs, where a verb together with one or more verbal adpositions forms a semantic unit whose meaning is not predictable from its parts. We can give a few examples with each of the adpositions *ú*, *kú*, *ká* and *lá* described in 4.5:

(79)   Phrasal verbs
     ú baró TR   'get used to'   cf. *baró* MID TR 'learn'
     ú fadhiisó INTR   'convene'   cf. *fadhiisó* MID INTR 'sit down'
     ú húr TR   'sacrifice'   cf. *húr* INTR 'be kindled, catch fire'
     kú àad TR   'be due on, fall on' cf. *àad* V TR 'go to, head for'
     kú wadó INTR   'be in the lead'   cf. *wadó* MID TR 'carry, drive for oneself'
     kú túf TR   'bless'   cf. *túf* TR 'spit (out)'
     ká áammus TR   'abstain'   cf. *áammus* INTR 'be silent'
     ká hél TR   'enjoy'   cf. *hél* TR 'find, get, receive'
     ká qáb TR   'owe'   cf. *qáb* TR 'hold, have'
     lá baxsó TR   'elope with'   cf. *báxsó* MID INTR 'escape, flee'
     lá hày TR   'be attracted to'   cf. *hày* TR 'have'
     lá tág TR   'steal'   cf. *tág* V TR 'go (to)'

5.4.3 Compound adjectives

Compound adjectives may be formed in similar ways to compound verbs, for example an adjective may join with a noun, which may act as a subject argument, as in (80), or an object argument as in (81):

(80)   (a)   *afgaabán*   'taciturn, reserved'
        from: *áf*   N M   'mouth'
               gaabán ADJ   'short'
     (b)   *foolxún*   'ugly'
        from: *fòol*   N M   'face'
               xún   ADJ   'bad'
     (c)   *qabwèyn*   'arrogant'
        from: *qáb*   N M   'pride'
               wèyn   ADJ   'big'

      (d)   *xeeldhèer*   'expert'
               from: *xéel*   N F   'experience, cunning'
                      dhèer   ADJ   'tall, deep'
      (e)   *xoogyár*   'weak'
               from: *xòog*   N M 'strength'
                      yár   ADJ   'small'
(81) (a)   *dulsaarán* 'placed on top of, staying with'
               from: *dúl*   N F   'top'
                      saarán ADJ 'placed on'
      (b)   *indhalá'*   'blind'
               from: *indhó*   N M 'eyes'
                      lá'   ADJ   'lacking'
      (c)   *warmòog*   'uninformed'
                from: *wár*   N M 'news'
                      má   NEG 'not'
                      óg   ADJ   'aware of'

As described earlier these examples are similar to constructions where adjectives like *badán* 'much, many' incorporate nouns. However the compounds in (80) and (81) are accentually marked as single words, containing a single High mora. It is also possible to find examples of phrasal adjectives, whose meaning once again is not transparent:

(82)   *ká cád*   'in poor shape, desperate'   cf. *cád* ADJ 'white'
      ugú wacán   'causing, responsible for'   cf. *wacán* ADJ 'good'

Chapter 6
Phrases

In this chapter we outline how the lexical categories described in chapter 4 combine to form phrases, the syntactic units that form the building blocks of sentences. Here we concentrate on their internal structure; in chapter 7 we will outline how they combine with other elements like sentence type markers and focus words to form complete sentences of different types. There are two main phrasal elements: the verbal group, where a verb combines with elements like adpositions and clitic pronouns, and the noun phrase. One feature of Somali phrasal syntax is that dependent elements often occur as clitics and in a fixed order. Apart from these two main types of phrase, it is possible to view the combinations of verbal adpositions and the noun phrases they govern as forms of discontinuous phrases, as we discuss in 6.3.

6.1   Verbal group (VGP)

The verbal group is the core element of a clause and consists minimally of a verb but may also include adpositions, clitic pronouns, adverbial clitics, and in auxiliary verb constructions, a dependent verb. We use the term *verbal group* rather than *verb phrase* since the latter is associated in languages like English with a syntactic unit comprising a verb and its phrasal complements, particularly noun phrases. As described in Saeed (1984) there is no syntactic evidence for such a unit in Somali: noun phrase objects, for example, do not form a constituent with the verb, nor occur in a fixed position relative to the verb, as we shall see in chapter 7. The verbal group on the other hand may contain clitic pronouns representing any argument of the verb, including subjects, and they occur in a fixed position relative to the verb and other elements.

6.1.1 The structure of the VGP

The verbal group is a head-final phrase with dependent elements occurring as clitics preceding the head verb. Its structure may best be described by a template:

(1) [$_{VGP}$ S PRO - O PRO I - ADP - ADV I - ADV II - O PRO II - V] [62]

The elements of the template are as follows, with the section of chapter 4 in which they are described given in parentheses:

(2)   Verbal Group elements
        S PRO: subject clitic pronoun   (4.2.1)
        O PRO I: object clitic pronoun - first series   (4.2.2)
        ADP: verbal adpositions   (4.5)
        ADV I: venitive *sóo* or allative *sii*   (4.10.3)
        ADV II: adverbials *wada* 'together' or *kala* 'apart' (4.10.3)
        O PRO II: object clitic pronoun - second series   (4.2.2)
        V: main verb or infinitive and auxiliary verb   (4.3)

These elements where present must occur in the order given above, with the exceptions noted in 6.1.7 below. We can provide some examples of verbal groups in (3a-7a), together with schematic analyses in (3b-7b):[63]

(3)   (a)   *uu   ú   sheegay*
                he   to   told
                'he told (it etc.) to (him etc.) '
        (b)   [$_{VGP}$ S PRO - ADP - V]
(4)   (a)   *ay   ku   keentay*
                she   you   brought
                'she brought you'
        (b)   [$_{VGP}$ S PRO - O PRO I - V]
(5)   (a)   *aannu   wada   tagnay*
                we (EXC)   together   went
                'we (EXC) went together'
        (b)   [$_{VGP}$ S PRO - ADV II - V]
(6)   (a)   *ay   iigá   sóo   iibsatay*
                she   me+for+from   VEN   bought
                'she bought (it etc.) for me from (him etc.)'
        (b)   [$_{VGP}$ S PRO - O PRO I + ADP + ADP - ADV I - V]

(7)   (a)   *aan   kúu   sóo   iibsán   lahaa*
                I   you+for   VEN   buy   would
                'I would buy (it etc.) for you'

        (b)   [$_{VGP}$ S PRO - O PRO I - ADP - ADV - V - AUX]

It is important to note that the above examples are not full sentences: they require additional elements such as sentence type markers, focus words, etc. to form sentences as described in chapter 7. One striking feature of the VGP is the rigidity of its clitic ordering: not only must the different types of clitic occur in the template order given in (1) above, but when more than one member of a category occurs their relative order is also fixed. We can take as an example of this the combinations of adpositions described in 3.3 earlier. In example (6) above we find the combination *iigá* which is formed by coalescence from the clitics *i* 'me' + *ú* 'for' + *ká* 'from'. In (6) we gloss this sequence as 'for me from (him etc.)', i.e. with the pronoun *i* 'me' governed by the adposition *ú* 'for'; however this sequence can also be interpreted as 'for (him etc.) from me', i.e. with *i* 'me' governed by *ká* 'from'. The order of clitics does not change to reflect the two different patterns of semantic government. We may conclude that the order of clitic elements in the VGP is not directly semantically motivated.[64]

6.1.2 Clitic pronouns in the VGP

Both subject and object clitic pronouns occur in the VGP but they differ in their behaviour. The occurrence of subject clitic pronouns is sensitive to the information structure of the containing sentence. Compare sentences (8) and (9) below, which show two important features of this sensitivity.

(8)   (a)   *Cali   wùu   yimi.*
                Ali   DM+he   came
                'Ali came.'

        (b)   *Cali   wàa   yimi.*
                Ali   DM   came
                'Ali came.'

(9)   (a)   *Cáli   bàa   yimí.*
                Ali   FOC   came
                'ALI came, It was ALI who came.'

        (b)   **Cáli   bùu   yimí.*
                Ali   FOC+he   came
                'ALI came, It was ALI who came.'

The difference between these sentences is that in (9) the subject *Ali* is in focus, shown by the focus word *bàa*.[65] The two sentences exhibit several morpho-syntactic differences, as we will discuss in chapter 7. For now we concentrate on the subject clitic pronoun behaviour. Taking non-focus sentences first: when the context makes the identity of the subject clear, as in the simple sentence (8), the subject clitic is optional.[66] In other words subject clitic 'doubling' is possible and seems to be governed by referential recoverability. In focus sentences on the other hand, the subject NP in focus cannot be doubled by a clitic in the VGP, as the ungrammaticality of (9b) shows.

If we turn to object clitic pronouns: they must obligatorily occur to fill the argument positions of the verb (and any adpositions). See for example the non-focus sentences below:

(10)   (a)   *Axmed   adiga   wùu   ku   arkay.*
             Ahmed   you (SG)   DM+he   you   saw
             'Ahmed saw you.'
       (b)   **Axmed   adiga   wùu   arkay.*
             Ahmed   you (SG)   DM+he   saw
             'Ahmed saw you.'

In the grammatical sentence (10a) we have object clitic doubling: outside the VGP is the independent pronoun *adiga* 'you (SG)' (4.1.3), which acts like a full NP. Within the VGP is a coreferential object clitic pronoun *ku* 'you (SG)'. Sentence (10b) without this clitic pronoun is ungrammatical. Similarly an adposition must have a clitic complement within the VGP:

(11)   (a)   *Axmed   adiga   wùu   káa   qaaday.*
             Ahmed   you (SG)   DM+he   you+from   took
             'Ahmed took it away from you (SG).'
       (b)   **Axmed   adiga   wùu   ká   qaaday.*
             Ahmed   you (SG)   DM+he   from   took
             'Ahmed took it away from you (SG).'

In (11b) the adposition *ká* 'from' is not interpreted as governing the independent pronoun *adiga* 'you (sg)' outside the VGP: the adposition has to have a clitic pronoun complement within the VGP.

In focus sentences, object clitics may double a focused nominal; see for example:

(12)   (a)   *Axmed   adíga   bùu   ku   arkay.*
                  Ahmed   you (SG)   FOC+he   you   saw
                  'Ahmed saw you.'
         (b)   **Axmed   adíga   bùu   arkay.*
                  Ahmed   you (SG)   FOC+he   saw
                  'Ahmed saw you.'

Here the focused object *adíga* 'you (SG)' cannot occur without a clitic in the VGP, paralleling (10).

One way of viewing this patterning, as suggested in Puglielli (1981: 16ff) and Saeed (1993b, 1996), is to say that, with the exception of focused subjects, all of the arguments of a clause are represented in the VGP as clitics, obligatorily for objects and optionally for subjects. The implication of this, discussed further in chapter 7, is that the traditional distinction between arguments and adjuncts is blurred in Somali.

One further feature merits mention. As noted earlier in 4.2.2, both series of object clitic pronouns have zero exponents for third person. To see how this works we could add a declarative sentence type marker *waa* to the verbal group in (3) to form the following full sentence:

(13)   *Wuu   ú   sheegay.*
          DM+he   to   tell
          'He told it/them to him/her/them.'

This sentence can be used where all the nominal referents are known or deictically available and the speaker does not choose to employ nominal focus; as a consequence no full lexical noun phrases are used. Consequently the complements of the verb *shèeg* 'tell' and the adposition *ú* 'to' will be interpreted as contextually determined third person pronouns.

Finally, it is important to note that full noun phrases cannot occur in the verbal group. We can for our example (13) add a NP outside the VGP to form either (14) or (15) below:

(14)   *Wárkíi   wuu   ú   sheegay.*
          news-the   DM+he   to   told
          'He told the news to him/her/them.'
(15)   *Wuu   ú   sheegay   wárkíi.*
          DM+he   to   told   news-the
          'He told the news to him/her/them.'

However the NP *wárkíi* 'the news' cannot interrupt the elements of the verbal group; consequently the sentences below are ungrammatical:

(16)   (a)   * *Wuu   wárkíi   ú   sheegay*
        (b)   * *Wuu   ú   wárkíi   sheegay*

The observation is that lexical noun phrases must occur outside the verbal group.

6.1.3 Infinitival constructions

The head of the verbal group may be a simple verb or a verb governing an infinitive dependent verb, as in:

(17)   (a)   *W*-[$_{VGP}$ *aan   tégi   doonaa*]
                DM-   I   go:INF   will:PRES
                'I will go'
        (b)   *W*-[$_{VGP}$ *aan   tégi   jiray*]
                DM-   I   go:INF   be:PAST
                'I used to go'
        (c)   *W*-[$_{VGP}$ *aan   tégi   lahaa*]
                DM-   I   go:INF   have:PAST
                'I would go'

In these examples the auxiliary verbs *dóon, jír* and *lahaa* carry the inflection and the verb *tág* 'go' is in an invariable non-finite form, the infinitive *tégi*. As outlined in 4.3.2, there are two types of infinitival constructions. The sentences in (17) form the first type, which we can term the grammatical auxiliary construction. In this type these three auxiliary verbs act as part of the tense-aspect-mood (TAM) system: the future marked by *dóon* 'will'; the past habitual with *jír* 'be, exist'; and the conditional with the fused adjective and verb form *lahaa* 'was possessing'. In these constructions the auxiliary verbs have lost their independent lexical content and function as markers of the relevant TAM distinction. Each auxiliary verb itself can only occur in a single TAM form in these constructions: *dòon* occurs in a present general form; *jír* in a past simple form; and *léh* occurs with the contracted past tense forms of *yahay* 'be'. A fourth grammatical auxiliary construction has only a limited distribution synchronically: the progressive forms with *hày* 'hold, have', as described in 4.3.4-6.[67]

The second type of infinitival construction involves a small number of verbs, which like the grammatical auxiliaries take infinitival verb complements. However, unlike the auxiliary verbs, these verbs maintain their lexical content. The most important of these are *kár* 'can, be able', *wàa* 'fail, be unable', and *gàadh* 'just fail, almost do'. As described in 8.4.2, we identify these as main verbs taking infinitival complements. They are also distinguished from grammatical auxiliaries by being able to repeat or 'stack' to form chains within the verbal group, for example:

(18) *oo*   [_{VGP} *ay*   *lá*   *órdi*   *kári*   *weydey*]   [SHS19.32]
     and   she   with   run:INF   can:INF   failed
     'and she could not run with it' (lit. 'and she failed to be able to run with it')

These constructions are interesting because they form the only fully non-finite verbal complements in Somali. As described in chapter 8, most clauses acting as verb complements are finite.

6.1.4 Predicative adjectives

The position of the verb in the VGP may be taken by a predicative adjective. As described in 4.4 adjectives may only occur as complements of the copula verb *yahay* 'be', for example:

(19) *W*-[_{VGP} *uu*   *wanaagsán* *yahay*]
     DM-   he   good   is
     'He is good'

In colloquial speech the copula's root *ah* may be elided and the remaining elements cliticized onto the adjective, for example:

(20) *W*-[_{VGP} *uu*   *wanaagsányay*]
     DM-   he   good-is
     'He is good'

In the past tense and in relative clauses this contraction is an obligatory rule, as the following show:

(21) (a)   **Wuu*   *wanaagsán* *ahaa*
           DM+he   good   was
           'He was good'

      (b)   *Wuu*   *wanaagsánaa.*
              DM+he   good-was
              'He was good.'

(22)   (a)   **nínkii*   *wanaagsán*   *ahaa*
              man-the   good   was:SUBORD
              'the man who was good'

      (b)   *nínkii*   *wanaagsánaa*
              man-the   good-was:SUBORD
              'the man who was good'[68]

Since this contraction rule elides the copula root *ah*, whenever this root is the total copula stem, as in present tense relative clauses or NP focus constructions, the verb disappears altogether, for example:

(23)   *nín*   *wanaagsán*
       man   good
       'a good man, a man who is good'

(24)   *Nínka* *ayàa*   *wanaagsán.*
       man   FOC   good
       'THE MAN is good.'

One result of this, as shown in the gloss above, is that there is no formal difference between a present tense relative clause with a predicative adjective and an attributive use of an adjective as a post-nominal modifier. Perhaps the simplest approach is simply to recognise all attributive adjectives as relative clauses, an approach discussed further in chapter 8.

     As described in 4.4.2, comparative and superlative adjectival constructions are formed with verbal adpositions in the VGP: *ká* for comparatives and the cluster *ugú* (←*ú* + *kú*) for superlatives:

(25)   *Áxmed*   *bàa*   *igá*   *dhèe.r*
       Ahmed   FOC   me+*ká*   tall
       'AXMED is taller than me.'

(26)   *Tanu*   *way*   *ugú*   *wanaagsán*   *tahay.*
       this:SUBJ DM+it:FEM SUPR   good   is
       'This one (FEM) is the best.'

6.1.5 Noun incorporation in the VGP

As mentioned in 6.1.3 full NPs cannot occur within the verbal group. However there is a process of noun incorporation where a noun is attached to a verb or adjective within the VGP in order to adjust its meaning. This combination of

incorporated noun and verb or adjective occupies the place of the verb in the
VGP template in (1) earlier. We can identify two types; in both types only a bare
noun may be incorporated, not a noun phrase, and the process is strongly
lexically restricted. The first type is where a verb incorporates one of a small set
of spatial nouns including *hór* N F 'front', *dhéx* N F 'middle', *díb* N F 'back', *ág* N
F 'proximity', etc. These incorporated nouns may be described as adverbial in
function because they modify the meaning of the verb, for example *dhéx* adds
the meaning 'between' or 'through'. In this type of incorporation the lexical
restriction is on the selection of noun, though the verb is also typically restricted
to a verb of movement or location. Examples are:

(27) *Cáli   iyo   Máxamed   coláad   bàa   dhéx   martáy.*
      Ali   and   Mohamed   emnity   FOC   between   passed
      'EMNITY grew between Ali and Mohamed.'

(28) *Afadii   iyo   ninkeedii   baa   markaas   isa   soo   ag   fariistay*
      wife-the   and   man-her   FOC   time-that   RE   VEN   proximity   sat
      'Then THE WIFE AND HER HUSBAND sat down near each other' [SHS 56.12]

We can see in (28) that the incorporated noun *ag* occurs closest to the verb,
after the venitive *sóo*.

    The second type of incorporation occurs with certain adjectives. In
particular the adjective *badán* 'much, many', as noted in 5.1, may incorporate a
noun to express the range of the quantification, for example:

(29) *Wuu   dád   badán   yahay.*
      DM+it:MASC   people   many   is
      'It is crowded.' (lit. 'It is people-many')

(30) *Way   xòog   badán   yihiin.*
      DM+they   strength   much   are
      'They are very strong.' (lit. 'They are strength-much')

We can see that the noun forms a unit with the adjective *badán* because the
verbal adpositions precede the combination, as in the comparative:

(31) *Way   ká   qurúx   badnayd.*
      DM+she   ADP   beauty   much-was
      'She was much more beautiful.'

This *badán* incorporation construction is a common way of intensifying
attributes, as in the following relative clause examples:

(32)   *inán*   *garàad*   *badán*
        girl   intelligence much
        'a girl who is very intelligent, a very intelligent girl'

(33)   *xooló*   *tiró*   *badán*
        livestock number   many
        'livestock that are very numerous, very numerous livestock'

Similarly, the adjective *lá'* 'lacking' may also incorporate a nominal, for example:

(34)   *yaxaasna*   *weli*   *wuu*   *carrab*   *la'yahay*
        crocodile-and yet   DM+he   tongue   lacking-is   [FS 29.11-12]
        'and crocodile still lacks a tongue' (lit. 'and crocodile is still tongue-lacking')

6.1.6 Permutations in the Verbal Group

The principal variation in the template of the VGP occurs with subject clitic pronouns and follows from their relationship to sentence type markers and the main clause negative word *má* 'not'. Classifiers, like declarative *waa* and interrogative *ma*, occur immediately before the verbal group and in this position attract the subject pronouns, which cliticise on to them, for example:

(35)   *Wùu*   *íi*   *kéenayaa.*
        DM+he   me+for   bring:PRES PROG
        'He is bringing (it etc.) for me.'

(36)   *Miyùu*   *íi*   *kéenayaa?*
        QM+he   me+for   bring:PRES PROG
        'Is he is bringing (it etc.) for me?'

When the negative word *má* occurs it occupies a position within the VGP following the adpositions, for example:[69]

(37)   *Iimá*   *kéenayó.*
        me+for-NEG   BRING:PRES PROG:NEG
        'He is not bringing it for me.'

Subject clitic pronouns are less commonly used in negative sentences than positives but when they do occur they attach to the negative word, for example:

(38)   *Iimá*   *uu*   *kéenayó.*
        me+for-NEG   he   BRING:PRES PROG:NEG
        'He is not bringing it for me.'

Comparing the position of *uu* in (35)-(36) and (38), we can see that subject clitic pronouns occur in different positions in negative and positive sentences. Note that this does not happen with the non-specific subject pronoun *la* which remains in the position given in (1) earlier in both positive and negative sentences, for example (where *lóo* ← *la* 'one' + *ú* 'in'):

(39) *Mùuska   sidáas   lóo   má   cunó.*
      banana-the   way-that   one+in   NEG   eats:NEG
      'One doesn't eat banana like that.'

Similarly, in subordinate clauses subject clitic pronouns follow the complementiser, for example *ín* 'that', or the head NP of a relative clause, which as described in 8.4.1, immediately precedes the clause, for example:

(40) *Wuu   dóonayaa   ín   ay   sóo   iibsató.*
      DM-he   wants   that   she   VEN   buys
      'He wants her to buy it.' (lit. 'He wants that she buys it.')

This behaviour means that in subordinate clauses the subject pronouns may be separated from the VGP by NPs, for example:

(41) *Afádii   wáxay   kú   tashatay   ín   **ay**   gabdháha   iyo   adóogóod*
      wife-the   what-she   on   decided   that she   girls-the   and   father-their
      *isku   dirtó   óo   **uu**   gúriga   ká   eryó   iyága* [SHS 24.16-18]
      RE-on   turn   and   he   house-the   from   throw:out   them
      'The wife decided that she would turn the girls and their father against each other so that he would throw them out of the house'

In the subordinate clauses in (41) the subject pronouns *ay* 'she' and *uu* 'he' (shown in bold) are both separated from their VGPs by full NPs.

To summarise, subject clitic pronouns occupy different positions in negative and subordinate clauses. This variability of position within, and potential detachment from, the VGP differentiates subject clitic pronouns from other elements in the VGP and may justify excluding them from the VGP template in (1) earlier.

6.2   Noun phrase

Noun phrases in Somali are head-initial: the noun is followed by various dependent elements. Some, like determiners, attach to the noun as clitics while

others, like modifying clauses and adjectives, follow the noun as morphologically independent items. In addition noun phrases may be conjoined by *iyo* 'and', described in 4.9, to form larger noun phrases with several equal heads, e.g.:

(42)   (a)   *Xásan iyo Cáli*
              'Hasan and Ali'
       (b)   *hílib iyo barìis*
              'meat and rice'
       (c)   *hílib, barìis iyo caanó*
              'meat, rice and milk'

In this section we concentrate on the dependent elements which may occur with head nouns.

6.2.1 Nouns modified by determiners

Each of the determiners described in 4.6 may modify nouns and occur as suffixed clitics, for example with *gèed* N M 'tree' and *mindí* N F 'knife':

(43)   Definite articles: *gèedka*   'the tree (non-remote)'
                            gèedkíi   'the tree (remote)'
                            mindída   'the knife (non-remote)'
                            mindídii   'the knife (remote)'
(44)   Demonstratives: *gèedkán*   'this tree'
                            mindídán   'this knife'
                            gèedkáa(s)   'that tree' (further away from speaker)
                            mindídáa(s)   'that knife'   "
                            gèedkéer   'that tree' (in the middle distance)
                            mindídéer   'that knife'   "
                            gèedkóo   'that tree' (in the far distance)
                            mindídóo   'that knife'   "
(45)   Interrogatives: *gèedkée*   'which tree?'
                            mindídée   'which knife?'
(46)   Possessives: *gèedkàyga*   'my tree'
                            gèedkàaga   'your tree', etc.
                            mindídàyda   'my knife'
                            mindídàada   'your knife', etc.

As we noted in chapter 4, the possessives in (46) occur with the definite article: nouns for family members, body parts and by extension, close friends

occur without this final article, e.g. *aabbáháy* 'my father'. When combinations of determiners on a noun occur they are always in the following order:

(47)
$$\text{Noun} \quad + \quad \text{possessive} \quad + \quad \left\{ \begin{array}{l} \text{definite article} \\ \text{demonstrative} \\ \text{interrogative} \end{array} \right\}$$

6.2.2 Nouns modified by nominals: possessive noun phrases

Noun phrases expressing possession occur in two types: the first, as we saw in the last section, involves the use of a special possessive determiner on the possessed noun. The order of nominals is possessor - possessed:

(46) *Cáli   gúrigìisa*
      Ali:ABS   house-his-the:ABS
      'Ali's house' (lit. 'Ali his house')

Note that in the unmarked case both nominals occur in absolutive case. The second type of possessive noun phrase shows the order possessed - possessor, for example:

(49) *gúriga   Cali*
      house-the:ABS   Ali:GEN
      'Ali's house'   (lit. 'the house of Ali')

Here in the unmarked case the possessed nominal occurs in the absolutive case while the possessor occurs in the genitive. In typical expressions of possession either construction is possible: we assume that the choice is influenced by the information structure requirements of the discourse context, for example what the current topic is, or which elements the speaker wishes to foreground.

Each possessive construction also occurs in extended uses where they are not interchangeable. For example, as mentioned in 4.1.2, the genitive construction is used in genealogies and by extension in the form of proper names:

(50) *Xásan   Cali*
      Hasan:ABS   Ali:GEN
      'Hasan child of Ali'   (lit. 'Hasan of Ali')

The possessive determiner construction is not used in this way in names.

Another important use of the genitive construction is when numerals are used to count nouns, for example:

(51)   *labá*   *meelóod*
       two:F:ABS   place:F:PL:GEN
       'two places' (lit. 'two of places')

Here the numeral is the head of the construction and the counted noun is dependent: the whole noun phrase has the gender of the numeral and the numeral will attract any determiners: see for example the following feminine noun phrase:

(52)   *labáda*   *nín*
       two-the:F:ABS   man:M:SG:GEN
       'the two men' (lit. ' the two of man')

These two examples also show a distinction in the behaviour of the counted noun: all nouns except declension 1 occur in a singular form after all numerals, as in (52), while declension 1 nouns occur in plural genitive forms after numerals greater than one, as in (51). This genitive construction also occurs with some other quantifying nouns, for example *ímmisa* 'how many', *méeqa* 'how many?' and *dhàwr* 'several', for example:

(53)   *dhàwr*   *nín*
       several:ABS   man:SG:GEN
       'several men' (lit. 'several of man')

As described in 4.6.4, the possessive determiner construction has extended uses in expressing partitive quantification, as in (54), and spatial and temporal location, as in (55):

(54)   (a)   *dádka*   *sáddexdóod*
             people-the   three-their
             'three of the people'
       (b)   *dádka*   *gíddigóod*
             people-the   all-their'
             'all of the people'
(55)   (a)   *mìiska*   *hóostìisa*
             table-the   underneath-its
             'underneath the table'

(b)   *habeennimáda   hórtèeda*
       nightfall-the   front-its-the
       'before nightfall'   (lit. 'nightfall its front')

6.2.3 Other modifiers

Nouns and noun phrases may also be modified by the following dependent items.

6.2.3.1   Adjectives (4.4, 6.1.5)

Adjectives used attributively follow the head noun and suffixed determiners and are formally equivalent to restrictive relative clauses, as discussed in 6.1.5 above.[70]   Some adjectives agree with plural head nominals by a process of reduplication, described in 3.5.   Sequences of adjectives follow the rules for restrictive relatives clauses, i.e. are linked by *oo* 'and' on an indefinite head NP and by *ee* 'and' on a definite NP (8.1).   Some examples are:

(56)   (a)   *shúqul   adág*
               work   hard
               'hard work'
        (b)   *buuró   dhaadhèer*
               mountains   high
               'high mountains'
        (c)   *wìil   dhèer   óo   dhuubán*
               boy   tall   and   thin
               'a tall thin boy'
        (d)   *wìilka   dhèer   ee   dhuubán*
               boy-the   tall   and   thin
               'the tall thin boy'

6.2.3.2   Attributive adjectives (4.4.3)

Attributive adjectives are not used predicatively and do not mark plural agreement by reduplication.   When co-occurring with predicative adjectives or relative clauses they occur first after the head nominal.   There seems little justification to view these as a form of relative clause modification.   Examples:

(57)   (a)   *xáragga   saré*
               skin-the   topmost
               'the outer skin'
        (b)   *dhinac   kastá*
               side   every
               'every side'

(c)   *márkíi   horé   ee   aan   ku   arkáy*
        time-the first   and   I   you   saw:SUBORD
        'the first time that I saw you'

6.2.3.3   Restrictive relative clauses (8.1)

(58)   (a)   *nínkíi   keenáy*
               man-the   brought: SUBORD
               'the man who brought it'
        (b)   *biyáha   qófku   cabbayó*
               water-the   people-the   drink:SUBORD
               'the water which people drink'
        (c)   *méeshaad   tagtó*
               place+the+you   go: SUBORD
               'the place which you go to'

6.2.3.4   Appositive (non-restrictive) relative clauses (8.2)

(59)   (a)   *aabbáhéed   oo   múslin   áh*
               father-her   and   muslim   is:SUBORD
               'her father, who is a Moslem,'
        (b)   *isága   oo   aan   arkáyn*
               he   and   NEG   saw:NEG:SUBORD
               'he, who did not see it,', 'he, not seeing it,'

The details of relative cause structure are given in chapter 8. It is very common
for an attribute of a head nominal to be described by a relative clause containing
a noun as complement of *yahay* 'be' or the adjective *léh* 'having', for example:

(60)   (a)   *nín   xòog   léh*
               man   strength   has:SUBORD
               'a strong man' (lit. 'a man who has strength')
        (b)   *wáx   qíimo   léh*
               thing   value   has:SUBORD
               'a valuable thing' (lit. 'a thing which has value')
        (c)   *méel   khátar   áh*
               place   danger   is:SUBORD
               'a dangerous place' (lit. a place which is danger')
        (d)   *hádal   rún   áh*
               speech   truth   is:SUBORD
               'true speech' (lit. 'speech which is truth')
        (e)   *gèed   gàar   áh*
               tree   specialness   is:SUBORD
               'a special tree' (lit. 'a tree which is specialness')

(f)   *cunnáda   kelí   áh*
        food-the   solitude   is:SUBORD
        'the only food' (lit. 'the food which is aloneness')

6.2.4 Multiple modifiers

One rule is common to all nominal modifiers: after the first modifier all others must be introduced by a conjunction, *oo* or *ee*, both meaning 'and'. So for example if one adds a further nominal modifier to a genitive NP, it must be introduced by a conjunction:

(61)   [head *taariikhdíi*]   [mod *Soomáaliya*]   *ee*   [mod *beryíhíi horé*]
        [history-the ABS]   [Somalia GEN]   and   [times-the   early GEN]
        'the early history of Somalia' (lit. 'the history of Somalia and of early times')

Similarly with any further modification of a quantified nominal:

(62)   *labá   kóob   óo   shàah   áh*
        two   cup:GEN   and   tea   is:SUBORD
        'two cups of tea' (lit. 'two of cup and which are tea')

In examples (56) and (57) above we saw this same behaviour with ordinary and attributive adjectives. The distinction between the conjunctions *oo* and *ee* is discussed further in chapter 8. Basically: on indefinite heads only *oo* is possible and there it is a marker of appositive relative clauses; on definite heads *ee* marks restrictive modification and *oo* appositive modification. Genitive constructions necessarily involve restrictive modification so *oo* is never found in multiple genitives like:

(63)   [head *wakaaládda*]   [mod *warárka*]   *ee*   [mod *dálka   Ciráaq*]
        agency-the   news:PL-the   and   country-the   Iraq
        'the Iraqi News Agency' (lit. 'the agency of news and of the country of Iraq')

6.3   Discontinuous phrases

From the examples and analyses thus far in this chapter it may be clear that much of the information about the semantic relations of a clause is encoded in the verbal group. With one relation in particular this means that an important semantic relationship is not reflected syntactically: the governing of NPs by verbal adpositions. As described in 4.5, the adpositions *ú, kú, ká* and *lá* assign

various semantic roles to nominals outside the VGP including location, source, goal, instrument, manner, comitative, etc. One major function of this relationship between a verbal adposition and a nominal is to provide adverbial modification of the clause, in particular *ú* governing the noun *sí* N F 'manner, way'. See for example:

(64)   *Sidée   bàad   ú   sameysey?*
        way-which   FOC+you   in   did
        'How did you do you it?' (lit. 'In which way did you do it?')

(65)   *Sí   xún   bùu   ú   sameeyey.*
        way   bad   FOC+he   in   did
        'He did it badly.' (lit. 'in a bad way')

The noun *sí* may be modified in any number of ways by determiners, adjectives and relative clauses to form adverbial expressions and it is useful to refer to the discontinuous elements as an adverbial unit, e.g.:

(66)   *sidán...ú*   'thus, in this way'
        sidáas...ú   'like that, in that way'
        sí diirán...ú   'warmly'
        sí qaabdarán...ú   'improperly, incorrectly'
        sí uun...ú   'somehow, in one way or another'
        sí wanaagsán...ú   'well'
        sí badán áh...ú   'abundantly'
        sí cád áh...ú   'clearly, plainly'
        sí gàar áh...ú   'especially'

When talking of these adverbial units we are in fact identifying a discontinuous adverbial phrase, a strategy which is also useful in discussing the commonly used intensifying phrase *àad...ú* 'much, a lot', e.g.:

(67)   *Áad   bùu   ú   jecélyahay.*
        much   FOC+he   in   liking-is
        'He likes it very much.'

Similarly one can identify a number of adverbial expressions involving *ú* governing a noun, for example:

(68)   Adverbial phrases of direction
        díb...ú   'backwards'
        gées...ú   'sideways'

hóos…ú   'downwards'
hór…ú   'forwards'

(69)   Adverbial phrases of manner
        áal…ú   'slowly, gently'
        dégdeg…ú   'quickly'
        dawáali…ú   'directly, straightforwardly'
        sáhal…ú   'easily'

In each of these constructions the adposition remains in the position in the verbal group described in the template in (1) while the governed nominal is outside the VGP with its position subject to the information structure of the clause, as described in chapter 7.

Chapter 7
Simple sentences

In this chapter we examine the structure of simple sentences consisting of a main clause. The various ways of joining simple sentences by conjunctions were described in 4.9 earlier and we leave complex sentences containing subordinate clauses until the next chapter. When looking at sentences it is useful to distinguish between three semantic and pragmatic systems which influence their structure: sentence type, focus and negation. The interaction between these systems is not symmetrical, for example focus is restricted to declarative and interrogative sentence types. Moreover each system is sensitive to the distinction between main (root) clauses and subordinate clauses: sentence type differentiation and focus are restricted to main clauses, and negation shows different forms in main and subordinate clauses, as we shall see in chapter 8. Nonetheless the structure of main clauses is determined by these systems and we organise this chapter in terms of them. We begin with a brief discussion of sentence types, then go on to describe declarative and interrogative sentences, for each type distinguishing the effects of negation and of focus. Thereafter we move onto the other major sentence types.

7.1 Sentence types

The basic pattern is for sentence type to be marked by a combination of verbal inflection and a sentence type identifying morpheme, which for convenience we term a marker (STM) (4.8). As we shall see, not all sentence types are unique in both markers. The sentence types can be classified pragmatically by their conversational uses, for example allowing the speaker to make questions, requests, wishes, etc. or semantically, as making various distinctions of modality, for example expressing judgements of states of affairs as actual or

hypothetical. Grammatically, sentence types fall into two groups distinguished by their ability to mark tense:

(1)

```
                        SENTENCE TYPES
                _____/        _____
           TENSED                        UNTENSED
          /      \                      /    |    \
  declarative  interrogative    imperative optative potential
```
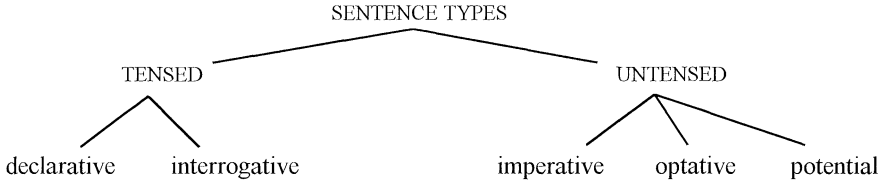
Declarative and interrogative sentences share the same verbal forms, and express the full range of tense and aspectual distinctions described in 4.3 earlier. The other moods each have a unique verbal inflection that does not show tense or aspectual distinctions. The distinction in (1) corresponds to other differences, for example, as mentioned above, nominal focus may be used in the tensed sentence types but not in the untensed types. We begin our discussion with the sentence types which show tense and aspect differentiation: declaratives and interrogatives. In each case there are major morpho-syntactic differences between sentences containing focus and those without.

7.2.   **Declaratives without focus**

Declarative sentences without focus are identified by the sentence type marker *waa*. In some earlier writing on Somali (for example Andrzejewski 1975a) *waa* was taken to be a verb focus particle in complementary distribution with nominal focus words *bàa*, *ayàa* and *wáxa(a)* (see 7.4 below). However as argued in Saeed (1984: 160-91) it seems preferable to view sentences with *waa* as neutral or unspecified for focus and to analyse *waa* as a sentence type marker (4.8). Amongst other factors, *waa* can occur in verbless equative sentences (7.3 below), and can be used in replies to questions with nominal focus. We discuss this issue in more detail in 9.1.3 below.

    As described in chapter 6 a declarative may be formed by adding *waa* to a verbal group (VGP). Since arguments are represented in the VGP by pronominal clitics (4.2) this forms a complete sentence, for example:

(2)   *Wàan   ku   ráacayaa.*
         DM+I   you   accompany:PRES PROG
         'I am going with you.'

The verb form in a declarative shows distinctions of tense (past, present, future) and of aspect (simple, progressive, habitual). The main modal use of declaratives is to describe actual states of affairs. The future tense is used to express a judgement that an event is certain to occur. This sentence type may also be used to express ability and permission by using the auxiliary verb *kár* 'can, be able', for example:

(3)   *Wàad   i   sóori   kartaa.*
        DM+you   me   feed:INF   can:PRES:2 SG
        'You can feed me.'

Typically, declaratives with *waa* are used when all the nominal referents are known and may be omitted. Noun phrases, when used, act like adjuncts or topics and may occur before or after the verbal group, for example:

(4)   *waanadii   waan   kaa   qaatay.*
        advice-the   DM+I   you+from   took
        'I took the advice from you'   [SHN 124. 22-3]
(5)   *waannu   is   mehersannay   aniyo   isagu.*
        DM+we:EXCL   RE   married   I-and   him
        'We married each other I and him'   [SHN 126. 23]

The position and order of these NPs is free and does not reflect their grammatical functions, unlike clitic pronouns within the verbal group which, as we saw in chapter 6, have a subject-object-verb (S-O-V) order. Wherever they occur such NPs are case marked; thus the NP *aniyo isagu* 'I and him' is subject marked, carrying the nominative case form of the definite article *-gu* on the final syllable.

Since *waa* declarative sentences are typically used when all the referents are known, they are common in casual conversation but much less common in narratives or the written prose of newspaper reports, letters, etc., where new information is more frequent and thus focus is used.

Negative declaratives without focus are distinguished by two characteristics: the sentence type marker *waa* is replaced by the negative word *má*; and the verb occurs in a special negative form. As described in 4.3, each verbal paradigm displaying a tense and aspectual distinction has both positive and negative forms. The negative past paradigms (past simple, past progressive and past habitual) are characterised by invariable forms which do not show person agreement. As mentioned in chapter 6, the negative word *má* 'not' occurs within the verbal group and often occurs without subject clitic pronouns, for example:

(6)   (a)   *Qálabkíi   má   jebín.*
              machine-the   not   break:CAUS:PAST:NEG
              'I/you/he/she etc. did not break the machine.'
       (b)   *Kumá   arkín.*
              you-not   saw:PAST SIMPLE:NEG
              'I/he/she/etc. did not see you.'

This suggests that negative declaratives are more context-dependent than corresponding positive declaratives. It is possible however to use subject clitic pronouns in negatives, where they follow *má* 'not':

(7)   (a)   *Ku   má   uu   garánéyn.*
              you   not   he   understood-NEG
              'He did not understand you.'
       (b)   *Kuu   má   aan   yeedhín.*
              you+to   not   I   called
              'I did not call you.'[71]

        Constituent negation may be applied to NPs by adding a following (optionally suffixed) *ba* '(not) any' (especially with inanimate referents) or the negative nominal suffix *-na*:

(8)   (a)   *Háwlba   má   uu   helín.*[72]
              work-NEG   not   he   found:NEG
              'He didn't find any work.', 'He found no work.'
       (b)   *Waxbá   má   aan   síin.*
              thing-NEG   not   I   gave:NEG
              'I didn't give him etc. anything.', 'I gave him etc. nothing.'
       (c)   *Qófna   má   arkin*
              person-NEG   not   saw-NEG
              'I/you/he etc. didn't see anyone', 'I etc. saw no-one'
       (d)   A: *Yàa   joogá?*   'Who is there?'
                  who   stays
              B: *Cídna.*   'No-one.'
                  person-NEG

7.3   Verbless declarative sentences

A special form of *waa* declaratives are the verbless sentences of the form 'A is B', used to express proper inclusion and equation. In these the copula verb *yahay* is not used; the sentence type marker *waa* occurs alone between the

subject NP and the predicate NP. See the examples of proper inclusion in (9) and equation in (10):

(9)   (a)   *Cali   waa   báre.*
                Ali   DM   teacher
                 'Ali is a teacher.'
    (b)   *Qálabkanu   waa   cagafcágaf.*
                machine-this   DM   bulldozer
                 'This machine is a bulldozer.'

(10)   (a)   *Mágacàygu   waa   Maxámed.*
                name-my-the   DM   Mohamed
                 'My name is Mohamed.'
    (b)   *Tani   waa   afádàydíi.*
                this   DM   wife-my-the
                 'This is my wife.'

As shown above, the subject NP is subject marked accentually, and segmentally where relevant, and the predicate nominal has the absolutive case found on all object nominals. Of course the predicate nominal may be complex, for example a relative clause:

(11)   *Carowdu waa cudur lays   qaadsiiyo oo ka   yimaadaa fiiruus*
          rabies-the DM disease one+RE transmits and from comes   virus
          'Rabies is a transmissible disease caused by a virus' [BCH 21.10]

An alternative form of verbless sentence occurs with what seems to be a contraction of *waa* and the copula verb *yahay*, for example (with optional forms in parentheses):

(12)   *Cali   báre   wèeye (wàaye, wàay, wàa'e)*
          Ali   teacher   DM
          'Ali is a teacher.'

While the above forms seem to be naturally derivable from the construction (13) below, the latter is not found:

(13)   **Cali báre   wàa   yahay.*
          Ali   teacher   DM   is
          'Ali is a teacher.'

These *weeye* forms seem to be optional alternatives to verbless *waa* constructions except where the predicate nominal is judged too long. Unlike *waa*, this *wéeye* form may occur following the focus word *wáxa(a)* (7.4.2) to focus the second element in a verbless sentence:

(14)   *Faarax wáxa   wéeyé **nín   géesi   áh**.*
        Farah   *wáxa* DM   man   brave:man   be
        'Farah (what he is) is **a brave man**.'

The subject NP in verbless sentences is often omitted, giving a contextually understood third person interpretation, for example:

(15)   *Waa áqal.*   'It's a house.'
             Waa Xabáshi.   'He is an Ethiopian.'
             Waa Xabashí.   'They are Ethiopians.'
             Waa Júmce.   'It's Friday.'
             Waa rún.   'It's true.' (lit. 'It is truth')
             Waa nabád.   'It is peace.' (reply to a greeting)

There are no corresponding negative verbless sentences. Negative versions must include a verb, for example:

(16)   *Rún   má   ahá.*
        truth   not   is:NEG
        'It is not true.' (lit. 'It is not truth')
(17)   *Anígu báre   má   ihí.*
        I   teacher not   am:NEG
        'I am not a teacher.'

There are a number of important idiomatic constructions using verbless *waa* sentences. For example a verbless declarative without a subject and with a complement clause (8.4.1) on *in* 'that' is commonly used to expresses obligation:

(18)   *Waa   ínaad   tagtó.*
        DM   that-you go
        'You must go.'   (lit. 'It is that you go')

With demonstratives acting as the predicate, the following expressions occur:

(19) (a) *Wàa   kán.*
           DM   this:M
           'He/it (masc.) is here.', 'Here he/it (masc.) is.'
      (b) *Wàa   tán.*
           DM   this:F
           'She/it (fem.) is here.', 'Here she/it (fem.) is.'
      (c) *Wàa   kúwán.*
           DM   these
           'Here they are.'
      (d) *Wàa   i   kán.*
           DM   me   this
           'Here I am.'

Similarly *waa* and predicate nouns with possessive determiners, or the determiners alone, may mark possession:

(20) (a) *Waa   rúntáa.*
           DM   truth-your
           'You are right.' (lit. 'It is your truth')
      (b) *Waa   rúntiis.*
           DM   truth-his
           'He is right.' (lit. 'It is his truth')
      (c) *Waa   kàyga/tàyda.*
           'It's mine (M/F).'
      (d) *Waa   kúwàyga.*
           'They're mine.'

There are also very common related idiomatic expressions, for example *Waa hagàag* 'All right, okay' (lit. 'It is straightness').

7.4   Declaratives with focus

Somali, along with other Cushitic languages like Rendille (Oomen 1978), Boni (Sasse 1981) and Dirayta (Hayward & Saeed 1984), employs special words to identify elements in focus.[73] We discuss the pragmatic functions of focus in chapter 9 where we review its relationship to discourse functions like introducing new information, contrastive emphasis, etc. We introduced the three focus words *bàa*, *ayàa* and *wáxa(a)* in 4.7 and outlined coalescence rules undergone by *bàa* in 3.3.2. In this section we examine the syntactic and morphological features of these focus constructions in declaratives, dealing first with *bàa* and *ayàa*, and then with *wáxa(a)* constructions. One major difference

between these is constituent order: elements focused by *bàa* and *ayàa* must occur before the verbal group while *wáxa(a)* constructions are postposing: the focused elements must occur after the verb, for example (where the focused element is shown in bold in both Somali and the English gloss):[74]

(21)　(a)　***Soomaaliyád***　*bàad*　　*tahay.*
　　　　Somali:FEM　　　*bàa*+you　are
　　　　'You are **a Somali woman.**'
　　(b)　*Wáxaad*　　*tahay*　***Soomaaliyád.***
　　　　wáxa+you　are　　Somali:FEM
　　　　'What you are is **a Somali woman**.'

7.4.1 Focus with *bàa* and *ayàa*

As described in 4.7, the words *bàa* and *ayàa* seem to be optional equivalents, with the choice possibly conditioned by the greater involvement of *bàa* in phonological rules of coalescence (3.3.2). We treat them as equivalent here. We can outline the major syntactic features of *bàa*/*ayàa* constructions before examining some of the morphological idiosyncrasies associated with them. *Bàa* and *ayàa* constructions can be viewed as left dislocation focus structures[75] and they may be used to focus any of the following constituents.

7.4.1.1 NPs *acting as arguments of the verb*

These may be subject as in (24) below, or object as in (23), and may be indefinite (24) or definite (23). The NP may consist of a simple noun, or name (22), a noun with suffixed determiner (23), or a noun modified by an adjective (24), or by a restrictive (25) or appositive relative clause (26):

(22)　***Cigaal***　*baa*　*geeloodii*　*la*　　*raaciyey.*
　　　Igal　　FOC　camels-their　one　accompany:CAUS:PAST
　　　'**Igal** was sent along with their camels.'　　[S5 14.1]
(23)　*Sayidkuna*　***gabaygan***　*ayuu*　*ka*　*tiriyey*　*dilkaas*　*Koofil*
　　　Sayid-the-and　poem-this　FOC+he　about　composed　killing-that　Corfield
　　　'And the Sayid composed **this poem** about that killing of Corfield.' [S1 27.12]
(24)　*Ka*　*dib*　***abaar***　***xun***　*bàa*　　*dhacday*
　　　more　after　drought　bad　FOC　　happened
　　　'Later **a severe drought** occurred.'　　[SHS 79.3]
(25)　*Markaana*　　***dad***　***badan***　***oo***　***reer***　***Masar ah***　*ayaa*　*qaatay*
　　　time-that-and　people　many　and　nation　Egypt　be　FOC　　took
　　　diinta　　　*islaamka.*
　　　religion-the　Islam-the

'And then **many Egyptians** converted to Islam.' (lit. 'And then **many people who were of the nation of Egypt** took the religion of Islam') [T4 22.16-17]

(26)   *Libaax   oo   ah   boqorkii   habardugaag* *baa beri bukooday   oo*
          lion   and is   king-the   beasts   FOC day   fell:ill   and

          awoodi   kari   waayey   in   uu   ugaarsado
          have:capacity   be:able   failed   that he   hunt

          'One day **Lion, who is the king of beasts**, fell ill and lost the ability to hunt.'
          [SHS 83.13-14]

7.4.1.2   *NPs acting as predicate nominals*

(27)   *Ninka   reerka   lihi*   **deeqsi**   *buu   ahaa*
          man-the family-the had   generous:person   FOC+he   was

          'The man who had the family was **a generous person**'   [SHS 56.8]

7.4.1.3   *NPs acting as adverbials*

(28)   *Arooska   weyn*   **had**   **dhow**   *bàan   dhigaynaa*
          wedding-the   big   time   near   FOC+we   arrange:PROG

          'We will be arranging the big wedding **soon**'   [SHN 72.12]

7.4.1.4   *The intensifier* aad   *(4.10.2)*

(29)   *Soomaalidu   guurka*   **aad**   *bay*   *u   tixgelin   jirtey*
          Somalis-the   marriage-the much   FOC+they   ADP   value   used

          'The Somalis used to value marriage **very highly**' [S2 1.12]

7.4.1.5   *Adverbial clauses (8.3)*

(30)   **Inta**   **qofku**   **nool yahay** *ayuu   awood   leeyahay oo*
          amount-the   person-the   alive is   FOC+he   potential has   and

          wax   qabsan   karaa
          thing   take:MIDDLE:INF can   [BCH 25.29-30]

          '**As long as a person is alive**, he has potential and can achieve something'

7.4.1.6   *In complement clauses (8.4.1)*

(31)   *Berri*   **inuu**   **soo**   **fufi**   **doono**   *baan   ku   fekeraa*
          tomorrow   that+it   VEN   sprout   will   FOC+I   on   think

          'I reflect **that it will sprout back again** tomorrow.'   [FS 41.8]

7.4.1.7   *Complete sentences as reported speech*

(32)   **Gacaliso   maxaad   iga   doonaysaa?**   -   *buu   weydiiyey*
          dear   what+you   me+from   want:PRES.PROG   FOC+he   asked

          yaxaaskii
          crocodile-the   [SHS 79. 24-5]

          ' "**Dear, what do you want from me?**" - the crocodile asked him.'

These focus words may not focus the verb or any element of the verbal group.[76]

The following are the major syntactic features of *bàa/ayàa* constructions:

(33)   (a)   The declarative sentence type marker *waa* does not occur with *bàa* or
               ayàa.
        (b)   Only one element may be focused by *bàa* or *ayàa* in a sentence.[77]
        (c)   The focused element must be a constituent in a main (root) clause: focus
               does not occur inside subordinate clauses.
        (d)   The focused element must be a maximal projection[78]: it is not, for example
               possible to focus part of a noun phrase, only the whole NP.
        (e)   In declaratives *bàa* and *ayàa* must occur with a verbal group. An element
               focused by *bàa/ayàa* cannot form an independent utterance.

In support of (33e) see B's possible answers to A's question below:

(34)   A: *Yàa   yimí?*   Who came?
                    who   came
            B: (a)   *Cáli   bàa   yimí.*   'Ali came'
                    (b)   * *Cáli   bàa*   'Ali'
                    (c)   *Cáli*   'Ali'

Example (34b) is not a possible reply.
    In addition to these features, *bàa/ayàa* constructions display a range of
morpho-syntactic behaviours which seem unusual when compared to other
sentences. These are particularly evident when the focused element is a subject.
We have mentioned some of these features before but we can summarise them
below:

(35)   (a)   A focused subject is not subject marked but occurs in the absolutive case.
        (b)   A verb agreeing with a focused subject shows a much reduced set of
               agreement markers (Andrzejewski's (1968) 'restrictive paradigm') and a
               distinct accentual pattern, AP1 (4.3.4.13).
        (c)   A focused subject may not be doubled by a clitic pronoun in the verbal
               group (6.1.2).
        (d)   Negative *bàa/ayàa* sentences employ the constituent negative word *áan*
               'not' rather than the main clause *má* 'not' (see below).

As has often been noted, this range of behaviours, while unusual for main
clauses, is exactly paralleled by relative clauses where the head nominal is also
subject of the clause, as we shall see in chapter 8. Since it is in just these two
contexts, relative clauses and focus, that share such effects, it seems plausible to
view focus structures in Somali as the result of a historical process involving the

grammaticalization of a relative clause-type construction as a focus strategy. Several analyses have been proposed to characterize the synchronic effects of this process but the issue is still controversial.[79]

Focus is less common in negative declaratives; when it occurs the negative word is *áan* 'not'. This negative word cliticises onto *bàa/ayàa*, for example:

(36)   ***Fáarax***   *báan*   *tegín.*
     Farah   FOC+not   went:NEG
     '**Farah** didn't go.'

The main use for this sentence type is for negative contrast: typically a presupposition of (36) above would be that someone went and the speaker uses this sentence to assert that it wasn't Farah. As discussed in Saeed (1993a: 234-5), the accentual pattern on *báan* above is unexpected: the usual rules of tonal sandhi would predict *bàan*. This may be to distinguish the combination of *bàa* + *áan* 'not' from *bàa* + *aan* 'I', thus distinguishing otherwise ambiguous sentences like:

(37)   (a)   ***Fáarax***   *báan*   *qabán.*
            Farah   FOC+not   catch:NEG
            '**Farah** didn't catch him'
     (b)   ***Fáarax***   *bàan*   *qabán.*
            Farah   FOC+I   catch:INF[80]
            'I will catch **Farah**.'

This unusual pattern only occurs when the focused element is subject of the negative and thus clitic subject pronouns do not occur. When these pronouns do occur they coalesce with the cluster of focus word and negative. The initial vowel of the pronoun is shortened if long and the expected accentual pattern occurs, as in (38) with an example in (39):

(38)   *bàa*   +   *áan* 'not'   +   *aan*   'I'   →   *bàanan*
     bàa   +   *áan* 'not'   +   *aad*   'you'   →   *bàanad*
     bàa   +   *áan* 'not'   +   *uu*   'he'   →   *bàanu*
     bàa   +   *áan* 'not'   +   *ay*   'she/they' →   *bàanay*
         etc.
(39)   *Saddexdiinna*   **mid**   *baanan*   *dhalin.*
     three-your   one   FOC+NEG+I   bear:CAUS:PAST:NEG
     '(Of) the three of you, **one** I didn't father.'   [SHS 50.37]

7.4.2 Focus with wáxa(a)

As mentioned earlier the word *wáxa* (*wáxaa*) forms a postposing focus construction. We compare the discourse functions of *wáxa(a)* and *bàa/ayàa* in chapter 9. The typical use of a *wáxa(a)* construction is to place into focus a long noun phrase or a clause occurring after the verbal group. Thus *wáxa(a)*, unlike *bàa/ayàa*, is not used to focus the intensifier *aad* (cf. (29) above) and indeed does not generally focus shorter, non-clausal adverbials. As with *bàa/ayàa*, *wáxa(a)* may not focus the verb, any element of the verbal group, or functional categories. The focused element may be a subject (40 below), an object (41), a predicate nominal (42) or an adverbial (43). It may be a noun phrase (40), an *in* 'that' clause (42), or a direct quotation (45). *Wáxa(a)* constructions are difficult to translate directly into English: in the following examples we sometimes use English *it*-clefts, sometimes *wh*-pseudo clefts and at other times existential *there*-constructions. The resulting structural parallels are often not very felicitous, especially in the case of adverbial clauses, like (41 below). In the following examples the focused elements are shown in bold:

(40)   *Waxaa keena **uun   aad   u   yar   oo   la   yiraahdo 'Fiiruus'***
        wáxaa brings creature INTENS ADP small and one calls   'virus'
        'What causes it is **a very small organism called a virus**' [BCH 10. 34-5]

(41)   *Waxa   la   wada   ogsoonyahay **nolosheennu in   ay   ku***
        wáxa one together aware-is   life-our   that it   to
        xidhantahay   roobka
        bound-is   rain-the
        'It is well known to all **that our life is bound to the rain**' [S2 19.29-30]

(42)   *Saahid Qamaan wuxuu   ahaa **nin gabayaa ah   oo   aad   ugu***
        Sahid Qaman *wáxa*+he was man poet   be   and INTENS ADP+ADP
        xeel   dheeraa   gabayada xikmadda, waanada iyo   duurxulka
        cleverness deep-was poems-the wisdom,   counsel   and oblique:language
        '**Sahid Qaman was a poet who was very ingenious at the poetry of wisdom, counsel and of veiled language**.'   [S1 12.3-4]

(43)   *Waxaa   la   doortay **goortii   Cismaan la   dilay***
        wáxaa one   chose   time-the   Osman   one   killed
        'He was chosen **when Osman was killed**' (lit. 'What one chose him *the time when one killed Osman*')   [T4 27.5-6]

(44)   *Kii   labaadna   wuxuu   yiri: "Aniguna   waxaan rabi lahaa*
        the:one second-and *wáxa*+he said: "I-and   *wáxa*+I like   would
        uubata   badan   oo   arigaada   marisa"
        wolves   many   and   goats-your   wipe:out
        'And the second one, what he said was: "And what I would like is many wolves to wipe out your goats." '   [FS 51.6-7]

Wáxa(a) constructions share the following features with *bàa*/*ayàa* constructions:

(45)   (a)   The declarative sentence type marker *waa* does not occur with *wáxa(a)*. Only one element may be focused by *wáxa(a)* in each sentence.[81]
      (b)   The focused element must be a constituent in a main (root) clause: focus does not occur inside subordinate clauses.
      (c)   The focused element must be a maximal projection: it is not, for example possible to focus part of a noun phrase, only the whole NP.
      (d)   *Wáxa(a)* must occur with a verbal group, except for the *wáxa wéeyé* construction described in 7.3 earlier.

The statement in (45a) needs two qualifications. Firstly we must distinguish between normal relative clauses on the noun *wáx* 'thing' and *wáxa(a)* focus constructions. The former have the normal distribution of relative clauses including occurring in verbless *waa* sentences (7.3), for example:

(46)   *Wúxuu*   *raadínayaa*   *waa*   *kìishkán.*
       thing+the+he   seeking   DM   bag-this
       'The thing he's looking for is this bag.'

We discuss this relationship below. Secondly, as described in 7.3, the focus word *wáxa(a)* may occur with the form *wéeye* to focus the second element in a verbless sentence:

(47)   *Faarax*   *wáxa*   *wéeyé*   **nín**   **géesi**   **áh**.
       Farah   *wáxa*   DM   man   brave:man   be
       'Farah (what he is) is **a brave man**.'

    Wáxa(a) constructions also show the same morphosyntactic 'irregularities' as *bàa*/*ayàa* constructions, that is to say they display features of relative clause grammar: a *wáxa(a)*-focused subject does not display subject case; an agreeing verb shows the reduced agreement pattern and AP1 tone pattern; and a subject may not be doubled by clitics in the verbal group. Negation too employs the constituent form *áan* 'not' rather than main sentence *má* 'not', for example:

(48)   *Wáxaanu*   *arkín*   **gabádhíi**   **huruddáy.**
       wáxa+NEG+he   saw:PAST:NEG   girl-the   sleep:PAST:SUBORD
       'It was **the girl who was asleep** he didn't see.', 'What he didn't see was **the girl who was asleep**.'

It seems clear from these characteristics that *wáxa(a)* constructions are a form
of cleft. Once again though there are different grammatical analyses. In Saeed
(1984) I suggested that *wáxa(a)* clefts could be derived, in what is probably best
seen as a historical process,[82] from verbless *waa* sentences containing a relative
clause on the word *wáx* 'thing' by omission of *waa* and a process of semantic
loss or 'bleaching' which would weaken the word *wáx-a* to a general expletive,
meaning 'what, who, when'. This approach can be represented schematically in
(49) and with the examples in (50), both of which are grammatical:

(49) [[REL.CLAUSE *wax-a…*] *waa* [PREDICATE …]] ⇒ [[CLEFT CLAUSE *wáxa(a)…*] [FOCUS…]]

(50) (a)   *Wáxaan   dóonayaa   waa   shàah.*
            thing-the+I   want:PROG   DM   tea
            'The thing that I want is tea.'
      (b)   *Wáxaan   dóonayaa   **shàah**.*
            thing-the+I   want:PROG   tea
            'What I want is **tea**'

The differences between a verbless sentences like (50a) and a cleft like (50b)
are, firstly, that the semantic restriction in (50a) that the complement be
inanimate (to match *wáx* 'thing') is relaxed in the cleft, where the focus element
can also be animate or human. Secondly, the NP *wáx-a* in (50a) is reanalysed as
a unit *wáxa* in (50b) and its head *wáx* no longer participates in nominal
morphosyntax, for example does not occur in the plural, or with other
determiners. Thirdly, though the original lexical item *wáx* is masculine, the
expletive does not control concord, so that the verb in the following shows
feminine concord with the following nominal:

(51) *Wáxa   timí   **gabádh   dhèer   oo   qurúx   badán**.*
      wáxa   came:F   girl   tall   and   beauty   much
      'There came **a tall, very beautiful girl**.'

Each of these differences seems consistent with the grammaticalization of
wáxa(a) from a lexical noun phrase to an expletive focus element.[83]

7.5   Interrogatives

There are three major types of interrogatives: polar questions, information
questions, and tag questions. We describe each in turn.

7.5.1 Polar Questions

Polar questions seek for confirmation or denial of a proposition and are often called yes/no questions since they set up an expectation for yes or no in the reply. This sentence type is marked in Somali by the question word *ma* which we will term an interrogative sentence type marker (QM). In polar questions without focus *ma* simply replaces the declarative sentence type marker *waa* (DM) and occurs in the same position, before the verbal group. The polar questions are not further distinguished from corresponding declaratives by word order. The clitic subject pronouns are more commonly omitted in polar questions than in declaratives. We can compare the following pairs of declarative and interrogative sentences:

(52) (a) *Aabbáháa   waan   lá   hadlay.*
           father-your   DM+I   with   spoke:3MS/1S
           'I spoke to your father.'
      (b) *Aabbáháa   ma   lá   hadlay?*
           father-your   DM+I   with   spoke:3MS/1S
           'Did he/I speak to your father?'

(53) (a) *Gabádha   kulá   shaqeeysá   wuu   yaqaan.*
           girl-the   you+with   works:3FS   DM+he   knows:3MS
           'He knows the girl who works with you.'
      (b) *Gabádha   kulá   shaqeeysá   ma   yaqaan?*
           girl-the   you+with   works:3FS   DM+he   knows:3MS
           'Does he know the girl who works with you?'

The NPs in the questions in (52b & 53b) are free to occur after the verbal group, just as in the declaratives (52a & 53a).

Clitic subject pronouns may occur in these questions; and they coalesce with the question word in two alternate ways. The first is as follows: *ma + aan* 'I' → *maan*, *ma + aad* 'you (SG)' → *maad*, *ma + uu* 'he' → *muu*, *ma + ay* 'she' → *may*, *ma + aannu* 'we (EXC)' → *maannu*, *ma + aynu* 'we (INC)' → *meynu*, *ma + aydin* 'you (PL)' → *meydin*, *ma + ay* 'they' → *may*. For example:

(54) *Muu   kúu   dhiibay?*
      QM+he   you+to   hand:PAST
      'Did he hand it to you?'

The second type of coalescence rule is as follows: *ma + aan* 'I' → *miyàan*, *ma + aad* 'you (SG)' → *miyàad*, *ma + uu* 'he' →*miyùu*, *ma + ay* 'she'

→ *miyày, ma + aannu* 'we (EXC)' → *miyàannu, ma + aynu* 'we (INC)' → *miyàynu, ma + aydin* 'you (PL)' → *miyàydin, ma + ay* 'they' →*miyày*. For example:

(55)   *Miyùu   kúu   dhiibay?*
       QM+he   you+to   hand:PAST
       'Did he hand it to you?'

Polar questions may contain the focus words *bàa, ayàa*, and *wáxa(a)* though their use is more restricted than in declaratives.   Firstly, as we shall discuss in chapter 9, there is a pragmatic asymmetry between focus in declaratives and in polar questions. The former is used in a wider range of contexts, introducing new information, contrastive emphasis etc., while the latter has a very specific use: *bàa/ayàa* structures in polar questions act as a form of constituent questioning, for example:

(56)   *Ma   Cáli   bàa   tegáy?*
       QM   Ali   FOC   went
       'Did ALI go?' 'Was it Ali who went?'

In (56) there would typically be a presupposition that someone went and the questions asks whether it was Ali. As a consequence *bàa/ayàa* in polar questions focus fewer types of constituent, typically noun phrases.

The question word *ma* behaves differently with the three focus words: with *bàa*, *ma* occurs before the questioned constituent, as in (57b) below:

(57)   (a)   **Xerádii**   *bàynu*   *nimi.*
             camp-the   FOC+we:INCL   came:to
             'We (INC) have come to **the camp**.'
       (b)   *ma*   **Xerádii**   *bàynu*   *nimi?*
             QM   camp-the   FOC+we:INCL   came:to
             'Have we (INC) come to **the camp**?'

By contrast, *ma* coalesces with *ayàa* in its position following the focused constituent, as follows: *ma + ayàa → miyàa*.   If the focused element is not subject, clitic subject pronouns occur, and then the question word, the focus word and the pronoun coalesce as follows: *ma + ayàa + aan* 'I' → *miyàan; ma + ayàa + aad* 'you (SG)' → *miyàad; ma + ayàa + uu* 'he' → *miyùu, ma + ayàa + ay* 'she' → *miyày, ma + ayàa + aanu* 'we (EXC)' → *miyàanu, ma +*

ayàa + aynu 'we (INC)' → *miyàynu*, *ma + ayàa + aydin* 'you (PL)' → *miyàydin*, *ma + ayàa + ay* 'they' → *miyày*. See for example (58b):

(58) (a) **Xerádii**   *ayàynu*   *nimi.*
            camp-the   FOC+we:INCL   came:to
            'We (INC) have come to **the camp**.'
      (b) **Xerádii**   *miyàynu*   *nimi?*
            camp-the   FOC+ QM+we:INCL   came:to
            'Have we (INC) come to **the camp**?'

With *wáxa(a)*, the question word *ma* occurs immediately before the expletive *wáxa(a)*, for example:

(59) *Dhakhtarnimadayda   ma   waxaad   moodaysaa   **bilaash**?*
       doctorhood-my-the   QM   *wáxa*+you   think   worthless:thing
       'Do you take my medical status to be **a worthless thing**?' [SHN 100.2-3]

       Focus words also occur in verbless polar questions of the form *ma* NP *bàa* or NP *miyàa*, both meaning 'is it/are they NP?':

(60) (a) *Ma   nabád   bàa?*
            QM   peace   FOC
            'Is it peace?'   (a greeting)
      (b) *ma   adigii   baa   dawacoy?*
            QM   you   FOC   jackal:VOC
            'Is it you, Jackal?'   [SHS 83. 27]
      (c) *noloshaadu   ma   nolol baa?*
            life-your-the   QM   life   FOC
            'Your life, is it a life?'   [SHS 78. 37]
(61) (a) *qáali   miyàa?*
            expense   QM+FOC
            'Is it expensive?'
      (b) *yartii   miyaa?*
            young:one:F   QM+FOC
            'Is she the girl?'   [SHN 80.26]

       In polar negative questions without focus we find two distinct forms: in neither does the question word *ma* co-occur with the negative word *má* 'not'. These two morphemes are formally distinct accentually and syntactically. The negative word *má* always has High tone while the QM *ma* has the seesaw pattern described in 3.4.3: if the following word contains a High tone, QM *ma* is Low,

while if the following word is all Low, QM *ma* is high. Syntactically the two morphemes occur in different positions: QM *ma* occurs in the pre-VGP sentence type marker position while negative *má* occurs within the VGP as described in 6.1.6. Finally negative *má* always occurs with a negative verb. In addition to these differences, as we shall see, the two morphemes never occur in the same clause.

The first type of negative question has the QM *ma* and the negative word *áan* 'not'. These questions usually include clitic subject pronouns and the three elements coalesce as follows: *ma* + *áan* 'not' + *aan* 'I' → *miyàanán*; *ma* + *áan* 'not' + *aad* 'you (SG)* → *miyàanád*; *ma* + *áan* 'not' + *uu* 'he' →*miyàanú*, etc[84]. For example:

(62)   *Miyàanú   sóo   noqónayn?*
        QM+not+he   VEN   return:PROG:NEG
        'Isn't he coming back?', 'Won't he come back?'

The second type of negative question uses a distinct question word *sòw* (*shòw*) which is a marker of speaker uncertainty also found in potential sentences (though with a distinct verbal inflection). This type of sentence has the negative word *má* 'not'; for example:

(63)   *Sòw   Fáarax   má   ihíd?*
        QM   Farah   not   be:NEG
        'Are you not Farah?'

This question type is used when the speaker expects an affirmative reply and its use extends naturally to seeking confirmation and to rhetorical questions, for example:

(64)   *Sow   nagu   ma   kalsoonid?*
        QM   us-in   not   trust:NEG:2SG
        'Do you not trust in us? [SHN 82.10]
(65)   *jaas   Soomaali   ihi   inuu   jiro   sow   ma   ogid?*
        jazz   Somali   be   that-it   exists   QM   not   know:NEG:2SG
        'Don't you know that Somali jazz exists?'   [SHN 164. 14-15]

This form of question also provides the tag element in the tag questions described in 7.5.3.

In polar negative questions with focus the negative word *áan* 'not' is added to the focus word *bàa* or *ayàa*; the verb must then be in a negative form. Compare for example:

(66)   (a)   ***Cáli***   *miyàa*   *tegáy?*
                  Ali   QM+*ayàa*   went
                  'Did **Ali** go?'
         (b)   ***Cáli***   *miyáan*   *tegín?*
                  Ali   QM+*ayàa*+not   went:NEG
                  'Did **Ali** not go?'
(67)   (a)   *Ma*   ***Cáli***   *bàa*   *tegáy?*
                  QM   Ali   FOC   went
                  'Did **Ali** go?'
         (b)   *Ma*   ***Cáli***   *báan*   *tegín?*
                  QM   Ali   FOC+not   went:NEG
                  'Did **Ali** not go?'
(68)   ma   **dalka**anad   ogayn?
                  QM   country-the+*bàa*+not+you   know:NEG
                  'Do you not know **the country**?'   [SHN 144.4]

7.5.2 Information questions

In their basic use information questions enable speakers to elicit new
information about part of a proposition while presupposing the rest. Thus the
question (69) below assumes the addressee went somewhere and asks for the
information 'where?':

(69)   *Xaggée*   *bàad*   *tagtay?*
         place+which   FOC+you   went
         'Where did you go?'

Given this distinction between new and presupposed information, it is not
surprising that, except in the verbless questions discussed a little later, question
words always attract focus, usually with *bàa* as in (69) or *ayàa*. *Wáxa* focus is
not normally used with question words, revealing a preference for a left
dislocation focus structure with information questions.

    In their syntax Somali information questions pattern more like declarative
sentences with focus than polar questions: they do not include, for example, an
independent interrogative marker like the question word *ma*.[85] The only marker
of the question is the presence of an interrogative noun phrase. We can identify
three main types of these. The first are NPs containing the interrogative
determiner *kée*/*tée* (4.6.3), as in (69) above and in the following:

(70)   (a)   *Sidée   bàad   ú   tagtay?*
                way-which   FOC+you   in   went
                'How did you go?'

        (b)   *Naagtée   bàad   aragtay?*
                woman-which   FOC+you   saw
                'Which woman did you see?'

        (c)   *Intée   bàad   joogtay?*
                amount-which   FOC+you   stayed
                'How long did you stay?

It is possible in fast speech for the focus word (with any attached elements) to coalesce with this *kée/tée* suffix. So paralleling (69) above we can find:

(71)   *Xaggèed   tagtay?*
        place+which+FOC+you   went
        'Where did you go?'

Here only the accentual pattern reveals the presence of the focus word. It is also possible in colloquial speech for the focus word, along with any coalesced elements, to be dropped altogether:

(72)   *Xaggée   tagtay?*
        place+which   went
        'Where did you go?'

The interrogative determiner may stand alone as a question word:

(73)   *Kée/tée/kuwée   bàad   dóonaysaa?*
        which:M/which:F/which:PL   FOC+you   want:PRES PROG
        'Which one (masc.)/(fem.)/which ones do you want?'

The second type of question words are nouns with interrogative *ma* as a suffix (4.6.3), for example:

(74)   *góorma?*   'when?'   (cf. *góor* N F 'time')
        méelma?   'where?'   (cf. *méel* N F 'place')
This suffix is also used with independent pronouns:

(75)   *idínma?*   'which of you (PL)?'
        innáma?   'which of us (INC)?'
        iyáma?   'which of them?'

See also the related forms:

(76)   *kúma?*   'who (MASC)?'
        túma?   'who (FEM)?'

The third type of question words are those with no easily discernible internal structure, for example *máxay?* 'what?', which also occurs with the verbal adposition *ú* 'for' as *máxay...ú* 'what for?', 'why?'. Other examples are *yàa* 'who?', *ímmisa* 'how much?, how many', and *méeqa* 'how much?, how many?'.

In addition to these three main types, there are a number of questions words with unusual distributions. Used alone or with a single NP the forms *mée* (*méeyey*) M, *méeday* F and *méeye* (*máye*) PL form a question meaning 'where?' for example:

(77)   *Méeyey?*   'Where is he?'
        Méeday?   'Where is she?'
        Méeye?   'Where are they?'
        Áxmed mée?   'Where is Ahmed?'

Similarly, used alone as one-word questions *Wàayo?* means 'Why?' and *Iyàa?* means 'What (did you say)? Pardon?'.

Information questions are not commonly used in negative forms. When they occur they have the same forms as negative declaratives with focus (7.4.1). There is however a strong tendency to use positive forms, for example by subordinating the clause under a verb with an inherently negative meaning:

(78)   *Maxàad*   *ú*   *tégi*   *weydey?*
        what+FOC+you   for   go:INF   failed
        'Why didn't you go?' (lit. 'Why did you fail to go?')
(79)   *Haddaad*   *saas*   *doonaysid*   *maxaad*   *anna*
        time-the+you   way+that   want   what+FOC+you   me
        Xajka   *iigu*   *dhigi*   *weyday,*   *Guuleedow?*
        pilgrimage-the   me+from+for arrange:INF   failed   Guuleed:VOC
        'If that's what you want, why didn't you arrange for me (to go on) the pilgrimage, Guuleed?' (lit. 'Why did you fail to arrange...')
                                    [SHN 49.27-8]

As discussed in Saeed (1984) there is no evidence to support an analysis of information questions involving a movement of the questioned element to its usual position at the front of the sentence, following any topics. Such a rule has

of course been proposed for other languages, including English. We might note several features briefly. As discussed in chapter 8, there are no subordinate interrogative clauses or 'indirect questions' in Somali. Paralleling main clause information questions we find relative clauses on non-interrogative nominals, for example:

(80)   *Goortaan   dhashay   ma   kuu   sheegaa?*
        time-the+I   born   QM   you+to   tell
        'Shall I tell you <u>when I was born</u>?' (lit. 'Shall I tell you <u>the time that I was born</u>?')   [SHN 204.29-30]

Secondly, when the main question clause contains an *in* complement clause (8.4.1), any argument in the complement clause that is coreferential with the question word is marked by a clitic pronoun rather than a gap, for example:

(81)   *Ninkée   bàad   sheegtay   ínuu   ku   caayéy?*
        man-which   FOC+you   said   that-he   you   insulted
        'Which man did you say (that he) insulted you?'

Finally, the relationship between the question word and such a coreferential pronoun is not subject to any of the constraints usually identified with movement rules; for example the pronoun may be within a relative clause:

(82)   *Naagtée   bàad   akhriday   bùuggii   ay   qortáy?*
        woman-which   FOC+you   read   book-the   she   wrote
        'Which woman did you read the book that she wrote?'

        Verbless information questions occur and are formed in the same way as verbless declaratives described in 7.2. The predicate nominal is a question word, or a NP morphologically questioned either by the interrogative determiner *-kée/-tée* or by a suffixed question word *ma*, for example:

(83)   (a)   *Xisáabtu   waa   ímmisa?*
               bill-the   DM   how:much?
               'The bill is how much?', 'How much is the bill?'
        (b)   *Dukàankii   waa   xaggée?*
               shop-the   DM   place-which
                   'The shop is where?', 'Where is the shop?'
        (c)   *Máanta   waa   ayáanma?*
               today   DM   day-what
               'Today is what day?', 'What day is it today?'

Again a missing subject will be understood from context:

(84)   (a)   *Waa   sannadkée?*
               DM   year-which
               'What year is it?'
        (b)   *Waa   kúma?*
               DM   who:M
               'It is who (M)?', 'Who (M) is it?'

As can be seen, in these sentences the questioned nominal does not occur clause initially.

7.5.3 Tag questions

Tag questions are used to seek confirmation of a proposition the speaker believes to be true. They are formed by adding to declaratives the negative question *sòw má aha*? 'Is it not so?'. This clause may be added to any declarative, for example:

(85)   *Adàa   arkáy,   sòw   má   ahá?*
        you+FOC   saw   QM   not   be:NEG
        'YOU saw it, didn't you?'
(86)   *Waa   rún,   sòw   má   ahá?*
        DM   truth   QM   not   be:NEG
        'It's true, isn't it?'

7.6   Other sentence types

In this section we outline the untensed sentence types shown in (1) earlier: imperative, optative, and potential. As mentioned, these sentence types are characterised by occurring with verbs in a single inflectional form that does not admit distinctions of tense or aspect. They also do not permit focus.

7.6.1 Imperative

Imperative sentences are used to direct the addressee(s) towards what the speaker wishes them to do: to make suggestions and give commands. In Somali society it is not considered impolite to use imperatives, even in formal situations and thus they are very common. Less typical uses exemplified below include greetings (87c) and proverbs (88). Imperatives occur with a special verb form that distinguishes between singular and plural addressees (4.3.4-6). No clitic

subject pronouns may occur.[86]   In positive forms there is no marker (STM) and
thus sentences may consist simply of a verbal group, for example:

(87)   (a)   *Sóo   noqó*!
                VEN   return:IMP:SG
                'Come back! (SG)

        (b)   *Íi   saméeya*!
                me+for   do:IMP:PL
                'Do (PL) it for me!'

        (c)   *Iská   wárran*!
                self+about   tell:news:IMP:SG
                'Tell (SG) the news about yourself!' (a greeting)

(88)   *Ama   táag   lahàw   ama   tamár   lahàw*.
        or   strength have:IMP:SG   or   cunning   have:IMP:SG
        'Either have strength or cunning.'   (Proverb)

Negative imperatives have the STM *ha* which occurs with negative verb forms.
Ha occurs before the verbal group, for example:

(89)   (a)   *Ha   sóo   noqón*!
                STM   VEN   return:IMP:NEG:SG
                'Don't (SG) come back!'

        (b)   *Ha   íi   samaynína*!
                STM   me+for   do:IMP:NEG:PL
                'Don't (PL) do it for me!'

When NPs occur they most commonly precede the verbal group as below:

(90)   (a)   *Dhárka   sóo   iibsáda*!
                clothes-the   VEN   buy:IMP:PL
                'Buy (PL) the clothes for yourselves!'

        (b)   *Diin   raadi   oo   degdeg   beerkiisa   iigu*
                tortoise track:IMP:SG and   haste   liver-his   me+to+in
                keen
                bring:IMP:SG   [FS 17.9-10]
                'Find (SG) Tortoise and quickly bring (SG) me his liver!'

        (c)   *ka   fiirso   oo   sida   xaqa   na   u   sheeg*!
                about   think:IMP:SG and   way-the   truth-the   us   to   tell:IMP:SG
                'Think (SG) about it and tell (SG) it to us truthfully!' [SHN 60.35]

Nominals may follow the imperative verb group for reasons of length or conversational rhythm: see below, where the last of four imperative clauses has this order:

(91)   *Hadda   nin   isu   beddeloo   reerkaas   tolkaa   ah*
         time-the man self-to change:IMP:SG+and homestead-that kin-your be
         u tagoo   cunno iyo biyo   weydiiso;
         to go:IMP:SG+and food   and water ask:IMP:SG
         xusuusnowse   wacadkaagii
         remember:IMP:SG+but   promise-your
         'Now turn yourself into a man and go to that homestead who are your kin and ask for food and water; but remember your promise.'   [SHS 91.29-30]

7.6.2 Optative

Optative sentences are used to express wishes and non-second person suggestions. They may also be used to express blessings and curses. The verb has a special form (4.3.4-6) and third person forms have a marker (STM) *há*. Other forms occur with subject clitic pronouns, which may be shortened:

(92)   (a)   *A(a)n tágno!*   'Let us go!'
        (b)   *Há   tago.*   'Let him go', 'May he go!'

Negative optatives have the marker *yàan* and occur with a negative verb form. *Yàan* occurs before the verbal group and may coalesce with subject clitic pronouns, as below (where optional forms are shown in parentheses):

(93)   (a)   *Yàynu   tégin*   'Let's not go'
        (b)   *Yàanu (yùu) imán*   'Let him not come!', 'May he not come!'

As with imperatives, there is a strong tendency for nominals to precede the STM and verbal group, for example:

(94)   *Alla   khàyr   há   ku   siiyo!*
         God   good:fortune   STM   you give
         'May God grant you good fortune!'

However it is possible for nominals to follow the verb; they can also come between the optative STM or pronoun and the verbal group, for example:

(95)   *War   aynu   lo'da   qaybsanno*
         hey   we   cattle-the   divide:OPT
         'Hey, let's share out the cattle!'   [SHS 85.16]

7.6.3 Potential

Potential sentences express the possibility of a proposition and are used for the description of possible situations and to ask the listener to entertain hypothetical situations. They usually occur with the marker *shòw* (*sòw*), which expresses uncertainty and also occurs in negative questions (7.5.1) and question tags (7.5.3). These sentences have a special verb form, described in 4.3.4-6 and do not occur with subject clitic pronouns:

(96)   (a)   *Shòw   sóo   noqotee.*
               STM   VEN   return:POT
               'Perhaps she will come back.', 'Suppose she comes back?'
        (b)   *Shòw   kúu   keenee.*
               STM   you+for   bring
               'Perhaps he'll bring it for you.', 'Suppose he brings it for you?'

When noun phrases occur, they are usually placed between *shòw* and the verbal group, for example:

(97)   *Shòw   dayuurád   raacee.*
        STM   aeroplane   take:POT
        'Perhaps he'll take an aeroplane,' 'Suppose he takes an aeroplane?'

There are no negative forms of potential sentences.

Chapter 8
Subordinate Clauses

Subordinate clauses are distinguished from main clauses in a number of ways, most notably:

(1) (a)   Sentence type distinctions are not made.
     (b)   Focus does not occur.
     (c)   Verbs are morphologically distinguished from main clause verbs.

We can briefly discuss each of these characteristics. Firstly, the range of sentence types found in main clauses (interrogative, imperative, optative, etc.) does not occur in subordinate clauses. The latter have a single type that most closely resembles declarative main clauses, except that the declarative marker *waa* does not occur. Indeed sentence type markers (STMs) do not occur at all in subordinate clauses. Thus, as we saw in the last chapter, there are no subordinate interrogative clauses, or 'indirect questions', in Somali. The function of these is performed by relative clauses on non-interrogative nouns of time, place, manner, etc., for example:

(2) *Weydìi   hálkuu   tégayó!*
     ask:IMP   place-the+he   is:going
     'Ask him where he is going!' (lit. 'Ask him the place to which he is going')
(3) *Má   garáneyó   inta   nín   hálkáa   kú   dhimatáy.*
     not know:NEG   amount-the   man   place-that   in   died
     'I don't know how many men died there.'
     (lit. 'I don't know the amount of men that died in that place')

Secondly, as described in chapter 7, the focus markers *bàa*, *ayàa* and *wáxa(a)* do not occur within subordinate clauses, though the whole clause may

be focused. Hence, for example, no NP within a subordinate clause may be focused.

Thirdly, as we saw in chapter 4, subordinate clause verbs are distinguished from main clause equivalents by a combination of accentual pattern and inflection. This is further emphasised when subordinate clauses are case marked for their grammatical function within the larger sentence.[87] In addition to these main features other differences occur: for example the negative word *má* 'not' does not occur in subordinate clauses, always being replaced by *aan* 'not'. Also, as we shall see, clauses acting as complements of certain verbs, like *dòon* 'wish, want', have a restricted range of tense and aspectual differences.

One striking feature of the system of subordinate clauses is the important role of relative clauses. Indeed it is possible to see all subordination as relative clause formation, though in some cases the head elements have been delexicalized and are no longer full nouns. In most cases though their historical relationship to nominals is clear. Because of this we begin with relative clauses proper in 8.1 and 8.2 before moving on to other kinds of subordinate clauses.

8.1   Restrictive relative clauses

8.1.1 Basic features

Restrictive relative clauses are used to help identification of the head noun from a larger set.[88] In the terminology of Keenan (1985), Somali restrictive relative clauses are post-nominal external relatives with the relativised position in the clause marked by a gap. See for example sentence (4a) below, which contains the head nominal and modifying relative clause (4b):

(4)   (a)   *Nínka   taló   xún   taliyaa   dhaxálkèeda   ayùu   helaa.*
               man-the   advice   bad   advises   legacy-its   FOC+he   finds
               'The man who gives bad advice receives its legacy.' (Proverb)
       (b)   *nínka   taló   xún   taliyaa*
               man-the   advice   bad   advises
               'the man who gives bad advice'

In (4b) the head noun is *nín* 'man' which attracts the determiner *-ka* 'the' and occurs before the relative clause. There is no relative pronoun in the clause. The head nominal always precedes the clause whatever position in the clause is relativised.

The relative clause resembles a main clause in syntax except that the tendency for verb final order is much stronger. In addition certain elements such

as subject clitic pronouns, like *uu* 'he, it' in (5) below, and the negative word *aan* 'not' in (6) are attracted to the head nominal and thus move away from the verbal group, for example:

(5)   *meeshuu   awal   sadcaalka   ka   bilaabay*
        place+the+it first journey-the from   began
        'the place from which the journey first began'   [BCH 9.38]
(6)   *nímáan   waxbá   oqóon*
        man+not   anything   know:NEG
        'a man who doesn't know anything'

Where the head NP ends in a vowel these elements may coalesce onto the head in the same way as with the focus word *bàa*, described in 3.1.2, for example:

(7)   *Wáxaanád   ahayni   ku   má   qabtàan.*
        thing-the+not+you   are:NEG   you   not   catch:NEG
        'The thing you are not does not catch you,' 'False rumour cannot harm you.'
              (Proverb)   (where *wáxaanád ← *wáx+a+aan+aad*)

Any argument position may be relativised, as in the following examples (where the head and relative clause are underlined where necessary): subject in (8), direct object in (9), and oblique arguments in (10) and (11):

(8)   *Timír   aan   láf   lahàyn*   *horàa   lóo   waayay.*
        dates   not   stone having:NEG   still+FOC   one+in failed:to:find
        'Dates which don't have stones still haven't been found.' (Proverb)
(9)   *Nín   daad   qaaday   xumbó   cuskéy.*
        man   flood   took   foam   grasped
        'A man who is taken by a flood grasps at foam.'   (Proverb)
(10)   *náagtii   aad   lá   hadasháy*
        woman-the   you   with   spoke
        'the woman with whom you spoke'
(11)   *méeshii   ay   ú   socdèen*
        place+the   they   to   travelled
        'the place to which they travelled'

Multiple restrictive relative clauses are joined by the conjunctions *oo* 'and' if the head nominal is indefinite and *ee* 'and' if the head nominal is definite, for example:

(12) *niman uu   yaqaan oo   u   socda degmadii   adoogiis   joogey*
      men   he   knows   and to   travel settlement-the father-his stayed
      'men whom he knows and who travel to the settlement where his father
      stayed'   [SHS 46.7-8]
(13) *cudurráda   halísta   áh   ee   qofku   ú   dhintó*
      diseases-the danger-the be   and   people-the from die
      'the diseases which are dangerous and which people die from' [BCH 25.32]

    If the head nominal is complex, for example consisting of a genitive or possessive construction, or a numeral followed by a dependent noun (6.2.2), relative clauses must be introduced by a conjunction:[89]

(14) *boqorkii   Bortaqiiska   ee   magaciisa la   oran   jirey Maanwel*
      king-the   Portuguese-the and   name-his   one   call   used   Manuel
      'the king of the Portuguese whose name was Manuel' [T5 23.24-5]
(15) *lixda   tuduc   ee   gabayga   ugu   horreysa*
      six-the   verse   and   poem-the   SUPR   begin
      'the first six verses of the poem' (lit. 'the six of verse which first start the
      poem')   [S6 15.19-20]

8.1.2 Case marking

The head nominal and its dependent clause act as a constituent, a noun phrase. Consequently, given the phrase final nature of case marking, the head nominal never receives main clause case marking and always occurs in the default absolutive case. The relative clause verb, as the final element of the phrase, carries the formal case marking. For example in (16a) below the head nominal *nínkíi* 'the man' is object of the adposition *lá* 'with' and the accentual marking of absolutive case applies to the relative clause verb *raacèen*. In (16b) the head is subject of the main clause and consequently subject marking applies to the verb, lowering its accentual pattern to *raaceen*, as described in 4.3.4.13.

(16) (a) *Nínkíi   ay   raacèen   ma   lá   hadashay?*
           man-the   they   accompanied   QM   with talked:2SG
           'Did you talk to the man they travelled with?'
      (b) *Nínkíi   ay   raaceen   wàa   ayó?*
           man-the   they   accompanied   DM   who
           'The man they travelled with is who?'

In other verbal forms subject marking involves a vowel change in addition to the accentual change; compare the case marking on the verb *wád* 'drive' in (17a) and (17b) below:

(17) (a) *Baabùurkuu wadó má aragtay?* [NON-SUBJ]
    truck-the-he drives QM saw:2SG
    'Have you seen the truck that he drives?'

   (b) *Baabùurkuu wadaa waa kán.* [SUBJ]
    truck-the-he drives DM this
    'The truck that he drives is this one.'

The effects on verbal morphology of this case marking are described in 4.3.4.13.

8.1.3 Subject-verb agreement

As described in 4.3, verbs normally show agreement with their subjects. However since the relativised position in a relative clause is marked by a gap, the relative verb may have this gap as its subject. When this happens a special, reduced form of agreement is marked on the relative clause verb. Compare for example the relative clause in (18a) below, where this happens, with that in (18b), where it does not (note that all of the nominals are masculine plural):

(18) (a) *nimánka buugágga keená*
    men-the books-the bring
    'the men who bring the books'

   (b) *buugágga nimanku keenàan*
    books-the men-the bring
    'the books which the men bring'

In (18a) above this reduced agreement, part of what Andrzejewski (1968, 1969, 1975a) calls the *convergent* paradigm, occurs because the subject of the relative clause verb *keená* 'bring' is a gap, coreferential with the head *nimánka* 'the men'. In (18b) the relative clause verb *keenàan* has an overt subject *nimanku* 'the men' and thus full agreement, part of what Andrzejewski (*op.cit.*) calls the *divergent* paradigm, occurs.

   This inflectional distinction survives under the main clause case marking described in 8.1.2. The subject-marked version of (18a) occurs in (19) below; and the subject-marked version of (18b) is in (20):

(19) *Nimánka buugágga keenaa waa kúwán.*
   men-the books-the bring DM these
   'The men who bring the books are these.'
   (i.e. 'These are the men who bring the books')

(20) *Buugágga nimanku keenaan waa kúwán.*
   books-the men-the bring DM these
   'The books which the men bring are these.'

(i.e. 'These are the books which the men bring')

The morphological details of reduced agreement forms are given in 4.3.4.13.[90]

8.1.4 Attributive adjectives

As discussed in 4.4 and 6.1.4, attributive adjectives seem closely related to restrictive relative clause constructions. They perform the same function semantically and significantly, it is not possible to form relative clauses with a copula verb and adjective; compare for example:

(21)   *nin*   *géesi*   *áh*
        man   brave:man   is
        'a man who is a brave man', 'a brave man'
(22)   (a)   *nín*   *dhèer*
               man   tall
        (b)   **nín   dhèer   áh*
               man   tall   is
               'a tall man', 'a man who is tall'

As (22b) shows, it is not possible to form a relative clause where the copula *yahay* 'be' occurs with an adjectival complement like *dhèer* ADJ 'tall', as is possible with a nominal like *géesi* NM 'brave man, hero'. Multiple attributive adjectives are joined by *oo* and *ee* in the same pattern as restrictive relative clauses:[91]

(23)   (a)   *hàad   wèyn   oo   madów*
               bird   big   and   black
               'a big black bird'
        (b)   *hàadkíi   wèyn   ee   madów*
               bird-the   big   and   black
               'the big black bird'

It is very common for nominal modification to involve restrictive relative clauses. One consequence is that combinations of attributive adjectives and relative clauses often occur, joined in the same way:

(24)   *shaati   cad   oo   gacma   gaab   ah*
        shirt   white   and   arms   shortness   be
        'a short-sleeved white shirt'   [MF 1.14]

(25)   *geeri   badan   oo   aan   loo   baahnayn*   [TD 168.21]
       deaths   many   and   not   one+for   need:NEG
       'many unnecessary deaths' (lit. 'many deaths for which one had no need')

(26)   *kabbadii   yarayd   ee   caanaha ahayd   ee   hashu   dhiiqday*
       mouthful-the small   and milk-the was   and   camel-the gave
       'the small mouthful of milk that the camel gave' (lit. 'the small mouthful which
       was milk and which the camel gave')   [SHS 71. 1]

A further consequence is that it is common for relative clauses to be stacked
inside each other, for example:

(27)   *silsilad   dheer   oo   xirmo   furayaal   ah   ka   lulata*
       chain   long   and   bunch   keys   be   from   hang
       'a long chain from which hangs a bunch of keys'   [MF 82.11-12]

Here the relative clause *xirmo furayaal ah* 'a bunch which is keys' is inside the
relative clause *xirmo furayaal ah ka lulata* 'from which hangs a bunch of keys',
which is the second dependent element on the head nominal *silsilad* 'a chain',
the first dependent element being the adjective *dhèer* 'long'.

8.2   Appositive relative clauses

Appositive, or non-restrictive, relative clauses are used to add new information
about the head nominal. They only occur on definite heads, typically names and
nouns with suffixed determiners. Since restrictive relatives do not occur on
names, this means that the two types of relative normally only co-occur on
nouns with determiners. Appositives are very similar in form to restrictives
except that the conjunction *oo* 'and' links the head NP and an appositive relative
clause, for example (where again the head and relative clause are underlined):

(28)   *Gabaygan   waxaa   curiyey   Sheikh Ismaaciil   Axmed   Barre oo*
       poem-this   FOC   composed   Sheikh Ismail   Ahmed   Barre and
       Daraawiishtii   Sayid   Maxamed   ka   mid   ah
       Dervishes-the   Sayid   Mohamed   from   one   be
       'This poem was composed by Sheikh Ismail Ahmed Barre, who was one of
       Sayid Mohamed's Dervishes.'   [S2 26.3-4]

Since with a single restrictive relative the clause attaches directly to the head,
this *oo* is a formal marker of appositives. When multiple relative clauses occur
this formal distinction is maintained since in the only context both can occur, on

nouns with determiners, restrictives are joined by *ee* 'and' (8.1) while appositives continue to employ *oo* 'and', for example:

(29)   *aabbáháa   oo   dá'   áh   oo   áan   welí   fásax   qaadán*
        father-your   and   age   is   and   not   ever   holiday   taken:NEG
        'your father, who is aged and who has never taken a holiday'

Appositive relative clauses have an important sub-type in which the appositive is attached to, and often coalesces with, an independent pronoun, for example:

(30)   *iságoo   hurdá*
        he+and   sleeps
        'he, who is/was asleep', 'he, being asleep'
(31)   *iyádoo   shaqéynaysá*
        she+and   is:working
        'she, who is/was working', 'she, working'

These clauses are untensed. Only one form is possible: the present progressive (4.3.4.7) for dynamic verbs (as in (31) above) and the present general (4.3.4.6) for stative verbs (as in (30)). There is however no present tense meaning associated with these forms, which we reflect by using English participle glosses above.

These untensed appositives function as absolute adverbial clauses, in the terminology of Thompson and Longacre (1985); that is they can be used to imply a range of adverbial relations between a backgrounded predication and the main predication. These include temporal relations such as 'during' or 'while' as in (32) below; a circumstantial or contributory factor as in (33); or simultaneity as in (34):

(32)   *Habeen   baa   isagoo lo'dii   soo xareeyay, nin   socota   ahi*
        evening   FOC   he+and cattle-the VEN penning   man   traveller be
        u   yimid.
        to   came   [FS 21.2]
        'One evening, while he was penning up the cattle, a traveller came to him.'
(33)   *iyagoon   cidi   u   doodin   baana   goobtii lagu   laayay*
        they+and+not person for argue:NEG   FOC+and site-the one+on killed
        'And nobody arguing for them, they were killed on the spot' [FS 23.16-17]
        (lit. 'And they and no-one arguing for them, one killed them on the site')

(34) *Markaas   buu   ninkii   guurdoonka ahaa doortay gabadhii*
      time-that FOC+he man-the   suitor-the was   chose   girl-the
      waxgaradka   ahayd, isaga oo   aan   eegayn
      intelligent:person-the was   he   and not   seeing:NEG
      quruxla-aanteeda,   baa la   yiri.
      lack:of:beauty-her   FOC one said
      'Then the suitor chose the intelligent girl, <u>not noticing her lack of beauty</u>, so it
      is told.'   [SHS 58.2]

In the examples so far the independent pronoun is coreferential with a
participant in the sentence. There is also an impersonal use of this construction,
using the feminine singular pronoun *iyáda* 'she, it (F)', for example:

(35) *Iyadoo berigii hore   uusan   jirin   dharku,   Soomalidu*
      it+and   time-the former   it+not be:NEG   cloth-the Somalis-the
      waxay   xidhan   jireen   hargaha   adhigooda.
      waxa+they   wear:INF used   skin-the   sheep:and:goats-their
      '<u>There being no cloth in early times</u>, the Somalis used to wear the skins of their
      sheep and goats.' [T4 64. 5-7]

Here the head of the clause is the feminine pronoun *iyáda* (translated above as
'there') even though the noun phrases in the clause, the adverbial *bérigii* 'the
time' and the subject *dhárku* 'the cloth' are both masculine. In this example the
appositive carries an implication of reason or cause.

8.3   Adverbial clauses

Somali employs a range of subordinate clauses to modify sentences adverbially.
In nearly all cases these adverbial clauses are relative clauses on noun phrases of
time, location, manner, etc. In this section we exemplify some major types,
dividing them into semantic categories.

8.3.1 Time

A number of temporal nouns are used to head relative clauses acting as sentence
adverbials. To convey the meaning 'when' nouns such as *már* N M 'time,
occasion', *kól* N M 'time, period', and *wàa* N M 'period, era' are used, with
suffixed definite articles *-ka/-ta*, *-kii/-tii* (4.6.1), where appropriate. Some
typical examples are given below:

(36) *Márkíi   uu   qólkíi   ká   baxáy   wáxaan   kú   idhi*
      time-the   he   room-the   from   went   wáxaa+I   to   said
      "Nabád   Gélyo".
      peace   enter:CAUS:OPT
      'When he left the room, I said "Goodbye" to him.'

(37) *Noloshii   miyigu   markay   ku adkaatay,   ayuu   magaalo   u*
      life-the   country-the time+the+it for grew:hard FOC+he   town   to
      soo   qaxay,   isagoo   18   jir   ah.
      VEN   moved   he+and   18   year   be
      'Rural life, when it got too hard for him, he moved to town, aged
      eighteen.'   [MF 67.1-2]

(38) *Kolkaan   sidaa   u   qaado   dhallinyaradoo   dhammi way*
      time+the+I   way-that in   lift   youth-the+and   all   DM+they
      dhacdaa.
      fall
      'When I lift them like that all the young men faint.' [SHN 204. 2-3]

(39) *waagii   ay   inanta   ahayd   Dhegdheer   way   qurux   badnaan*
      time-the   she   girl-the   was   Dhegdheer   DM+she   beauty   much+be
      jirtey
      used'   [SHS 12.15]
      'when she was a girl, Dhegdheer [Long-Ear] was very beautiful.'

A number of other temporal clauses occur. Clauses headed by the complex nominal *islá márka/islá márkii* (where *is* is the reciprocal/reflexive pronoun (4.2.4) and *lá* is the adposition 'with' (4.5)) convey the meaning 'as soon as', for example:[92]

(40) *Islá   márkíi   uu   tegáy   sháqàan   bilaabay*
      RE+with   time-the   he   went   work+I+FOC   began
      'As soon as he left, I began working.'

Relative clauses headed by *inta/intii* (cf. *ín* N F 'amount, extent, part') communicate the meaning 'while' or 'until', for example:

(41) *intuu   raggii   la   haasaabayay   baana   gabbalkii   ku*
      part-the+he   man-the   with   chatted:PROG   FOC+and   daylight-the on
      dhacay
      fell
      'And while he was chatting with the man, night fell.'   [FS 45.2-3]

When negative, the meaning of *inta/intii* clauses is 'before', for example:

(42) *Nebigu* <u>*intuunan*</u> <u>*dhiman,*</u> *ma macacaabin ninka*
prophet-the   part-the+he+not   died:NEG   not   name:NEG   man-the
dawladda   Islaamka   madax   ka   noqon   lahaa
government-the   Islam-the   head   at   become   would
'The Prophet, <u>before he died</u>, he did not name the man who would become head
of the leadership of Islam' (lit. 'while he had not died') [T4 19.9-10]

The phrase *tán iyo*, literally '(between) this and' is used to mean both 'since'
and 'until' with both simple nominals and relative clauses:

(43) *tán iyo góortáas*
this and   time-that
'since then, until then'

(44) *Wùu bukay* <u>*tán iyo íntuu*</u> *hálkán yimíd.*
DM+he   was:sick this   and   time-the+he   place-this   came
'He has been sick <u>since he came here</u>.'

The preposition *iláa* 'to, as far as, until', which as described in 4.5 is a
loan from Arabic and the only true preposition in Somali, can be used with
simple nominals to mean 'until':

(45) *iláa berríto*
'until   tomorrow'

Interestingly this word, though not a noun, may introduce a clause, with the
same meaning:

(46) *Wax bá ha qaníinin* <u>*iláa aad hubsató ínaad wada*</u>
thing   any   NIM   bite:NEG   until   you   ascertain   that+you   completely
<u>*líqaysó.*</u>
swallow
'Don't bite anything <u>until you are sure you can swallow it all</u>.'   (Proverb)

Iláa can also mean 'since' or 'until' when linked by *iyo* 'and' with relative
clauses on *ínta* 'the amount, extent':

(47) *Wùu bukay* <u>*iláa iyo íntuu*</u> *hálkán yimíd.*
DM+he   was:sick   until   and   time-the+he   place-this   came
'He has been sick <u>since he came here</u>.'

(48)   *Marka   waxay   labadii   kale   ku heshiiyeen   inay   indhaha*
       time-the *wáxa*+they two-the   other   on agreed   that-they eyes-the
       ka   xiraan habeenkii   marka   uu   guriga   joogo
       from   close evening-the   time-the he   house-the   stays
       ilaa   iyo   inta   ay   iyagu   ka   seexanayaan
       until   and   extent-the   they   they:IPRO from   go:to:sleep
       'Then the other two agreed to blindfold him when he was in the house <u>until they</u>
       <u>(left him to) go to bed</u>'   [SNR 14.4-5]

In addition to these types of clause, as we described in 8.2, tenseless appositive
relative clauses can be used to describe backgrounded events occurring
simultaneously with the main predication.

8.3.2 Location

Adverbial clauses of location are formed by relative clauses on nouns like *méel*
N F 'place', in a similar way to time clauses, for example:

(49)   <u>*Méeshaad   tagtóba*</u>   *wax*   *hubsó.*
       place+the+you   go-any   something   make:sure:of
       '<u>Wherever you go</u>, always make sure.'   (Proverb)
(50)   <u>*Méeshaanád   seexánáyn*</u>   *lagú   má   lulmoodó.*
       place+the+not+you   sleep:PROG:NEG   one+in   not   doze:NEG
       'One should not doze <u>where one is not planning to sleep</u>.' (Proverb)
       (lit. 'One does not doze in the place which you are not going to sleep in.')

8.3.3 Manner

Manner adverbial clauses are headed by the noun *sí* N F 'way, manner', with the
definite articles *-ta* and *-tii*, when appropriate, for example:

(51)   <u>*Sídii   uu   dóonayáy*</u>   *ayày*   *ú*   *noqotay*
       way-the   he   wanted   FOC+it   in   became
       'It happened <u>how/like he wanted it</u>.'
(52)   *Sidaas   daraaddeed   waa   loo   baahan yahay   in   la   yaqaanno*
       way-that   sake-its   DM   one+in   needing is   that one knows
       sida   qofkii   halaq   qaniino   loogu   gargaari   karo.
       way-the   person-the   snake   bites   one+in+for   help   can
       'Therefore one needs to know <u>how one can help someone bitten by a snake</u>.'
           (lit. 'the way in which one can help…') [BCH 35.18-19]

In this type of construction the relative clause is governed by the adposition *ú* in
the main clause, as described for adverbials more generally in 6.3.

8.3.4 Purpose

One important type of purpose adverbial clause is a relative clause headed by the undefined noun *sí* N F 'way, manner', governed by *ú* 'for' in the same subordinate clause:

(53) *waxaa   gurigoodii soo   galay   jiirarkii, si   ay   cunno*
      wáxa(a) house-their VEN entered mice-the, way they food
      uga   xadaan
      for+from steal
      'the mice went into their house <u>in order to steal food from it</u>' [SHS 80.40-1]

(54) *waxaan kaa   codsanayaa   inaad   i   amaahisid   carrabkaaga,*
      FOC+I you+from   ask   that+you me loan:2SG   tongue-your
      si   aan   ugu   mashxarado   arooska
      way   I   for+at ululate   wedding-the
      'what I ask from you is that you lend me your tongue, <u>so I may ululate at the wedding</u>' [SHS 79.27-8]

Another form of purpose clause consists of *ú* 'for' governing an *ín* 'that' complement clause (8.4.1):

(55) *Wúxuu   gúrigíi   ugá   tegay   ínuu   shaqó   doontó.*
      wáxa+he   house-the   for+from went   that+he   work   get
      'He left home in order to find work.' (lit. 'He left home for that he find work')

8.3.5 Reason

For the most part the expression of reason does not involve clausal subordination. The two major strategies are the juxtaposition of main clauses and the use of complex noun phrases without a clausal element. The former strategy involves main clauses like the question *maxàa yeeláy*? 'what caused it?' being used to connect two other main clauses, as in the following example. Note that the orthography does not always show the independence of the three clauses by punctuation, perhaps showing that speakers feel that the three grammatically independent clauses form a semantic unit:

(56) *waa in   degdeg qofka   loo   jiifiyaa   oo laga   dhawraa*
      DM that quickly person-the one+u lies:down and one+from prevents
      dhaqdhaqaaqa oo idil, maxaa yeelay dhaqdhaqaaqu wuxuu dedejiyaa
      movement-the and all   what did   movement-the FOC+it increases
      dhiigga   wareeggiisa
      blood-the   circulation-its

'it is necessary to lie the person down quickly and prevent him from moving at all because (lit. 'what caused it') movement increases the circulation of the blood'   [BCH 35.30-1]

As a link between main clauses the question *maxàa yeeláy*? 'What did it' may be replaced in other contexts by the question *maxàa jirá*? 'What is it?', the statement *sabábta wáxa wèeye* 'the reason it is (this)', or the conjunction *wàayo* 'because'.

The second strategy is to use nouns like *dár* N F 'sake, reason', *áw* N M 'reason', or *áwo* N F 'reason' in possessive constructions, with both simple and complex nominals, for example:

(57)   *báqdin   áwgéed*
        fear   reason-its
        'because of fear'
(58)   *sídáas   dártéed*
        way-that   sake-its
        'because of that, therefore'
(59)   *Tigídh la'áan áwadéed   ayáan   ú   tégi   waayey.*
        ticket   lacking sake-its FOC+I   to   go   failed
        'Because I didn't have a ticket, I didn't go.' (lit. 'Lacking a ticket its sake I didn't go')[93]

Sentence (59) is an example of nominalization, which we discuss in 8.4.3 below.

8.3.6 Conditional

Conditional clauses are headed by the noun *hád* F 'moment, point in time'[94] suffixed with the definite article *-tíi* to form *haddii*. *Haddii* coalesces with subject clitic pronouns and the negative word *aan,* as in the following examples:

(60)   *Haddáan   geeriyi ku helín   gábow   bàa   ku   héli.*
        time-the+not death   you find:NEG old:age FOC   you   find:INF
        'If death does not find you, old age will.'   (Proverb)
(61)   *Haddaad   dhir iyo dhagax dooneyso badi baad   u helaysaa.*
        time-the+you trees and   stone   want   plenty FOC+you in find
        'If you want trees and stone, you will find them in plenty.'   [TD 85.2-3]

The clause may be preceded by the intensifier *xátaa* 'even':

(62)   *Xátaa haddùu   roob da'ó, waa wákhti dambé ímminka ín   la*
        even   time-the+it rain   falls, DM   time   late   now   that one

nabad geliyó   midháha
peace   enter:CAUS crops-the
'Even if it rains, it's too late now for the crops to be saved.'

This type of clause is used as a counterfactual when the clause verb is in the conditional paradigm (4.3.4.9):

(63) *haddii   dalku   beero   falan,   guryo iyo adduun badan lahaan*
      time-the country-the farms   ploughed houses and wealth much have:INF
      lahaa dan   bay   kuu   ahaan lahayd inaad   u
      had   advantage   FOC+they you+for be   have   that-you for
      dagaalantaa.
      fight.
      'if the country had ploughed farms, houses and great wealth, there would be profit for you in fighting for it.'   [TD 84.5]

8.3.7 Concession

Clauses headed by the phrase *ín kastá oo*, or the coalesced form *ín kastòo*, carry the meaning 'although, though'. The phrase may be analysed as an appositive relative clause on the NP *ín kastá* 'any extent, each amount', though the whole expression seems to form a semantic unit in this context. See for example:

(64) *Ín   kastá oo   uu gaajóonayáy   wúxuu   diiday   cuntádii.*
      amount   any   and he hungered   *wáxaa*+he refused food-the
      'Although he was hungry, he refused the food.'
(65) *In   kasta oo   dalka   baddiisu dheer   tahay   kalluuna*
      amount   any   and country-the sea-its   long   is   fish-and
      hodan ku   yahay, haddana dadweynaha Soomaaliyeed intiisa
      rich   in   is   still   public-the   Somali:GEN   part-its
      badani   mallayga   ma   cunto.
      much   fish-the   not   eat:NEG
      'Although the country's shoreline is very long and rich in fish, nonetheless most of the Somali public do not each fish.' [J3 9.2-3]

8.3.8 Comparison

Comparison clauses are formed by restrictive relative clauses headed by *sí* N F 'way, manner', with suffixed definite articles *-ta* and *-tii*; for example:

(66) *sida   la   yeelo   wiyisha   markii   la   cararayo*
      way-the   one   does:to rhino-the time-the one   runs:away
      'as one does with a rhino when one is running away'   [SHS 13.27-8]

(67)   *Sidaan   u   joojiyey   rabshadii   Burcadda*   si   la   mid ah
       way-the+I   in   stopped   disturbance-the   Burad   way   with   one be
       yeel.
       do:IMPER
       'Do it <u>like I stopped the Burad disturbance</u>.' (lit. '<u>The way that I stopped the</u>
       <u>Burad disturbance</u>, do it in a way that is one with it')   [TD 153.18]

8.4   Verb complement clauses

In this section we discuss constructions where clause or clause-like elements
occur as complements of verbs. The great majority of such complements are
finite clauses introduced by the word *ín* which we gloss in English as 'that'.
There are in addition a very small number of verbs that allow infinitival verb
forms as complements. Finally in this section we briefly discuss nominalizations
which function as equivalents of clausal complements.

8.4.1 Finite ín *clauses*

Clauses acting as arguments of a verb are introduced by the word *ín* which is
probably historically related to the noun *ín* N F 'amount, extent'. However
synchronically *ín* is grammaticalised as a complementiser and is semantically
empty; see for example:

(68)   *Ínuu   imánayó   ayàan   ógahay.*
       that+he   is:coming   FOC+I   aware-am
       'I know that he is coming.'
(69)   *Ínuu   imánayaa   waa   hubáal.*
       that+he   is:coming   DM   certainty
       'That he is coming is certain.'
(70)   *Wáxaan   dóonayaa   ínaan   tagó.*
       wáxa+I   am:wanting that+I   go
       'I want to go.' (lit. 'I want that I go')

These examples show *ín* clauses acting like nominal arguments: the clause is
object in (68) and subject in (69) and is case marked accordingly. The clause is
marked as in focus by *ayàa* in (68) and by *wáxa* in (70).

        The word *ín* acts like the headword of a relative clause in that in attracts
certain elements out of the verbal group, notably subject clitic pronouns, as in
the examples above, and the subordinate negative word *áan* 'not' as in (71)
below. As with relative clause heads, when both elements occur they coalesce:

(71)   *Wúxuu*   *sheegay*   *ínaanu*   *garánéyn*
        wáxa+he   reported   that+not+he   understand:NEG
        'He said that he didn't understand it.'

Ín clauses have a similar syntax to relative clauses except that, as we might expect, *ín* never occurs as the subject of the clause verb and thus the reduced (or 'convergent') verb forms described in 4.3.4.13 and 8.1 never occur.

    The choice of main clause, or controlling, verb affects the tense and aspect of *ín* clauses. For many controlling verbs, the *ín* clauses have the full range of tense and aspectual distinctions found in main clauses; compare for example:

(72)   *Wáxa*   *la*   *íi*   *sheegay*   *ínuu*   *akhriyéy.*
        wáxa   one   me+to   reported   that+he   read
        'I was told that he read it.'
(73)   *Wáxa*   *la*   *íi*   *sheegay*   *ínuu*   *akhrínayéy.*
        wáxa   one   me+to   reported   that+he   was:reading
        'I was told that he was reading it.'
(74)   *Wáxa*   *la*   *íi*   *sheegay*   *ínuu*   *akhriyó.*
        wáxa   one   me+to   reported   that+he   reads
        'I was told that he reads it.'

In these examples there is a contrast between the past simple, past progressive and present general in the *ín* clause controlled by the verb *shèeg* 'report'. With certain classes of controlling verbs, however, no tense/aspect distinctions are possible. With these controlling verbs the *ín* clause verb must occur in the present general form but no present tense meaning is communicated. One class of these controlling verbs includes *dòon* 'wish', *goosó* 'decide', *weydiisó* 'request', *táli* 'advise', *ámar* 'command', *isku...dèy* 'try', *súg* 'wait for'. These verbs seem to share the semantic feature that their complement predicates are unreal or potential. Compare (75) & (76) below with (72) - (74) above:

(75)   *Wúxuu*   *dóonayaa*   *ínuu*   *tagó*
        wáxaa+he   is:wanting   that+he   go:PRES GEN
        'He wants to go.' (lit. 'He wants that he goes.')
(76)   *Wúxuu*   *dóonayay*   *ínuu*   *tagó*
        wáxaa+he   was:wanting   that+he   go:PRES GEN
        'He wanted to go.'   (lit. 'He wanted that he goes.')

No tense distinction is possible in an *ín* clause controlled by *dòon* 'want, wish' and other verbs of this class, like *goosó* 'decide', for example:

(77) *markuu*   *goostay*   *inuu*   *diinta*   *Kiristaanka*
      time-the+he   decided:PAST   that+he   religion-the Christianity-the
      galo
      enter:PRES GEN
      'when he decided to join the Christian religion'   [T4 5.2]

A second class of verbs which similarly controls tenseless *in* clauses includes the verbs *bilàw* 'begin' and *jóoji* 'stop', for example:

(78) *dabadeedna*   *inuu*   *isku*   *fidneeyo*   *billaabay*
      after-its-and   that-he   self+in   cause:trouble:PRES GEN   began:PAST
      'and then he began to cause trouble between them.' (lit. 'and then he began that he cause trouble between themselves'.)   [FS 71.10-11]

Unlike the first class of verbs, here the action of the clause is reported as actual, not unknown or unreal.

One important use of *in* clauses is as complements in verbless *waa* declaratives (7.3), where they express obligation and necessity:

(79) *Markuu*   *soo* *baraarugo* *waa* *in*   *la*   *siiyaa* *qaxwe* *ama* *shaah*
      time-the+he VEN comes:to   DM   that one gives   coffee   or   tea
      kulul
      hot   [BCH 32.28]
      'When he regains consciousness one must give him coffee or hot tea.'
(80) *Dawladda*   *Ingiriisku*   *waa*   *iney*   *ogaataa*   *in*
      government-the   English-the   DM   that-they   learn   that
      hadalladaasu   *ay*   *yihiin*   *been*   *iyo*   *khiyaamo.*
      talks-those   they   are   lies   and   deceit   [TD 200.25-6]
      'The English government have to learn that those talks are lies and deceit.'

8.4.2 Infinitival complements

As we have seen, most clausal subordination consists of the types of relative clause discussed in 8.1-2 and the closely related *in* 'that' complement clauses described in the last section. However a handful of verbs control infinitival complements. As discussed in 4.3.2 and 6.1.3 these divide into two types. The first are the grammaticalized auxiliaries *dòon*, *jír*, *léh* and *hày* whose role in the tense-aspect-mood system (TAM) bears no direct relation to the meaning and grammatical behaviour these verbs have as independent items. The verb *dòon* for example governs an infinitive verb when part of the TAM system signifying future tense, but requires an *in* complement clause when it is an independent verb meaning 'wish, want'. The second type is a small group of verbs that are

distinguished from the auxiliaries both grammatically and semantically. In infinitival constructions these verbs retain the same meaning as elsewhere and unlike auxiliaries they may be 'stacked' or concatenated. They are also morphologically independent in that they bear the whole range of TAM distinctions, unlike auxiliaries, which occur in a single positive and negative form.

These verbs include *kár* 'can, be able', *wàa* 'fail, miss, lack', and *gàadh* 'just fail, almost do'. The infinitival construction these verbs occur in bears little resemblance to main clauses: the infinitival verb is positioned immediately before the controlling verb within the verbal group (6.1), for example:

(81)　(a)　*Gaadhígii　ayàan　sóo　iibsán　karay.*
　　　　　car-the　　FOC+I　VEN　buy:INF　can:PAST SIMPLE
　　　　　'I was able to buy the car.'

　　　(b)　*Gaadhígii　ayàan　sóo　iibsán　waayey.*
　　　　　car-the　　FOC+I　VEN　buy:INF　fail:PAST SIMPLE
　　　　　'I failed to buy the car.'

　　　(c)　*Gaadhígii　ayàan　sóo　iibsán　gaadhay.*
　　　　　car-the　　FOC+I　VEN　buy:INF　almost:did:PAST SIMPLE
　　　　　'I almost bought the car.'

Such verbs may of course also occur with grammatical auxiliaries, for example:

(82)　*dhaqso*　[VGP *isma　ay　u　rogi　kari　jirin*　]
　　　quickly　　　　self-not　she　at　turn:INF　can:INF　used:NEG
　　　kolkay　ordayso
　　　time-the+she　running　　　　　[SHS 13.25]
　　　'she didn't use to be able to turn quickly when she was running'

Here the verbal group contains a grammatical auxiliary *jír* marking the past habitual, governing an infinitival controller *kár* 'can', which in turn governs the verb *róg* 'turn'.

8.4.3 Nominalized complements

Nouns derived from verbs and adjectives may function as equivalents to complement clauses. These may be intransitive as in (83) and (84) below, or control other nominals as arguments as in (85):

(83)　*Baxsadkoodii　bay　sii　wadeen.*
　　　escaping-their　　FOC+they　ALL　carried:on　　[SHS 20.14-15]
　　　'They kept on running away.' (lit. 'They continued their escaping away.')

(84)   *Maroodigii   orodkii   buu   ka   daaley oo   meel   istaagey.*
        elephant-the   running-the   FOC+he   from   tired   and place stood
        'The elephant grew tired of running and stood still.' (lit. 'The elephant grew
        tired of the running and stood in a place.') [SHS 90.6]

(85)   *Ninkii   waraabe buu   isu   rogey oo ratiga   kuruskiisii*
        man-the   hyena   FOC+he self+in turned and male:camel-the hump-his
        cunid   ku   bilaabay
        eating   on   began   [SHS 93.13-14]
        'The man changed into a hyena and began biting at the male camel's hump.'

The derived nouns in these examples are: *báxsad* N M from *baxsó* V MID 'escape,
flee'; *órod* N M 'running' from *órod* V 'run', and *cuníd* N F 'eating, consuming'
from *cún* V 'eat'.

Chapter 9
Discourse

In this chapter we look at some features of discourse that are most intimately linked with the grammatical form of sentences. We begin with information structure: the linguistic marking of speakers' decisions about how to tailor their message to the context of discourse. Such decisions are clearly reflected in the morphosyntax by the presence of focus words, choices of word order, cleft constructions, etc. Indeed Somali has been termed a 'discourse configurational language' (Kiss ed. 1995) because of the importance to its grammar of pragmatic functions like focus and topic. In this chapter we also look at the linguistic resources available for reference assignment: the means speakers choose to identify and track entities in a discourse. Finally in 9.3 we apply some of these notions in an analysis of a short extract from a traditional story.

9.1   Information structure

As we have seen in earlier chapters Somali has the overt markers of focus *bàa*, *ayàa* and *wáxa(a)*, and allows full NPs which are known or given to occur freely in any order, especially in *waa* sentences, as shown below:

(1)   (a)   *Cali   warqáddii   wuu   ii   dhiibay*
                 Ali   letter   DM+he   me+to   passed
                 'Ali passed the letter to me.'
(2)   (a)   *Warqáddii Cali wuu ii dhiibay*
        (b)   *Wuu ii dhiibay Cali warqáddii*
        (c)   *Wuu ii dhiibay warqáddii Cali*
                 'Ali passed the letter to me.'

All of the sentences in (1) & (2) are grammatical and the grammatical relations are very clear: the accentual patterns on the nominals mark *Cali* as subject and *warqáddii* as object, while in the verbal group the verb agrees with the 3:MASC:SG subject clitic pronoun *uu*. These and related features have led many observers to suggest that word order in Somali is primarily governed by pragmatic principles of information structure or 'information packaging'. In this section we discuss the related notions of focus and topic, and try to outline some of their major discourse functions, beginning with focus.

9.1.1 Bàa/ayàa focus

A traditional view of focus in Somali may be somewhat simplified as follows:

(3)   (a)   Every sentence must have one element (and only one element) in focus.
        (b)   Either a nominal or the verb must be focused.
        (c)   *Bàa*, *ayàa* or *wáxa(a)* mark nominal focus.
        (d)   *Waa* marks verbal focus.

While this view bears some relationship to the facts it oversimplifies the processes of information structure. Firstly, we have already seen some qualifications that need to be added:

(4)   (a)   Focus is restricted to the two sentence types (declarative and interrogative) that may bear tense distinctions.
        (b)   Focus is restricted to main clauses.
        (c)   Focus has a more restricted role in negative and interrogative sentences.
        (d)   *Bàa*, *ayàa* or *wáxa(a)* may focus a wider range of constituents than simply NPs.

In what follows we explore other changes to the simple outline in (3). In order to examine focus in more detail we might adopt a typology of focus suggested by Lambrecht (1994):
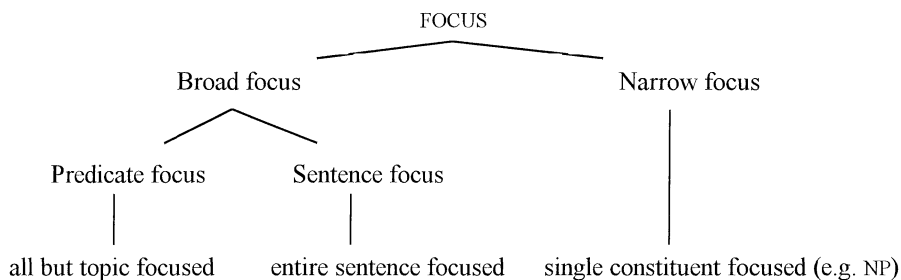
These distinctions may be shown by using questions to set up various contexts of speaker and hearer knowledge. For simplicity, we concentrate for the moment on positive declarative sentences and also leave *wáxa(a)* focus to section 9.1.2. As an example of predicate focus, a question like (5) below identifies *wìilkii* 'the boy' as topic, and asks a polar question without constituent focus. In the contextually appropriate reply (6a) *waa* is used; the reply in (6b) with nominal focus is grammatical but inappropriate (marked #):[95]

(5)　　*Wìilkii　má　matagay*?
　　　　boy-the　QM　vomited
　　　　'Did the boy vomit?'

(6)　(a)　*Hàa,　(wìilkii)　wùu　matagay*
　　　　　Yes,　(boy-the)　*waa*+he　vomited
　　　　　'Yes, (the boy) he vomited'

　　(b)　*#Hàa,　wìilkíi　ayàa　matagáy*
　　　　　Yes,　boy-the　FOC　vomited
　　　　　'Yes, THE BOY vomited'

　　　As an example of sentence focus we might use a question like (7) below, assuming this to be the first utterance in an exchange and therefore carrying no linguistically marked presuppositions about the entities in the situation.[96] The replies in (8) report an event: (8a) is the most appropriate; (8b) and (8c) are less appropriate (again marked #).

(7)　　*Maxáa　dhacáy*?
　　　　what+FOC　happened
　　　　'What happened?'

(8)　(a)　*Cáli　bàa　Fáarax　caayéy.*
　　　　　Ali　FOC　Farah　insulted
　　　　　'ALI insulted Farah.'

　　(b)　*#Cáli　Fáarax　bùu　caayey.*
　　　　　Ali　Farah　FOC+he　insulted
　　　　　'Ali insulted FARAH.'

　　(c)　*#Cali　Fáarax　wùu　caayey.*
　　　　　Ali　Farah　DM+he　insulted
　　　　　'Ali insulted Farah.'

It seems that the normal information structure for this kind of event report where all the nominals are new is for the subject to be focused: focusing the object or having the subject as topic is pragmatically less appropriate.[97]

assign particular grammatical roles like subject or object to an entity reflects a speaker's choice of viewpoint.[99]

Bàa/ayàa focus may also be used in situations where all the arguments are activated, as in the following examples:

(12) *Naagihii   saddex   ka   mid   ah   buu   ninkii   guursaday.*
      women-the   three   from   one   be   FOC+he   man-the   married
      '(Of) the three women, the man married ONE.'   [SHS 69.20]

(13) *Tól   iyo   fardó,   tól   bàan   doortay.*
      kinfolk   and   horses,   kinfolk FOC+I   chose
      '(Of) kin and horses (i.e. wealth), I chose KIN.'   (Proverb)

(14) *Libàax   yeedháy   iyo   libàax   aammusáy,   libàax   aammusáy   bàa   xún.*
      lion   roared   and   lion   kept:silent lion   roared   FOC   bad
      '(Of) a roaring lion and a silent lion, A SILENT LION is worse.' (Proverb)

Sentences like these involve contrastive focus: the speaker assumes or provides a finite set or list and uses focus to highlight one or more members. Of course the list may be potential and infinite rather than actual, as in one possible context for the sentence:

(15) *Háshán   anàa   léh*
      camel:F-this   I+FOC   own
      'I own this female camel,' 'This female camel belongs to ME'

A similar use of focus may occur in comparative constructions: here of course the list usually consists of just two members and once again the subject seems to be treated as more salient than the object/standard and thus often attracts focus:

(16) *War   hasha   labaad baa ka   caano badan midda   hore…'*
      sir   camel:F-the   second FOC than milk   much   one-the first
      'Sir, THE SECOND CAMEL has more milk than the first…' [SHS 70.36]

We can summarise these uses by saying that a second major function of *bàa/ayàa* focus is to place contrastive emphasis on one part of activated information.

More sophisticated extensions of these basic uses of focus are found in longer stretches of discourse such as narratives. As expected, new characters in a narrative may be introduced in focus, as in for example the first line of a story, *Bela Habreed* 'The Danger of an Old Woman' from the collection in Siyaad (1985):

As an example of single constituent focus we might use a question like (9) below, which identifies *inántii* 'the girl' as topic, and which questions her location. The question word in (9) and the relevant information supplied in the appropriate reply in (10a) are both focused; as (10b) and (10c) show, to focus the topic in the reply, or to use a *waa* sentence would both be inappropriate.

(9)   *Inántii   hálkày   joogtaa?*
        girl-the   place-which+FOC+she   stay
        'Where is the girl?', lit. 'The girl, which place is she in?'

(10)   (a)   *Qólkáasày   kú   jirtaa.*
               room+that+FOC+she   in   is
               'She's in THAT ROOM.'

        (b)   *#Inántii   bàa   qólkáas   kú   jirtá.*
               girl-the   FOC   room-that   in   is
               'THE GIRL is in that room.'

        (c)   *#Inántii   qólkáas   way   kú   jirtaa.*
               girl-the   room-that   DM+she   in   is
               'The girl is in that room.'

Examples like these show that Somali has different focus constructions for different matchings between given and new information. We will adopt a distinction discussed by Prince (1981) and others, between two types of given information. The first is variously called pragmatic presupposition, shared knowledge or common ground (Stalnaker 1974) and constitutes the assumptions and beliefs shared by the interlocutors. The second type is composed of activated or salient entities. The idea here is that certain items may be raised to salience in a linguistic exchange, form part of a frame of attention, and then fade over time.[98] Focus is a linguistic mechanism that plays an important role in this process of bringing new items into this frame of attention.

The examples so far allow us to make the following observations:

(11)   (a)   In an information question, the questioned element is focused by *bàa/ayàa*, and in the reply the relevant new information is similarly focused.

        (b)   In declarative sentences where all the nominals are activated and unfocused, *waa* occurs.

        (c)   In all-new, event reporting sentences the subject attracts *bàa/ayàa* focus.

We can summarise these by saying that one function of *bàa/ayàa* focus is to identify new information; and when more than one element is new, it attaches to the most salient element. We assume that independently of focus the decision to

(17)   *Waa baa waxaa   belo   isugu   faanay   libaax, good iyo*
        time FOC *wáxa(a)* danger self+in+about boasted   lion   asp   and
        habar.
        old:woman   [FS 37.1]
        'ONCE a lion, an asp and old woman boasted about how dangerous each
        was.'[100]
        (lit. 'Once there boasted about their danger a lion, an asp and an old woman')

Here the three characters are introduced by a *wáxa(a)* focus construction
(9.1.2). In the three following sentences these characters, though activated,
receive what Ajello (1995) describes as 'narrative focus':

(18)   (a)   *Libaax baa   hor   hadlay   oo   yiri:...*
                lion   FOC   first   spoke   and   said
                'LION spoke first and said...'
        (b)   *Good   baa   xigay,   oo   yiri...*
                asp   FOC   followed and   said
                'ASP followed on and said:...'
        (c)   *Habartii   baa   hadashay, oo   tiri...*
                old:woman-the   FOC   spoke   and   said
                'THE OLD WOMAN spoke and said...'   [FS 37. 2-6]

As each character moves centre stage in the narrative and acts, it attracts focus.
We can perhaps see this use of focus for narrative dynamism as an extension of
the placing of contrastive emphasis.
        Another narrative function of *bàa/ayàa* focus is also important: the
adverbials that set the time or location of the events described often attract
focus. We can take as examples the first lines of two other stories from Siyaad
(1985): *Dhaartii Xiidxiito* 'The Oath of the Ringed Plover' in (19) below and
Hawa Adduun 'Desire for Wealth' in (20):

(19)   *Beri ayaa iyadoo abaar   xumi jirto, dugaag   iyo haad   oo*
        time FOC it+and drought bad exist wild:animals and raptors and
        idli   shireen.
        all   met
        'ONE TIME when there was a bad drought, the wild animals and all the birds of
        prey held a meeting.'   [FS 23. 1]
(20)   *Maalin baa nin socota   ahi, shabeel waddada dhex   bilqan*
        day   FOC man traveller be   leopard path-the across spread
        la   kulmay...
        with met

'ONE DAY a man who was a traveller came across a leopard stretched across the path…'   [FS 69. 1-2]

In these two sentences it is the time adverbial which attracts focus despite the fact that all the nominals represent new information.

As we mentioned earlier, in such cases sometimes a second focus construction occurs in the same sentence (contrary to (3a) above). This is often a *wáxa(a)* focus construction, as for example in (17) above and in the first line of the story *Madax Bannaani* 'Independence' from the same collection:

(21)   *Waa   baa   waxaa   jiri   jiray,   qoys   laba   dameer   oo*
         time FOC   *wáxa(a)* be   used:to household two   donkey   and
         mataana   ah   leh.
         twins   be   have   [FS 33:1]
         'ONCE there used to be a household which had two donkeys which were twins.'

Here again the time adverbial attracts *bàa/ayàa* focus while a *wáxa(a)* construction introduces the other new information. As we shall see in 9.3, the management of narrative sequence, or what we might call the timeline in narratives, is an important function of *bàa/ayàa* focus.

Having established these functions, it must be noted that the use of *bàa/ayàa* focus is more restricted in negative sentences and in polar questions. In both of these *bàa/ayàa* focus only has the contrastive function described above. We can compare the following pair of negative sentences, for example:

(22)   (a)   *Cali   má   bixín.*
                Ali   not   left:NEG
                'Ali didn't leave.'
        (b)   *Cáli   báan   bixín.*
                Ali   FOC+not left:NEG
                'ALI didn't leave.'

Sentence (22b) is typically used against the background of a presupposition 'Someone left', a context that (22a) does not require. Focus in (22b) marks a contrast between Ali and the presupposed person(s) who did leave. We find a similar contrast in polar questions like the following pair:

(23)   (a)   *Cali   má   baxay?*
                Ali   QM   left
                'Did Ali leave?'

(b)   *Ma   Cáli   bàa   baxáy?*
       QM   Ali   FOC   left
       'Did ALI leave?'

Again (23b) is typically used against a presupposition that someone left and compares Ali to this person, asking if they are the same. Sentence (23a) does not require such a presupposition.

The negative polar question in (24) below has an even more specific licensing context:

(24)   *Ma   Cáli   báan   bixín?*
        QM   Ali   FOC+not   left:NEG
        'Didn't ALI leave?'

Here the typical context contains two presuppositions: 'Someone left' and 'the someone = Ali'. The question asks for confirmation of this second presupposition, i.e. of the identity of Ali with the leaver.

Thus *bàa/ayàa* focus has a narrower and more specific function in negative sentences and polar questions, requiring the support of more contextual assumptions. We can see examples of this asymmetry when negative and positive clauses are conjoined; in the same context the positive clause often has focus while the negative clause does not, for example:

(25)   *Maroodigu   takarta   ku   joogta   ma   arka   ee   tan   kale*
        elephant-the   gadfly-the   on   is   not   sees:NEG   but   the:one   another
        ku   joogtuu   arkaa.
        on   is+FOC+it   sees   [S1 88.6-7]
        'The elephant does not see the gadfly which is on itself but sees THE ONE ON ANOTHER (elephant).'   (Proverb)

Here the noun phrase *takarta ku joogta* 'the gadfly which is on it' does not receive focus in the negative clause while the contrasted noun phrase *tan kale ku joogtaa* 'the one which is on another' receives focus in the positive clause.

9.1.2 Wáxa(a) focus

Wáxa(a) cleft structures were introduced in chapter 7. They may be described as presentational structures[101] in which a lexically empty expletive *wáxa(a)*, related to the noun *wáx* 'thing', occurs before the verb and focused elements occur after the verb. This structural arrangement is significant because of the strong tendency, discussed in 9.1.4. below, for only known information to occur

after the verb. Typically post-verbal NPs, for example, are part of the background presuppositions of a sentence and may be described as after-thought topics. *Wáxa(a)* constructions allow the speaker to counter this general tendency by allowing new and contrasted information in post-verbal position. One important factor in the choice of *wáxa(a)* focus is the length or 'weight' of a focused element. *Wáxa(a)* clefts provide a mechanism for postposing long or 'heavy' constituents to a position after the verb. It is thus an important device for introducing quotations (26 below), lists (27), or long clausal arguments, for example relative clauses (28) or complement clauses (29). In these examples angle brackets < > show the limits of the cleft focus in the Somali, while small capitals again show *bàa/ayàa* focus:

(26) *Markaasuu   wuxuu   ku   yidhaahdaa,   <yartaneey,*
      time-that-FOC-he *wáxa*+he you   says   young:girl:VOC
      adigaan   kula   hadlayaa   ye gurigiinu ma kaasaa?>
      you+FOC+I you+with   talk:PRES.PROG and house-your QM   that+FOC
      'THEN he says to you: "Young girl, I'm talking to YOU, and is this YOUR
      HOUSE?"'   [SHN 178. 21-22]

(27) *Soomaalidu   waxay   ka   gabyi   jirtay   <dhinacyo badan sida*
      Somalis-the *wáxa*-they about poetise used:to   subjects   many like
      dhinac   ammaanta, cayda,   habaarka,   baroorta,   falsafadda,
      subject   praise-the insult-the curse-the   mourning-the philosophy-the
      xigmadda,   jacaylka iyo   kuwo kale   oo   badan>.
      wisdom-the   love-the and   ones other   and   many
      'The Somalis they used to compose poetry on many subjects including the
      subjects of praise, insults, curses, mourning, philosophy, wisdom, love and
      many others.'   [S2 22.29-31]

(28) *Magacaas   waxaa   la   siiyaa   <cudurro   ka   yimaada*
      name-that   *wáxaa*   one   gives:to   diseases   from   come
      bahallo   yaryar   oo   badidood   caloosha   gala>
      creatures   small   and   most-their stomach+the   enter
      'That name is given to diseases which are caused by small organisms, most of
      which enter the stomach'   [BCH 13.30-31]

(29) *waxayna   ogaatey <in naagtu   ka   adagtahay ninka,*
      wáxa+she+and   realised that woman-the ADP   tough-is   man-the
      inkasta   oo   uu ka   xoog   badan   yahay iyada>
      although   and   he than strength   more   is   her
      'and what she realised was that woman is tougher than man, although he is far
      stronger than her'   [SHS 80.29-30]

Unlike *bàa/ayàa*, *wáxa(a)* is not used to focus adverbial expressions or questioned constituents in information questions. As described in Saeed

(1984:154), a question like (30) below is avoided except with a pejorative implication:

(30)   *Wáxa   keenáy   áyo?*
        wáxa   brought   who
        '(The one)Who brought it is who?' 'Who brought it?'

We also do not find questions like (31a) paralleling the *bàa* sentence in (31b):

(31)   (a)   *Wáxay   sameeyeen   máxay?*
               wáxa+they   did   what
               'What they did was what?' 'What did they do?'
        (b)   *Maxáy   sameeyeen?*
               what+FOC+they   did
               'What did they do?'

The form of a cleft may be given to information questions but by a circumlocution which retains the question word to the left of the main verb, for example:

(32)   *Waxay   taageen   muxuu   ahaa?*
        what+they   raised   what+FOC+it was?
        'What they raised, what was it?'   [S1 35.32]

Another difference from *bàa/ayàa* noted by Lecarme (1991) is that *wáxa(a)* is not used to postpose empty indefinites like *wáx* 'thing, something' which must be used with transitive verbs to avoid a default third person pronoun interpretation. Thus we can find a *bàa* sentence like (33) but not a *wáxa(a)* sentence like (34):

(33)   *Wáx   bàan   akhrínayaa*
        thing   FOC+I   read:PRES.PROG
        'I'm reading something.', 'I'm reading.'[102]
(34)   **Wáxaan   akhrínayaa   wáx.*
        wáxa I   read:PRES.PROG   thing
        'I'm reading something.', 'I'm reading.' (lit. 'What I'm reading is something')

To summarise: *wáxa(a)* clefts are typically used to postpose and place in focus long NPs, clauses and quotations. They differ from *bàa/ayàa* focus in that they are not used with questioned elements, adverbials,[103] and lexically empty indefinites.

9.1.3 The status of waa

The problem of *waa* is part of the more general question of whether there is overt verb focus in Somali. This has often been assumed to be the case and the lexically empty word *waa* has been identified as a verb focus marker. The motivation for this is clear: as we have seen, *waa* is used in sentences where all the nominal arguments are activated and *waa* does not occur in sentences with *bàa/ayàa* or *wáxa(a)* focus. This at least suggests that *waa* is consistent with verb focus, as in examples (1-2) earlier.

However there are difficulties with identifying *waa* as a specific marker of verb focus. One problem is that it has a very restricted distribution: it does not occur in polar questions with the question word *ma*, nor in negative sentences with the negative word *má* 'not'. Neither of these grammatical morphemes seems to be a candidate for a verb focus word. This would mean that verb focus, if identified, is restricted to positive declarative sentences.

A more serious problem is that *waa* occurs in verbless sentences like those in (35-6) and (37b) below, described in 7.3 earlier:

(35) *Axmed   waa   askári.*
      Ahmed   *waa*   soldier
      'Ahmed is a soldier.'
(36) *Wàa   gabdhó.*
      waa   girls
      'They are girls.'
(37) (a)   *Ma   nabád   bàa?*
              QM   peace   FOC
              'Is it peace?'   (a greeting)
      (b)   *Wàa   nabád.*
              'It is peace.'   (the reply)

Firstly it seems unlikely that a verb focus particle would allow deletion of the verb since, as we have seen, focus is a device for marking salience. Secondly these *waa* verbless sentences seem to focus the complement NP. We saw earlier that the locus of *bàa/ayàa* focus in questions is matched in the appropriate replies; yet in (37) above *nabád* 'peace' is focused by *bàa* in the question (37a) and occurs with *waa* in the reply (37b). Similarly, as we saw in 9.1.1, the questioned element in information questions attracts *bàa/ayàa* focus yet we commonly find verbless information questions with *waa*, like the following:

(38) (a)   *Wàa   maxáy?*
              'What is it?'

(b) *Waa   kúma?*
     'Who is it?'
(c) *Waa   xaggée?*
     'Where is it/are they?'

Discussing these and similar problems I argued in Saeed (1984) that there is no overt verb focus word in Somali and that *waa* should be viewed, like the question word *ma*, as merely consistent with verb focus. In this approach *waa*, like the question word *ma*, and the optative word *há*, is analysed as a sentence type marker (STM), in this case identifying positive declarative sentences.   This has the advantage that it explains why *waa* is replaced by other markers in other sentence types, and by the negative word *má* in negative declaratives.[104]

      Of course in this approach we must explain why *waa* does not occur with *bàa/ayàa* and *wáxa(a)* in positive declaratives with focus. We might take note here of Ajello's (1995) observation that focus can be viewed as part of the modality system in Somali. He suggests that nominal focus asserts the existence of the referent: "In altre parole, anche quando la focalizzazione verte su un sintagma nominale, essa rappresenta l'asserzione che un tale sintagma nominale non è virtuale, ipotetico, ma reale e che esso ha un certo ruolo all'interno della predicazione principale." (Ajello 1995: 16). *Waa* sentences on the other hand involve the assertion of the whole predication. This identification of a functional contrast   may   help   explain   the   complementary   distribution   of   the constructions.[105]

9.1.4 Word order and topics

As we have seen, the description of word order, or more properly constituent order, has to make reference to both syntactic status and information structure. Inside phrases the order of elements is generally fixed, and the constituents of verbal groups (6.1) and noun phrases (6.2) can be described by templates. The relative order of noun phrases and verbal groups, however, is dependent on information structure. It is possible to discern an underlying preference for the verb final clause structure described for other Cushitic languages. One piece of evidence is the rigid order of subject-object-verb we find in clitic pronouns in the VGP (6.1.1). Another is the fact that the basic focus strategy, *bàa/ayàa* focus, requires focused elements to occur before the verb. Speakers must use a *wáxa(a)* cleft to position focused elements after the verb. This is related to a further significant feature: the tendency for post-verbal NPs to be known and activated. Indeed it is reasonable to view post-verbal NPs, other than in clefts, as afterthought topics. As Gebert (1986) points out, in an all-new context provided by a question like (39a) below, reply (39c) is inappropriate:

(39)   (a)   *Maxàa   dhacáy?*
                   what+FOC   happened
                   'What happened?'

         (b)   *Cáli   bàa   Maryám   diláy.*
                   Ali   FOC   Mariam   hit
                   'ALI hit Mariam'

         (c)   *#Cáli bàa   diláy   Maryám.*
                   Ali   FOC   hit   Mariam
                   'ALI hit Mariam'

The nominal *Maryam* can appropriately be new in (39b) but not in (39c).

    This observation needs qualification, however, since there are a number of cases where this tendency does not hold. Adverbial expressions for example can occur post-verbally even when new. Gebert (1986) provides the following examples:

(40)   (a)   *Maxaa   dhacay?*
                   'What happened?'

         (b)   *#Cali   baa   dilay   Maryam.*
                   Ali   FOC   hit   Maryam
                   'ALI hit Maryam.'

         (c)   *Cali   baa   dhintay shaley.*
                   Ali   FOC   died   yesterday
                   'ALI died yesterday.'

Here the adverbial *shaley* 'yesterday' but not the object *Maryam* is appropriate as new information in post-verbal position. Similarly, locative NPs governed by verbal adpositions may occur post-verbally when new; to use Gebert's examples again:

(41)   (a)   *Wax   cusubi ma   jiraan?*
                   thing   new   QM   are
                   'What's the news?'

         (b)   *#Cali   baa   tegay   Marka.*
                   Ali   FOC   went-to Marka
                   'ALI has gone to Marka.'

         (c)   *Cali   baa   ku   dhintay   Marka.*
                   Ali   FOC   in   died   Marka
                   'ALI has died in Marka.'

Representing new information in (41), the town name *Marka* is appropriate post-verbally when governed by the adposition *kú* 'in' but not when it is an argument of the verb *tág* 'go to'.

Interestingly this seems to be more dependent upon the function role of the governed nominal than the simple syntactic fact of government by an adposition, since in the following exchange (42c) is less appropriate than (42b):

(42)   (a)   *Wáx   cusubi   má   jiraan?*
                thing   new   QM   are
                'What's the news?'
        (b)   *Cáli   bàa   gabádh   lá   baxáy*
                Ali   FOC   girl   with   went
                'ALI has eloped with a girl'
        (c)   *#Cáli   bàa   lá   baxáy   gabádh.*
                Ali   FOC   with   went   girl
                'ALI has eloped with a girl.'

The NP *gabádh* 'girl' is new information in (42b-c), and though governed by the adposition *lá* 'with', is more appropriately positioned pre-verbally.

This tendency to avoid new information in post-verbal position is also relaxed when the new argument is a clause, for example a complement clause:

(43)   *Nin   doob   ah   baa   beri   damcay   in   uu   guursado*
        man   young:bachelor be   FOC   day   decided   that   he   marry
        inan   qurxoon   oo   uu   jeclaaday...
        girl   beautiful   and   he   loved
        'A YOUNG BACHELOR decided one day to marry a beautiful girl that he
        loved...'   [SHS72. 11]

We can perhaps see this as part of a general strategy to shift long items to the post-verbal position.

This general tendency towards verb-final word order is also disguised by the positional freedom of activated nominals. In order to discuss the positional options of these it is helpful to adopt a notion of markedness: a less marked, or more basic, word order may be defined in a number of ways. One of the most important is appropriateness in a wider number of contexts and this of course may be related to a higher numerical frequency. Gebert (1986) identifies two unmarked or basic orders in *bàa/ayàa* sentences (where +F = focused; *pro$_i$* = coreferential subject clitic pronoun):

(44) Unmarked order #1
      SUBJ$_i$ - OBJ[+F] + *pro$_i$* - V,
      which occurs when the subject is topic, for example:
      Durwaa   Ilaah   buu   u   gargelyooday
      Hyena   God   FOC+he   to   complained
      'Hyena, he complained to GOD'   [FS 21.15]

(45) Unmarked order #2
      SUBJ[+F] - OBJ - V
      which occurs in all-new information contexts, for example in the first line of a text:
      Raage Ugaas baa gabadh la   odhan jirey   Cabban Cilmi
      Raage Ugaas FOC girl   one call   used-to Abban Ilmi
      Hagoog   jeclaaday
      Hagoog   love-INCH-PAST   [S1 8.1-2.]
      'RAAGE UGAAS fell in love with a girl who was called Abban Ilmi Hagoog'

Each of these unmarked orders has several more marked versions, for example more marked versions of (44) are:

(46)   OBJ[+F] + *pro*   -   V   -   SUBJ
             where the subject is an afterthought topic.

(47)   OBJ [+F] + *pro*   -   SUBJ   -   V
             which reduces the topicality of the subject.

Similarly, more marked versions of (45) are:

(48)   OBJ   -   SUBJ[+F]   -   V
             where the object is activated.

(49)   SUBJ[+F]   -   V   -   OBJ
             which is an even more marked pattern in which the object must be activated, and acts as an afterthought topic.

We shall see examples of some of these word order patterns in section 9.3 where we examine a short text.

To summarise, we can contrast the fixed constituent order inside the NP and the VPG with the great positional freedom of NPs around the VPG. Within this freedom of order we can detect a bias towards an underlying verb-final order which may be overridden by a number of factors. Important among these is the positional freedom of known or activated nominals that are represented as clitics within the VPG. Such varying orders are appropriate in different discourse contexts. Another important factor is the presentational use of *wáxa(a)* clefts to place focused elements after the verb.

Although the study of these discourse processes in Somali has scarcely begun we might venture a couple of simple hypotheses which might be tested in future work. We can take as two basic sentence types: (a) focus sentences with *bàa*/*ayàa* and (b) non-focus sentences with *waa*. We can identify in each a core, as follows:

(50)   Focus sentences with *bàa*/*ayàa*
         X   -   [$_{CORE}$ FOCUS…Z…VPG]   -   Y

(51)   Non-focus sentences with *waa*
         X   -   [$_{CORE}$ *waa* - VPG]   -   Y

In both sentence types typical participant NPs can only occur in positions X and Y if they are activated and represented by clitic pronouns in the core.[106] We can call position X the topic and position Y the afterthought topic, respectively, understanding that more than one NP can occur in either position. As we have seen certain unactivated or new elements may occur in these positions without being topics: adverbials and locative NPs can occur in both X and Y, and 'heavy' elements like complement clauses can occur in position Y. In the focus sentence pattern (50) position Z may be occupied by either new or activated nominals but the latter do not act as topics in this position.

This is only the slightest sketch but allows us to reflect that a combination of overt focus markers (*bàa*/*ayàa*), word order, and clitic pronouns linguistically signal certain important pragmatic roles, particularly of focus and topic. It will require the study of different sentence types in texts across of range of different genres to fill out these few tentative suggestions.

It is perhaps worth noting that the relationship between a topic and the core predication may be of various types. Most frequently, the topic is coreferential with a core argument expressed by a pronoun, as in (52) below:

(52)   *Anigu   geel,   ido   iyo   riyo   badan   baan   haystaa.*
        Me   camels sheep and   goats many   FOC+I have
        'Me, I have MANY CAMELS, SHEEP AND GOATS.'   [TD 84.33]

Though this coreferential element is most commonly a clitic pronoun, it may also be another element with a referential function, for example a possessive determiner as in:

(53)   *Hooyo,   aniga   ilmahaygu   wacal   ma   aha.*
        mother   me   child-my   bastard not is
        'Mother, (as for) me, my child is not a bastard.'   [SHN 106.1]

In this example, since the topic *aniga* 'me' is not coreferential with the subject of the sentence *ilmahaygu* 'my child', it is not subject marked.

The topic may also be in a set-subset, or partitive, relationship with an argument in the core, as in (54) below:

(54) *hase yeeshee   saddexdii   bakheyl   mid   ka   mid   ah*
      however   three-the   miser   one   from   one   be
      ayaa   diiday   inuu   wax   ka   bixiyo   qiimihii   feynuuska.
      FOC   refused   that-he   thing   from   pay   cost-the   light-the
      'however, the three misers, ONE OF THEM refused to pay anything towards the
      cost of the light'   [SNR 14. 2-3]

9.2   Reference assignment

There is a range of nominal forms available to speakers to help manage reference to entities within a discourse. Indefinite nominals, which are often used to introduce new entities into the discourse are unmarked, i.e. there is no indefinite marker or article in Somali. See for example the first line of a story:

(55) *Waa   baa   waxaa   israacay   laba nin   oo   saaxiib*
      time   FOC   *wáxaa*   RE-accompanied   two   man   and   friend
      ahaan   jirtey
      be:INF   used-to
      'ONCE there travelled together two men who were friends.'   [MCI 12.17]

In this sentence the two referring nominals *waa* 'time', and *laba nin* 'two men' (lit. 'two of man') are new and undefined.

When the first mention has been made an entity is activated as part of the frame of attention and may be referred to by a range of definite nominals and pronouns. In this story the two men are referred to in the second sentence by the subject clitic pronoun *ay* 'they' (4.2.1), and the possessive determiner -*kood*/-*tood* 'their' (4.6.4), both marked in bold below:

(56) *Goor   **ay**   labad**oodu**   meel   marahayaan   **bay**   waxay*
      time   they   two-their   place   pass:PROG   FOC+they   *wáxa*+they
      iswarsadeen   in   nin   walba   wuxuu   jecel   yahay   in
      RE-asked that   man   each   thing-the+he   liking   be   that
      Eebbe   ugu   roonado   sheegto.
      God   to+for   grant   tell.
      'As the two of them were passing along, they agreed that each man should say
      what he would like God to grant him.' [MCI 12.18-19]

As can be seen, within this sentence the pattern is repeated. A new quantified nominal *nin walba* 'each man' is introduced by an indefinite NP and then this nominal is anaphorically referred to by a subject clitic pronoun *uu* 'he' and then later in the sentence by the zero third person object pronoun (4.2.2).

In subsequent sentences the men, singly or together, are referred to by noun phrases containing the remote definite article *-kii/tii* 'the' (4.6.1) which is used for referents distant in space or temporally in the past and which in discourse has an anaphoric function. Thus in successive sentences we find *nin-kii kale* 'the other man', *laba-dii nin* 'the two men', as well as further uses of the subject clitic pronouns *ay* 'they', *uu* 'he' and zero third person object pronouns.

In the middle of the story another character is introduced by an indefinite noun *nin rati subag uu u saaran wata* 'a man driving a camel loaded with clarified butter':

(57)   *Markaas muddo   yar dabadeed baa   iyagoo   meel   marahay*
        time-that period   little after-its   FOC   they-and place pass:PROG
        waxaa   ka   hor   yimid   nin   rati   subag   uu   u
        wáxaa   from   front came   man   camel   clarified:buttter he   on
        saaran   wata.
        loaded   drive.
        'A short time after that, while they were passing a place, there came before them a man driving a camel loaded with clarified butter.' [MCI 13.9-11]

In the next clause following (57) the two men are referred to again, this time by a noun phrase containing a demonstrative determiner *-kaas/-taas* 'that, those' (4.6.2) used deictically for objects in the middle distance, *laba-daas nin* 'those two men'. The new character is referred to with the remote definite article *-kii/-tii*: *nin-kii safar-ka ahaa* 'the man who was a traveller'. In this way the demonstrative picks out the entities, the two men, 'on the other side', so to speak, of the newly introduced character. The two types of determiner are used to operate a form of discourse deixis.

Thus we can see a series of nominal forms being used to manage reference to entities within the discourse: indefinite nominals, clitic pronouns, nominals with the remote article *-kii/tii*, and with the middle distance demonstrative -*kaas/-taas*. One other feature of this system is worth noting: clitic pronouns must occur within the verbal group; therefore when a pronoun is selected elsewhere in a sentence, to act as a topic, or head of a relative clause, or in focus, then an independent pronoun must be used. Thus in (56) above we find

the appositive relative clause *iyagoo meel marahay* 'they, passing a place' headed by the independent pronoun *iyága* 'they, them' (4.1.3).

9.3 Analysis of a text

To look at the features of information structure in a little more detail we may examine some brief extracts from a traditional story *Guurdoon* 'Choosing a Bride' in Axmed Xartan Xaange (1988:49-50.). This story is given in its entirety in Appendix II. The first sentence of the story is (58):

(58) *Waa baa waxaa jirey nin iyo afadiis, waxayna*
     time FOC *wáxaa* was man and wife-his *waxa*+they+and
     haysteen xoolo ay dhaqdaan.
     had livestock they cared:for
     'Once there was a man and his wife and they had some livestock which they
       raised.' (i.e. 'and they raised livestock for their living')

In this sentence, the initial element, the time adverbial *waa* 'one time, once' is focused by *bàa*, while the first characters, the man and his wife are introduced in a *wáxa(a)* cleft. A second cleft introduces a relative clause on an indefinite NP *xoolo* 'livestock'. The second sentence is:

(59) *Waaggii dambe baa wiil u dhashay reerkii, ninkiina*
     time+the after FOC boy to was:born family-the man-the-and
     aad buu ugu rayreeyey wiilka u dhashay.
     much FOC+he ADP+ADP be:happy boy-the to was:born.
     'Some time later a boy was born to the family, and the man was very happy
     with the boy born to him.'

In this sentence the initial time adverbial, *waaggii dambe* 'some time later', is again focused by *bàa*. The new entity *wiil* 'a boy' is introduced by an indefinite noun without focus in the first clause and then referred to in the second clause by a NP containing the non-remote definite article *-ka/-ta* (4.6.1): *wiil-ka* 'the boy'. The second clause has the activated NP *ninkii* 'the man' (containing the remote definite article *-kii*) as a topic, coreferential with a clitic pronoun *uu* 'he' in the core of the predication. The focus of this clause is the intensifier *aad* 'much'.

     If we move on four or five sentences, which describe the boy growing up and deciding to choose a bride with his father's approval, we find the sentence:

(60)   *Maalintii   dambe baa   wiilkii   soo   watay   gabar dheer oo*
        day-the   next   FOC   boy-the   VEN   came-with   girl   tall   and
        cas   oo   qurxoon   oo   yiri:
        copper-coloured   and   beautiful   and   said:
        'The next day the boy brought along a tall, beautiful, copper-coloured girl and
        said:'

Here it is the initial time adverbial which receives focus with *bàa*. The new
information in the first clause occurs without focus and post-verbally,
presumably because it is a long NP, here a relative clause 'a girl who is tall and
copper-coloured and beautiful'. The second clause, beginning *oo yiri* 'and said',
introduces a quotation, which is the independent sentence (61) below:

(61)   *Waa   tan,   aabbe,   inantaan   rabaa inaan   guursado.*
        DM   this   father   girl-the+I   want   that-I   marry
        'Here she is, father, the girl I want to marry.'

This sentence has no overt focus markers, though the verbless construction *waa
tan* 'it is this one, here she is' can be seen to place its complement *tan* 'this one
(F)' in some form of focus, as suggested in 9.1.3.
     The father sets his son the task of lifting a heavy stone and his subsequent
failure deters the girl from marrying him. Four other girls are described as
undergoing the same test. They are introduced by an indefinite nominal with *bàa*
focus:

(62)   *Afar gabdhood oo kale buu   wiilkii hor   keenay aabehiis*
        four girls   and other FOC+he boy-the before brought father-his
        'Four other girls the boy brought before his father' [SHS 49.38-9]

In the context of the story before (62) both NPs *wiilkii* 'the boy' and *abbehiis*
'his father' are activated, but the latter is given higher topicality as an
afterthought topic. Indeed the next clause in the story has the father as subject:

(63)   *isaguna   dhagaxii   in   uu   soo   qaddo   buu   sharuud*
        he+and   stone-the   that he   VEN   lift   FOC+he   condition
        kaga   dhigay wiilkiisa
        ADP+ADP   set   son-his   [SHS 49.39-40]
        'And he set on his son as a condition that he lift the stone'
        (lit. 'And he, the stone, that he lift it, he set as a condition on him, his son')

In (63) the complement clause *in uu soo qaddo* 'that he lift it' represents given or old information and is focused by *bàa*. There are two topics: the first is the external or main clause topic, the independent pronoun *isagu* 'he', referring to the father. The second is the NP *dhagaxii* 'the stone' which is a topic associated with the complement clause. A third nominal *wiilkiisa* 'his son' is presented as an afterthought topic to the main clause predication. These three activated nominals all occur outside the core predication.

Finally a fifth girl is introduced by an appositive relative clause in focus by *bàa*, shown in (64) below. Interestingly, although this is the first mention of this person the nominal is definite, containing the remote definite article *kíi/tíi*.[107]

(64) *Gabadhii   shanaad   oo   aan   u   qurux   badnayd   siddii*
      girl-the   fifth   and   not   in   beauty   much:was   way-the
      afartii   hore   baa   wiilkii   u   keenay   adoogiis.
      four-the   before   FOC   boy-the   to   brought   father-his.
      'The boy brought to his father the fifth girl who was not as beautiful as the first four.'   [SHS 50.4-5]

This girl assists the son and passes the father's test:

(64) *Markii   gabadhu   aragtay   in   wiilku   qaadi   kari   waayey   dhagaxii*
      time-the girl-the saw   that   boy-the   lift   can   failed   stone-the
      weynaa   bay   la   qaadday oo   soo   agdhigeen   odayga
      big   FOC+she   with   took   and   VEN   near-put   father-the
      cagihiisiii,   isaguna   eegayo   sida   ay   isu   kaashadeen
      feet-his   him-and   watching way-the they each:other-ADP helped
      wiilka   iyo   gabadhu.
      boy-the   and   girl-the.
      'When the girl saw that the boy wasn't able to lift the big stone, she lifted it with him and they put it at the father's feet, while he watched them help each other, the boy and girl.' [SHS 50. 5-7]

In this long sentence *bàa* focus is placed once again on an initial time adverbial, here a restrictive relative clause, *markii gabadhu aragtay in …weynaa* 'when the girl saw that...stone' (8.3.1). The sentence ends with an appositive relative clause acting as an absolutive adverbial, as described in 8.2. In this appositive relative clause the NP *wiilka iyo gabadhu* 'the boy and the girl' occurs post-verbally as an afterthought topic, which is coreferential with the preverbal clitic pronoun *ay* 'they' and is subject marked.

The end of the story, including the lesson of the test, can be read in Appendix II. These extracts exhibit some of the main features of discourse

structure discussed in previous sections. We see focus being used in its main functions: primarily here to establish a clear timeline to the narrative, via adverbial phrases and clauses, for example in (58), (59), (60) and (65). We also see focus used to introduce new information, for example (62) and (64); and to provide emphasis on already activated information, for example (63). We can also see the positional flexibility of activated nominals, which can occur as topics, as in (59) and (63), afterthought topics, for example in (65), or of course within the main predication, for example *wiilkii* 'the boy' in (60). We can also see, for example in (60), how long or 'heavy' nominals may override the tendency for only activated nominals to occur after the verb.

Appendix I
Data Sources

The following are the principal published sources for the Somali examples used in this book. They are cited in the text by square brackets containing the initials below followed by page and line numbers, for example [SHS 25.17-19], meaning lines 17-19 on page 25 of *Sheekoxariirooyin Soomaaliyeed.*

BCH   *Buugga Caafimaadka iyo Hufnaanta. Dadka Waaweyn.* Muqdisho: Wasaaradda Waxbarashada. Guddiga Afka Soomaaliga, 1973. [*The Book of Health and Hygiene for the General Public.* Mogadishu: Ministry of Education. Somali Language Commission.]

FS   Ciise M. Siyaad. 1985. *Favole somale.* (Studi Somali 6) Rome: Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

HS   Muuse Haaji Ismaa'iil Galaal and Andrzejewski, B.W. 1956 *Hikmad Soomaali.* [Somali Wisdom] London: Oxford University Press.

J3   *Juqraafi. Fasalka Saddexaad.* Xamar: Wasaarada Waxbarashada iyo Barbaarinta. Xafiiska Manaahijta. (1977). [*Geography. Class Three.* Mogadishu: Ministry of Education and Training. Curriculum Dept.]

MF   Maxamed Daahir Afrax. 1993. *Maana Faay.* Second edition. Stockholm: African Triangle.

MCI   Muuse Cumar Islaam. 1973. *Sheekooyin Soomaaliyeed.* Muqdisho: Wasaaradda Hiddaha iyo Taclinta Sare [*Somali Stories.* Mogadishu: Ministry of Culture and Higher Education].

SHN   Hassan Sheikh Mumin.   1974.   *Leopard among the Women.*
            Shabeelnagood.   Translated   with   an   introduction   by   B.W.
            Andrzejewski. London: Oxford University Press.

SHS   Axmed Cartan Xange.   1988.   *Sheekoxariirooyin Soomaaliyeed.*
            Folktales from Somalia.   Uppsala: Somali Academy of Arts and
            Sciences in co-operation with the Scandinavian Institute of African
            Studies.

SNR   Abdullahi. A. Issa and John D. Murphy.   1984. *A Somali Newspaper*
            Reader. Kensington, Maryland: Dunwoody Press.

S1   *Suugaan. Dugsiga Sare. Fasalka Kowaad.* Xamar: Wasaarada
            Waxbarashada iyo Barbaarinta.   Xafiiska Manaahijta. (not dated).
            [*Literature. Secondary School. Class One.*   Mogadishu: Ministry of
            Education and Training. Curriculum Dept.]

S2   *Suugaan. Dugsiga Sare. Fasalka Labaad.* Xamar: Wasaarada
            Waxbarashada iyo Barbaarinta.   Xafiiska Manaahijta. (not dated).
            [*Literature. Secondary School. Class Two.*   Mogadishu: Ministry of
            Education and Training. Curriculum Dept.]

S5   *Suugaanta. Fasalka Shanaad.* Xamar: Wasaarada Waxbarashada iyo
            Barbaarinta.   Xafiiska Manaahijta. (1976). [*Literature. Class Five.*
            Mogadishu: Ministry of Education and Training. Curriculum Dept.]

TD   Jaamac Cumar Ciise.   1976.   *Taariikhdii Daraawiishta iyo Sayid*
            Maxamed Cabdulle Xasan (1895-1921). Muqdisho: Akadeemiyaha
            Dhaqanka. [The Historyof the Dervishes and Mohamed Abdullah
            Hasan (1895-1921). Mogadishu: Academy of Culture]

T4   *Taariikh. Fasalka Afaraad.*   Xamar: Wasaarada Waxbarashada iyo
            Barbaarinta. Xafiiska Manaahijta. (not dated). [*History. Class Four.*
            Mogadishu: Ministry of Education and Training. Curriculum Dept.]

T5   *Taariikh. Fasalka Shanaad.*   Xamar: Wasaarada Waxbarashada iyo
            Barbaarinta.   Xafiiska Manaahijta. 1977. [*History.   Class Five.*
            Mogadishu: Ministry of Education and Training. Curriculum Dept.]

Appendix II
Texts

1. Animal fable: *Qaybtii Dawaco* 'The Jackal's Division'.
2. Traditional story: *Guurdoon* 'Choosing a Bride'.
3. Extracts from a letter written in 1903 from Sayid Maxamed Cabdulle Xasan to the British authorities.
4. Extract from a school text *Taariikhdii Soomaaliya ee Hore ee Tii Waayadii Dheexe* 'The History of Early and Medieval Somalia.'
5. The poem *Alleyl Dumay* 'Night fell' by Raage Ugaas.
6. Extract from the poem *Gammaan waa magac guud* 'Horse is a general term' by Cali Bucul.

Text 1: Animal fable. *Qaybtii Dawaco* 'The Jackal's Division' (Siyaad 1985: 57)

Beri	*baa*	*waxaa*	*wada*	*ugaarsaday*	*libaax,*	*dawaco,*	*dhurwaa*	*iyo dugaag*
time	FOC	*wáxa(a)*	together	hunted	lion	jackal	hyena	and animals
kale.	*Waxay*	*heleen*	*qaalin*	*baarqab*	*ah,*	*dhurwaa*	*ayaana*	
other	*wáxa(a)*+they	got	camel:foal	male:camel	be	hyena	FOC+and	
libaax	*u*	*xilsaaray*	*inuu*	*qaalinka*	*qaybiyo.*	*Waraabe*	*wax*	*ka*
lion	to	entrusted	that+he	camel:foal-the	divide	hyena	things	about
faan	*jecele*	*isagoo*	*xilkii*	*boqorku*	*saaray*	*la*	*dawaqsan,*	
boasting	liking	he+and	responsibility-the	king-the	put:on	about	confused,	
buu	*qaybtii*	*ku*	*dhawaaqay*	*oo*	*yiri:*	*"Qaalinka*	*bar*	*boqorka ayaa*
FOC+he	division-the	on	pronounced	and	said:	"camel-the	half	king-the FOC
leh,	*barka*	*kalana*	*aniga,*	*dawaco*	*iyo*	*dugaagga*	*kale*	*ayaa qaybsaneyna"*
has	half-the	other-and	I,	jackal	and	animals-the	other	FOC are:sharing
	Libaax	*gartaas*	*gurracan*	*uma*	*dulqaadan*	*karine,*	*inta*	*dharbaaxo*
	lion	verdict-that	crooked	with-not	patient:be	could:NEG	while	blow

il iyo goon fujiyay, buu isagoo indhihii cara la guduuteen oo
eye and cheek struck FOC+he he+and eyes-the anger with became:red and
calyeynaya dawo ku jeestay oo ku yiri: 'Adigu qaybi qaalinka."
salivating jackal to turned and to said: you divide:IMP camel-the
Wax ka fiira dheere, dawo wey garatay danta libaax
thing ADP sight long+be jackal DM+she understood aim-the lion
leeyahay; inaan masiibadii dhurwaa qabsatay oo kale ku dhicin
had that-not misfortune-the hyena got and other to happen:NEG
beyna goosatay. Dabadeedna inta libaax hor kadaloobsatay bey ku
FOC+she-and decided then-and while lion front squatted:down FOC+she to
tiri: "Qaalinka bar waad ku qadeyn, waaxi waa cashadaadii,
said: camel-the half DM+you on lunch quarter DM dinner-the-your
fallarna waad ku quraacan; inta soo hartana dugaagga
eighth-and DM+you on breakfast portion-the VEN remain-and animals-the
kale ayaa qaybsanaya."
other FOC share

    Kolkay qaybti dhammeysay bey laba tallaabo qaadday oo
    time-the+she division-the finished FOC+she two steps took and
qoladii ay qadoodiga badday dhex tu'atay. Libaax oo farax la
clan-the she hunger-the forced middle squatted lion and pleasure with
dhoola caddeynaya baa dabadeed yiri: 'Yaa ku baray, qaybta
front:teeth whitened FOC then said who you taught division-the
sidaan u wanaagsan."
way-this in good

    Seyte: "Daankii waraabe oo dunsanaa baa i baray."
so:she:said jaw-the hyena and collapsed FOC me taught
Sidaas baana dawo iyo dugaaggii kalaba daan dunsan, qadoodi
way-that FOC-and jackal and animals-the other-all jaw collapsed hunger
uga doorteen.
in+over chose

Once, Lion, Jackal, Hyena and some other animals went hunting together. They killed a young male camel and Lion entrusted Hyena to divide up their prey. Hyena liked boasting about things and although he was unsure about the responsibility that the king had placed on him, he announced the division, saying: "Half the camel belongs to the king and the other half will be shared by me, Jackal and the other animals."

    Lion lost his temper at this unfair decision and mauled the eye and cheekbone of the hyena with his claw. He turned to Jackal and, salivating and with his eyes turning red with anger, said: 'You divide the camel!' Jackal, who was very astute, understood Lion's intentions and resolved that the misfortune which Hyena suffered would not happen to her. So then, settling down in front of Lion, she said to him: "Half of the camel is for your lunch, a quarter is for your dinner, and an eighth is for your breakfast; the remaining portion can be shared by the other animals."

When she finished the division she took two steps back and sat down among the companions whom she had condemned to hunger. Lion, smiling with pleasure, then said: 'Who taught you such a good way of dividing it up?'

So she replied "The mangled jaw of the hyena taught it to me." In this way Jackal and all the other animals chose hunger rather than a mangled jaw.

Text 2: Traditional Story. *Guurdoon* 'Choosing a Bride' (Xaange 1988:49-50)

Waa   baa   waxaa   jirey   nin   iyo   afadiis,   waxayna
time   FOC   *wáxa(a)* was   man   and   wife-his   *wáxa(a)*+they+and
haysteen xoolo   ay   dhaqdaan.   Waggii   dambe baa   wiil   u
had   livestock   they   cared:for   time+the after   FOC   boy   to
dashay   reerkii,   ninkiina   aad   buu   ugu   rayreeyey   wiilka   u
was:born family-the man-the-and much FOC+he in+with   be:happy   boy-the to
dhashay. Sannoyin ka   dib   wiilkii   waa weynaadey oo   noqday nin
was:born years   from after boy-the DM big:grew   and became man
dheer, xoog   badan leh   oo geesi   ah, walidkiisna   jecel yihin.
tall   strength much possessing and brave:man be parents-his-and loving are
"Abbow, inaan   guursadaan   rabaa,   haatan waan   weynaadaye,"
father,   that+I   marry+FOC+I   want   now   DM+I   big:grew-and
wiilkii   baa   yiri berigii   dambe. "Waa yahay, mandhe;   haseyeeshee,
boy-the   FOC said time-the later   DM   is   son   however
marka   hore   i   soo   tus   inantaad   guursanaysid," buu   ku
time-the first me   VEN show girl-the+you   are:marrying   FOC+he to
jawaabay aabbehii. "Yeelay,   aabbe," buu   yiri   wiilkii.
replied   father-his   accepted   father   FOC+he   said boy-the.

Maalintii dambe baa wiilkii soo watay   gabar   dheer oo
day-the   next   FOC boy-the VEN came-with girl   tall   and
cas   oo   qurxoon oo yiri: "Waa tan, aabbe, inantaan
copper-coloured and beautiful and said: DM   this father girl-the+I
rabaa inaan guursado." "Waa yahay, maandhee; bal   ii   soo   qaad
want that-I marry   DM is   son   now me+to VEN carry
dhagaxaas weyne   halkaaas yaal," buu   yiri   aabbehii.
rock-that   big-and   place-that is   FOC+he said   father-his
Dhagaxu aad   buu   cuslaa,   wiilkiina   si   kasta yeel
rock-the much FOC+it heavy-was boy-the-and way   any   did
oo qaadi kari waa. "Naa ninkaas dhagaxa yar   qaadi kari
and lift   can   fail   girl man-that rock-the small lift   be:able
waayey miyaad   guursanaysaa?" aabbehii baa   weydiiyey   inantii.
failed   QM+you   are:marrying   father-his FOC asked   girl-the

"Haba yaraatee maya," bay tiri oo iska tagtay.
OM-yet small:INCH:POT no FOC+she said and self+from went

Afar gabdhood oo kale buu wiilkii hor keenay aabehiis,
four girls and other FOC+he boy-the before brought father-his

isaguna dhagaxii in uu soo qaddo buu sharuud kaga dhigay
he+and stone-the that he VEN lift FOC+he condition on+as set

wiilkiisa, ka hor inta uusan gabadhaas middoodna ka oggolaan
son-his ADP before case-the he+not girls-those one-their-and from accept

in uu guursado.Wiilkii si kasta yeel oo qaadi kari waa dhagaxii
that he marry boy-the way any did and lift be:able fail stone-the

cuslaa, gabdhihii sidaas awgeed bay u wada diideen in ay
heavy girls-the way-that reason-its FOC+they for together refused that they

xilo u noqdaan ninkaas tabarta liita oo waxmataraha ah.
wife to become man-that strength-the weak:be and uselessness be

Gabadhii shanaad oo aan u qurux badnayd siddii afartii hore
girl-the fifth and not in beauty much:was way-the four-the before

baa wiilkii u keenay adoogiis. Markii gabadhu aragtay in wiilku
FOC boy-the to brought father-his. time-the girl-the saw that boy-the

qaadi kari waayey dhagaxii weynaa bay la qaadday oo soo agdhigeen
lift be:able failed stone-the big FOC+she with took and VEN near-put

odayga cagihiisiii, isaguna eegayo sida ay isu kaashadeen
father-the feet-his him-and watching way-the they each:other-ADP helped

wiilka iyo gabadhu.
boy-the and girl-the.

"Waa taas gabadha u qalanta inaad guursato, maandhow; gabdhihii
DM that:one girl-the for fit:be that+you marry son:VOC girls-the

kale kulama qaadin dhagaxa, tanise way kula qaaday.
other you-with-not lifted stone-the this:one-but DM+she you-with lifted

Iskaashigu waa gundhigga nolosha reerka." buu yidhi aabbehii
mutual:help-the DM foundation-the life-the family-the FOC+he said father-the

oo isu dhisay wiilkii iyo inantii.
and each:other:to gave boy-the and girl-the

Once there was a man and his wife and they raised livestock for a living. Some time later a boy was born to the family, and the man was very happy with the child born to him. Some years later the boy grew up and became a tall, strong and brave man, much loved by his parents. "Father, I want to get married now that I am grown,' the boy said to his father one day. "All right, son; however first you must show me the girl you are going to marry," his father answered him. "Agreed, father," said the son.

The next day the boy brought along a tall, beautiful, copper-coloured girl and said: "Here she is, father, the girl I want to marry.' "Fine, son; now just carry over to me that big rock over there," said his father. The rock was very heavy, and try as he might, the son could not lift it. "Now girl, are you going to marry this man who can't lift this little rock?" his father asked the girl. "Certainly not," she replied and went off.

The boy brought four other girls to his father and each time he set as a condition on his son that he had to lift the rock, before he would allow him to marry any one of the girls. However hard he tried the son could not lift the rock and, as a result, all of the girls refused to marry such a weak and useless man.

The boy brought to his father a fifth girl who was not as beautiful as the first four. When the girl saw that the boy wasn't able to lift the big stone, she lifted it with him and they put it at the father's feet, while he watched the boy and girl help each other.

"This one is the right girl for you to marry, my son; the other girls didn't help you lift the stone but she did. Mutual support is the foundation of family life," said the father and he gave his blessing to the young man and woman.

Text 3. Extracts from a letter written in 1903 from Sayid Maxamed Cabdulle Xasan to the British authorities. From Ciise (1976: 84-5).

Sanad baynu   haatan dagaallameyney. Inaan dalkayga   u   taliyo   oo
year   FOC+we(INCL) now   are:fighting   that+I country-my over govern and
dhiintayda   daafaco baan   doonayaa …
religion-my defend   FOC+I want
Anigu qalcado ma lihi,   guryana   ma lihi,   dalna   ma lihi.
I   fortresses not own:NEG houses-and not own:NEG country-and not own:NEG.
Sancaaniin ma lihi…   Haddii dalku   beero falan,   guryo iyo adduun
craftsmen   not own:NEG if   land-the farms ploughed houses and wealth
badan lahaan lahaa dan   bay   kuu   ahaan lahayd inaad   u dagaallantaa.
great have   would profit FOC+it you+for be   would that+you for fight
Dalku   waa wada duur,   faaiidana adiga kuuma   laha.   Hadaad
country-the DM all   wilderness use-and you   you+for-not has:NEG if+you
dhir   iyo dhagax dooneyso badi   baad   helaysaa. Dundumooyin
trees   and stone   want   plenty FOC+you find:PROG ant:hills
badanna   waa leeyahay. Qorraxdu aad   bay   u   kulushahay. Waxaa
many-and DM   has   sun-the   very FOC+it ADP   hot-is   thing-the
kaliya oo   aad iga   heli kartaa waa dagaal, wax   kale   ma leh.
only   and you me+from find can   DM   war   thing else   not have
Raggaagii dagaal baan   kula   kulmay waana   laayey. Taasna aad
men-your fighting FOC+I in+with met   DM-and   killed   that-and much
baan   ugu   faraxsanahay. Raggaygii dagaalka lagu   laayay
FOC+I   in+about happy-am   men-my   battle-the one-in killed
jannay   tageen. Eebbe ayaa noo   dagaallamaya. Annagu waa ku
paradise+FOC+they went   God   FOC us-with is:fighting   we   DM you
dili,   adna   waad   na dili.   Annagu amar Eebbaan   ku
kill:INF you-and DM+you us kill:INF we   order God+FOC+we by

dagaallamaynaa.   Runtii   waa taas.   Naxariista   Eebbaannu   dalbaynaa.
are:fighting   truth-the DM   that   blessing-the   God+FOC+we ask:for
Eebbaa   ila   jira   markaan   tan   kuu   soo   qoraayo.
God+FOC me-with is   time-the+I   this   you+to VEN write
Haddaad dagaal   dooneyso   waan ku   faraxsanahay;   haddaad   nabad
if+you   war   want   DM+I with   happy-am   if+you   peace
dooneysana raalli   baan ka   ahay; laakiin haddaad   nabad rabtid
want-and   satisfaction FOC+I with am   but   if+you   peace want
dalkayga   ka   tag   oo   kaagii   tag. Haddaad dagaal rabto halkaaga
country-my from go   and   yours   go   if+you   war   want place-your
joog.   Hadalkayga   maqal.
stay:in:IMP   speech-my   listen:to:IMP

We have been fighting each other for a year now. I want to govern my own country and to defend my religion... I own no fortresses, no houses and no land. I have no craftsmen... If the country had ploughed fields, houses and great wealth, there would be some point in your fighting for it. The land is all wilderness and has no use for you. If you want trees and stone you will find them in plenty. It has many anthills. The sun is very hot. The only thing you can get from me is war, nothing else. I have met your men in battle and killed them. I am greatly pleased by that. My men who have been killed in battle have gone to paradise. God is fighting for us. We will kill you, and you will kill us. We are fighting on God's command. That is the truth. We ask for God's blessing. God is with me as I write this to you. If you want war I am happy with that; and if you want peace that pleases me; but if you desire peace, go from my country and return to your own. If you want war, stay where you are. Listen to my words.

Text 4. Extract from a school text *Taariikhdii Soomaaliya ee Hore ee Tii Waayadii Dheexe* 'The History of Early and Medieval Somalia.' From the source [T4.64-5] in Appendix I.

Taariikhdii Soomaaliya ee   berihii   hore waa mid xiiso   u leh. Tan iyo
history-the Somalia   and times-the early DM one   interest in has. this and
muddo aad u fog   baa dhulka   waxa   ku noolaa aabbayaashii
period much in distant FOC land-the *wáxa(a)* in live   fathers-the
Soomaaliyeed. Waxay   jirtey   in Soomaalidii hore ay   xoolo   dhaqan
Somalis:GEN   *wáxa(a)*+they existed that Somalis-the early they livestock tend
jireen, oo ay cuni jireen hilibka xoolaha,   isla   markana   cabbi
used:to and they eat used meat-the livestock-the self-with time-the-and drink
jireen caanaha xoolohooda.   Iyadoo   berigii   hore uusan jirin   dharku,
used:to milk-the livestock-their they+and time-the early it+not be:NEG cloth-the

Soomaalidu waxay   xidhan jireen hargaha adhigooda.   Marka
Somalis-the *wáxa(a)*+they wear   used:to skins-the sheep:and:goats-the time-the
naftooda iyo   adduunkooda   la   soo weeraro waxay   Soomaalidii
life-their and   belongings-their one VEN attacked *wáxa(a)*+they Somalis-the
hore iyagoo   isku   duuban ku   dagaallami jireen hootooyin iyo
early they+and self-with unite   with fight   used:to spears   and
gashaamo. Gaadiidka   u   dheereeya ee   Soomaalidu lahayd wuxuu   ahaa
shields   transport-the for go:quickly and Somali-the had   *wáxa(a)*+it was
fardaha.
horses-the.

    Dalka   Soomaalidu wuxuu   ku yaalla geesta qorrax ka   soo
    country-the Somalis-the *wáxa(a)*+it on is   side-the sun   from VEN
baxa   Afrika ugu   baalleysa. Bad baa dhinac kasta kaga   wareegsan
comes:out Africa most be:beside sea FOC side   each around circles
galbeedmooye. Waqooyiga waxa   ka xiga   Gacanka Cadmeed, barina
west   excepted north-the   *wáxa(a)* to be:next gulf-the Aden:GEN east-and
BadweyntaHindiya, galbeedna Itoobiya. Dadka   Soomaalidu wuxuu   ku kala
ocean-the India   west-its-and Ethiopia. people-the Somalis-the *wáxa(a)*+it in apart
faafsan yahay dhul lagu   qiyaasay hal malyuun oo   mayl oo   laba jibbaaran.
spread is   land one-in estimated one million and mile and two doubled.
Macayo kala duwan   baa shisheeyuhu u   yiqiin dhulka   Soomaalida.
names   apart different FOC foreigners-the for know land-the Somalis-the
Masaaridii   hore oo   ay u talin jireen   boqorradii lagu magacaabi jirey
Egyptians-the early and they by ruled used:to kings-the one+as name used:to
Fircoonno, waxay   u   yiqiinneen, 'Bunt' ama 'dhulkii   Udgoonka',
Pharaohs *wáxa(a)*+they as knew   Punt or   land-the   perfume-the
'Beeyada   iyo foolka-marodiga'. Waxa   kale oo ay u yiqiinneen
incense-the and tusk-the-elephant-the thing-the other and they as   knew
'Tiraa' oo   ah dhulkii eebbeyaasha ama dhulkii daahirka ahaa. Giriggii
 Tiraa   and be and-the gods-the   or land-the sacredness was Greeks-the
iyo Roomaaniyiintii hore   waxay   ku magacaabi jireen 'Barbaryihiin'.
and Romans-the   ancient *wáxa(a)*+they as called   used:to Barbarians
Carbtii hore ee   dalmarka ahayd magacaas hore ayey   in   yar
Arabs early and explorers were   name-that early FOC+they amount small
doodiyeen   oo ka dhigeen 'Dhulkii Baarber' ama 'Ber-el-Cajam'.
deviated:from and as put   land-the Berbers   or   *Ber-el-'Ajam*
Magacaasna   haddeer waxa   loo   yaqaan Gobolka Boosaaso;
name-that-and time-that *wáxa(a)* one+as know   region-the Bosaso
wuxuuna   ka   tusinayaa xiriirkii ay Carabtu la   lahayd dadka
wáxa(a)+it-and about shows   links-the they Arabs-the with had   people-the
Soomaaliyeed. Reer Yurubkii   Afrika qabsaday waxay   dhulka
Somalis:GEN   people Europe-the Africa seized   *wáxa(a)*+they land-the
Soomaaliyeed ku magacaabeen 'Geeska Afrikada Bari', magacaas oo tan
Somalis:GEN as   named   horn-the Africa-the east   name-that and this

iyo   haddeer   loo   yaqaan
and   time-that   one+as   knows.

The history of ancient Somalia is one of interest. The ancestors of the Somalis have lived in this land since distant times. The ancient Somalis were pastoralists and lived by eating the meat and drinking the milk of their livestock. Since at that time there was no cloth, Somalis used to wear the skins of their sheep and goats. Whenever their lives or herds were threatened the early Somalis united amongst themselves and defended themselves with spears and shields. They rode horses for speed.

   The country of the Somalis is on the easternmost side of Africa. Sea surrounds each side except the west. To the north lies the Gulf of Aden, to the east the Indian Ocean, and to the west, Ethiopia. The Somali people are spread over an area estimated as one million square miles. Foreigners have known their country by different names. The ancient Egyptians, ruled by their Pharaoh kings, knew it as 'Punt', 'The Land of Perfume', or '(The Land of) Incense and Ivory'. They also knew it as *Tiraa*, the 'The Land of Gods' or 'The Sacred Land'. The ancient Greeks and Romans called it '(The Land of) Barbarians'. The old Arab explorers changed that name a little and turned it into 'The Land of the Berbers' or *Ber-el-Cajam*. The region of Bosaso was known by that name in early times, revealing the links that existed between the Arabs and Somalis. The Europeans who colonized Africa described the Somali lands as 'The Horn of East Africa' and they have been known by that name ever since.

Text 5. The poem *Alleyl Dumay* 'Night fell' by Raage Ugaas, published in the source [S1.8] in Appendix I.

Raage Ugaas was a nineteenth century poet and pastoralist. This poem laments the poet's broken betrothal to a woman called Cabban Cilmi Hagoog and has the metrical form of the *gabay*, the alliterative poetical form used for the most serious subjects. This poem shows vowel alliteration.[108]

1.   *Alleyl   dumay,   albaabadoo   xidhan uunku   wada   seexday*
     night   fell   doors-the+and   closed living:creatures together   slept
2.   *Onkod yeedhay   uugaamo   roob   alif   banaadiiq   ah*
     thunder called   drumming   rain   thousand   rifles   be
3.   *Iihdaydabixi   baa   libaax iman la   moodaaye*
     lament-my-giving   FOC   lion   come one   mistook
4.   *Raggase   adhaxda   iyo   ooftu   waa   udub   dhexaadkiiye*
     men-the-but spine   and   ribs   DM   post   centre-its
5.   *Labadii   wax   laga   eegi   jirey   waan   ka   awdnahaye*
     two-the   thing one+by see   used:to DM+I   from   blocked:am
6.   *Halkaan   aa   ka   leeyahay Ilaah keliya   uun   baa   og*
     place-the+I complaint from   have   God   only   just   FOC   aware

7. *Aboodigu ma lalo garab hadduu iin ku leeyahaye*
   vulture-the not flies wing if+it damage in has
8. *Orod uma hollado ooglihii adhaxda beelaaye*
   running in-not tries horse-the spine-the loses
9. *Ma aarsado il iyo oof ninkii iimi kaga taale*
   not take:revenge eye and ribs man-the damage in+from be
10. *Aroos uma galbado nimuu wadnaha arami jiifaaye*
    wedding in-not go:home:married man-he heart-the wound sleeps
11. *Geeluba kolkuu oomo waa olol badnaadaaye*
    camels-the-all when-they thirst DM roars many:are
12. *Sidii inan yar oo hooyadeed aakhiro u hoyatey*
    way-the girl small and mother-her afterlife to went
13. *Oo aabaheed aqal mid kale meel illin ah seexshey*
    and father-her house one other place entrance be sleep:CAUS
14. *Hadba waxaan la urugoonayaa uurkutaallada eh*
    time-every *waxa(a)*+I with worry anxieties-the be
15. *Ninkii ooridiisii rag kale loo igdhaan ahaye*
    man-the wife-his man other one+to married:CAUS+I am
16. *Ninka ila biyo leh soo arkoo ooman baan ahaye*
    man-the springs water have VEN see+and thirst FOC+I am
17. *Nin ugaas walaalkiis yahoo eeday baan ahaye*
    man chief brother-his is+and turned:against FOC+I am
18. *Afdhabaandhow aayar ninkaa aamusaan ahaye*
    mouth-gagged slowly man-that silent+FOC am

1. Night fell, doors were closed and all creation slept,
2. Thunder roared and the drumming of the rain was a thousand rifles,
3. My howls of lament were mistaken for an approaching lion.
4. The spine and ribs are the central pillar of a man,
5. Closed are my eyes which used to see.
6. The source of my sorrow only God himself knows.
7. The vulture does not fly with a damaged wing,
8. A broken-backed horse will not try to run,
9. A man damaged in eye and rib cannot take revenge,
10. A man with a wounded heart cannot make a bridegroom,
11. Loud are the cries of camels in thirst.
12. Like a small girl whose mother passed away,
13. And whose father leaves her to sleep in the doorway of another woman's house,
14. Time and again I wear myself with grief.
15. I am a man whose wife has been given to another man,
16. I am a man who has seen springs full of water and must thirst,
17. I am a man whose brother is a king who has turned against him,
18. I am a man who is gagged and now must be silent.

Text 6. Extract from the poem *Gammaan waa magac guud* 'Horse is a general term' by Cali Bucul, published in the source [S2.32-3] in Appendix I.

This poem by the nineteenth century poet Cali Bucul, is in praise of his horse *Guulside* ('Victory-Bearer'). It is a *geerar*, a form with shorter lines and length than the *gabay*. Traditionally associated with war and challenges, *geerar* poems were originally recited on horseback. This particular poem alliterates in *g*, echoing the horse's name.

1.   *Guulsidow   faraskayga*
     Guulside-VOC   horse-my
2.   *Haddii   aan   gallaaddiisiyo*
     if   I   goodness:his-and
3.   *Gedihiisa   tilmaamo*
     character-his   describe
4.   *Waa   gal   maaxanaya   oo*
     DM   pool   refills:MID   and
5.   *Gaadhimayno   xogtiisa   oo*
     we:reach:NEG   secret-his   and
6.   *Waan   ka   gaabsanayaaye*
     DM+I   from   fall:short
7.   *Rag   allow   gef   miyaa?*
     Men   whatever:VOC   fault   QM
8.   *Almis   gooddiyadeediyo*
     Almis   slopes-its-and
9.   *Galihii   harawaadiyo*
     gorge-the   Haraw:GEN:and
10. *Gureys   beelaha   yaalla*
     Gureys   camps-the   be
11. *Galab   buu   isku   maraaye*
     evening   FOC+he   self-in   passes
12. *Ma   daruur   gudgudaa?*
     QM   cloud   night:rain+FOC
13. *Gabangoobi   abaar   ah   oo*
     desert:plain   drought   be   and
14. *Dadku   guunyo   ka   quustay*
     people-the   livestock   from   despair
15. *Geel   buu   soo   tu'iyaayee*
     camels   FOC+he   VEN   drives
16. *Ma   gob   reera   dhacdaa?*
     QM   nobleman   camps   raids+FOC
17. *Galuhuu   miranaayo*
     field-the+he   grazes:at:night
18. *Gurxan   dhawrta   fogeeyiyo*
     whinnying   wild:animals-the   makes:distant

19. *Gooh   weyn   baa   ka   galoosee*
    neigh   great   FOC   from   leaves+and
20. *Ma   aar   goosan   wataa?*
    QM   male:lion   pride   leads+FOC
21. *galuhuu   ku   xidhnaana*
    field-the+he   in   tethered:is
22. *Nacab   soo   gardarraystiyo*
    enemy   VEN   wrong-doing-and
23. *Gaas   laxaadliyo   guuto*
    Warrior:troop strength-having-and   raiding:band
24. *Guluf   weerar   ku   qaadiyo*
    Horsemen   raid   on   heading
25. *Gaade   goonya   wareegiyo*
    Gangs   corners   roaming
26. *Guhaad wuu   ka   xijaabay oo*
    anger   DM+he   from   protects and
27. *Rugtiisa   gaajo   ma   geydo*
    home-his   hunger   not   reaches:NEG
28. *Saanaa   loo   gama'yaa*
    security+FOC   one+in   sleeps
29. *Wedka   uunka   gumaadee*
    death-the creatures-the   destroys
30. *Laga   giigay   miyaa?*
    one+from   recoils   QM+FOC
31. *Gudcur roob   gibil   saaray'oo*
    darkness rain   carpet   puts:on-and
32. *Gufaacadiisiyo   caydhku*
    howling-its-and   gale-the
33. *Nafleydii   gilgileen   oo*
    living:things-the   shake   and
34. *Guuxu   maanka   gantoobay*
    uproar-the   mind-the   divides
35. *Tuuga   soo   gabbanaayuu*
    thief-the   VEN   skulking+FOC+he
36. *Dabayshiisa   gartaayee*
    scent-his   recognises-and
37. *Ma   kuhaan   garabdaar   leh   oo*
    QM   soothsayer   intuition   has   and
38. *Faalka   guun   ku   noqdaa?*
    divination:by:beads-the hidden:thing in   returns+FOC
39. *Korkiisoo   aan   gaboobin*
    body-his-and   not   grown:old:NEG
40. *Halgaraadka guntiisiyo*
    neck-the   base-its-and

41. *Guudki   buu   ka   caddaadaye*
     mane-his   FOC+he   on   grows:white
42. *Ma   galool   ubax   laa?*
     QM   acacia:tree   flowers   has+FOC

1.   O my horse *Guulside*!
2.   If his goodness
3.   And his nature I try to describe,
4.   He is a pool that refills itself
5.   And I cannot plumb his secret,
6.   I fall short.
7.   Tell me, people, am I at fault?
8.   The slopes of Mount Almis
9.   And the Haraw ravine,
10. The camps on Mount Gureys,
11. He passes them in one evening.
12. Is he the cloud that brings rain in the night?
13. From the drought-bound desert,
14. When people despair of their flocks,
15. He drives the camels home.
16. Is he a noble warrior raiding enemy camps?
17. Grazing his field at night,
18. His whinnying keeps wild animals at bay,
19. A great neighing cry leaps from him.
20. Is he a male lion leading his pride?
21. Tethered in his field,
22. A threatening enemy and
23. Strong warrior troops and raiding bands,
24. Plundering horsemen,
25. Roaming gangs,
26. His anger protects from them all,
27. And hunger does not reach his home.
28. People sleep in safety.
29. Is he not the death which destroys all creatures
30. And from which we recoil?
31. In a dark night carpeted in rain,
32. When the roaring of the gale
33. Makes all living things tremble,
34. And the uproar divides their wits,
35. The skulking thief
36. His scent he still catches.
37. Is he an inspired soothsayer
38. Capturing the unseen by divination?
39. Although his body is not grown old,

40.   The base of his neck
41.   And his mane have grown white.
42.   Is he an acacia tree in bloom?

Notes

1. See Appendix II for a brief note on poetry and selected poems. For accessible introductions to Somali culture and poetry, see Loughran et al. (eds. 1986) and Andrzejewski and Andrzejewski (1993).

2. For characteristic accounts, very different in kind, see Burton (1894), and Laurence (1954, 1963).

3. Zorc and Osman (1993), for example, list around seventeen hundred Arabic loans in their twenty six thousand-entry Somali-English dictionary.

4. There are some relatively minor regional variations within the Northern dialect, or Standard Somali, and this study is largely based on the speech of speakers from the north-western part of its range, near the towns of Hargeysa and Burco (Burao). It is clear from published materials that the Standard Somali used in Mogadishu during the Somali Democratic Republic (i.e. until 1990) was influenced by the local Benadir dialects. It seems likely that without the political upheaval a Northern-Benadir *koine* might have formed the future standard variety, at least in the south.

5. See Lamberti (1986a, 1986b) for a proposed classification of these dialects.

6. For description of Benadir and other southern dialects, see Moreno (1955). May dialects are described in von Tiling (1922, 1925) and Saeed (1982, 1992).

7. Note though that Orwin (1994) is an unpublished PhD dissertation on aspects of Somali phonology.

8. The use of accents to mark tone is discussed in detail in section 2.4. Briefly, three tones are marked: an acute accent marks High tone (*á*); no mark is used for Low tone (*a*) and a grave accent marks Falling tone (*à*). Tones on long vowels and diphthongs are only marked on the first vowel, e.g. *áa, àa*.

9. The last correspondence between northern *dhdh* and southern *dh* suggests that southern dialects represent a more extended version of the northern rule which weakens the retroflex plosive *dh* sound to a corresponding flap intervocalically. We may posit a process where both intervocalic and final *dh* became a flap which then collapsed with

the alveolar trill *r*. Thus in positions where northern dialects have a double intervocalic consonant *dhdh*, the southern weakening process results in single *dh* rather than a flap and subsequent *r*. Note that in writing, speakers of the northern variety usually simplify a double *dhdh* to *dh*.

10.   There are a few exceptions to this generalisation in loans from Arabic, for example *fekrád* N F 'idea', *maktabád* N F 'library', *bikrád* N F 'virgin'. Syllable final *k* sometimes also occurs in loans from other languages, e.g. *mikroskòob* N M   'microscope' from English, though this is often pronounced *mikoroskòob* to avoid this.

11.   Note that Orwin (1996: 61-5) discusses some complications in the phonological weight of diphthongs. In polysyllabic words with a final open diphthong, for example *árdey* (N M) 'student', the final diphthong is short for purposes of accent assignment though it may be counted as long or short in the metrical system. Closed syllable diphthongs in mono- and poly-syllabic words are always long. Note that the long diphthongs *aay, aaw*, etc. only occur in open syllables.

12.   See Johnson (1979, 1984, 1996), Guuleed (1980), Andrzejewski (1982), Banti & Giannattasio (1996), and Orwin & Riiraash (1997) for discussions of poetic scansion; and Andrzejewski and Lewis (1964) and Andrzejewski and Andrzejewski (1993) for anthologies of poems translated into English. Andrzejewski (1985) is an overview of Somali literature.

13.   Although sometimes a slight rising tone may be heard on these syllables.

14.   This statement, which seems to have been true, may now need qualification. Younger Somali speakers, perhaps under the influence of foreign languages, seem to happily produce questions by intonation alone, i.e. turning the negative declarative *Má yeelín*, 'You didn't accept it' into a question by giving it rising intonation, *Má yeelín?* 'Didn't you accept it?', unlike the more traditional strategy employing an interrogative sentence type marker, e.g. *Miyàanád yeelín*? 'Didn't you accept it?' See chapter 4 for discussion of sentence type markers and chapter 7 for negative questions.

15.   The choice of *k* and *t* as the base form of the initial consonant is suggested by the form of the sandhi rules described in this section, but confirmed by their presence in independent non-suffixed forms where *ka/ta*, *kíi/tii* are used pronominally to mean 'the one' e.g. *kíi kalé* 'the other one (M)', *tíi kalé* 'the other one (F)'

16.   With the one notable exception of the word *gèel* 'camel, camels' which deletes the initial *k*, i.e. *gèela* 'the camel, camels'.

17.   As mentioned earlier, *dhdh* is usually simplified to *dh* in print; so this nominal would appear as *gabadha* 'the girl'.

18.   Noun stems under discussion include noun roots like *áabbe* 'father', roots with plural *-o* suffixes like *maalmo* 'days' (cf. *maalín* 'day') and verbal roots nominalised by an agentive suffix like *báre* 'teacher' (cf. *bár* V 'teach').

19.   Which here undergo the same sandhi rules as other determiner suffixes, i.e. *-kay* → *-hay* and *-tay* → *-day*).

20. One of the problems of Somali phonology is how to explain the various forms of these lexical affixes. This causative affix, for example, surfaces in most verbal forms as *–i*. In some derived nouns it is *-is* (e.g. *jóoji* V CAUS 'stop', *jóojis* N M 'act of stopping something'). When occurring between a verb root and a middle affix it occurs as *-s-*, for example *joogsaday* '(I) stopped myself'. This has the morphological structure: *joog* (root, 'stop') + *s* (causative affix) + *at* (middle affix) + *ay* (past simple inflection). One explanation which is based on historical and comparative evidence is that the affix was previously *-*ish*, which then developed into *-i(y)* and *-is* and *-s* forms in different environments. See Sasse (1979) and Lamberti (1986a) for detailed discussion.

21. Once again this is an affix which has different forms: a basic form *-at-* (usually surfacing as *-ad-* or *-t-*, as will be described here), *-ó* in imperative singular forms, e.g. *joogsó* 'Stop yourself! Halt!', and *-a-* in the infinitive, e.g. *joogsán* 'to stop oneself', which has the morphological structure: *joog* (root, 'stop') + *s* (causative affix) + *a* (middle affix) + *n* (infinitive inflection).

22. As here, a noun phrase focused by one of the focus words will be written in small capitals in the translation.

23. One important exception is the noun phrase *haddíi* 'the occasion, time', which as described in 8.3.6, introduces conditional clauses. Subject verbal pronouns regularly coalesce with this nominal, for example *haddàad tagtó* 'if you go', although the uncoalesced form is also possible: *haddii aad tagtó*.

24. See 7.4.1 for a discussion of the unusual tone pattern on *báan* here.

25. The missing forms are because, as described in 4.1.1.4, there are several forms of vocatives. The vocative marked by AP change alone is normally used with names, while vocative forms with suffixes are used for both common nouns and names. Thus the predicted vocative forms by AP change alone, *díbi*, *bísad* and *cárruur* would normally not occur, being replaced by the suffixed forms *díbiyohow* (or *díbiyòw), bisádyahay* (or *bisádèey*) and *carrúuryahay* (or *carrúurèey*). Note though that the use of the different vocatives reflects a speaker's attitude and might vary for example between real life situations and when animals address each other in fables.

26. The changes from Saeed (1993a) arise from following Banti's (1988b) analysis of the lowering and shifting of High tones. The consequences include recognising that most plural nouns and the irregular nouns of declension 2B have AP1 inherently, which becomes AP3 in some contexts, rather than the other way around; and identifying declension 6 and 7 nouns as inherently AP1 in the singular, becoming AP2 in some contexts.

27. Note though that it is possible for the name suffixes to occur with Low rather than Falling tone, e.g. *Xásanow* rather than *Xásanòw*. This seems to be an optional pronunciation.

28. Adpositions, as described in 4.5, 6.1, and 6.3, also occur in the verbal group.

29. Two points are perhaps relevant here: firstly some stative verbs do occur in imperatives, e.g. *yahay* 'be', described in 4.3.6. Secondly, stative root verbs may become dynamic when lexical affixes are attached. Thus the stative verb *húb* 'be certain', has a derived

causative verb *húbi* 'make certain, ascertain' and a middle voice verb *hubsó* 'make certain for one self, be certain of' both of which are dynamic and may occur in imperatives.

30. Many writers have commented on the egalitarian nature of traditional Somali society, for example Lewis (1961). There has been little written on sociolinguistic topics like politeness, but it seems clear that the ready use of imperatives, along with the lack of polite pronouns, honorifics, etc. reflect the lack of social hierarchy felt amongst Somalis.

31. Though some speakers do form a negative auxiliary construction paralleling the positive forms. This consists of the infinitive followed by an invariable negative form of the auxiliary *léh*: e.g. *má súgi laháyn* 'I/you/he etc. would not wait for it'.

32. Distinguished from the negative imperative sentence type marker *ha* by its accentual pattern.

33. See for example the contrast between sentences (90) and (91) below.

34. See 8.1 for further details and examples.

35. This adjective also occurs as *feejíg* and *foojíg*, while some speakers use the associated derived adjectives *feeyigán* (*feeyignaa*), *feejigán* (*feejignaa*), *foojigán* (*foojignaa*).

36. These languages form, with Somali, the Omo-Tana subgroup of Eastern Cushitic as shown in chapter 1.

37. But note that Central Somali dialects, which are quite distinct from the Standard Somali varieties described here, do have as described in Saeed (1982: 89-90) a singulative suffix *ko/to* which belongs to this set of determiners, e.g. *náŋ* 'man' *náŋko* 'a (single) man'.

38. We assume that the accentual patterns (APs) on the possessives in (125) are the category's inherent APs. The absolutive and subject case markings are the phrasal exponents of nominal case marking. We will assume that possessives, unlike nouns, do not have as their default an absolutive case form because they are not nouns and therefore do not have inherent case.

39. There is variation in the length of the final vowel of *wáxa/wáxaa* that has not yet been satisfactorily explained. It seems in part to be regional since in north-eastern varieties the short final vowel version often occurs while in southern varieties, as reported by Svolacchia and Puglielli (1995), it is rare.

40. The difference between the two forms of negative question are discussed in chapter 7, where we also discuss the morphosyntactic structure of sentences like *Miyàanú sugín?* 'Didn't he wait for it?' For now we treat *miyàanú* as a fusion of *ma* (QM) + *áan* ('not') + *uu* 'he'.

41. The conjunctions *oo* and *ee* are also used to mark distinctions between types of relative clauses, as discussed in chapter 8. Note that this distinction may not occur in this way in all dialects: the use of conjunctions has not been compared across dialects.

42. This is from a *jiifto* by the poet and Dervish leader Maxamed Cabdille Xasan.

43. As will be clear from the examples in this section, the meaning of these verbs is lexically weakened or generalised in ideophone constructions. The verb *yidhi* 'say', for example, takes as its object not only ideophones describing sounds but also those describing light and movement. As mentioned by Dhoore and Tosco (1998) the use of verbs meaning 'say' with ideophones is an areal feature of the Horn of Africa. Similarly *sii* 'give' is weakened to a general causative or factitive meaning. Compare this to the grammaticalisation of this verb as a causative affix, discussed in 5.1.5 below.

44. Many of Dhoore and Tosco's (1998) examples in this category are of actions involved with ingestion, for example *díg* used to describe eating violently, or wolfing something down; *daláq* for swallowing something small, and *julúq* for swallowing without chewing. These might possibly be seen as subtypes of their first category, movement.

45. Some speakers pronounce the imperative allomorph as *aw* rather than *ow*, e.g. *adkàw* 'become hard'. This pronunciation makes these morphological conditioned exponents suspiciously like parts of the stem of *yahay* 'be', shown in bold in the following; for example the imperative ah**aw**!, the infinitive ah**aan**, and the forms found in optative, negative conditional, and optative paradigms, e.g. the potential ah**aad**-ee /ah**aat**-ee 'perhaps he/she is'. We have already seen in 4.4 the similarity between forms of *yahay* beginning with *ah-* and inflection on adjectival stems; it is possible that this distribution of inchoative exponents is part of the same overlap.

46. In some varieties of Standard Somali this affix occurs as *oo* in the imperative singular, e.g. *dhaxamòo* 'feel cold', *gaajòo* 'feel hungry', etc.

47. The alternation *a~e* in these forms, *ayn/eyn*, *ays/eys*, seems at least in part the result of contextual assimilation rules but the details of the rule seem subject to regional variation within Standard Somali in ways that have not been described. We simply list the alternant forms here.

48. The verb *furmay* in (b) is from **furamay* by the rule of Stem Contraction (3.1.4).

49. I am grateful to the late B.W. Andrzejewski for this form of the prayer. The suffix *àay*, glossed EXCL, is a marker of exclamations.

50. We mentioned in chapter 4 the difficulty in deciding on a reference form for this affix, whose principal forms are *is, i,* and *s*. On historical grounds, and for the best account of word derivation, as we shall see in the section on derived nouns (5.3.1), the form *is* seems more basic and we use this as the reference form here. It is, however, difficult to find good synchronic phonological motivation for all of the occurrences of the *i* and *s* variants. We leave this problem to future phonological studies.

51. We assume that *hindhisíisáy* is phonologically derived from **hindhissíisáy* because of the phonotactic constraint against geminate fricatives.

52. The verb *qadèe* 'eat lunch' is formed from the noun *qádo* F 'lunch' by the affix *ays* described later in this section.

53. A perhaps obvious but important point is that the existence of this affix means that its absence from a verb also communicates meaning. Thus we can compare the meanings of the following root and middle verbs:

(a)   *Baabùur   má   wat-aa?*
      car   QM   drive-2SG
      'Do you drive a car?' i.e. 'Do you drive for a living?'
(b)   *Baabùur   má   wad-at-aa?*
      car   QM   drive-MID-2SG
      'Do you drive a car for yourself?' i.e. 'Do you have a car?'

54.   Other colour terms are nouns which do not form adjectives, for example *bèey* N M 'ash greyness (of livestock)', *ásfar* N M 'yellow gold', *hurúud* N F 'saffron, yellow'.

55.   Puglielli (1984) proposes the following derivations for these forms: *daajis+n →daajín* and *beerat+lo →beerashó*, with the underlying forms exhibiting the basic forms of the causative affix *is* and the middle affix *at*. Presumably this approach would also include a rule *dhammays+n →dhammáyn* for the factitive affix *ays*.

56.   One of the few papers devoted to compounds is Puglielli (In press).

57.   Unlike for example the compound adjective *tirabadán* 'numerous, many' formed from *tiró* N F 'number, quantity' and *badán* ADJ 'much many'. See 5.4.3. for some other compound adjectives.

58.   And other adjectives participate in this incorporation process; see for example:

Odaygu   wuu   dumar   jeclaa
old:man-the   DM+he   women   liking+was
'The old man was women-liking', 'The old man was a womaniser' [SNR 83.5]

59.   Depending on the analysis of *le*, discussed in 5.3.3.2, it is possible to view this example as a compound composed of two nominals: *baqó* N F 'retreat' and the derived noun *beenlé* N M 'liar' (lit. 'one who possesses lies'), i.e. literally 'retreat of liar'.

60.   If *aqóon* can be analysed as a deverbal noun these compounds would become examples of the next type, deverbal or synthetic compounds.

61.   One question about these compounds concerns their internal structure. Is *xoghaye* 'secretary', for example, best bracketed as [[*xog*] – [*hay-e*]] or [[*xoghay*] *–e*]] ? In other words, is this a compound of a noun and an agentive verbal noun or do we have a compound verb which is then nominalised by an agentive suffix? From the limited information available, it seems that the relevant compound verbs do not occur and so the first analysis seems preferable, though the evidence is not very strong.

62.   In this template, for simplicity, we treat the impersonal subject pronoun *la* as a subject clitic pronoun and the reflexive/reciprocal *is* as a first series object clitic pronoun. It is highly unlikely that all of these elements would occur together in a single verbal group. The template shows the maximum potential group and their relative ordering.

63.   As described in 4.2.2 and 6.1.2, object clitic pronouns have a third person gap in their paradigm. Consequently lack of any other object clitic pronoun will result in a contextually appropriate third person reading. For brevity, in the glosses below we reflect this by arbitrarily placing one interpretation in parentheses and adding 'etc.'. Thus '(it etc.)' represents a contextual determined choice between 'it, him, her, them'.

64. It is interesting however that the pronominal clitics maintain a SUB-OBJ-V (SOV) order that is commonly postulated as basic in Cushitic languages, even though the order of full noun phrases in sentences does not reflect grammatical functions but pragmatic processes of information packing, as we shall see in later chapters.

65. The syntax of nominal focus is described in chapter 7 and its use in discourse in chapter 9.

66. This optionality also depends on the type of sentence and the distinction between main and subordinate clauses. For example, as described in chapter 7, negative main clauses occur very frequently without subject pronouns, as do yes-no questions. This means that simple main clauses may occur without a subject as in (1):

(1) (a) *Má yaqaan?*
         QM know:3MSG
         'Does he know (it)?'
     (b) *Hàa, wàa yaqaan.*
         yes DM know:3MSG
         'Yes, he knows it.'
     (c) *Máya, má yaqáan.*
         no NEG know: 3MSG:NEG
         'No, he doesn't know it.'

However subordinate clauses must occur with an overt subject; compare:

(2) (a) *Wúxuu   dóonayaa   ínuu   tagó.*
         what+he wants   that-he go:SUBORD

     (b) **Wúxuu dóonayaa   ín   tagó.*
         what+he wants   that   go:SUBORD
         'What he wants is to go,' He wants to go' (Lit. 'What he wants is that he go')

Sentence (2b) is ungrammatical because there is no subject in the complement clause. Finally, it is perhaps also worth noting that the degree of optionality of subject pronouns in declarative sentences seems to vary across dialects in ways that remain to be documented.

67. In this progressive auxiliary construction, merging of the main verb and the auxiliary has taken place to varying degrees in different Somali dialects. The resulting coalesced forms are no longer transparently auxiliary constructions.

68. As discussed earlier, the fact that this cliticization process attaches the inflection of the copula verb *yahay* onto the adjectival stem has led some linguists to analyse adjectives as a type of (defective) verb. There does seem to be an ongoing process of merger of adjectives and verbs by this elision of the copula. This merging is exemplified by the item *leeyahay* 'have'. This form is clearly a merging of the adjective *lé(h)* 'having' and the copula *yahay* 'be' and all of its forms may be predicted by such a rule. Nonetheless, many descriptions by native speakers (e.g. Mohamed Diriye Abdullahi 1996:111) treat the item as a verb 'have' and it may well be that for speakers the form has been re-analysed as a simple verb.

69.   Banti (1985:34) reports that in Benadir and May dialects this negative word occurs in the same position as the sentence type markers i.e. immediately preceding the VGP.

70.   An interesting exception is the adjective *dhán* 'all' which presumably because of its meaning can only occur as a post-nominal modifier in an appositive relative clause, e.g. *dádka óo dhán* 'all the people'.

71.   See 6.1.6 for further discussion.

72.   Before *uu* 'he', *má* 'not' often assimilates to *mú*, for example the common pronunciation: *Háwlba mú uu helín* 'He found no work'.

73.   Important studies of Somali focus include Hetzron (1965, 1971), Andrzejewski (1975a), Antinuci and Puglielli (1980), Antinucci (1980), Saeed (1984), Lecarme (1991, forthcoming), Ajello (1995) and Svolacchia et al. (1995).

74.   In this section the focused element is shown in bold; elsewhere in the book we use small capitals just in the English gloss when discussing focus.

75.   Though unlike many left dislocation structures cross-linguistically, *bàa/ayàa* structures do not allow coreferential clitic pronouns when the focused element is subject, as we shall see. As will also be clear from examples, left-dislocated elements focused by *bàa/ayàa* may be preceded by topic elements. We discuss the difficult question of word order in chapter 9.

76.   In addition they may not focus functional categories like classifiers, negative words, conjunctions, etc. The fact that the intensifier *aad* 'very, much' may be focused by *bàa/ayàa* suggests that like many adverbials this word is nominal in origin.

77.   This constraint is very strong. In addition *bàa/ayàa* constructions do not normally occur in the same sentence as *wáxa(a)* focus but, as described in chapter 9, the combination sometimes occurs, especially in narratives when a sentence initial adverbial may be focused by *bàa/ayàa* and later in the same sentence *wáxa(a)* is used to focus another sentence element, for example (where again focused elements are in bold):

Beri   *baa*   *waxaa*   *sheekeystay*   **hal,**   *sac*   *iyo*   **ri**.
time   FOC   *wáxa(a)*   conversed   camel:F   cow   and   goat:F
Once there conversed **a female camel, a cow and female goat**.'   [FS 41.2]

Another context for this doubling, as in the example below sent to me by Martin Orwin, is news broadcasts, where a subject may be introduced by *bàa/ayàa* focus and what they say is then introduced by a *wáxa(a)* focus construction, e.g.:

Wakiilo   *ka*   *tirsan*   *kooxaha*   *mucaaradka*   *ee*   *waddanka*   *Ciraaq*
delegates   from   counted   parties-the   opposition-the   and   country-the   Iraq
ayaa   *waxay*   *sheegeen* **inay**   **ku**   **heshiiyeen**   **inay**...
FOC   *wáxa(a)*+they   reported   that-they   on   agreed   that-they...
'**Representatives of the Iraqi opposition parties** reported **that they had agreed to**...'
(BBC Somali Service, Evening news bulletin, 23 November 1998)

78. By maximal projection of a head, we mean the largest (or most inclusive) phrase of which it is the head; or in co-ordinate structures, the unit consisting of all of the co-ordinates.

79. For example Antinucci and Puglielli (1980) suggested an analysis where relative clauses underlyingly contain focus words on their heads which are then obligatorily deleted. This was criticised in Saeed (1984) where I proposed the opposite line: *wáxa(a)* focus is derived from relative clauses headed by *wáx-a* 'the thing, what' and *bàa/ayàa* constructions are derived from *wáxa(a)* focus by a fronting rule. I no longer consider this plausible as a synchronic rule and indeed Lecarme (forthcoming) shows that asymmetries between *wáxa(a)* and *bàa/ayàa* constructions make this derivation unconvincing. In this study Lecarme proposes an analysis within a Chomskyan Minimalist approach. This seeks to capture the similarities by assigning *bàa/ayàa* focus and relative clauses a similar phrase structure with the focused element and the relative clause head as sentence external elements co-indexed with an empty argument within the sentence. No doubt other proposals will emerge to give a synchronic characterisation of what is probably an ongoing historical process.

80. Here *qabán* is a short form of the future *qabán doonaa* 'I will catch' (4.3.3.8).

81. Of course, as we saw in (42), if a direct quotation is introduced by *wáxa(a)* it may itself include a *wáxa(a)* construction. This is also true of *bàa/ayàa* constructions. We must therefore distinguish between direct quotation, which may contain independent main clauses, and clausal subordination.

82. Though in Saeed (1984) this is described as a synchronic syntactic rule.

83. For an alternative analysis see Lecarme (1991) which proposes that *wáxa(a)* constructions are a form of extraposition rule involving an expletive *wáx* focused by *bàa*.

84. As reported by Banti (1985: 30) central varieties of Standard Somali combine the negative word *áan* and the subject pronouns in the opposite order, i.e. *ma + aan* 'I' + *áan* 'not' → *miyàanán*; *ma + aad* 'you (SG)' + *áan* 'not' → *miyàadán*; *ma + uu* 'he' + *áan* 'not' → *miyùusán*, etc. Here we note that the vowel of *áan* 'not' is shortened and the 3MS pronoun *uu* 'he' occurs as *uus*.

85. Though as we see below, *ma* may occur suffixed to the questioned element.

86. Though independent pronouns may be used for emphasis and will be subject marked, for example:

    Adigu   hayla   hadlin!
    you:SG:SUBJ   *ha*+me+with speak:NEG:SG
    'You, don't speak (SG) to me!'   [SHN 92.7]

87. Consequently in southern dialects in which subject marking is reduced or non-existent, subordinate clauses are less distinguished from main clauses: see Lamberti (1983) and Banti (1985) for some details.

88. Thus restrictive relative clauses do not occur modifying proper nouns (names).

89. We can recognise here the general principle, described in 6.2.4, that nominals may only support one modifier before co-ordination is triggered.

90. See Antinucci and Puglielli (1980) and Saeed (1984) for discussion of this phenomenon. It seems reasonable to suppose that the gap in subject position inside the clause is not sufficient to trigger full subject agreement on the verb.

91. Note that as usual with co-ordination, a list of conjuncts may just have a conjunction between the last two items, for example:

Waa baa waxaa jirey nin gaabaan, faruuran   oo   aad   u   foolxun
time FOC *wáxa(a)* was   man short   hare-lipped and intens ADP ugly
'Once there was a man who was short, hare-lipped and very ugly'   [SHS 53.37]

92. We can gloss *islá márka* as 'the same time'; see for example a comparable expression *islá sánnadkáa* 'that same year'. The syntax of these expressions is unusual since, as we have seen, nominal modifiers usually follow the head NP in Somali and clitic clusters such as *islá* are more typical of the verbal group (6.1).

93. In (59) *la'áan* is a feminine noun 'state of being without' derived from the adjective *lá'* 'without, lacking, not possessing', as described in 5.3.2. Here the derived noun incorporates the nominal object of the adjective to form the complex nominal *tigídh la'áan* 'the state of being without a ticket'. This type of incorporation by adjectives is discussed briefly in 5.4.

94. Perhaps a better English gloss for this particular usage would be 'case' or 'instance'. Compare the related use of the word in the expression *haddii kalé* 'otherwise' (lit. 'the other case').

95. Here, as in the rest of this section, focused elements are shown in small capitals.

96. Of course the use of names like 'Ali' in this example means that the identity of the individuals has to be known. As discussed later in this section, we must make a distinction between knowledge that the interlocutors might have and knowledge that is 'activated' in the conversation and thus linguistically marked in some way.

97. These all-new sentences are of course rather uncommon in ordinary conversation. Support for this observation that subjects in them attract focus comes from another context in which all-new sentences are used: the beginning sentences of news stories and bulletins. So for example in the fifty-one Somali news stories in Issa and Murphy (1984), twenty seven have first sentences containing *bàa/ayàa* focus, rather than *wáxa(a)* constructions, etc. Of these twenty-seven, twenty-three have *bàa/ayàa* focus on the subject compared with four with focus on the object. Moreover in two of the latter the subject is a clitic pronouns which cannot receive focus.

98. Or in an alternative metaphor, become foregrounded in consciousness and then fade into the background. For discussions of this notion of salience/activation see Chafe (1976), Prince (1981, 1985), Givón (1983), and Dryer (1996).

99. For various views of the factors governing subject assignment see Kuno (1987), Wilkins (1988), Givón (1990), Williams (1994).

100. Note that the time adverbial *waa* 'once' receives *bàa/ayàa* focus as described later in this section.

101. Or 'presentative' structures to use the term proposed by Hetzron (1971).

102. As Lecarme (1991, forthcoming) notes, these sentences are a problem for a simple and coherent view of *bàa* focus. While it is clear that *wáx* here is an indefinite NP, and thus formally a candidate for 'new' information, it seems odd that an almost lexically empty element like this should receive focus. This is especially so since this sentence is a possible answer to the question *Maxáad saméynaysaa*? 'What are you doing?'

103. Including, as we saw in chapter 7, the adverbial intensifier *àad* 'much' 'many'.

104. Remembering that, as outlined in 4.8, sentence type markers are usually sensitive to the difference between positive and negative sentences.

105. It is worth noting that the problem of identifying the functions of preverbal grammatical morphemes is a common one in Cushitic languages. See for example the discussion of 'preverbal selectors' in Arbore in Hayward (1984). As Hayward notes, part of these are clearly sentence type identifiers. We might also note the discussion of Dahalo 'selectors' in Tosco (1991), and Oromo 'preverbal elements' in Stroomer (1987).

106. For discussions of the syntactic status of this core predication see Lecarme (forthcoming) and Svolacchia, Mereu and Puglielli (1995).

107. We may speculate that the formalised structure of the story and its sequence of girls allow the existence of a further girl to be pragmatically implied or expected.

108. For general discussions of Somali poetry see Andrzejewski and Lewis (1964), Andrzejewski (1985), and Banti (1987). For descriptions of poetic genres and metrical forms, see Johnson (1979, 1984, 1996), Guuleed (1980), Andrzejewski (1982), Banti & Giannattasio (1996), and Orwin & Riiraash (1997). Briefly, the various poetic types are characterised by different metrical structures. The *gabay*, for example, may consist of more than a hundred lines, each of which contains twenty moras, typically divided into four equal feet of five moras each. As mentioned in chapter 2, only vowels are counted in establishing moras. The details of the various mappings between moras, feet, syllables and lines in each poetic form are described in the works cited above. In addition the major poetic types are characterised by alliteration, which may be of two types: consonantal and vowel. In consonantal alliteration, one consonant must be used to alliterate throughout the poem. In vowel alliteration all vowels are considered to alliterate with each other, whatever the quality. Text 6 has consonantal alliteration in the sound *g*. Since this poem is a *geerar*, its line structure requires that one word in each line must alliterate. Text 5 has vowel alliteration. Since it is a *gabay*, its line structure requires that each half line must contain at least one alliterating word.

References

Abdullahi, Mohamed Diriye. 1996. *Parlons Somali*. Paris: Editions L'Harmattan.

Adan, Hussein Mohamed (ed.) 1980. *Somalia and the World. Proceedings of the International Symposium held in Mogadishu on the Tenth Anniversary of the Somali Revolution, October 15-21, 1979*. Mogadishu: State Printing Press.

Agostini, F., A. Puglielli and Ciise Moxamed Siyaad. (eds.) 1985. *Dizionario somalo-italiano*. Rome: Gangemi.

Ajello, Roberto. 1981. 'La funzione del raddoppiamento nel sistema verbale somalo,' *La bisaccia dello Sheik*. Venice: Università degli studi di Venezia.

Ajello, Roberto. 1984. 'Il focus nell'idioma degli Ashraaf di Shingaani,' in Puglielli (ed.), 135-46.

Ajello, Roberto. 1995. 'La focalizzazione in Somalo,' in R. Ajello and S. Sani (eds.) *Scritti Linguistici e Filologici. In Onore di Tristano Bollelli*. Pisa: Pacini editore, 1-28.

Andzrejewski, B.W. 1955. 'The Problem of Vowel Representation in the Isaaq Dialect of Somali'. *BSOAS*, 17.3: 567-580.

Andrzejewski, B.W. 1956. 'Accentual Patterns in Verbal Forms in the Isaaq Dialect of Somali'. *BSOAS*, 18.1: 103-129.

Andzrejewski, B.W. 1960. 'Pronominal and Prepositional Particles in Northern Somali', *African Language Studies* 1: 96-108.

Andzrejewski, B.W. 1964. *The Declensions of Somali Nouns*. London: School of Oriental and African Studies, University of London.

Andrzejewski, B.W. 1967. 'Plural and Sub-plural Forms of Somali nouns,' Unpublished manuscript, School of Oriental and African Studies, University of London.

Andrzejewski, B.W. 1968. 'Inflectional Characteristics of the So-called 'Weak Verbs' in Somali,' *African Language Studies*, 9: 1-51.

Andrzejewski, B.W. 1969. "Some Observations on Hybrid Verbs in Somali". *African Language Studies*, 10: 47-89.

Andrzejewski, B.W. 1970. 'The Role of Tone in the Borana Dialect of Galla,' in *Proceedings of the Third International Congress of Ethiopian Studies, Addis Ababa, 1996, Vol 2*. Addis Ababa: Institute of Ethiopian Studies, Haile Sellassie I University, 88-98.

Andrzejewski, B.W. 1975a 'The Role of Indicator Particles in Somali,' *Afroasiatic Linguistics* 1.6: 123-191.

Andrzejewski, B.W. 1975b. 'Verbs with Vocalic Mutation and their Significance for Hamito-Semitic Comparative Studies, ' in James and Theodora Bynon (eds. 1975) *Hamitico-Semitica*. The Hague: Mouton, 361-376.

Andrzejewski, B.W. 1979. 'The Case System in Somali,' mimeo, School of Oriental & African Studies, University of London.

Andrzejewski, B.W. 1982. 'Alliteration and scansion in Somali oral poetry and their cultural correlates,' in Görög-Karady, V. (ed), 68-83.

Andrzejewski, B.W. 1985. 'Somali Literature,' in B.W. Andrzejewski et al. (eds.) *Literatures in African Languages: Theoretical Issues and Sample Surveys*. Cambridge: Cambridge University Press, 337-407.

Andrzejewski, B.W. and Sheila Andrzejewski. 1993. *An Anthology of Somali Poetry*. Bloomington: Indiana University Press.

Andrzejewski, B.W. and Lewis, I.M. 1964. *Somali Poetry An Introduction*. London: Oxford University Press.

Antinucci, Francesco. 1980. 'The Syntax of Indicator Particles in Somali. Part Two: The Constructions of Interrogative, Negative and Negative-Interrogative Clauses,' *Studies in African Linguistics* 11.1: 1-37.

Antinucci, Francesco and Annarita Puglielli. 1980. 'The Syntax of Indicator Particles in Somali: Relative Clause Constructions,' *Afroasiatic Linguistics* 7.3: 85-102.

Appleyard, David L. 1990. 'Prepositional Particles in Somali and their Cognates in Other Cushitic Languages', *African Languages and Cultures* 3.1: 15-32.

Armstrong, L.E. 1934. "The phonetic structure of Somali". Mitteilungen des Seminars für Orientalische Sprachen zu Berlin no. 37. (Republished in 1964 by Gregg International Publishers Limited, Farnborough, England.).

Bakker, E. 1994. 'Voice, Aspect and Aktionsart: Middle and Passive in Ancient Greek,' in Barbara Fox and Paul J. Hopper (eds.) *Voice: Form and Function*. Amsterdam: John Benjamins, 23-48.

Banti, G. 1985. 'Lineamenti di fonologia, morfologia e sintassi del somalo e dei suoi dialetti,' Rome: University of Rome.

Banti, G. 1988a. "'Adjectives' in East Cushitic". In Bechhaus-Gerst and Serzisko (eds. 1988), 203-59.

Banti, G. 1988b. 'Two Cushitic Systems: Somali and Oromo Nouns', in van der Hulst and Smith (eds. 1988), 11-49.

Banti, G. 1994. 'Parti del Discorso Controverse,' in *Atti dei Convegni Lincei 107*. Rome: Academia Nazionale dei Lincei, 53-82.

Banti, G. and F. Giannattasio. 1996 'Music and Metre in Somali Poetry,' in Hayward and Lewis (eds.), 83-127.

Bechhaus-Gerst, M. and Serzisko, F. (eds.) 1988. *Cushitic-Omotic: Papers from the International Symposium on Cushitic and Omotic Languages Cologne, January 6 -9, 1986*. Hamburg: Helmut Buske.

Bell, C.R.V. 1953. *The Somali Language*. London: Longmans, Green & Company. (Reprinted in 1968 by Gregg International, Farnborough, England)

Bender, M. L. (ed.) 1976. *The Non-Semitic Languages of Ethiopia*. Monograph No. 5 Occasional Papers Series Committee on Ethiopian Studies. Michigan: African Studies Center, Michigan State University.

Biber, D. 1984a. 'The Diachronic Development of Preverbal Case Markers in Somali', *Journal of African Languages and Linguistics* 6: 47-61.

Biber, D. 1984b. 'Pragmatic Roles in Central Somali Narrative Discourse,' *Studies in African Linguistics* 15: 1-26.

Biber, D. 1995. *Dimensions of Register Variation: a cross-linguistic comparison*. Cambridge: Cambridge University Press.

Bruno, Biancamaria. 1984. 'Note sui verbi di derivazione nominale in somalo,' in Puglielli (ed.), 115-131.

Burton, Richard. 1894. *First Footsteps in East Africa, or An Exploration of Harar*. Two volumes. London: Tylston and Edwards. [Republished as one volume in 1987 by Dover Publications, Mineola, New York] .

Bybee, J., R. Perkins and W. Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Bynon, Theodora. 1990. 'Pronoun and Verb Agreement in Typological and Diachronic Perspective,' *SOAS Working Papers in Linguistics* 1: 97-11.

Caney, John Charles. 1984. *The Modernisation of Somali Vocabulary, with Particular reference to the Period from 1972 to the Present*. Hamburg: Helmut Buske.

Carboni, Fabio (ed.) 1983. *Bibliografia somala. (Studi somali 4)* Rome: Ministero degli affari esteri - Dipartimento per la cooperazione allo sviluppo, Comitato tecnico linguistico per l'università nazionale somala.

Cardona, G.R. 1981. 'Profilo fonologico del somalo,' in Cardona, G.R. & F. Agostini (eds. 1981), 5-26.

Cardona, G. R. and F. Agostini. (eds.) 1981. *Fonologia e lessico. (Studi somali 1)* Rome: Ministero degli affari esteri - Dipartimento per la cooperazione allo sviluppo, Comitato tecnico linguistico per l'università nazionale somala.

Cerulli, Enrico. 1957-64. *Somalia*. Three volumes. *Vol 1.* (1957), *Vol. 2* (1959), *Vol 3* (1964). Rome: A cura dell'Amministrazione Fiduciaria Italiana.

Ciise, Jaamac Cumar. 1974. *Diiwaanka Gabayadii. Sayid Maxamad Cabdulle Xasan. Uruurintii Kowaad*. Muqdisho: Wasaaradda Hiddaha iyo Taclinta Sare. [Collected Poems of Sayid Maxamad Cabdulle Xasan. Volume 1. Mogadishu: Ministry of Culture and Higher Education.]

Ciise, Jaamac Cumar. 1976. *Taariikhdii Daraawiishta iyo Sayid Maxamed Cabdulle Xasan (1895-1921)*. Muqdisho: Akadeemiyaha Dhaqanka. [The Historyof the Dervishes and Mohamed Abdullah Hasan (1895-1921). Mogadishu: Academy of Culture]

Chafe, Wallace. 1976. 'Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View,' in Charles N. Li (ed.) *Subject and Topic*. New York: Academic Press, 25-55.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Dhoorre, Cabdulqadir Salaad and Mauro Tosco. 1998. '111 Somali Ideophones,' *Journal of African Cultural Studies* 11.2: 125-146.

Dowty, David R. 1991. 'Thematic Proto-Roles and Argument Selection,' *Language* 67: 547-619.

Drysdale, John. 1994. *Whatever Happened to Somalia?* London: Haan Publishers.

Dwyer, Matthew S. 1996. 'Focus, Pragmatic Presupposition, and Activated Propositions,' *Journal of Pragmatics* 26: 475-523.

Ehret, C. & Mohamed Nuuh Ali. 1984. 'Soomaali classification,' in T. Labahn (ed. 1984), 201-269.

El-Solami-Mewis, C. 1987. *Lehrbuch des Somali*. Leipzig: VEB Verlag Enzyklopädie.

Farnetani, E. 1981. 'Dai tratti ai parametri: introduzione all'analisi strumentale della lingua somala', in Cardona and Agostini (eds. 1981), 27-108.

Ferrario, B. 1914-1915. 'L'Accento in Somalo,' *Rivista degli Studi Orientali*, Vol.6: 961-967.

Ferrario, B. 1916-18. 'Note di fonologia Somala,' *Rivista degli Studi Orientali*, Vol.7, Part 1: 199-217.

Foley, William A. and Robert D. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Galaal, Muuse Haaji Ismaa'iil and B.W. Andrzejewski. 1956 *Hikmad Soomaali*. London: Oxford University Press.

Galboran, F.L. and S. Pillinger 1990. *An Initial Rendille-English Dictionary and English Rendille Dictionary*. Nairobi: Bible Translation and Literacy (East Africa).

Gebert, Lucyna. 1986. 'Focus and Word Order in Somali,' *Afrikanistische Arbeitpapiere* 5: 42-69.

Givón, Talmy. 1975. 'Serial Verbs and Syntactic Change: Niger-Congo,' in Li (ed. 1975), 47-112.

Givón, Talmy. (ed.) 1983. *Topic Continuity in Discourse: a Quantitative Cross-Language Study*. Amsterdam: John Benjamins.

Givón, Talmy. 1984. *Syntax: a Functional-Typological Introduction. Volume 1*. Amsterdam: John Benjamins.

Givón, Talmy. 1990. *Syntax: a Functional-Typological Introduction. Volume 2*. Amsterdam: John Benjamins.

Givón, Talmy. 1994. *Voice and Inversion*. Amsterdam: John Benjamins.

Görög-Karady, V. (ed.) 1982. *Genres, forms and meanings: Essays in African Oral Literature*. Journal of the Anthropological Society of Oxford Occasional Papers No. 1, Oxford.

Guuleed, Cabdullaahi Diiriye. 1980. "The scansion of Somali poetry". In Hussein Mohamed Adan (ed.), 132-140.

Hayward, R.J. 1978. 'Bayso revisited: some preliminary linguistic observations-I,' *B.S.O.A.S.*, 41.3: 539-570.

Hayward, R.J. 1979. 'Bayso revisited: some preliminary linguistic observations-II,' *B.S.O.A.S.*, 42.1: 101-132.

Hayward, R.J. 1984. *The Arbore Language: a First Investigation*. Kuschitische Sprachstudien Band 2. Hamburg: Helmut Buske.

Hayward, R.J. (ed.) 1990. *Omotic Language Studies*. London: School of Oriental and African Studies, University of London.

Hayward, R.J. 1991. 'Tone and accent in the Qafar noun,' *York Papers in Linguistics*, 15: 117-137.

Hayward, R.J. and I. M. Lewis (eds.) 1996. *Voice and Power: the Culture of Language in North-East Africa. Essays in Honour of B. W. Andrzejewski*. African Languages and Cultures, Supplement 3. London: School of Oriental and African Studies, University of London.

Hayward, R.J. and J.I. Saeed. 1984. 'NP Focus in Somali and Dirayta: a comparison of *baa* and *pa*.' in T. Labahn (ed. 1984) *Proceedings of the Second International Congress of Somali Studies, UNiversity of Hamburg, August 1-6, 1983*. Hamburg: Helmut Buske.

Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. 1991. *Grammaticalization: a Conceptual Framework*. Chicago: University of Chicago Press.

Hetzron, Robert. 1965. 'The Particle *baa* in Northern Somali,' *Journal of African Languages* 4.2: 118-130.

Hetzron, Robert. 1971. 'Presentative Function and Presentative Movement,' *Studies in African Linguistics, Suppl*. 2: 79-105.

Hopper, Paul J. and Elisabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Hyman, L. 1981. 'Tonal accent in Somali,' *Studies in African Linguistics*, 12.2: 169-203.

Issa, Abdullahi A. and John D. Murphy. 1984. *A Somali Newspaper Reader*. Kensington, Maryland: Dunwoody Press.

Johnson, J.W. 1979. 'Somali prosodic systems,' *Horn of Africa*, 2.3: 46-54.

Johnson, J.W. 1984. 'Recent researches into the Scansion of Somali Oral Poetry,' in T. Labahn (ed.), 313-331.

Johnson, J.W. 1988. 'Set theory in Somali poetics: structures and implications". In Puglielli, A. (ed.), 123-32.

Johnson, J.W. 1996. 'Musico-Moro-Syllabic Relationships in the Scansion of Somali Oral Poetry,' in Hayward and Lewis (eds.), 73-82.

Keenan, Edward L. 1985. 'Relative clauses,' in Shopen (ed.), Vol. II, 141-170.

Kemmer, S. 1994. 'Middle Voice, Transitivity and the Elaboration of Events,' in Barbara Fox and Paul J. Hopper (eds.) *Voice: Form and Function*. Amsterdam: John Benjamins, 179-230.

Keenadiid, Yaasiin Cismaan. 1976. *Qaamuuska Af Soomaaliga*. [Dictionary of the Somali Language] Mogadishu & Florence: Le Monnier.

Kiss, Katalin E. (ed.) 1995. *Discourse Configurational Languages*. Oxford: Oxford University Press.

Klaiman, M.H. 1991. *Grammatical Voice*. Cambridge: Cambridge University Press.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago: University of Chicago Press.

Labahn, T. (ed.) 1984. *Proceedings of the Second International Congress of Somali Studies. Volume I: Linguistics and Literature*. Hamburg: Helmut Buske.

Laitin, David. D. 1977. *Politics, Language, and Thought: The Somali Experience*. Chicago: The University of Chicago Press,.

Laitin, David D. and Said S. Samatar. 1987. *Somalia: a Nation in Search of a State*. Boulder, Colorado: Westview.

Lamberti, Marcello. 1986a. *Die Somali-Dialekte*. Kuschitische Sprachstudien Band 5. Hamburg: Helmut Buske..

Lamberti, Marcello. 1986b. *Map of Somali Dialects in the Somali Democratic Republic*. Hamburg: Helmut Buske.

Lamberti, Marcello. 1988. *Die Nordsomali-Dialekte: Eine synchronische Beschreibung*. Heidelberg: Carl Winter Universitätsverlag.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Laurence, Margaret. 1954. *A Tree for Poverty: Somali Poetry and Prose*. Published for the British Protectorate of Somaliland. Nairobi: Eagle Press. Reprinted, Dublin: Irish UP, and Hamilton: McMaster University, 1970.

Laurence, Margaret. 1963. *The Prophet's Camel Bell*. Toronto: McClelland & Stewart.

Lecarme, Jacqueline. 1991. 'Focus en Somali: syntaxe et interprétation,' *Linquistique africaine* 7:33-63.

Lecarme, Jacqueline. 1996. 'Tense in the Nominal System: the Somali DP,' in J. Lecarme, J. Lowenstamm and U. Shlonsky (eds.) *Studies in Afroasiatic Grammar*. The Hague: Holland Academic Graphics.

Lecarme, Jacqueline. forthcoming. 'Focus in Somali,' in L. Rehuschi and L. Tuller (eds.) *The Grammar of Focus*. Amsterdam: John Benjamins.

Lewis I. M. 1961. *A Pastoral Democracy: A Study of Pastoralism and Politics among the Northern Somali of the Horn of Africa*. London: International African Institute. Reprinted New York: Africana Publishing Company for the International African Institute, 1982.

Lewis, I. M. 1980. *A Modern History of Somalia Nation and State in the Horn of Africa*. Revised edition. London: Longman.

Lewis, I. M. 1993. *Understanding Somalia: Guide to Culture, History, and Social Institutions.* London: Haan Associates.

Li, Charles N. (ed.) 1975. *Word Order and Word Order Change.* Austin: University of Texas Press.

Loughran, Katheryne S., John L. Loughran, John William Johnson and Said Sheikh Samatar. (eds.) 1986. *Somalia in Word and Image.* Bloomington: Indiana University Press.

Mansur, Abdalla Omar. 1988. *Le lingue Cuscitiche e il Somalo.* Italian edition edited by Carlo Serra Borneto. . *(Studi somali 8)* Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

Mansur, Abdalla Omar and Annarita Puglielli. 1999. *Barashada Naxwaha Af Soomaaliga. A Somali School Grammar.* London: Haan Associates by arrangement with il Dipartimento di Linguistica, Università degli Studi di Roma Tre.

Mansur, Abdalla Omar and Annarita Puglielli. 1999. *Furaha Layliyada Barashada Naxwaha Af Soomaaliga. Answers to Exercises in a Somali School Grammar.* London: Haan Associates by arrangement with il Dipartimento di Linguistica, Università degli Studi di Roma Tre.

Marantz, A. 1982 . 'Re Reduplication,' *Linguistic Inquiry*, 13.3: 435-482.

Maxamed, Cabdi Maxamed. 1986. *Dictionnaire Français-Somali. Qaamuus Fransiis-Soomaali*, (2 vols.) Paris: Institut national des langues et civilisations orientales.

Mohamed, Abdi Mohamed (ed.) 1993. *Anthropologie Somalienne: Actes du IIe Colloque des Études Somaliennes (Besançon - 8/11 octobre 1990)* Annales Littéraires de l'Université de Besançon, 495.

Moreno, M.M. 1951. 'Brevi notazioni di ǧiddu,' *Rassegna di studi etiopici* 10:99-107.

Moreno, M.M. 1953. 'Il dialetto di Ašraf di Mogadiscio,' *Rassegna di studi etiopici* 12: 107-38; 13: 5-19.

Moreno, M.M. 1955. *Il somalo della somalia. Grammatica e testi del benadir, darod e dighil.* Rome: Istituto poligrafico dello stato.

Morin, Didier. 1989. 'A Propos du Système Nominal du Somali,' in Taddese Beyene (ed.) *Proceedings of the Eighth International Conference of Ethiopian Studies, Univeristy of Addis Ababa, 1984. Volume 2.* Addis Ababa: Institute of Ethiopian Studies, 521-530.

Moxamud, Ibrahim Xashi and Moxamed Xashi Xidhkayome. 1998. *Qawacidka Luuqadda Somaaliga.* Stockholm: African Triangle.

Mumin, Hassan Sheikh 1979. *Leopard Among the Women. (Shabeelnaagood)* translated by B.W. Andrzejewski. London: Oxford University Press.

Nichols, Johanna. 1986. 'Head-Marking and Dependent-Marking Grammar,' *Language* 62: 56-119.

Oomen, A. 1978. 'Focus in the Rendille Clause,' *Studies in African Linguistics* 9: 35-65.

Orwin, Martin. 1993. 'Phonation in Somali Phonology'. In Mohamed Abdi Mohamed (ed.) (1993), 251-257.

Orwin, Martin. 1994. 'Aspects of Somali Phonology,' Unpublished Ph.D. dissertation, School of Oriental and African Studies, University of London.

Orwin, Martin. 1995. *Colloquial Somali : a complete language course.* London: Routledge.

Orwin, Martin. 1996. 'A Moraic Model of the Prosodic Phonology of Somali,' in Hayward and Lewis (eds.), 51-71.

Orwin, Martin and Maxamed Cabullahi Riiraash. 1997. 'An Approach to Relationships between Somali Metre Types,' *African Languages and Cultures* 10.1: 83-100.

Owens, Jonathan. 1980. 'Observations on Tone in the Booran Dialect of Oromo (Galla),' *African Language Studies* 17:141-96.

Owens, Jonathan. 1982. 'Case in the Booran Dialect of Oromo. *Afrika und Übersee* 65: 43-74.

Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Kuschitische Sprachstudien Band 4. Hamburg: Helmut Buske.

Panza, Bruno. 1974. *Af Soomaali: Grammatica della lingua somala con piccolo vocabulario in appendice*. Florence: Le Monnier.

Pillinger, O.S. 1989. 'Accent, tone and prosodic structure in Rendille,' Unpublished Ph.D. thesis, School of Oriental and African Studies, University of London.

Prince, Ellen F. 1981. 'Towards a Taxonomy of Given-New Information,' in Peter Cole (ed.) *Radical Pragmatics*. New York: Academic Press, 223-55.

Prince, Ellen F. 1985. 'Fancy Syntax and "Shared Knowledge",' *Journal of Pragmatics* 9: 65-81.

Puglielli, Annarita. 1981. 'Frase dichiarativa,' in A. Puglielli (ed. 1981), 3-46.

Puglielli, Annarita. 1984. 'La derivazione nominale in somalo,' in A. Puglielli (ed. 1984), 3-52.

Puglielli, Annarita. In press. 'Compounds in Somali,' in G. Banti (ed.), *Proceedings of the Second International Symposium on Cushitic and Omotic Languages*. Rome: University of Rome La Sapienza.

Puglielli, Annarita. (ed.) 1981 *Sintassi della lingua somala. (Studi somali 2)* Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

Puglielli, Annarita. (ed.) 1984 *Aspetti morfologici, lessicali e della focalizzazione. (Studi somali 5)* Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

Puglielli, Annarita. (ed.) 1988. *Proceedings of the Third International Congress of Somali Studies*. Rome: Il Pensiero Scientifico Editore.

Puglielli, Annarita and Ciise Mohamed Siyaad. 1984. 'La flessione del nome,' in Puglielli (ed.), 53-112.

Puglielli, Annarita. (ed.) 1998. *Dizionario Italiano-Somalo*. Rome: Carocci editore.

Raabi, Maxamed Xaaji Xuseen. 1977. *Codaynta Af Soomaaliga*. Mogadishu: Jaamacadda Ummadda Somaaliyeed, Kulliyadda Waxbarashada Lafoole.

Saeed, John Ibrahim. 1982. 'Central Somali - A grammatical outline,' *Afroasiatic Linguistics*, 8/2: 1-43.

Saeed, John Ibrahim. 1984. *The Syntax of Focus and Topic in Somali*. Kuschitische Sprachstudien Band 3. Hamburg: Helmut Buske.

Saeed, John Ibrahim. 1988. 'An Argument for the Category Adjective in Somali', in Bechhaus-Gerst and Serzisko, (eds), 567-579.

Saeed, John Ibrahim. 1992. 'Dialect Variation in Somali,' in H.A. Adam and C.L. Geshekter (eds.) *Proceedings of the First International Congress of Somali Studies*. Atlanta, Georgia: Scholars Press, 464-490.

Saeed, John Ibrahim. 1993a. *Somali Reference Grammar*. Second Revised Edition. Kensington, Maryland: Dunwoody Press.

Saeed, John Ibrahim. 1993b. 'Adpositional Clitics and Word Order in Somali,' *Transactions of the Philological Society* 91.1: 63-93.

Saeed, John Ibrahim. 1995. 'The Semantics of Middle Voice in Somali,' *African Languages and Cultures* 8.1: 61-85.

Saeed, John Ibrahim. 1996. 'Head-Marking and Pronominal Clitics in Somali,' in Hayward & Lewis (eds. 1996), 37-49.

Saeed, John Ibrahim. In press. 'Morphological Causatives and Verbal Argument Structure in Somali', in G. Banti (ed.) *Proceedings of the Second International Symposium on Cushitic and Omotic Languages*. Rome: University of Rome La Sapienza.

Samatar, Abdi Ismail. 1989. *The State and Rural Transformation in Northern Somalia 1884-1986*. Madison: Universityof Wisconsin Press.

Samatar, Ahmed. I. 1988. *Socialist Somalia: Rhetoric and Reality*. London: Zed Books.

Sasse, H.-J. 1979. 'The consonant phonemes of Proto-East-Cushitic (PEC): A first approximation,' *Afroasiatic Linguistics*, 7/1: 1-67.

Sasse, H-J. 1981. '"Basic Word Order" and Functional Sentence Perspective in Boni', *Folia Linguistica* XV/3-4: 253-290.

Serzisko, Fritz. 1992. 'Collective and Transnumeral Nouns in Somali,' in H.A. Adam and C.L. Geshekter (eds.) *Proceedings of the First International Congress of Somali Studies*. Atlanta, Georgia: Scholars Press, 516-30.

Shopen, Timothy. (ed.) 1985. *Language Typology and Syntactic Description*. Three volumes. Cambridge: Cambridge University Press.

Sim, R.J. 1981. 'Morphophonemics of the Verb in Rendille,' *Afroasiatic Linguistics*, 8.1: 1-33.

Simons, A. 1995. *Networks of Dissolution. Somalia Undone*. Boulder, Colorado: Westview.

Siyaad, Ciise M. 1985. *Favole somale*. *(Studi Somali 6)* Rome: Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

Squillacciotti, Massimo. (ed.) 1995. *Produzione e riproduzione nel gruppo domestico. (Studi somali 10)* Turin: L'Harmattan Italia.

Stalnaker, Robert C. 1974. 'Pragmatic Presuppositions,' in M.K. Munitz and P.K. Unger (eds.) *Semantics and Philosophy*. New York: New York University Press, 197-213.

Stroomer, Harry. 1987. *A Comparative Study of Three Southern Oromo Dialects in Kenya*. Kuschitische Sprachstudien Band 6. Hamburg: Helmut Buske.

Svolacchia, Marco, Lunella Mereu and Annarita Puglielli. 1995. 'Aspects of Discourse Configurationality in Somali,' in K.E. Kiss (ed.) *Discourse Configurational Languages*. Oxford: Oxford University Press, 65-98.

Thompson, Sandra A. and Robert E. Longacre. 1985. Adverbial Clauses,' in Shopen (ed.), Vol. II, 171-234.

Tiling, Maria von. 1922. 'Die Sprache der Jabarti, mit besonderer Berücksichtigung der Verwandschaft von Jabarti und Somali,' *Zeitschrift für Eingeborenensprachen* 12: 17-52; 97-162.

Tiling, Maria von. 1925. 'Jabarti Texte,' *Zeitschrift für Eingeborenensprachen* 15: 50-158.

Tosco, Mauro. 1989. *Schizzo grammaticale del dialetto Karre di Qoryooley . (Studi somali 9)* Rome: Ministero degli afari esteri-Dipartimento per la cooperazione allo sviluppo.

Tosco, Mauro. 1991. *A Grammatical Sketch of Dahalo*. Kuschitische Sprachstudien Band 8. Hamburg: Helmut Buske.

Tosco, Mauro. 1997. *Af Tunni. Grammar, Texts, and Glossary of a Southern Somali Dialect*. Kuschitische Sprachstudien Band 13. Köln: Rüdiger Köppe.

van der Hulst, H. & Smith, N. (eds.) 1988. *Autosegmental Studies on Pitch Accent*. Foris Publications, Dordrecht.

Vycichl, Werner. 1956. 'Zur Tonologie des Somali,'. *Rassegna di Studi Orientali*, 31:221-227.

Wilkins, W. (ed.) 1988. *Syntax and Semantics. Volume 21: Thematic Relations*. New York: Academic Press.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Xaange, Axmed Cartan. 1988. *Sheekoxariirooyin Soomaaliyeed. Folktales from Somalia*. Uppsala: Somali Academy of Arts and Sciences in co-operation with the Scandinavian Institute of African Studies.

Xaange, Axmed Cartan. 1998. *Sheekooyiin. Favole somale raccolte da Axmed Cartan Xaange*. Edited by Annarita Puglielli. . *(Studi somali 11)* Turin: L'Harmattan Italia.

Zorc, R.D. and Issa, A.A. 1990. *Somali Textbook*. Kensington, Maryland: Dunwoody Press.

Zorc, R. David. and Madina M. Osman. 1993. *Somali-English Dictionary with English Index*.
    Third Edition. Kensington, Maryland: Dunwoody Press.

Index

The letter n before a number indicates a note.