ARABIC GRAMMARS OF TURKIC

STUDIES IN SEMITIC
LANGUAGES AND LINGUISTICS

EDITED BY

T. MURAOKA AND C.H.M. VERSTEEGH

VOLUME XXVIII

ARABIC GRAMMARS OF TURKIC

ARABIC GRAMMARS
OF TURKIC

The Arabic Linguistic Model Applied to Foreign Languages & Translation
of 'Abū Ḥayyān al-'Andalusī's
Kitāb al-'Idrāk li-Lisān al-'Atrāk

BY

ROBERT ERMERS

BRILL
LEIDEN · BOSTON · KÖLN
1999

Voor mijn ouders
Jan en Julia Ermers

CONTENTS

PREFACE

This study discusses the way Arabic grammarians analyse Turkic lan-
guages, or the application of the Arabic linguistic model and concepts
to a foreign language. The first part is a thorough revision of my Ph.D.
dissertation, which was financed with a four-year's grant from the
Netherlands' Organisation for Scientific Research (NWO), and which I
defended at the University of Nijmegen in 1995. Part Two contains a
translation of one of the most important sources, i.e. ꞌAbū Ḥayyān al-
ꞌAndalusī's *Kitāb al-ꞌIdrāk li-Lisān al-ꞌAtrāk*.

The subject and the outline of the present study are discussed in
Chapter One. Chapter Two contains a sketch of the sources, their con-
tents and internal arrangement and references to the most important
relevant publications. In this chapter statements are made about the
author and the Turkic language which is the subject of the description,
and presents my assumptions in regard to mutual relations of the
sources, and the way they are related to their basic material.

Chapter Three gives an account of the way morpho(no)logical fea-
tures of Turkic are discussed. Most of the sources contain a short sec-
tion with indications regarding the 'correct' pronunciation of Turkic.
It shall be seen how these descriptions are related to Arabic phonetic
and phonologic terminology. Moreover, it will be shown that in many
cases the remarks and labels that are usually understood as referring to
the phonetic qualities of Turkic words are intended as instruments to
determine which type of suffixes the verbal stem takes, i.e., suffixes
with either 'velar' or 'palatal' consonants.

Chapter Four, Cases and Markers, is intended as an introductory
chapter to Chapters Five and Six. It discusses the Arabic concept of
case and declension (ꞌiꞏrāb) and the possible consequences of the ap-
plication of this concept to Turkic languages. My assumption is that
the Arabic analytical method of dividing speech into separate elements
could, to some extent, be suitable for a segmentable language like
Turkic. The application of the concept of ꞌiꞏrāb to Turkic is the main
theme of Chapters Five and Six.

Chapter Four gives a preliminary study of the ways to express
synonymy between Arabic and Turkic on different levels. Chapter Five
discusses how various Arabic constructions that involve a genitive case

(ǧarr) were translated and analysed into Turkic; it deals with the Turkic
equivalents of Arabic prepositional and possessive constructions.

Chapter Six treats instances in which in Arabic the accusative case is
used. Unlike Arabic, objects in Turkic may be marked with several
types of markers. In some instances there is no marker at all, in others
it is not the accusative case but the dative case that is used. This chapter
analyses how Arabic grammarians dealt with these features of Turkic.
Finally, Chapter Seven summarises the findings of the research.

It is impossible to write a study like this all by oneself, especially if one
is still in the process of forming one's own scholarly concepts and
opinions, and apart from the scholarly feedback one also needs moral
support. Em. Prof. Dr Barbara Flemming, Dr Everhard Ditters and
Prof. Dr Erik-Jan Zürcher were involved in this study when it was still a
Ph. D. project; Dr James Kelly kindly sent me copies of microfilms in
his possession that were very important for my research. I am very
much indebted to Mr Ane Nauta, Mr Hans Nugteren and Ms Marti
Roos who provided much of the turcological feedback which was neces-
sary for this book. I felt inspired by discussions with Michel Limpens
on Arabic linguistics, and the fine atmosphere among the collegues at
TCMO, Nijmegen.

Prof. Dr Kees Versteegh's enthousiasm for, and thorough knowledge
of the Arabic linguistic tradition induced me in the first place to en-
gage in this fascinating subject. Later on, as a very inspiring supervisor
and mentor, he carefully guided my research activities which culmi-
nated in the completion of my Ph. D. dissertation. I am very grateful
to him for all the time and energy he has invested in my formation as a
scholar. A special word of gratitude is due to Prof. Dr Rafael Talmon,
who not only most cordially hosted me as a post-doc researcher at the
Department of Arabic of the University of Haifa, but also gave me
ample opportunity to work on the present book, for which I feel much
indebted. I wish to thank the anonymous reader for his elaborate and
constructive comments to an earlier version of the manuscript, and
publishing house Brill for accepting it in their series of *Studies in
Semitic Languages and Linguistics*.

In the period I was preparing my Ph.D. dissertation, my friends
Nanneke, Turan & Zöhre, Karin, Mark & Madeleine were always there
when I most needed them. My brothers and uncles and aunts sup-
ported me morally when we all went through very difficult days, thus
giving extra dimensions to our family ties.

As fate wants it, I found myself separated from my beloved wife Zifa and our daughters Janel and Rianne both while working on my Ph.D. dissertation and the present revision. Despite the distance that separated us I felt a magic inspiration radiate from them which has never ceased to reach me.

PART ONE

ARABIC GRAMMARS OF TURKIC

THE ARABIC LINGUISTIC MODEL APPLIED TO FOREIGN LANGUAGES

CHAPTER ONE

SUBJECT AND OUTLINE

INTRODUCTION

The subject of the present study is the method used by Arab grammari-
ans to describe Turkic languages. I shall use concepts from traditional
Western grammar to describe how Arabic grammarians applied their
grammatical system to a language other than Arabic. Here I outline the
subject of this study and the way I intend to deal with it.

1. GENERAL OBSERVATIONS IN REGARD TO
THE SUBJECT OF THIS STUDY

This study deals with the application of a foreign linguistic model to
the description of language. In the course of history, many languages
have been described with a linguistic model that had been developed on
the base of another, even a typologically very different language. In this
way, the Greek model formed the basis for a description of Latin,
which, in its turn, served as the base for the description of many lan-
guages. With some major adaptations and alterations this model is still
used as the main descriptive linguistic tool for people in the Western
world. As examples one may refer to so-called missionaries' grammars
of the languages of African and American peoples, set up according to
the rules and concepts that their compilers knew from Greek and Latin
grammar (cf. Auroux and Queixalos 1984). Notions and concepts
originally inherent in the descriptive model were taken over and
applied to the language in question in order to clarify its rules with the
use of a familiar model. Similarly, the grammatical principles of the
Greek/Latin tradition were, with some adaptations, applied by early
orientalists to both Arabic and Turkish/Ottoman in order to elucidate
the rules of these languages to learners and students in the West. As a
result, even the oldest Western grammars of Arabic and Turkish /
Ottoman do not differ significantly from modern manuals.
    There cannot be such a thing as one ultimate descriptive model for
all languages. Apart from advantages, each model will have its short-

comings. Ideally, the advantages and shortcomings of several descrip-
tive models should be analysed objectively. In practice, however, such a
comparison is not an easy task. One has to understand fully the no-
tions and concepts of all models involved, in addition to having a
thorough knowledge of the language that is the subject of description.
In this respect, it is obviously wide of the mark to assume that certain
concepts of the model one is most familiar with are universal or basi-
cally obvious to anyone confronted with a certain language. It seems
equally wrong to consider a certain categorisation of phonemes or ele-
ments as, objectively speaking, the most suitable.

Of course, there are many examples of this. If, for instance, the ver-
bal form *ḍaraba* 'he hit' according to the Arabic model contains a hid-
den *(mustatir)* agent, this does not mean that it does so *objectively
speaking*. In the Arabic model this is only assumed in order to facilitate
the analysis of certain features related to word order in Arabic.
According to the Greek/Latin model, for example, there is no implica-
tion of an agent included in this form whatsoever. Similarly, the inter-
pretation of the Turkic ending **-dan** as an ablative is no better or worse
than regarding it as a particle or a postposition.

To give yet another example, in Russian manuals Turkic languages
usually have seven cases. This seventh case, the instrumental, consists
of various versions of the word **bilā** 'with' (see for example Kazakh:
SKJa 1962: 176ff). In Western grammatical practice, however, all
Turkic languages have six cases only, and here one finds **bilā** ranged
among the category of 'postpositions'. One could argue that the
Russian analysis is 'wrong', because **bilā** *is not* a case ending. On the
other hand, this analysis very well serves the needs of Russian students
of Turkic languages, who are accustomed to the concept of an instru-
mental case. In all instances, thus, each explanation is satisfying for
those who can place it in an overall linguistic concept, and the result of
any analysis depends on the extent to which one is familiar with this
type of analysis.[1]

Many traps lie in wait when giving a survey of a description of a lan-
guage one is familiar with by means of an unfamiliar model. This is es-
pecially true if the model that is the object of analysis has not been

---

[1] The same holds for the application of the Arabic system of verbal patterns along
with their associated meanings to modern Arabic languages, e.g. Egyptian. Although the
patterns may show a resemblance, the meanings associated with them may differ (the
Vth verbal pattern, /tafaᶜᶜala/, for example, frequently used to denote a reflexive action in
Classical grammar, often denotes a passive in modern Arabic languages). Furthermore,
new patterns have developed in these modern Arabic languages, which do not exist in
Classical grammar.

fully adapted to the language that is being described. In such a case one would compare his own fully developed method with a partially undeveloped system; one that is as yet incapable of giving a satisfying analysis of the new language's features. In this respect one could also compare a modern description of Turkish by, e.g., Lewis (1984 [1967]), with one of the ancient eighteenth century Western grammars of Ottoman.

Although both descriptions are based on the same principles, modern turcology has developed an innovative terminology derived from existing concepts (e.g., the terms 'aorist' and 'gerund' stand in original Greek/Latin terminology for similar, but different notions, and in the Fundamenta the new term 'Konverb' is proposed to replace 'gerund') in order to match certain features of Turkish. In the study of linguistic concepts the terminology is only one aspect.

If one were to base a study of the feasibility of describing Turkic with the Western model solely on ancient grammars of Ottoman, the conclusion would be obvious: the model is not particularly adequate for such a description. But, in doing so, one would disregard the possibility that at a later stage the same model could yield a more adequate description. In this sense, it is important to note that in regard to the descriptions of Turkic with the Arabic model, the grammarians, too, probably stood at the beginning of a tradition in which certain unfamiliar concepts had not yet been thoroughly worked out.[2]

Another trap is the possibility that one could find concepts that are apparently similar to concepts in his own system, and attribute them in this way to the second descriptive model. An example of this is the attribution of the concept of, for example, 'preposition' to Arabic grammatical theory, although the mere fact that Arabic linguistics recognises a group of elements that resemble in their syntactical posi-

---

[2] In this respect it is perhaps interesting to refer to Kerslake's (1994) analysis of two 19th century Ottoman grammars of Turkish. In her article Kerslake takes Western concepts of case, mood, and certain specific verbal endings like gerunds in their special context for Turkic linguistics as starting points for her analysis. Although she recognises the fact that both grammars were in essence set up according to the principles of Arabic linguistic theory, her article reflects a feeling of dissatisfaction with the way the grammar of Turkish is elucidated. Typical of this attitude is the following passage: "The corresponding discussion in the *Mikyās* is much more ambitious, but basically misguided. In one of the most tortuous, opaque passages of the book, Fevzī Efendi struggles desperately to force Turkish sentence structure into an Arabic straitjacket" (158). In this context the expression 'basically misguided' is especially striking, for it seems to convey the opinion that, whatever the efforts of the grammarians, it is *per se* impossible to give an accurate description of Ottoman using the Arabic linguistic model.

tions and effects the class of prepositions in the sense of Western grammar does not mean that the two concepts are essentially the same in all aspects.

A different issue is whether it is legitimate to call those elements prepositions, or to consider certain verbs, such as *kāna*, copula. I think that it is legitimate to do so, if it helps understanding aspects of Arabic, but one should of course not argue that Arabic theory had the same concept of preposition or copula.[3]

2. THE SUBJECT OF THIS STUDY

Usually studies of the Arabic linguistic system concern themselves — quite legitimately, of course — with the description of the Arabic language by the Arabic grammarians. The metalanguage used in these studies is based on the traditional Greek/Latin linguistic system. In this sense they deal with two descriptive models and one language that is subject of both descriptions.

The subject of the present study is to give an insight into the way Arabic grammarians applied their model to Turkic. In order to accomplish this task, two steps are required. First, some aspects of Arabic and Arabic terminology must be described within the frame of the traditional Western grammatical system. Second, it is necessary to explain in 'Western' terms the grammatical features of the language that is described by the Arabic grammarians. Only then is it possible to give a description of the way Arabic grammarians approached Turkic languages.

Schematically the subject of the present study can be reflected as follows:

---

[3] Here we may refer to the translation of the Arabic term *istiqbāl* (e.g, MS 366,16) with 'aorist' instead of 'future tense' in Dankoff and Kelly's translation (1982-5) of *Dīwān Luġāt at-Turk*. The term 'aorist' is a turcological term for the verbal ending -Ir, and therefore a direct interpretation of the Turkic material rather than a translation of the term *istiqbāl* (or its synonym *ġābir*) (This is recognised by Dankoff and Kelly 1985 III: 313). The term *istiqbāl* in *Dīwān* alternates with *ġābir* (MS 18,14) — Dankoff and Kelly's emendation for *ʿābir* 'past tense' is probably correct; cf. Versteegh 1993:24; also Biesterfeldt 1990; Talmon 1997]).

Greek/Latin Linguistic System

| |

serves to describe   *serves to describe*

|

Arabic Linguistic System |

|

serves to describe |

| |
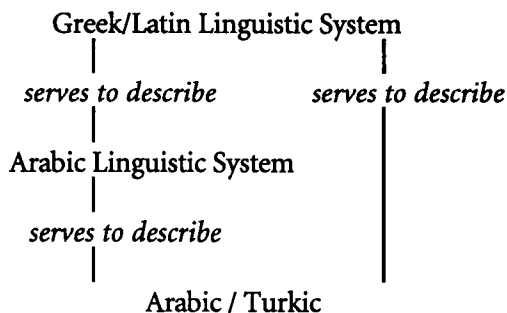
Arabic / Turkic

It follows from the discussion in the previous section that the fact that
I have chosen the Western grammatical system as an instrument to de-
scribe Arabic grammatical theory, Arabic, and Turkic languages does
not imply that I consider it the best or most suitable model for such a
description. The reason for this choice is merely the fact that the meta-
language and the concepts of the Greek/Latin system are familiar both
to the reader and myself. As a matter of fact, a similar study could be
carried out by putting another linguistic system in the place of the
Greek/Latin model.

3. RELATION TO OTHER DISCIPLINES

Two disciplines are directly involved in the present study — i.e. the
study of the Arabic grammatical theory and, to some extent, Turkic
linguistics. Furthermore, its results may be also be of interest from the
perspective of general linguistics. I propose to sketch out below the im-
portance of these disciplines in regard to this study.

3.1 *Relation to Arabic linguistics*

The relationship of the present study and Arabic linguistics is obvious.
The descriptions of Turkic are based on Arabic linguistics and stand
firmly in this tradition. However, since this study is an analysis of the
application of Arabic concepts to Turkic languages, there will be rela-
tively little possibility for a discussion of details. Moreover, the sources
themselves seem to stand quite isolated within Arabic linguistics and
there are no traces of discussion of certain topics related to Turkic
grammar between the authors. Therefore, the aim of the study cannot
be to give surveys of different opposing opinions: instead, it will be
limited to general sketches of the theoretical background of certain

ideas, in order to understand the analyses the authors of our sources
give of the Turkic data.

Since most of the sources — the exceptions being *Dīwān Luġāt at-
Turk* (469/1077) and *Šuḏūr* (1028/1619) — were compiled in or about
the 14th century, the discussion of subjects related to Arabic linguistics
will be limited to the opinion of so-called 'late grammarians' in the
Arabic linguistic tradition, with special interest in the views of ʾAbū
Ḥayyān al-ʾAndalusī, the author of *Kitāb al-ʾIdrāk li-Lisān al-ʾAtrāk*,
which will be one of the major sources for this study.

The study will point out how flexible the Arabic system is for a de-
scription of Turkic. In some instances the grammarians seem to get
confused[4] by some unfamiliar aspects of Turkic, whereas in others they
find apparent agreements with Arabic which they do not fail to inter-
pret in a similar way. In Chapters Five and Six I intend to give an ac-
count of this based on the statements the Arabic grammarians make
on declensional endings in Arabic and Turkic.

In his studies on Arabic linguistics, Owens (1988 and 1990) argues
that Arabic grammatical theory is a type of dependency grammar. In a
dependency grammar the elements of speech are hierarchically struc-
tured within a sentence and can govern each other. According to
Arabic grammar the governed elements show by means of a set of
markers *(ʾiʿrāb)* both the fact that they are governed and, further, by
which type of governor. The results of this study will give more evi-
dence for the assumption that the Arabic grammatical theory is a de-
pendency grammar based on characteristics of Arabic itself. (Further
discussion in Section 3.3.)

3.2 The contribution of turcological studies

As I have pointed out above, this study relates in the first place to
Arabic linguistics. In view of the fact that it deals with Turkic as well, it
is related to turcological research.

The interest of turcologists in regard to the sources used in this study
has been concentrated on the compilation of the language material
and the reconstruction of the languages described in the sources, rather
than on their methodological context and the model with which the

---

[4] This 'confusion' is illustrated in Chapters Five and Six, in which I point out that, for
example, the dative case ending ḳä/ġä is analysed in two ways. In the first place it is
regarded as an equivalent to the Arabic particle *(ḥarf)* ʾilā 'to', and in the second place it
is a 'marker of the object' *(ʿalāma al-mafʿūl)*. To my opinion, this 'confusion' is not the
result of being 'basically misguided' (see quotation from Kerslake [1994] in note 2
above).

language in question is described. This may be illustrated by statements of Telegdi, who, in his article on *Qawānīn* (1938:282), criticises the anonymous author for his erroneous categorisation of the material:

> An einer... Stelle erfahren wir, dass ähnliche Werke schon vor dem seinen vorhanden waren... In diesen älteren Schriften mögen die irrigen Regeln und Beobachtungen über türkische Grammatik gestanden haben, die im *Qawānīn* mehrmals erwähnt und berichtigt werden.

Although a generation later Dankoff and Kelly display quite a different approach in their thorough and valuable studies on Kāšgarī's *Dīwān*, they, too, limit themselves to occasional references to Western studies on Arabic grammar.

In this study the analysis of the Turkic language material will be of secondary importance. Instead, the starting point is to consider the sources in the first place as documents of Arabic linguistic thinking, written by specialists, who had a profound insight into the rules of the Arabic language and Arabic linguistic theory. In spite of possible mistakes and, perhaps, an occasional case of imperfect knowledge of Turkic, there is no reason to doubt their scientific accuracy or their good intentions. They depended, like every scholar, on the reliability of their sources and their own intellectual capacities.

Within this framework I shall not attempt to verify all statements of the grammarians in regard to the features of the Turkic languages they describe in contemporary studies on Turkic languages, nor elaborate on questions concerning the type of language (Qipčaq-Oğuz),[5] the possible Turkic provenance of the authors, or engage in a discussion on specific features such as vowel length in Turkic. Likewise, apart from occasional references in Chapter Two, the historical context of this type of sources included in this study can only in a very general sense be the subject of our discussion. In regard to the Turkic language material, our principle is that it is basically sound, leaving the historical reconstruction to others.

In each chapter of this study the relevant aspects of Turkic languages will be analysed in a general sense, but based specifically on data provided by the sources. These analyses are meant to give non-turcologists an insight into the morphological and syntactic structure of the Turkic data. In this respect, all Turkic and Arabic examples alike will be followed by a word-by-word analysis. I shall concentrate on the descriptions of Turkic rather than on the Turkic data itself. In Chapter Three,

---

[5] For a discussion and a bibliography of Qipčaq and Oğuz manuscripts see Flemming (1977a and 1977b).

for example, I make some suggestions as to the interpretation of the labels for velarisation and palatalisation. The distribution of these labels can be explained with the aid of principles that were elaborated in Arabic linguistics. Chapter Three also deals with the question of why the Arabic grammarians regarded it as necessary to label Turkic words for either velarisation or palatalisation, and shows that the markers are put in in the first place for morphological and semantic reasons, rather than for purely phonetic purposes, and my interpretations are indications rather than attempts to give an exact phonetic reflection in each instance

I have opted for a rather basic transcription/transliteration of Turkic, which is sufficient for the description of morphological and syntactic features. This transcription system is based on the information given in the sources themselves, and remains close to the way the consonants and vowels are reflected in Arabic script. For example, the sources use the same grapheme for both [g] and [k]; they are occasionally distinguished with three diacritical dots.[6] In the basic transcription used in this study [k] is represented by k, and [g] by ḳ, respectively.

The limited scope of this study does not allow for discussions in regard to the respective places of the sources and the data they contain in other studies. References to philological works will be made as indications for further study to the reader, and will only occasionally be included in discussions in the present framework.

3.3 *General Linguistics*

Since two languages and two descriptive systems are involved in our study, it may also yields some results in regard to general linguistics. My main hypothesis is that some concepts of the Arabic linguistic theory are based on the perception of some features of Arabic itself, in particular their conception of case and declension. Furthermore, this study discusses the distinction between semantic and syntactic case which was formulated earlier for Russian and its results for Arabic and Turkic.

3.3.1 Phonology and phonetics

Chapter Three deals with phonological and phonetic descriptions of Turkic as perceived by the Arab grammarians. It consists of two main parts, the first part discusses the vowels, the second deals with the con-

---

[6] This approach is similar to that of Devereux in his articles on Nawā'ī's *Muḥākama al-luġatayni* (1964 and 1965).

sonants. The discussion of the vowels is quite complicated, since in Arabic theory the quality of the vowel is to a large extent made depen-dent of the neighbouring consonants.

The main subject in the first part of this chapter is how the Arab grammarians distinguished between back and front vowels, and why they did so. The latter of these questions is of interest too, for what seem reasonable or obvious points, such as, e.g., vowel harmony, in modern western linguistic thinking might not have been so in the framework of Arabic theory.

In general my conclusion is that the Arab grammarians apply the phonetic and phonological concepts they had elaborated for Arabic, which in itself is not surprising. The range of 'standard' phonemes the Arab grammarians posit for Turkic include, apart from the regular Arabic consonantal phonemes, some allophones they find in non-standard pronunciations of Arabic. To this range they add a number of phonemes that they produced by elaborating on notions — such as ve-larisation — which they for Arabic applied in a very limited phonologi-cal context. Chapter Two further gives an interpretation of labels the Arab grammarians use for indicating velarisation and palatalisation of consonants. The Arab grammarians applied these labels in the same way they were used for Arabic, which has consequences for the pho-netic interpretation of Turkic words that contain r, l — and, possibly, z — written in Arabic script. The principles the Arab grammarians give with regard to the pronuciation of words, in particular the distinction between front and back words, were probably not meant as an accu-rate describe Turkic phonemes, but rather intended as a set of criteria for the morphological problem of selecting the right verbal and nomi-nal suffixes whose shape differs according to the phonological context.

Topics that relate to vowels, such as vowel harmony, are only marginally touched upon. The phonetic descriptions of vowels, such as [ö] and [o] are mainly given in terms of the consonantal context.

3.3.2 Syntax: A functional or a formal approach?

The Chapters Four, Five and Six deal with the question why five of the seven endings of Turkic case are not incorporated in the Arabic con-cept of case, whereas two others seem to fit their criteria. In order to give an impression of the direction which my argumentation takes in those chapters, I shall summarise the main hypotheses here.

It shall be seen that the Arab grammarians treat Turkic case endings in two ways. In the first place, the accusative is dealt with in more or less the same way as Arabic case. The Turkic accusative suffix is consid-

ered to have the same or a very similar function as its Arabic equiva-
lent, i.e. a 'marker'. Second, Turkic genitive, dative, locative and abla-
tive are interpreted as particles, i.e. governing elements. The only case
that is not explicity described is the nominative (which in Turkic takes
zero ending and in Arabic -*u*; my hypothesis is that it is treated similar
to Arabic nominative).

One could conceive of this approach as a functional one. This point
of departure correlates quite naturally with the respective functions of
the Turkic case endings, and appears to elude any association with
Arabic case. One could assume, therefore, that the Arab grammarians
took in account only the superficial functional resemblances between
Turkic case and Arabic particles, considering the Arabic model merely
partly applicable to Turkic.

The choice for a functional approach rather than a formal one, how-
ever, presupposes three conditions which, I believe, are not pertinent
here. In the first place, it demands the availability of different ap-
proaches, paired to an awareness that it is at all possible to choose or
shift between them. Secondly, it presupposes the possibility to see the
formal aspects of grammar, i.e. the hierarchical relationships between
the syntactic elements, detached from their functions. In the third
place, and related to the former two conditions, it presupposes that
certain key concepts, such as that of underlying structure and gover-
nance, which dominate Arabic linguistic thinking, for some reason
should not apply to Turkic.

It is my contention that the approach of the Arab grammarians is
not based on a free scholarly choice between a formal or a functional
approach, but the consequence of their perception of certain features
of the Arabic language which inevitably are incorporated in their con-
cept of how a language should be analysed. In this respect I intend to
show that the Arab grammarians considered their concepts basically
universally applicable. In concrete terms this means that on theoretical
level they posited the same hierarchical relationships of governance for
the Turkic phrases as the ones they had worked out for Arabic.

An important lead in my hypothesis is the distinction between seman-
tic and syntactic case and the typology of Turkic and Arabic case. In
Western theory case covers both types of case. This broad concept not
surprisingly matches the typological characteristics of Greek and Latin
which have a mixture of semantic and syntactic case. From these two,
semantic case contributes to the sentence's semantic load, whereas
syntactic case only serves to indicate the function of a word in a sen-
tence. More specifically, syntactic case only occurs if the appropriate

governor is also present in the phrase. Syntactic case may drop out due
to prosodic or morphonological reasons; it even can be posited on a
theoretical level for languages that normally do not have case in surface
structure, such as English. Semantic case, on the other hand, does not
need a governor in order to appear. It only occurs in languages with
case in surface structure and cannot be omitted without an inevitable
distortion in the meaning of the sentence.

It makes therefore sense to assume that the traditional western the-
ory does not distinguish between the two types of case because Greek
and Latin case is of both types. Western theory, therefore, developed a
rather broad concept of case.

Let us suppose that a given language only, or basically posesses syn-
tactic case and other declensional endings that can be accounted for in
a similar way. If scholars would develop a theory of case solely based on
that language, their concept of case is likely to cover syntactic case
only. If such a theory subsequently is applied to a language with seman-
tic case, it would not recognise that as 'case'.

In this way many features of Arabic theory can be accounted for. As
their theory is built upon the syntactic cases of Arabic, the Arab gram-
marians' concept only admits syntactic case. Arabic case is basically
syntactic, and the most suitable way to account for syntactic case is by
means of a theory in which governance plays a role. The Arabic theory
of case (and verbal inflection) explains all occurrences of case, includ-
ing the nominative, in terms of hierarchical relationships of governing
elements, and case endings are basically conceived of as markers of
governance. In essence, *ʾiʿrāb* is set up around a concept of declen-
sional endings as a sort of syntactic case.

Quite naturally, the five semantic cases of Turkic (GEN, DAT, LOC
and ABL) hardly match this particular concept of case. In most in-
stances, though, their functions neatly agree with those of the Arabic
'particles', and they are assigned similar governing functions.

For these reasons, I do not think that the Arab grammarians had a
functional rather than formal approach toward Turkic. This does not
mean, though, that they never had to adapt their describing system
when they applied it to Turkic. There is a number of instances in
which they do, and three of them are discussed in this book. We shall
see, for example, that issues such as whether a the Turkic equivalent of
'with' *(maʿa)* should be regarded as a noun or a particle, Turkic case +
postposition, and the indirect object cannot entirely be solved within
the Arabic concept. In these instances one could, perhaps, speak of an
innovating and, in some respects, a functional approach.

4. TRANSLATION OF TERMINOLOGY

The translation of Arabic linguistic terminology should, like any trans-
lation, follow the original as closely as possible. In practice this is not
always possible. There are two basic options. First, one may use terms
already existing in Western grammar. The use of an existing term as a
translation for notions and concepts from Arabic linguistics carries
with it the danger, apart from being imprecise, of identifiying the
translation with the translated term, which may lead to confusion and
misinterpretations (I have already touched upon this above). The ad-
vantage of this approach is that the translation immediately engenders
a desired association with an existing concept. As long as the translat-
ing term is incorporated in an analytic context, the advantages will
outweigh the disadvantages.

The other option, the introduction of new terms (such as, for ex-
ample, 'operates on' for *yaʿmalu* [cf. Carter 1981]), has the advantage
of being closer to the text, but the disadvantage of being more opaque.
There is also a third option, i.e. refraining from a translation, giving a
description instead. This option, however, is possible in a few instances
only, e.g., a-inf for 'a-inflection', i.e. accusative or subjunctive mood,
both of which end in *a* (Owens 1990).

In this study I have chosen the former option and adopt existing
terms to express Arabic notions. In this way, I translate the Arabic
term *yaʿmalu* as 'governs' and *ẓarf* as 'locative'. Further, I have chosen
to translate *taqdīr* as 'underlying level', by which I mean to express
similarity to, rather than equivalence with the term 'deep structure' in
Generative Linguistics. I am, on the other hand, reluctant to use the
Western term 'preposition' as a translation for *ḥarf ǧarr*, because this
might entail the inclusion of *ẓurūf* 'locatives'. The terms *rafʿ*, *naṣb* and
ǧarr are translated as 'nominative case', 'accusative case' and 'genitive
case', respectively, as far as they are applied to nouns.

In this book the term 'Turkic' will be used frequently as an indis-
criminate reference to the languages described in our sources. In this
study it translates the term *turkiyya*, the term used in the sources for
their particular variant of Turkic. As a term, Turkic comprises all
Turkic languages, old and modern alike, covering both a long period in
time and a large geographical area. The languages range from the eighth
century Runic inscriptions, to modern languages such as, for example,
Turkish, Saryġ Yuġur and Tuva. In regard to the sources used for this
study, the term *turkiyya*, covers 11th century Ḥaqānī-Turkic, variants
of Qipčaq and Oġuz, and also early 17th century Ottoman. (One could

assert that the modern term 'Turkic' comprises all Turkic languages and none in particular, and that as such 'Turkic' cannot be described.) In this sense the application of the term *turkiyya* in the sources is much like that of *ʿarabiyya*. This term refers to an ideal, in some ways perhaps an artificial language, i.e. Arabic or Classical Arabic, but nevertheless one single language which can be described and for which rules have been elaborated.

5. NOTES ON TRANSCRIPTION

For both Arabic and Turkic the following transcription is used:

ء	*hamza*	ʾ	ض	*ḍād*	ḍ	Non-standard		
ب	*bāʾ*	b	ط	*ṭāʾ*	ṭ	consonants:		
ت	*tāʾ*	t	ظ	*ẓāʾ*	ẓ	چ	č	
ث	*ṯāʾ*	ṯ	ع	*ʿayn*	ʿ	پ	p	
ج	*ǧīm*	ǧ	غ	*ġayn*	ġ			
ح	*ḥāʾ*	ḥ	ف	*fāʾ*	f			
خ	*ḫāʾ*	ḫ	ق	*qāf*	q			
د	*dāl*	d	ك	*kāf*	k			
ذ	*ḏāl*	ḏ	ل	*lām*	l			
ر	*rāʾ*	r	م	*mīm*	m			
ز	*zāy*	z	ن	*nūn*	n			
س	*sīn*	s	ه	*hāʾ*	h			
ش	*šīn*	š	و	*wāw*	w			
ص	*ṣād*	ṣ	ي	*yāʾ*	y			
			ا	*ʾalif*	"			

Vowels:

ـَ	*fatḥa*	a
ـُ	*ḍamma*	u
ـِ	*kasra*	i

The Arabic *hamza al-qaṭʿ* is reflected in the transcription, e.g., *ʾiʿrāb;* whereas the phonological *hamza (hamza al-waṣl)* is not, e.g., *al,* and

ism. The Arabic article *al-* assimilates in transcription with the so-called *ḥurūf šamsiyya* in both nouns and proper names, e.g., az-Zamaḫšarī.

The sequence /uw/ is transcribed as *ū*, /iy/, as *ī*, and /a"/ as *ā*, respectively.

In Chapter Three Turkic words from the sources are always in boldface, Arabic in italics. Other Turkic items are plain. Phonetic transcriptions are put between square brackets [ ]. Phonetic vowel length is occasionally indicated with a semicolon [:]. Morpho(no)logical reconstructions are put between slashes //. Capitals refer to (archi)phonemes with different phonetic realisations. For example, in Turkic languages the (archi)phoneme G can be realised as [ġ] or [g], respectively.

For further details on transcription and the interpretation of the Arabic alphabet for Turkic, the reader is referred to Chapter Three.

CHAPTER TWO

THE SOURCES

INTRODUCTION

This chapter presents the ten sources upon which this study is based. In
the first section I attempt to shed some light upon such matters as the
date of compilation, author, size, content, the edition used in this
study, other editions of the text, primary sources and the Turkic lan-
guage that is described. In the second section I develop the data on the
works that served as primary material for the sources. In this respect, I
distinguish between a. direct evidence, i.e. references to both primary
Turkic material and primary Arabic material, and b. indirect evidence,
i.e. the internal structure of the sources. Finally, in the third section, I
evaluate the findings of the preceding sections and decide which of the
sources fit best within the framework of the present research. In this
respect I point out that it is important to determine the tradition to
which each source belongs, i.e. the Arabic grammatical or the lexico-
graphical tradition. In order to be able to study the way Arabic gram-
marians described Turkic languages, it is essential that the works on
which I have based this study belong to the same tradition.

1. THE SOURCES

Arabic linguistic treatises on Turkic languages have long been the sub-
ject of Turcological studies. This type of study, however, is often associ-
ated with grammars and glossaries of Qipčaq-Turkic that date from
Mamlūk times. Here I wish to approach the subject from a different
angle: I intend to analyse the way grammarians who were educated in
the Arabic linguistic tradition described Turkic languages. For this
study I have examined ten treatises compiled at different times and
under different circumstances, ranging from a large lexicon compiled
in 11th century Bağdād down to a word list compiled in the Ottoman
Empire at the beginning of the 17th century. The ten works which
form the core of this study are:

1 *Dīwān Luġāt at-Turk*
2 *Kitāb Ḥilya al-Insān wa-Ḥalaba al-Lisān*
3 *Kitāb al-ʾIdrāk li-Lisān al-ʾAtrāk*
4 *Kitāb at-Tuḥfa az-Zakiyya fī l-Luġa at-Turkiyya*
5 *al-Qawānīn al-Kulliyya li-Ḍabṭ al-Luġa at-Turkiyya*
6 *Kitāb Tarǧumān Turkī wa-ʿArabī wa-Muġalī*
7 *aš-Šuḏūr aḏ-Ḏahabiyya wa-l-Qiṭaʿ al-ʾAḥmadiyya fī l-Luġa at-Turkiyya*
8 *Kitāb Bulġa al-Muštāq fī Luġa at-Turk wa-l-Qifǧāq*
9 *ad-Durra al-Muḍīʾa fī l-Luġa at-Turkiyya*
10 The Margin Grammar.
The sources are described in greater detail below.

1.1 Dīwān luġāt at-turk

Dīwān Luġāt at-Turk, compiled by Maḥmūd bn al-Ḥusayn bn Muḥammad al-Kāšġarī between 464/1072 and 469/1077, is the oldest known and most famous dictionary of Turkic. The unique manuscript of *Dīwān*, a copy made from the autograph by Muḥammad bn ʾAbī Bakr (Damascus 664-5/1266), is kept in Istanbul in the Millet Genel Kütüphanesi (Fatih - Ali Emiri No. 4189). Apart from numerous lexical entries, it contains much ethnological and geographical information about the world as Kašġarī knew it and, moreover, it contains the oldest map known of Central Asia.

Of the numerous publications on *Dīwān* I mention only a few. *Dīwān* was first edited in print by Kilisli Muʿallim Rifʿat (Bilge) (1917-1919). Brockelmann based his *Mitteltürkischer Wortschatz* (1928) on this edition. In 1939-1941 Atalay published a translation into Turkish which was followed in 1942 by a barely legible facsimile edition of the MS. In the Soviet Union, Mutallibov published a translation into Uzbek (1960-1963). Clauson's *Etymological Dictionary of Pre-Thirteenth-Century Turkish* (1972; hereafter EDT) depends to a large extent on Rifʿat's edition of *Dīwān*. Between 1982 and 1985, Dankoff and Kelly edited an elaborate translation of the manuscript into English. Finally, in 1990 the Turkish Ministry of Culture published a new, full colour facsimile edition of the MS, which has contributed very much to the accessibility of the work. All references to *Dīwān* in the present study are based on this edition.

Kāšġarī was of Qarahānid descent, the ruling class in Central Asia and Transoxania at the time.[1] He was born in Barsġān[2], near Issık-kul, but lived in Baġdād. Pritsak (1953), who attempted to reconstruct Kāšġarī's genealogy, concludes that the year of his birth would have been some time between 420/1029 and 429/1038. With regard to Kāšġarī's position between his Islamic education and Turkic roots we cite Dankoff and Kelly (1982 I:4):

> Being a Turk of noble stock and a Muslim deeply educated in the Arabic humanities, he was able to understand both the native tribal tradition and the court Islamic tradition; his aim, indeed, was to interpret the former in terms of the latter... Since Diwan is dedicated to the caliph al-Muqtadī ([467-487 AH] 1075-94) [cf. Dīwān 3,10-17], its immediate purpose was perhaps to explain to the Abbasid court at Baghdad the language and customs of their Seljuk overlords.

In Dīwān (18,10) Kāšġarī refers to a work he wrote on Turkic gram-mar, entitled Ǧawāhir an-Naḥw fī luġāt at-Turk which, unfortunately, has not been preserved. He tells us that before compiling Dīwān he had already treated morphological subjects like the plural (ǧamʿ), the elative form (tafḍīl), and diminution (tasġīr) in this work.

Kāšġarī's aim in writing Dīwān is clear enough; he intends to show that the Turkic languages deserve to be studied properly. He even com-pares Turkic and Arabic metaphorically to two racehorses that try to keep up with each other (cf. Dīwān 595).

Further Kāšġarī tells us he travelled among the Turkic peoples and learned their languages and dialects (Dīwān 3,2-3). Dīwān's immense vocabulary has inspired the language reformers of the thirties and for-ties of the present century in the Turkish Republic in their search for new Turkish equivalents for Arabic and Persian loan-words.[3]

Dīwān's internal structure resembles the structure of an Arabic lexicon rather than a grammar. Kāšġarī mentions al-Ḫalīl's famous dictionary Kitāb al-ʿAyn:[4]

---

[1] Cf. Hazai "Kāšġarī "in EI².

[2] An indication for this may be the fact that he situated Barsġān and not Kāšġar in the middle of the famous map in Dīwān (22-23). At the same time Kāšġarī depicts the Turks of Barsġān in quite negative terms (cf. Dankoff 1972: 26).

[3] The respect for Dīwān has led many scholars, especially in Turkey, to suppose that Dīwān has been the basis for later Arabic grammars of Turkic languages (Cf. Ülkütaşır 1946: 28; Tomanov 1965). In fact, there is no evidence to sustain this assumption. I shall discuss this topic in Section 2 of this chapter.

[4] Al-Ḫalīl or Ibn 'Aḥmad ʿAbd ar-Raḥmān al-Ḫalīl al-Farāhīdī (d. 175/791) who is said to have compiled the first dictionary of Arabic (cf. Wild 1965; Talmon 1997). The entries in Kitāb al-ʿAyn are arranged according to the respective places of articulation of

"It was in my mind to arrange my book like al-Ḫalīl arranged Kitāb al-
ʿAyn…". (wa-laqad taḥālaǧa fī ṣadrī ʾan ʾabniya l-kitāb kamā banā l-ḫalīl
Kitāb al-ʿAyn…, Dīwān 4,17).

Nevertheless, there is no doubt that Kāšġarī based the structure of his
work on the Arabic lexicon Dīwān al-ʾAdab fī Bayān Luġa al-ʿArab
compiled by ʾAbū ʾIbrāhīm ʾIsḥāq bn ʾIbrāhīm al-Fārābī (d. 350/961).[5]
The influence of Dīwān al-ʾAdab on Dīwān Luġāt at-Turk is so over-
whelming that the impact of Kitāb al-ʿAyn, if any, must be considered
as secondary. We discuss this point in the second section of this chap-
ter.

Dīwān Luġāt at-Turk consists of eight parts, each of which bears the
title 'book' (kitāb), the first six bear the same titles as their equivalent
parts in Fārābī's work. The titles of the eight 'books' are as follows: 1.
Kitāb al-hamz (MS 29-91): words with initial hamza (i.e. the glottal
stop /ʾ/); 2. Kitāb as-sālim (159-406): words having sound consonants;
3. Kitāb al-Muḍāʿaf (406-445): words containing a geminate conso-
nant or two identical consonants; 4. Kitāb al-Miṯāl (445-493): words
having an initial weak consonant, i.e. wāw or yāʾ; 5. Kitāb ḏawāt aṭ-
Ṭalāṯa (493-535): words having a medial weak consonant, i.e. wāw,
yāʾ or ʾalif; 6. Kitāb ḏawāt al-ʾArbaʿa: (535-599): words having a final
weak consonant.[6] Finally, Kāšġarī adds two titles to the existing struc-
ture, namely 7. Kitāb al-ġunna (599-622): words with /ŋ/ or /nč/
(78,8: al-ġunna al-ǧīmiyya) and 8. Kitāb al-ǧamʿ bayn as-sākinayn
(622-638): words containing clusters of consonantal sounds that do
not exist in Arabic.[7]

The titles of the chapters 4. Kitāb al-miṯāl, 5. Kitāb ḏawāt aṯ-ṯalāṯa
and 6. Kitāb ḏawāt al-ʾarbaʿa, seem incomplete. The term al-miṯāl
originally means 'example', 'model' or perhaps 'pattern'. The term
ḏawāt (plural of ḏāt fem. of ḏū—meaning both 'self' and 'owner') can
theoretically be conceived as a term for 'radicals' in general but in itself
none of these terms expresses anything with regard to weak conso-
nants.

---

the consonants of the root pattern. According to this structure the ʿayn is the first conso-
nant in the sequence, because it is pronounced in the back part of the troat.
    [5] Bergsträsser (1921) was the first to point out that Kāšġarī had applied Fārābī's
system.
    [6] In Fārābī's work the sequence of the 'books' is: 1. Kitāb as-sālim; 2. Kitāb al-
muḍāʿaf; 3. Kitāb al-miṯāl; 4. Kitāb ḏawāt aṯ-ṯalāṯa; 5. Kitāb ḏawāt al-ʾarbaʿa; 6. Kitāb
al-mahmūz.
    [7] For a detailed discussion of each chapter see Dankoff & Kelly (1982: 31-40) and
Kelly (1976).

Fārābī explains the meaning of these terms (*Dīwān al-ʾAdab* I 76,10[8]): *miṯāl: mā kānat fī ʾawwalihi wāw ʾaw yāʾ* "that of which the first radical is *wāw* or *yāʾ*"; *ḏū aṭ-ṯalāṯa:* "that of which the middle radical is a weak consonant" *(mā kānat al-ʿayn min-hu ḥarfan min ḥurūf al-madd wa-l-līn)* and, finally, *ḏū l-ʾarbaʿa:* "that of which the third radical is likewise" [viz. a weak consonant] *(mā kānat al-lām min-hu kaḏālika)*. This calls for further explanation. The terms *ḏawāt aṭ-ṯalāṯa* and *ḏawāt al-ʾarbaʿa* are Kūfan synonyms for the known terms *ʾaǧwaf* and *nāqiṣ*, respectively. *ḏawāt aṭ-ṯalāṯa* is applied to the so-called 'hollow' verbs (*ʾaǧwaf*) in which the middle weak consonant is reduced to a short reflex of the original glide when the verb is conjugated for the first person singular. An example of this is the verb *qāma* 'he stood up' (>* *qawama*) - *qumtu* 'I stood up'. The middle short vowel *u* is a reflex of a supposed *wāw* in the underlying structure (cf. Bohas 1982:430). The other term, *ḏawāt al-ʾarbaʿa*, is applied to verbs whose last radical is weak (*nāqiṣ* 'lacking'). An example of this category is the verb *ḥakā* 'he told' - *ḥakaytu* 'I told'.

Like *Dīwān al-ʾAdab*, each book—except book eight—in *Dīwān Luǧāt at-Turk* consists of two parts; the first part treats the nouns (*ʾasmāʾ*) and the second part the verbs (*ʾafʿāl*). The Turkic entries are further divided into uniradical, biradical, triradical words and so on, up to septiradical.[9] Within each book the lexical stock is arranged according to the Arabic system of consonant and vowel patterns, e.g. *faʿl*, *fuʿl*, *fiʿl*, *faʿal*, *faʿul*, *fuʿal* (in which the consonants *f*, *ʿ* and *l* stand for any consonant). Within each chapter, words are arranged by the final consonant. Within the group of words with the same final consonant, they are arranged according to the first and following consonants.

We can illustrate this with an example. Let us try to find, for example, the word **kumuš** (kümüš) *al-fiḍḍa* "silver". **Kumuš** consists of three consonants and follows the pattern *fuʿul*. Its final consonant is š. Since its structure is not subject to irregularities, **kumuš** is most likely to be found in the chapter *Kitāb as-Sālim*, the sound words. Within *Kitāb as-Sālim*, the heading *ʾAbwāb al-mutaḥarrika al-ḥašw* (178,7) introduces words with a vocalised middle consonant (*ḥašw*), for which there are three main patterns: *faʿal*, *faʿul* and *faʿil*. I find **kumuš**, on 186,10 after the entry **kališ**. We shall deal with this in greater detail in Section 2.2.2.

---

[8] References in this book to *Dīwān al-ʾAdab* are based on ʾAḥmad Muḫtār ʿUmar's 1984 edition.

[9] For more details see Dankoff & Kelly (1982 I: 33).

In his work, Kāšġarī mentions as many as twenty Turkic peoples
(20,16-21) that neighbour *(tuġārī)* the Arabs (516,5). He states where
each tribe lives and he adds a map of Central Asia to illustrate the in -
formation. Kāšġarī even discusses the marks the Oġuz used to brand
their animals (40-41).[10]

As for the language variety, Kāšġarī seems to have had a preference
for one particular dialect:

> The language described is called 'Turkiyya' (here translated 'Turkic'); it is
> basically the dialect of the important Čigil tribe, belonging to the
> Qarakhanid confederation.... In particular, Kāšġarī gives equal weight to
> two main dialect groups: that of the 'Turks' (including Čigil, Tuxsi, etc.)
> and that of the 'Turkmān' or Oguz ('Ghuzziyya', often including the dialect
> of Qifčāq, etc.)... (Dankoff & Kelly 1982 I: 4-5).

A reason for his preference for the Oġuz dialect may have been the fact
that the Selġuq sultan in Baġdād at that time was of Oġuz descent.
Further, Kāšġarī considered the language of people who remained unaf-
fected by other languages (Persian) and who kept away from the cities
to be the purest (*'afṣaḥ*, *Dīwān* 24,1-27).[11]

1.2 Kitāb ḥilya al-insān wa-ḥalaba al-lisān

Ḥilya was written by Ġamāl ad-Dīn Ibn al-Muhannā in the fourteenth
century AD. It consists of three parts: Arabic-Persian, Arabic-Turkic
and Arabic-Mongolian. It was first edited in 1900[12] by Melioranski
and based on five different manuscripts, three of which were kept in
Oxford (Bodleian Library), one in Berlin (Royal Library) and one in
Paris (Bibliothèque Nationale), respectively. One of the manuscripts
lacked the Persian, and a second the Mongolian part. Kilisli Muʿallim
Rifʿat (Bilge) found a sixth manuscript at the Müze-i Hümayūn Library
in Istanbul which contains all three parts, and which he edited in 1921.
All references in this book to the text of *Ḥilya* are based on the Rifʿat
edition. In an article about the origins of Ibn al-Muhannā, the author
of *Ḥilya*, Doerfer (1976) refers to Sergei Malov, who showed that

---

[10] Dankoff (1972) discusses the tribal organisation of the Turkic peoples based on
data from *Dīwān*.

[11] For a discussion see Dankoff and Kelly (1982: I 20-21 and 44-48). Interesting to
note is that Kāšġarī does not mention mingling with speakers of Arabic as a negative in-
fluence. This may be due to the fact that he did not consider the influence of Arabic as
'bad', or that he was not aware of such an influence.

[12] *Arab Filolog o Turetskom Jazyke*, St. Petersburg.

Rifʿat's copy is the oldest and therefore the most original one.[13] Accor-
ding to Rifʿat, the manuscripts Melioranski used for his edition were
poorly legible, and contained no mention of the author; they even
have different titles, like *Tarǧumān Turkī wa-ʿArabī* and *Tarǧumān
Muǧalī wa-ʿArabī*.[14]

Rifʿat, in the preface to his edition of *Ḥilya* (p. *ḥāʾ*), suggests that
Ibn al-Muhannā be identified as ʾAḥmad bn ʿAlī bn Ḥusayn bn ʿAlī
bn Muhannā bn ʿAnba al-ʾAṣġar, also known as Ǧamāl al-Milla wa-d-
Dīn, who died at Karmān in 825/1424 (cf. GAL II 199). This Ibn al-
Muhannā was known as the author of a work entitled *ʿUmda aṭ-Ṭālib
fī ʾAnsāb ʾĀl ʾAbī Ṭālib*.

With regard to Ibn al-Muhannā's origins, Doerfer (1976: 251)
concludes that he must have originated from "dem Zentrum des alten
Chorasan..."[15] rather than Azerbaijan. In the same article Doerfer
summarises S. Širvani Yusif-Zia's earlier findings with regard to the pe-
riod the work was compiled. Led by a remark in the text with regard to
the existence of the Mongolian Empire (736/1336 - 758/1357) at the
time of its compilation, Yusif-Zia came to the same conclusion as
Melioranski, namely that *Ḥilya* must have been compiled in the 13th
century or the first half of the 14th century.[16] Doerfer agrees with this
assumption: "Nun hat aber die Mongolenherrschaft im vorderasiatis-
chen Raum nur bis 1336 bestanden bzw., die letzten Ausläufer, die
aber keineswegs mehr einen so «unwiderstehlichen» Eindruck mach-
ten, mitgerechnet, bis 1357" (Doerfer 1976:243).

There is another interesting point of discussion with regard to *Ḥilya*
and its author, namely the fact that Ibn al-Muhannā knew the Uygur
alphabet, at least two centuries after it had become obsolete for most
western Turkic languages (cf. *Ḥilya* 72). Other sources from the same
period do not mention the existence of the Uygur alphabet.[17] One
wonders why Ibn al-Muhannā would have taken the trouble to learn it

---

[13] Doerfer (1976:244) refers to 'Ibn-Muchanna o tureckom jazyke' *Zapiski Kollegii
Vostokovedov pri Aziatskom Muzee*, III, Leningrad [St. Petersburg] 1928: 221-248.

[14] These titles suggest a resemblance with *Tarǧumān Turkī wa-ʿAǧamī wa-Muǧalī*,
the Houtsma MS, see Section 1.6.

[15] Another indication of Ibn Muhannā's origin may be the fact that the names of Iraq
and Ḥurāsān occur in some exemplary sentences: "from Iraq to Ḥurāsān" (*min al-ʿIrāq
ʾilā ḥurāsān*, 91,13); "Our chief went to Ḥurāsān" (*ʾamīrunā mašā ʾilā ḥurāsān*,92,5).

[16] Doerfer refers to S. Širvani Yusif-Zia: "Nekotorye zamečanija otnositel'no Ibn-
Muxanny i ego sočinenija"in: *Struktura i istorija tiurkskich jazykov* Moskwa 1971. The
passage itself is as follows: *hāḏihi d-dawla l-muǧūliyya al-qāhira* "this strong Mongol
Empire" (*Ḥilya* 186).

[17] The only other source that does is *Dīwān* which dates from the fifth/eleventh cen-
tury.

and dedicate so much space to it. There may be two reasons for this. In the first place one might suppose that the work was written in an area in which the Uygur alphabet was still known. This may have been somewhere near Anatolia, where the Uygur alphabet remained in use until the sixteenth century (cf. Sertkaya 1973; 1975).[18] A second option is to suppose that Ibn al-Muhannā was a scholar, probably of Turkic descent, originating from Khurasān, who lived somewhere outside Anatolia, and who had learned the Uygur alphabet for scholarly reasons only.

Ḥilya's Turkic part in Rifʿat's edition comprises 120 pages (71-191). Ibn al-Muhannā divided his work into two parts (nawʿ): Part one tamhīd "introduction" (72), which also contains an elaborate phonetic introduction to the sounds of Turkic. It contains the following chapters: "declensions of the nouns" (taṣrīf al-ʾasmāʾ, 81); "particles" (al-ʾadawāt al-ḥarfiyya, 91); "verbs in the past tense" (al-ʾafʿāl al-māḍiya, 98); and, finally, a list of Turkic words (102-116). Part two deals with the "simple meanings" (al-musammayāt aṣ-ṣādiǧa, 127). It is divided into twenty-four sections; an Arabic-Turkic word list with the entries arranged according to semantic categories. Of special interest are section 21, on history (185), and a short Turkic-Arabic word list (188-191).

Ibn al-Muhannā had used several primary sources for his work: in one instance he refers to the author of Kitāb Nādir ad-Dahr ʿalā Luǧa Malik al-ʿAṣr (78,13). On two occasions Ibn al-Muhannā quotes Muḥammad bn Qays, the author of a work on Turkic dedicated to Ǧalāl ad-Dīn Ḥwārezm-šāh (cf. 93,16-94,2 and 101,15), who was the last ruler of the Ḥwārezm dynasties (d. 628/1231).[19] Muḥammad bn Qays may be the author of either one of the three works mentioned in Ḥilya or, perhaps, of a fourth one, since we do not know the title of the work he wrote (this may have been Tibyān al-Luǧāt at-Turkī [sic!] ʿalā Lisān al-Qanqlī; cf. Wittek 1928:174). Further, Ibn al-Muhannā mentions a work on Turkic entitled Kitāb Yaḥyā al-Malik (96,5). In his foreword to the edition of Ḥilya Rifʿat notes that in Melioranski's earlier edition the title Yaḥyā al-Malik was read Ḥilā al-Malik instead. In our view, an argument in favour of reading Ḥilā al-malik instead of Yaḥyā al-Malik would be the fact that Ibn al-Muhannā called his work

---

[18] For a brief history of the use of the Uygur alphabet, see Clauson (1962:175ff). Sertkaya (1973:5) transcribes and translates some poems composed in the 16th century by Ottoman poets who wrote in Čaġatay Turkic in both the Uygur and the Arabic alphabets. See also Mansuroğlu 1954:256 and art. 'Turks' by Samoylovitch in EI[1] VIII 911.

[19] Cf. art. Djalāl ad-Dīn Khᵂārazm-shāh in EI[2] (Boyle).

Kitāb Ḥilya al-ʾInsān wa-Ḥalba al-Lisān, literally "The jewel of man and the racehorse of the language". In this sense *Ḥilya* can be understood as a reference to *Ḥilā*, since *ḥilya* 'jewel' is the singular of *ḥilā* 'jewels'.

On 129,3 Ibn al-Muhannā refers to the author of a work on Turkic entitled *Tuḥfa al-Malik*[20] as saying: "The author of the book *Tuḥfa al-Malik* said: I asked the Turkic *ʿulamāʾ*... but they did not answer anything" (*qāla ṣāḥib kitāb tuḥfa al-malik saʾaltu ʿulamāʾ at-Turk... falam yuǧībū bi-šayʾ*).[21] From the passage quoted above it is possible to deduce that, although there were scholars of Turkic descent in the region in which *Tuḥfa al-Malik* was compiled, they did not seem to occupy themselves very intensively with their language, at least not in the way the Arabic language was studied.

Further, Ibn al-Muhannā, when explaining the Turkic animal calendar (185,13), refers to *Kitāb Ṭabāʾiʿ al-Ḥayawān*, compiled by Šaraf az-Zamān aṭ-Ṭabīb al-Marwazī (for which see Iskandar [1981]).

Ibn al-Muhannā refers to the Turkic language as *al-luġa at-turkiyya* (cf. 72,12; 73,12; 91,12; 119,1,3) or *luġa at-Turk* (118,9) but also quite frequently as *al-luġa at-turkistāniyya* "the Turkestanian language" (73,16). Ibn al-Muhannā often describes the people whose language he describes as *ʾahl bilādinā* "the people of our country" or *ʾAtrākunā* "our Turks". This expression may lead to the assumption that he was of Turkic origin himself. The esteem with which he writes about Turkic and the way he compares it with Arabic, much like Kāšġarī did before him, sustain this suggestion :

> "know that its origin is directly from the people of Turkestan like Arabic [comes] from the Ḥiǧāz" (*ʾIʿlam ʾanna l-luġa t-turkiyya manšaʾuhā bi-l-ʾiṣāla ʿan ʾahli turkistān kamā l-ʿarabiyya ʿan al-ḥiǧāz...*, Ḥilya 73,12)

Similarly, in agreement with the term *turkistāniyya* which he uses for the Turkic language, Ibn al-Muhannā calls the Turks *Turkistaniyyūn* "Turkestanians" (cf. 78,17, 79,16 and 135,6). Although this seems very precise at first sight, it is in fact not easy to determine the geographical situation of *Turkistān*; its exact place seems to have been subject to changes related to the various perceptions people had of political and

---

[20] The exact transcription of *m-l-k* is unclear, since it was left unvocalised in the printed text. The use of the word *malik* in the title may refer to a Ḥwārazmi-ruler: "Khᵂarazmi, G̲h̲aznawid and Salḏj̲ūq rulers called themselves *malik*, usually in combination with honorific adjectives, e.g. *al-kāmil, al-ṣāliḥ, al-ʿādil*" (Ayalon EI², *malik*).

[21] The question referred to was why ġil and kil were both used with the imperative form and why ġil sometimes means 'possessor' (not equivalent to the suffix -liq of equipment; cf. *Fundamenta* I).

social circumstances in history. Summarising Barthold's findings (in
EI[1]), in the sixth century AD *Turkistān* was thought to begin immedi-
ately north of the Oxus. Later on, in the fourth/ninth century, when
the Turks were driven far back to the north, Arab geographers situated
Turkistān north of the area of Arab culture, i.e. north of the Sir Daryā.

1.3 Kitāb al-ʾidrāk li-lisān al-ʾatrāk

There are three extant manuscripts of *ʾIdrāk*. The first is preserved in
the Beyazıt National Library (Istanbul), registered Veli ed-Dīn No.
2896. It consists of 132 pages, and bears the date Thursday, *Ramaḍān*
20th, 712/January 15th, 1313. In the Veli ed-Dīn MS, the Turkic
words are not always vocalised and the reader is handicapped by many
poorly legible glosses both in the margin and between the lines of the
text.[22] The glosses of this MS were edited by İzbudak (1936) but
unfortunately in transcription only. A second MS is at the Library of
the University of Istanbul, registered Hālis Efendi No. 6597. This MS
consists of 194 pages and was copied in the city of Laḏiqiyya by
ʾAḥmad aš-Šāfiʿī on *Ǧumādā l-ʾAwwal* 4th, 805/November 30th,
1402. A third manuscript is kept in the Dār al-Kutub Library at Cairo
(cf. al-Ḥadīṭī 1967: 176ff and 552).[23]

ʾIdrāk was first edited in 1892 by Mustafa Beg, but based solely
upon the Veli ed-Dīn MS. Mustafa Beg's edition was severely criticised
by Huart (1892) and Bouvat (1906). In 1931, Ahmet Caferoğlu pub-
lished a new edition of the text, this time based on both Istanbul MSs,
along with a translation into Turkish. He dedicated much space to the
word list, which he both transcribed and elaborated.

Much detailed information is available about the author of *ʾIdrāk*. It
was compiled by the Andalusian grammarian and theologian ʾAbū
Ḥayyān al-ʾAndalusī (654/1256 - 745/1345)[24] who lived and worked in
Cairo. Before compiling *ʾIdrāk*, ʾAbū Ḥayyān had already devoted at
least three other studies to Turkic entitled *Kitāb al-ʾafʿāl fī lisān at-
Turk* (cf. *ʾIdrāk* 120,19), *Zahw al-mulk fī naḥw at-Turk* and a book

---

[22] The text in the margins of *Idrāk* contains an additional grammar of a Turkic lan-
guage, whose exact contents have not been published previously. Some features of this
work are discussed in Section 1.10 and further througout this book.

[23] Al-Ḥadīṭī (1967) describes the contents of *ʾIdrāk* based on the Cairo MS. For the
present study and the translation in Part Two, I had access to the two Istanbul
manuscripts.

[24] Cf. art. ʾAbū Ḥayyān in EI[2] by Glazer.

about the Turks: *Tuḥfa al-musk fī sīra at-Turk.*[25] ʾAbū Ḥayyān must
have been very interested in the phenomenon of language since he was
one of the few Arabic grammarians ever to refer to private knowledge
of other languages than Arabic, as he did in in his *Manhaǧ as-Sālik li-
ʾAlfiyya Ibn Mālik* (Cf. *Manhaǧ* 230). ʾAbū Ḥayyān also wrote treatises
about Persian, Ethiopian, and Coptic *(Bašmūr)*, but his fame rests
mainly on a large number of treatises on linguistic and theological sub-
jects, which were not limited to the Islamic sciences only but also
comprised a study of the Torah.[26]

ʾAbū Ḥayyān's views on Arabic were well-known in his time; his
reverence for the principles of Sībawayh, written down in his *Kitāb*,
the oldest known grammar of Arabic and the ultimate reference for all
Arabic grammarians, became proverbial among Arabic linguists. ʾAbū
Ḥayyān was also known for his interest in languages other than Arabic,
which was quite an exceptional phenomenon in his day. The following
fragment from *ʾIdrāk* is an example of his views on language in gen-
eral, at the same time revealing the basis for the structure of both
ʾIdrāk and *Irtišāf.*

> "The certainty of every language is obtained by knowledge of three things:
> the first is the meanings of all simple words, which is called 'lexicology'.
> The second one is the rules of those simple words before their
> construction, which is called 'morphology'. The third is the rules in case of
> a construction, which is called by those that speak about the Arabic
> language: 'syntax'." *(fa-ʾinna ḍabṭa kull luġa yaḥṣilu bi-maʿrifati ṯalāṯat
> ʾašyāʾ ʾaḥaduhā madlūl mufradāt al-kalim wa-yusammā ʿilm al-luġa, wa-
> t-ṯānī ʾaḥkāmu tilka l-mufradāt qabla t-tarkīb wa-yusammā ʿilm at-taṣrīf
> wa-t-ṯāliṯ ʾaḥkāmuhu ḥāla t-tarkīb wa-yusammā ʿinda l-mutakallimīn
> ʿalā l-lisān al-ʿarabī ʿilm an-naḥw, ʾIdrāk 5,9-12.)*

ʾIdrāk consists of two large parts, the first part being a large Turkic-
Arabic word list and the second part a large grammatical section in
which most aspects of a specific Turkic language are discussed.[27] In
Caferoğlu's edition, the Turkic-Arabic word list consists of exactly 100
pages, whereas the grammatical section takes up pages 101-155.[28] See

---

[25] Cf. Caferoğlu's (1931: X) attribution of *ad-Durra al-muḍīʾa fī l-luġa t-turkiyya* to
ʾAbū Ḥayyān cannot be verified, since he does not mention the sources upon which it is
based. However, no such work is mentioned in the Arabic biographical literature on
ʾAbū Ḥayyān.

[26] Al-Ḥadīṯī (1967) lists a total of 65 works, most of which are lost. Famous are ʾAbū
Ḥayyān's *Manhaǧ as-sālik li-ʾAlfiyya Ibn Mālik* and his exegesis to the Qurʾān, *al-Baḥr
al-muḥīṭ.* Also a work on the Torah is mentioned (68).

[27] *Dīwān* and *ʾIdrāk* are the only ones of our sources to contain a Turkic-Arabic word
list instead of the usual Arabic-Turkic lists.

[28] In the Veli ed-Dīn MS this is 2ʳ16-32ʳ.

also the outline of the contents of the sources in the appendix to this chapter (55 ff).

'Abū Ḥayyān conducted his own research, in which he was helped by native speakers (cf. 6). In his word list he mentions a number of times his master Šayḫ Faḫr ad-Dīn (cf. 14; 27; 57), probably 'Abū Ṭāhir 'Ismāʿīl bn 'Aḥmad bn 'Ismāʿīl bn Burtuq bn Buzġuš al-Miṣrī (also Ġalāl ad-Dīn, d. 715/1315; in Ṣafadī [IX No. 4001] mentioned as an acquaintance of 'Abū Ḥayyān; cf. also Suyūṭī I No. 906) from whom he learned the seven readings of the Qur'ān (cf. Ḥadīṯī 1967:69). Šayḫ Faḫr ad-Dīn seems to have known Turkic, since he is always referred to in the context of the meaning of a Turkic word, e.g. at, 'iškik 'oar' (al-miġdāf) "our master Faḫr ad-Dīn did not know this word" (wa-lam yaʿrif šayḫunā Faḫr ad-Dīn hāḏihi l-lafẓa, 'Idrāk 14).

Apart from this living source, 'Abū Ḥayyān used a written source compiled by a certain Baylik, to which he refers in both the grammatical section ('Idrāk 133,13 and 146,11) and in numerous instances in the word-list.[29] Pritsak (1959:75) attempts to identify this Baylik:

> ...Besonders wichtig muß al-'Anwār al-muḍīʿa des Kiptschaken ʿAlā ad-Dīn Beilik al-Qifġāqī gewesen sein, das eine der Hauptquellen folgender, uns erhaltener späterer Werke war: 'Abū Ḥaiyān (bei ihm heißt es schlicht al-kitāb "Buch"), Tuḥfat az-zakiyya und Bulġat al-muštāq.[30] Der Verfasser ist m. E. mit dem bei Brockelmann (GAL 12, 652) genannten Beilik (Beilaq b. ʿAbdallāh al-Qibġāq) identisch, der ca. 1250 bis nach 1282 in Kairo tätig gewesen war.[31]

Another possibility is Bīlīk al-Ḥaznadār (d. 676/1277) who was known for his knowledge of foreign languages, apart from being a student of history and ḥadīṯ (cf. Haarmann 1988:99. This same Baylik is also mentioned in Bulġa [see below Section 1.8]).

'Abū Ḥayyān refers once to al-mawlā Tāǧ ad-Dīn ('Idrāk 136,4), probably Tāǧ ad-Dīn bn Maktūm who lived between 682/1284 and 749/1348 (cf. Ḥadīṯī 1967: 502).

The name sangar, which occurs in numerous exemplifying sentences throughout the work, may not be a random choice but rather intended as a reference to Sangar ad-Dawādārī (d. 699/1299-1300) who was famous for his knowledge of the science of tradition, and

---

[29] Cf. 'Idrāk 27; 32; 39; 46; 57; 62; 67; 69; 72; 73; 75; 82; 85; 86; 90; 92; 95; 98; 99.

[30] I could not find any reference to Baylik in Tuḥfa. See the respective remarks about Tuḥfa and Bulġa in the following sections.

[31] For Bulġa, see Zajączkowski (1958: XI).

who was the superintendent of the Ibn Ṭulūn mosque (where—from 698/1298—ʾAbū Ḥayyān taught grammar; cf. Haarmann 1988:97).[32]

The Turkic language that ʾAbū Ḥayyān describes in ʾIdrāk is called *turkiyya*. ʾAbū Ḥayyān sometimes distinguishes it from *Qibğāqiyya* 'Qipčaq' (cf. 105, 138, 147) and *Turkmāniyya*, here interpreted as Oğuz (on this see also Doerfer 1976:246) (cf. 105, 128, 129, 130, 132, 135, 144, 147, 154).[33] Under the entry *Turk* in *Idrāk* one finds the following definition: "a tribe of the non-Arabs; they are the people of this language" (*qabīla min al-ʾaʿāğim wa-hum ʾahl hāḍā l-lisān*, *Idrāk* 37). *Turkmān* is defined as follows: "a tribe of the non-Arabs too" (*qabīla min al-ʾaʿāğim ʾayḍan*, 38), as if he considers them a separate tribe. Apart from numerous references to both Oğuz *(Turkmānī)* and Qipčaq, he also refers to *Tatar* (63), *Bulğār*[34] (9), *Tuqṣubā*[35] (15), *Čapnī* (41),[36] "the Ḥwarezm language" (*luğa ḥwārizmiyya*, 91), and finally, to "the language of *Turkistān*" (*luğa Turkistān*, 26), which is, of course, a rather vague indication (see remarks in Section 1.2). The *Uyğur* (37; 67) are described as a people living in *Turkistān* and so is the tribe of the *Bulğār* (9). A characteristic of the speech of this tribe seems to have been the change of y into ḍ. An example of this is ʾaḍaq 'foot' instead of more common ʾayaq and further "ʾuyi-dī 'he slept' and the change of the y for a ḍ is a Turkestanian variant. They say ʾuḍidī." (*ʾuyidī nāma wa-ʾibdāl al-yāʾ ḍālan luğa Turkistāniyya qālū ʾuḍidī*, ʾIdrāk 26).

Apart from these references to Turkic peoples, earlier research on the language material of ʾIdrāk has indicated that it deals with a mix-ture of Oğuz and Qipčaq (on this see also Flemming 1977a). There are indications that the language in ʾIdrāk is close to the Western-Qipčaq

---

[32] Another possibility is that the name Sanğar refers to the Salğūq (Oğuz) ruler Sanğar bn Mālikšāh. Sanğar governed Ḥurāsān and the north eastern provinces of the Selğūq Empire between 512/1118 and 552/1157 (cf. EI[1] art. *Saljūḳ*). The occurrence of the name Sanğar in ʾIdrāk, therefore, may be an indication of Oğuz influence during its compilation. (See also Part Two, Translation n. 37). I thank Dr Erica Gilson for drawing my attention to the possibility that this name probably was not a random choice.

[33] In his word list to ʾIdrāk (29), ʾAbū Ḥayyān refers to a small Qipčaq tribe or clan **barlī** *qabīla min al-qibğāq* "a tribe of the Qipčaq".

[34] Cf. art. 'Bulgẖār' in EI[2] by Hrbek.

[35] The ethnonym Ṭuqṣubā may be interpreted as **toquz opa** 'nine villages' and hence as a reference to a conglomerate of nine different tribes, perhaps Oğuz, since it is an Oğuz word (*Dīwān* MS 55,17). I thank Prof. Dr Talât Tekin for this suggestion. Kudayberdy-uli (1990:71) mentions a Qipčaq tribe called **toksaba** in works of the Arab historiographers ʾAmīr ad-Dīn and Ibn Ḥaldūn. (I checked Ibn Ḥaldūn's *Muqaddima* on this but could not find this reference.)

[36] "Tribe from the Turks" *(qabīla min at-turk)*. In *Dīwān* (MS 41,2; also Dankoff and Kelly I 102) listed as a clan of the Oğuz.

languages (such as, e.g., Karaim, Karačay-Balkhar), but especially Crimean-Tatar, which I shall point out here briefly.

In the first place the Qipčaq language in *'Idrāk* does not show the typical ǧ-/ž- instead of y-, like Kazakh, Karakalpak, Nogai and certain Tatar dialects (Berta 1989). This phenomenon is described by Kāšġarī (MS 26,4) for Oġuz [sic]/Qipčaq tribes (e.g. **ǧinǧä** for **yinǧä** 'pearl'). This implies that the phenomenon of initial y- < ǧ-/ž- was already existent in the 11th century, about two centuries before *'Idrāk*, with which the possibility that it was due to a later development is ruled out. Secondly, the ending -**mayin** 'without ...ing' (*'Idrāk* 138,12) seems to survive in Karaimian (dialects of Troki and Halicz) and Karačay-Balkhar only (*Fundamenta* I 335 and 362; also Musaev 1964, 22, 302). In the third place, the ending -**kinǧä**/-**ǧinǧä** (*Idrāk* 150-1) survives in Qipčaq languages (cf. *Fundamenta* I). This form is especially prevalent in Karaim GInčA (cf. Musaev, 1964, 301), whereas other Turkic languages have -günce/-gunca (also Karačay-Balkhar, -ġıncı[4]; cf. K-B -Russkiy slovar' 1989:823). Two of the three above-mentioned features—except -**mayin**— are also described for Crimean Tatar by Doerfer (1959), and with the meanings given by 'Abū Ḥayyān.

In the last place we could add, albeit not without a firm caveat in regard to the known problems related to Turkic ethnonyms, Ibn Muḥammad Ṣāliḥs seventeenth century witness report of the language described by 'Abū Ḥayyān as similar to that spoken by the inhabitants of the Crimea. To this language Ṣāliḥ refers as *tatariyya* (cf. below 1.7).

Other scholars also pointed to the possibility that the language in the 14 century manuscripts could be a form of Tatar (e.g., Von Gabain [1959:48], Hattori [1979-80] and Mahmutova [1982]). In all instances, though, they refer to the Turkic language in the 14th century Codex Comanicus rather than the Arabic sources.

1.4 Kitāb at-tuḥfa az-zakiyya fī l-luġa at-turkiyya

Tuḥfa is an anonymous grammatical treatise on Turkic, preserved in only one MS which is kept in Istanbul, at the Beyazıt Library, coded Veli ed-Din No. 3092. It consists of 90 folios, written down in a single hand. The Turkic words are written in red ink. *Tuḥfa* has never been edited in print so far, apart from a legible facsimile published in 1942 by Tibor Halasi Kun, which forms the basis for the present study. It was followed in 1945 by Besim Atalay's translation into Turkish along with an edition of quite poorly legible facsimile copies of the text. In

his translation, Atalay transcribes the Turkic words in such a manner that the original spelling cannot be retrieved.[37]

There is little information about the place and time *Tuḥfa* was compiled. *Mamlūk*-ruled Egypt or Syria is generally taken as its place of compilation.[38] It must have been written before 829/1426, which is the date of a gloss on the title page (Cf. Atalay 1945: xxiii). Fazılov, in an article on *Tuḥfa* (1976: 335), made some suggestions as to the meaning of the word *zakiyya* in the title, according to which it may have referred to someone called Zakī, perhaps the author himself, or to the word *zakī* 'intelligent' in a poem on the last page of the MS (90ᵛ).[39]

Tuḥfa, like most other grammars of Turkic, contains an Arabic-Turkic word list (3ʳ9-38ᵛ11) starting with *hamza*. Each section is divided into two subsections, the first of which treats nouns and the second one verbs. The second part of *Tuḥfa* (38ᵛ12-90ᵛ13) contains grammatical information, which is divided into two sections: *ṣarf* "Morphology" (38ᵛ12) and *al-ʾaḥkām at-tarkībiyya* "The rules of construction" (65ᵛ5).[40]

The author of *Tuḥfa* does not refer to any direct sources for his work, but he mentions the name of ʾAbū Ḥayyān al-ʾAndalusī, and he quotes his principles (see 1.3) with regard to the three basic compo- nents of the study of language, namely lexicology, morphology and syntax (2ʳ12-2ᵛ1-6).[41] Apart from this passage, there are many other resemblances between *Tuḥfa* and *ʾIdrāk* as far as the general structure is concerned, and to a far lesser extent in terms of the content. I shall discuss this in greater detail in the second part of this chapter.

The language described in *Tuḥfa* is generally called *turkiyya* or, sometimes, *qibǧāqiyya*.[42] It is occasionally opposed to *tatariyya*[43] and, more especially, to *turkmāniyya*, mostly introduced by the expression

---

[37] Atalay's translation was a severely criticised by Tibor Halasi Kun, at the time liv- ing and working in Ankara, who apparently had had the idea of a similar publication (cf. Halasi Kun 1947, 1948, and Atalay 1948).

[38] Fazylov (1976: 335) asserts that *Tuḥfa*'s author had spent some time in Syria or was born there.

[39] Fazylov (1976) deciphered the meaning of some other lines on the last page that announced the death of Nāṣir Muḥammad Qaytbay, the son of ʾAšraf Qaytbay on Wednesday *Rabīʿ al-ʾAwwal* 15th, 904 (=October 31th, 1498). He further succeeded in reading a name on the title page, i.e. ʾAbū l-Qāsim bn ʾAḥmad bn Muḥammad bn Muhannā al-Ḥanāfi.

[40] For a list of contents, see the Appendix.

[41] ʾAbū Ḥayyān's name is rendered somewhat differently: Šaraf ad-Dīn instead of ʾAṯīr ad-Dīn, but his *kunya* is right. There are no direct references to Baylik, a main source for *ʾIdrāk*, in spite of Pritsak's (1959: 75) remarks quoted above.

[42] Cf. 2ʳ2; 47ʳ1; 62ʳ12.

[43] Cf. 43ᵛ12; 51ᵛ1; 63ᵛ13; 71ʳ13.

qīla only: 'I have not referred to Turkmānī [Oġuz] unless there was a need; [in that case] I say 'it is also said'" *(wa-qīla) (wa-lam ʾaḏkur at-turkmāniyya ʾillā ʿinda ḍ-ḍarūra fa-ʾaqūlu wa-qīla* 2ʳ2).[44]

1.5 al-Qawānīn al-kulliyya li-ḍabṭ al-luġa at-turkiyya

Qawānīn is another grammar of Turkic. It was compiled about the 14th or 15th century AD, most likely in a *Mamlūk*-Turkic dominated region, probably Cairo. An indication for this may be found at 71,1, where the author refers to Turkic soldiers who had come to Cairo after having served in Tamerlane's armies.

The only existing copy of *Qawānīn*, registered Şehid Ali Paşa 2659 in the Süleymaniyye Library at Istanbul, is a well-preserved MS consist-ing of 169 pages (1-85ʳ), written in a very neat and legible hand. On the introductory page the work is attributed to ʾAbū Ḥayyān.[45] In 1928, Kilisli Muʿallim Rifʿat (Bilge) edited the text with an introduc-tion by Mehmet Fuat Köprülüzade. Ten years later, in 1938, Szigmund Telegdi published an article about *Qawānīn* in which he rearranged both the Turkic grammatical material and the word list of approxi-mately 500 entries.

The printed edition of *Qawānīn* consists of 76 pages (3-79). The largest part is dedicated to grammatical subjects, after a brief introduc-tion (3-58). The author divided *Qawānīn* into three large chapters: I "the verb and what is attached to it" (*al-fiʿl wa-mutaʿalliqātuhu wa-lawāḥiquhu*, 6,11); II "the noun and what is attached to it" (*al-ʾism wa-mutaʿalliqātuhu*, 25,16) and, finally, III "the particles" (*al-ḥurūf*, 68,6). The Arabic-Turkic word list, arranged according to semantic cat-egories, takes up about six pages (58-64) within the chapter on the nouns.[46]

In regard to his chapter division, the author of *Qawānīn* refers to a work of Ibn al-Ḥabbāz an-Naḥwī, probably Šams ad-Dīn bn al-Ḥusayn al-ʾIrbilī al-Mawṣilī (d. 637/1239; cf. Suyūṭī *Buġya* I, 304; also in ʾAbū Ḥayyān's *Irtišāf* III 657 [index]):

> "Ibn al-Ḥabbāz an-Naḥwī said: 'The limitation of the word to three cate-
> gories is not restricted to the language of the Arabs, since the argument for

---

[44] *Qīla* is used in the following places: 2ʳ2; 7ᵛ5; 8ʳ10; 9ʳ3; 47ʳ4; 65ʳ11; 67ʳ7 69ʳ9; 73ʳ12; 77ʳ5; 81ʳ1; 88ʳ2; 88ᵛ11;13; 89ʳ1 and numerous occurrences in the word list. *Wa-qīla* is the usual way to indicate alternative expressions.

[45] Karamanlıoğlu (1962) remarks that ʾAbū Ḥayyān cannot have been *Qawānīn*'s author, because he was not alive (d. 765/1345) by Tamerlane's time.

[46] In fact, the word list is longer, but ends with a discussion of the numerals (pp. 64-68), which is a part of the grammatical section. For a full list of contents, see Appendix.

it is rational and rational matters do not differ with the difference of lan-guages'." (*qāla ibn al-ḥabbāz an-naḥwī: lā yaḥtaṣṣu inhiṣār al-kalima fī l-ʾanwāʿ aṭ-ṯalāṯa bi-luġa al-ʿarab li-ʾanna d-dalīl alladī dalla ʿalā ḏālika ʿaqlī wa-l-ʾumūr al-ʿaqliyya lā yaḥtalifu bi-ḥtilāf al-luġāt, Qawānīn* 6,1.)

In accordance with this principle the author of *Qawānīn* divided his work into three parts, as explained before.

Although the MS contains a reference to ʾAbū Ḥayyān al-ʾAndalusī on the preliminary pages, it is generally accepted that he is not *Qawānīn*'s author. The anonymous author was most likely not of Tur-kic origin himself (cf. 3,11-4,1-2). He tells us he wrote *Qawānīn* because many of his friends had asked him to do so (cf. 1,6-7). He must have done some research himself among people of Turkic de-scent. Indications for this assumption may be found in the use of the first person sg. in sentences like: "I did not hear it from them other-wise" (*wa-lam ʾasmaʿ-hu min-hum ʾillā ka-ḏālika*, 14,19; 30,5; 49,11).

The Turkic language that is described in *Qawānīn* is called *turkiyya*, as in most of the other sources or, occasionally, *luġa al-ʾAtrāk* (7,16). It is opposed to *Turkmānī* (Oġuz), against the use of which the author warns severely:

> "The language of the Oġuz is not Turkic... it is held in contempt by them and whoever speaks it, is despised by them." (*luġa Turkmān laysat turkiyya... fa-ʾinnahā mustahġana ʿindahum wa-l-mutakallim bihā muḥtaqar ladayhim, Qawānīn* 7,15; cf. also Telegdi 1938.)

In this fragment the Arabic *-hum* 'them' refers to the Turks who spoke pure Turkic, the same way *hum* in grammars of Arabic refers to the Arabs, the Bedouins, of the Arabic Peninsula who were said to speak the purest Arabic (cf. Ditters 1992). On 20,21 the author specifies this in a particular context: "the pure speakers among them" (*al-fuṣaḥāʾ minhum*).

With regard to the Turkic material, the author had access to certain written sources he does not specify. With regard to Oġuz influence, he calls for careful use of this material:

> "Most of those who compiled a *tarġumān* on this language depend on it [sc. on Oġuz-*Turkmānī*] and they occupy themselves very little with the language of the Turks" (*wa-ġālib man ṣannafa fī hāḏihi l-luġa tarġumānan ʾinnamā yaʿtamidu ʿalayhā wa-lā yuʿarriġ ʿalā luġa al-ʾatrāk ʾillā fī n-nazar al-qalīl, Qawānīn* 7,16.)

Of course, it is not known which sources the author may have meant. With regard to the Arabic primary material he used, he quotes Ibn al-Ḥabbāz an-Naḥwī, to whom I have already referred above.

1.6 Kitāb tarǧumān turkī wa-ʿarabī wa-muġalī

Tarǧumān is an anonymous work of which only one copy exists in the
library of Leiden University (the Netherlands), cod. 517 Warner. It
consists of 76 folios. Its date of compilation is known exactly due to a
registration in the epilogue of the text: Sunday, *Šaʿbān* 27th,
743/January 25th, 1343. Houtsma (1894) had read this date as *Šaʿbān*
27th, 643 which he converted to January 28th, 1245. This unfortunate
mistake was copied by all referents to *Tarǧumān*, until Flemming
proved convincingly in an article that the date Houtsma proposes can-
not possibly be correct (cf. Flemming 1968).[47] In the same article
Flemming also proposes Ḫalīl bn Muḥammad bn Yūsuf al-Qunawī as
the author of *Tarǧumān*, rather than its copyist.

    Tarǧumān consists of two parts: 62 folios of Turkic-Arabic text and
14 folios Mongolian-Persian. According to Houtsma, these texts were
compiled by the same person who, accordingly, must have had a fair
knowledge of at least three languages. In the printed edition, the
Turkic part takes up 57 pages. After an introduction (2-4), the author
divides the material into four main parts: I "the nouns" (*al-ʾasmāʾ*,
5,1), which contains an Arabic-Turkic list of nouns, arranged accord-
ing to semantic categories; II "the verbal nouns and the imperative [of
the verbs]", which is an Arabic-Turkic list of verbs (*maṣādir al-ʾafʿāl
wa-l-ʾamr bihā*, 33,8); III "conjugation of the speech and the verbs"
(*taṣrīf al-kalām wa-l-ʾafʿāl*, 44,11) and, finally, IV "Basic rules of the
speech and obligatory elements" (*ḍawābiṭ al-kalām wa-mā lā budd
min-hu*, 50,4), in which both Turkic and Arabic particles are treated.[48]

    Like most other sources, the author calls the language he describes
turkī 'Turkic', "the pure Turkic langugage" (*al-luġa at-turkiyya al-
ḫāliṣa*, 2,15) or "the Turkic-Qipčaq language" (*al-lisān at-turkī al-
qifǧāqī*, 2,11) and he clearly distinguishes it from *turkmānī* e.g.:

> "I specified in which [elements] the difference occurs between the pure
> Turkic language and the Oġuz *(turkmānī)* language and what is borrowed
> now from Persian and other [languages] by the people" (*wa-qad ʿayyantu
> mā waqaʿa al-ḫulf fīhi bayn al-luġa at-turkiyya al-ḫāliṣa wa-bayna l-luġa
> t-turkmāniyya wa-mā huwa mustaʿār bayn an-nās al-ʾān min al-luġa al-
> fārisiyya wa-ġayrihā, Tarǧumān* 2,15).[49]

---

[47] When Flemming wrote her article she did not have access to Dozy's catalogue of
Leiden Manuscripts (1851: I 109; no. ccxii), in which the correct date, 743, was already
given (p.c. July 1995).

[48] E.g. *lafẓa* **mū**, the particle of interrogation; *lafẓa* **dā** the locative suffix but also
some Arabic locatives, e.g. *fawqa* 'above'; *bayna* 'between'; *al-wasaṭ* 'middle'.

[49] See also 21,14; 23,12; 41,19 and numerous other places.

This proves that the author not only distinguishes between 'pure Turkic' and its variants (e.g. *Turkmānī* - Oġuz) but that he is also aware of the existence of Persian loan-words.

In *Tarǧumān* no direct references are made to its primary sources; the author only quite vaguely mentions the existence of some works on Turkic:

> "But I present my book, following with it the trail of those who preceded me in presenting the books on the translation of the Turkic language..." (*'ammā ba'd fa-'innī waḍa'tu kitābī hāḍā muqtafiyan bihi 'aṯara man taqaddamanī mimman waḍa'a l-kutub fī ·tarǧama al-luġa at-turkiyya...* 2,5-7.)

The nature of these unspecified sources is, of course, not known.

1.7 aš-Šuḍūr aḍ-ḍahabiyya wa-l-qiṭa' al-'aḥmadiyya fī l-luġa at-turkiyya

Šuḍūr is an Arabic-Turkic word list of which several MSs of various lengths exist. There are six manuscripts with this title in the Bibliothèque Nationale de Paris (Supplément Arabe No. 4329, 4330/31/32/33/34; cf. Blochet 1932) and two in the Süleymaniye Library at Istanbul (Kılıç Ali Paşa 1021/2 and Laleli 3539). Further, Zajączkowski (1965: 41) mentions a MS in Uppsala. Rossi (1935) and Rieu (1888) mention several MSs in the Vatican and the British Museum in London, respectively. In this study all references to *Šuḍūr* are based on the Parisian MS Supplément Arabe No. 4333 (old code: 1385), which consists of 33 folios. In 1949 *Šuḍūr* was translated into Turkish by Besim Atalay, who based his translation on two manuscripts from his private library.[50]

Šuḍūr was composed by Mawlāh Ibn Muḥammad Ṣāliḥ in 1029/1619.[51] According to Atalay, in his preface to the Turkish transla- tion of *Šuḍūr*, Ṣāliḥ was of Turkic descent and a professor at the Madrasa of al-Mālik al-'Ašraf at Cairo. Ibn Muḥammad Ṣāliḥ states that he wrote the book in the first place to teach Turkic to 'Aḥmad, the son of an Egyptian *qāḍī* (cf. *Šuḍūr* 2ᵛ4 and 26ᵛ18)[52], which is reflected in the title of the book by the expression *'aḥmadiyya*. The

---

[50] Both Nissman (1969:5) and Pritsak (1959) range *aš-Šuḍūr aḍ-ḍahabiyya*, a source that dates from 1619 and that describes Ottoman Turkic, among the Qipčaq glossaries.

[51] This date is taken from Atalay (1949) who based it on a secondary source, which is unavailable to me. According to the dates of Ṣāliḥ's sources, which I shall discuss below, this date may be right.

[52] This *qāḍī* was, according to Atalay, a certain Šarīf Muḥammad al-Ḥusaynī al-Ǧumāḍā.

second reason for writing his book was the fact that many of his
friends and relatives had asked him many times to write a book about
Turkic:

> "For a long time I have had it in mind to write a treatise on the Turkic lan-
> guage; I was asked many times and my relatives and friends kept repeating
> it [their request] to me all the time." (*ṭāla mā ǧāla fī ḫaladī 'an 'aḍaʿa
> risāla fī l-luǧa at-turkiyya wa-suʾiltu bi-ḏālika mirāran wa-takarrara
> ʿalayya min 'aqribāʾī wa-'aṣdiqāʾī tikrāran, Šuḍūr* 2ʳ10.)

As for the internal structure of the work, Ṣāliḥ divided it into four
'bases' (*'arkān*, sg. *rukn*). The first 'base' deals with "as many deriva-
tions as possible" (*ar-rukn al-'awwal fī bayān al-muštaqqāt bi-qadr al-
'imkān*, 5ʳ) and it is divided into eight chapters: 1. the verbal noun
(*bayān al-maṣdar*, 7ᵛ); 2. the imperative (*bayān al-'amr*, 7ʳ); 3. the
prohibition (*n-nahy*, 8ᵛ); 4. the past tense (*bayān al-māḍī*, 10ᵛ); 5. the
present tense (*bayān al-muḍāriʿ*, 12ᵛ); 6. the active participle (*bayān
'ism al-fāʿil*, 14ʳ); 7. the marker of the negation (*bayān ʿalāma al-nafy*,
15ᵛ); 8. the marker of the plural (*bayān ʿalāma al-ǧamʿ*, 16ʳ); II the
nouns (*bayān al-'asmāʾ*, 20ʳ4); III the pronouns (*bayān aḍ-ḍamāʾir*,
25ʳ16); IV words that occur in both Arabic and Turkic (*bayān al-
kalimāt al-muštarika bayn al-ʿarabiyya wa-t-turkiyya*, 28ʳ2). The
fourth *rukn* is followed by an additional chapter, *ḫātima* (29ʳ20), in
which one finds a number of Arabic sentences translated into Turkic.

   Although these chapter headings may suggest that a large body of
grammatical data is given, each *rukn* and *bāb* consists of no more than
a mere list of words of the category mentioned, the only exception be-
ing the eighth *bāb* of the first chapter about the marker of the plural
form, in which the author explains the different ways to indicate plu-
rality in Turkic.

   With regard to the sources he used, Ṣāliḥ tells us he possesses a large
number of books in several languages of which he also mentions the
languages in which they were compiled. In the MS at our disposal, the
name of each language is separated from the next one by a thick dot,
but since some languages—especially *turkī*—are mentioned twice or
even three times, the list most likely contains combinations of lan-
guages: *ʿarabī wa-turkī* 'Arabic-Turkic'; *ʿarabī wa-fārisī* 'Arabic-
Persian'; *turkī wa-fārisī* 'Turkic-Persian'; *turkī wa-nawāʾī* 'Turkic-
Nawāʾī' and, finally, *turkī wa-bahlawī* 'Turkic-Pahlawī'. With regard to

Turkī, Ṣāliḥ defined it explicitly as Ottoman Turkic; I shall return to this point below. Pahlawī is a term used to denote Middle-Persian.[53]

Things are different for *nawāʾī* since this is not a term used to de-note a language but rather the *nisba* of the Turkic poet ʿAlī Šīr Nawāʾī (845/1441-906/1501), the outstanding 15th century Čaġatay poet who had a great impact on all Turkic literary languages. Proof of the tremendous interest in his words are the many specialised dictionaries that were compiled on the basis of his works.[54]

After having listed the languages used in his sources, on pages 3ʳ-4ᵛ, Ṣāliḥ mentions as many as 25 authors and works. Among them are two Arabic sources. First, Ṣāliḥ tells us he has a copy of *ad-Durra al-Muḍīʾa fī l-Luġa at-Turkiyya* which he attributes to ʾAbū Ḥayyān. Ṣāliḥ quotes ʾAbū Ḥayyān from *Durra* as saying:

> "I arranged it according to the principles of Arabic and called it *ad-durra al-muḍīʾa fī l-luġa at-turkiyya.*" (...*ġaʿaltuhu ʿalā qawāʿid al-ʿarabiyya wa-sammaytuhu bi-d-durra al-muḍīʾa fī l-luġa at-turkiyya, Šuḏūr* 3ᵛ15.)

However, I found no utterance of this kind in my copy of *Durra*. Moreover, ʾAbū Ḥayyān did not compile any work entitled *ad-Durra al-Muḍīʾa fī l-Luġa at-Turkiyya*. We shall discuss this point in Section 2.1.

In the second place he refers several times (e.g. on MS 3ᵛ27) to Saʿd ad-Dīn at-Taftazānī (722/1322-793/1390; cf. Suyūṭī *Buġya* II, 285; GAL I 354 and II 278), author of *Šarḥ at-Talḫīṣ*, who was a "celebrated authority on rhetoric, logic, metaphysics, theology, law and other subjects" (Storey on Taftazānī in EI¹). *Šarḥ at-Talḫīṣ*, compiled in 748/1347, also known as *Muḥtaṣar Šarḥ Talḫīṣ al-Miftāḥ*, is a com-mentary on al-Qazwīnī's work on rhetoric, *Talḫīṣ al-Miftāḥ*. Taftazānī was not unfamiliar with foreign languages, and he seems to have known Turkic. Ṣāliḥ had access to a manuscript of *Šarḥ at-Talḫīṣ*, the margins of which were full of Turkic words:

> "...I am pleased with the commentary to *at-Talḫīṣ* by *šayḫ* Saʿd ad-Dīn at-Taftazānī of which the margins were filled in Turkic from the first [page] until the last [page] in the writing of a teacher." (... *yasurru lī* [sic!] *šarḥ at-talḫīṣ li-š-šayḫ saʿd ad-dīn at-taftazānī muḥassan min ʾawwali-hi ʾilā ʾāḫiri-hi bi-t-turkī bi-ḫaṭṭ mawlā min al-mawālī, Šuḏūr* MS Paris No. 4334, 4ᵛ13-5ʳ3.)

Ṣāliḥ (*Šuḏūr* 3ʳ19ff) lists further the following works in this sequence:

---

[53] See Henning (1958) and Boyce (1968). By the term 'middle Persian' is meant pre-islamic Persian.

[54] See art. 'Mīr ʿAlī Sh̲ir Nawāʾī' by M.E. Subtelny in EI²; also Devereux (1964) and (1965) on *Muḥākama al-Luġatayni*.

1. *ʾAḫterī al-Kebir*, (also called *Luġat-i Aḫterī*) that is the Arabic-Turkish dictionary compiled in *Raġab* 952/May 1545 at Kütahya by Muṣliḥ ad-Dīn Muṣṭafā bn Šams ad-Dīn al-Qaraḥiṣārī al-ʾAḫterī (d. 968/1559-1560).[55]

2. *Mirqāt*,[56] by which may be meant *Mirqāt al-Luġa*, an anonymous Arabic-Turkic dictionary compiled between 796/1394 and 936/1530.[57]

3. Niʿmat Allāh (d. 969/1561), who compiled *Luġat-i Niʿmatullāh*, a Persian-Turkish dictionary.[58]

4. Ḥalīmī Čelebī, that is Luṭf Allāh bn Abī Yūsuf Ḥalīmī (d. 951/1544) the author of *Luġat-i fārisī wa-turkī* (also: *Luġat-i Ḥalīmī*), a Persian-Turkish dictionary, written or completed in 981/1573/4.[59]

5. *Ṣiḥāḥ al-ʿAǧam,* a Persian grammar in Arabic, attributed to Niʿmat Allāh bn ʾAḥmad bn Mubārak bn Muḥammad ar-Rūmī.[60]

6. *Ṣiḥāḥ al-Bayān.*

7. *Ǧawāhir al-ʿUqūd,* that is, probably, the small Arabic-Persian dictionary in verse entiteld *ʿUqūd al-Ǧawāhir* by ʾAḥmad-i Dāʾī, dedicated to Sultan Murād II (1421-1451). This manuscript is described in Blaškovič (1961:401-2; no. 551). For Dāʾī see Alpay (1973:19).

8. Ibn Melek, the Arabic form of the name ʿAbd al-Laṭīf ʿIzz ad-Dīn Firišteoġlī (d. 795/1392) who wrote *Luġat-i Firišteoġlu,* a rhyming Arabic-Turkish dictionary.[61]

9. *Tuḥfa al-ʾAdab.*

10. *Silsila aḏ-ḏahab,* attributed to Nūr ad-Dīn ʿAbd ar-Raḥmān bn ʾAḥmad al-Ǧāmī (d. 898/1493).[62]

---

[55] See Sohrweide (1974: no. 220), where more MSs are listed; Götz (1979: no. 466); Cf. TÜYATOK I (Antalya) p. 225, no. 391; also *Aḫteri-i Kebir* TÜYATOK II Merzifonlu Kara Mustafa Paşa Library (Ist.) p. 157, no. 384.

[56] The tekst reads *m-r-q-z-t.*

[57] Cf. Dozy 1851: 91. The Leiden MS, cod. 237 Scal., was copied in 955/1548. Sohrweide (1974: no. 194) describes the work as follows: "Arabisch-türkisches Wörterbuch, vermutlich aus dem frühen 16. Jahrhundert."

[58] Cf. Götz (1979: no. 473); TTS II xxxviii; Karatay (1961: II, 27 no. 2063).

[59] Cf. Götz (1979: no. 471) and Sohrweide (1981: TÜYATOK I Ali Nihat Tarlan Dermesi Istanbul p. 55, no. 142; and Karatay (1961: II 20 no. 2039).

[60] It may also be identical with *Ṣiḥāḥ al-ʿaǧamiyya* by Muḥammad bn Pīr ʿAlī Muḥyī ad-Dīn al-Birkawī (Birgili) 929/1523-981/1573 (Cf. GAL II 442, GAL Suppl. II 657; cf. also Ḥāǧǧī Ḥalīfa VI, 362 and Karatay (1961: II p. 22, no. 2045 and no. 2046), who lists the same title but another author: Hindūšāh Naḫǧawānī (900/1494). Also Leiden University Library: cod. 781 Warner.

[61] Cf. Akün EI² II 423; and TS II xxiv and TTS I xxix. There exists also a MS in the library of Leiden University, no. 857 Warner.

[62] *Silsila aḏ-ḏahab fī s-sulūk wa-l-ʾadab* cf. Ḥāǧǧī Ḥalīfa III, 607 and GAL II, 446.

11. *Dānistan*, a Persian-Turkic word list composed by Muḥammad bn Ḥāǧǧī ʾIlyās.[63]

12. *Niṣāb aṣ-Ṣibyān*, an Arabic-Persian vocabulary in 350 verses by Badr ad-Dīn ʾAbū Naṣr Masʿūd bn ʾAbī Bakr bn Ḥusayn bn Ǧaʿfar al-Farāhī (VIIIth century).[64]

13. A commentary to *Niṣāb aṣ-Ṣibyān* by al-Ǧurǧānī.[65]

14. *Niṣāb al-Fityān*, by which may be meant *Naṣīb al-Fityān* by Ḥusām ad-Dīn Ḥasan bn ʿAbd al-Muʾmin ʾAbī l-Ḥawabbī the poet, that consists of 350 verses in Persian (cf. Ḥāǧǧī Ḥalīfa VI 350).

15. A copy of *Šāhidī*, probably *Tuḥfa-i Šāhidī*, a rhyming Persian-Turkish dictionary written in 921/1515 by ʾIbrāhīm Dede Šāhidī (d. 957/1550).

16. *Tuḥfa Ḥusām*, perhaps *Tuḥfa-i Ḥusāmī*, an anonymous Persian-Turkish dictionary.[66] According to Dozy (1851:102) it formed the basis for *Luġat-i Šāhidī*.

17. The *Gulistān*[67] in several translations and studies, even in Arabic.

18. *Luġa Turkiyya wa-Nawāʾiyya* (3ᵛ1). This may be understood as a reference to a work called *ʾAbūšqa* or *al-Luġāt an-Nawāʾiyya*, a Čaġatay-Ottoman Turkish wordlist based on Nawāʾī's works.[68]

19. *Luġāt al-Matnawī li-Mawlā Ḥunkār al-ʿArabiyya* "The Arabic word lists to the *Matnawī* by master Ḥunkār", one of the many explanatory lists of words to the *Matnawī*, which was written by Ǧalāl ad-Dīn ar-Rūmī, who lived between 604/1207 and 672/1273-4.[69] Ṣāliḥ possibly refers to the commentaries by Sūdī, Bosnalı (d. 1005/1596-97) and by Sūrūrī (see below).

20. Commentaries on *Gulistān*, one by Sayyidī ʿAlī, Muṣṭafā bn ʾAhmad bn ʿAbd al-Mawlā Čelebī (948/1541-1008/1600).

---

[63] Cf. Dozy 1851: 97. One of the MSs is preserved in the library of Leiden University, cod. 167 Golius.

[64] Cf. GAL II, 193 (p. 246) GAL Suppl. 258. MS extant in Leiden 112/3. Cf. Sohrweide (VOHD XIII,3 1974: 301, no. 345).

[65] Cf. GAL S I 342.

[66] Cf. Götz (1979: 447, no. 472). Cf. also to TÜYATOK III (07) p. 11 no. 1632.

[67] The *Gulistan*, 'Rosegarden' was originally composed in Persian by Saʿdī bn ʿAbdallāh aš-Šīrāzī (d. 691/1291) in 656/1258. Cf. art. Saʿdī by Davis in EI².

[68] Nawāʾī also wrote *Muḥākama al-Luġatayn* (904/1499), a treatise in which Persian is compared to Čaġatay-Turkic. Cf. Devereux (1964 and 1965). *Muḥākama al-Luġatayn* has been edited by Quatremère, E. *Chrestomathie en Turc oriental* Paris 1841. Cf. also Menges (1963: 238) and Karatay (1961: II no. 2103).

[69] *Mawlā Ḥunkār* is another name for Mawlāna Ǧalāl ad-Dīn ar-Rūmī (Cf. art. Mollā Hünkâr in İA). For a list of Turkish commentaries on the Matnawī, see N. Pekolcay (1967: 109-110), *İslâmi Türk Edebiyatı*, İstanbul.

21. "and the other by Sururī Čelebi,[70] the teacher of Prince Muṣṭafā, the son of Sultān Sulaymān, peace be on them",[71] meaning the poet Sururī, Muṣliḥ ad-Dīn Muṣṭafā bn Šaʿbān (d. 969/1561) au thor of commentaries on the Maṯnawī, Ḥāfiẓ, Gulistān, and Bustān.

22. A commentary to Šayḫ Saʿdī's (d. 691/1291-2)[72] Dībāǧa-i Gulistān, probably referring to the commentary written in 917/1511-12 by Muḥammad bn ʿUṯmān bn ʿAlī, also called Lāmiʿī (d. 938/1531-2).[73] This Lāmiʿī was the author of Tuḥfa ʾāl ʿUṯmān ḏū an-Nūr as-Sātiʿ, to which Ṣāliḥ refers twice (cf. Šuḏūr 4ᵛ4 and 4ᵛ11).

And other books.

Further references in Šuḏūr.

23. Ṣāliḥ (25ᵛ17) refers to a work which he calls the Tarǧama of Kamāl Pāšā (tarǧama kamāl pāšā) by whom he probably means the well-known Ottoman scholar Ibn Kamāl Pāšā (d. 940/1533).[74] The term tarǧama may refer to any lexicographical work compiled by this scholar.

Ibn Muḥammad Ṣāliḥ is very explicit about the language he de-scribes in Šuḏūr. He calls it turkiyya (5ʳ2) and specifies the area where it is spoken as "the lands of Ottoman Turkic" (bilād at-turkiyya al-ʿuṯmāniyya, 3ᵛ22) and "the state of our time is the Ottoman state and their language is the sultanic Turkic language" (dawla zamāninā hāḏā d-dawla al-ʿuṯmāniyya wa-luǧatuhum at-turkiyya as-sulṭāniyya, 2ʳ7).

At the same time he considers the language described in Durra a form of Tatar: "The Turkic of his [sc. ʾAbū Ḥayyān's] time was Tatar, it is not used anymore in our time in the lands of Ottoman Turkic" (wa-ʾaydan turkī zamāni-hi tatarī matrūk fī zamāni-nā fī bilād at-turkiyya al-ʿuṯmāniyya, 3ᵛ23-24). Ṣāliḥ says that he went to Kafa[75] himself and had found its people speaking like ʾAbū Ḥayyān had de -scribed (cf. 3ᵛ26).

---

[70] Cf. Ḥāǧǧī Ḥalīfa V 230 and art. 'Sururī' by Babinger in EI[1]. Berthels described Sururī's vocabulary as following: "[Sururī's] work... is divided into three parts: verbs, particles and inflection, nouns. His sources were... Ṣiḥāḥ-i ʿaǧam " (Berthels, art. 'Sururī' EI[1]; cf. also art. by Akün in İA). There is also a MS of this work in the library of Leiden University, cod. 164 Golius.

[71] Muṣṭafā was sultan Sulaymān's (936/1530 – 973/1566) eldest son. He died at Konya in 960/1553.

[72] Cf. also Haig and Kramer's article on Saʿdī in EI[2].

[73] Ṣāliḥ calls this Saʿdī: "Šayḫ Saʿd ad-Dīn". Karatay (1961: II no. 2203), lists Saʿdī as Maḥmūd bn ʿUṯmān instead of Muḥammad bn ʿUṯmān. See Flemming on Lāmiʿī in EI[2].

[74] Cf. art. 'Kemālpašazāde' by V.L. Ménage in EI[2], and GAL II 597. See also Sawaie 1991.

[75] Kafa or Kefe lies at the south eastern coast of the Crimea. It was called 'Theodosia' in Roman and 'Feodosia' in Russian times (cf., for example, art. Kefe by Orhonlu in EI[2]).

Although Ṣāliḥ's reference is made to *Durra*, whose material is not sufficient as to permit far-reaching conclusions, there is some evidence for this in at least *ʾIdrāk* and *Qawānīn*. A brief comparison with Tatar as described in Doerfer's article on Crimean Tatar (1959) yields evidence that the language described in *ʾIdrāk* and *Qawānīn* may be related to Turkic languages spoken on the Crimea. The closeness of the Crimea to the Ottoman Turks and the fact that in that place also Ottoman was spoken would account for the many instances in which both Qipčaq and Oġuz features appear in the language material. The language in *ʾIdrāk* and *Qawānīn* may be characterised as Western Qipčaq, which could very well match 14th century Tatar. Furthermore, the Crimea may have been a strategic place for the tradesmen who shipped boys to Cairo. However, the exact determination of the origin of the languages in the sources is outside the scope of the present study, and must be left to further research (see also Section 1.3).

Šuḏūr was written in Ottoman times and based mainly on Ottoman sources.

1.8 Kitāb bulġa al-muštāq fī luġa at-turk wa-l-qifġāq

Bulġa is an Arabic-Turkic vocabulary, of which only one MS exists in the Bibliothèque Nationale, Paris, registered as Suppl. Turc No. 293. It consists of 71 folios, of which are missing (viz. after 12 *verso*, 13 *verso* and at the end). The lexical material is not put side-by-side, but arranged quite awkwardly in slanting lines. *Bulġa* was written in the 14th century AD by Ǧamāl ad-Dīn ʾAbū Muḥammad ʿAbdullāh at-Turkī, probably in Mamlūk-ruled Egypt. This Ǧamāl ad-Dīn at-Turkī could not be identified further.

Zajączkowski edited the text of *Bulġa* in two separate publications. In 1954 he published Part II on the verbs, and Part I, on the nouns, in 1958. Both editions were provided with quite poor facsimile copies of the relevant parts of the MS, and with a full edition of the text in print.[76]

After a brief introduction (7ʳ-7ᵛ) the lexical material is divided into four parts, which are categorised according to semantic category: I "the name of God and [his] servants in the sky and others" (*fī ismi llāh taʿālā wa-mā fī s-samāʾ min musaḥḥarin wa-ġayrihi*, 8ʳ3), II "chapter about the earth, the minerals and other [elements]", with several sections (*al-ʾarḍ wa-mā fī-hā min al-maʿādin wa-ġayrihā*, 9ʳ1); III "the locative" (*bāb aẓ-ẓarf*, 17ʳ), including a section about the pronoun;

---

[76] Reviewed by Pritsak (1957).

and finally, IV "the verbs" (*bāb fī l-ʾafʿāl*, 20ᵛ). *Bulġa* contains almost no grammatical information.

The title of the work, *Kitāb Bulġa al-Muštāq fī Luġa at-Turk wa-l-Qifǧāq*, suggests that the author distinguishes between Turkic on the one hand and Qipčaq on the other, but he does not give any details.

The author of *Bulġa* mentions several of the sources he used, among which are a work called *al-ʾAnwār al-muḍīʾa* written by ʿAlāʾ ad-Dīn Baylīk al-Qibǧāqī and one called *aṣ-Ṣaḥīḥ Min ad-Durra al-Muḍīʾa* by ʿImād ad-Dīn Dāwad[77] bn ʿAlī bn Muḥammad al-Warrāq al-Miṣrī (MS 7ᵛ) (cf. Pritsak 1959:75). Thus *Bulġa*, along with *ʾIdrāk* and, perhaps some other works, forms a part of the chain of which the afore-mentioned Baylīk (Section 1.3) was the initiator.

1.9 ad-Durra al-muḍīʾa fī l-luġa at-turkiyya

In 1963 Ananiasz Zajączkowski discovered the manuscript of an Arabic-Turkic vocabulary in the Medicea Laurenziana Library of Florence, registered under No. Orient 131.[78] The manuscript consists of 24 folios and probably originates from a *Mamlūk*-governed area, most likely 14th century Syria.[79] In the sixties of this century Zajączkowski published four articles in which he edited and discussed most of the lexical material he found in *Durra*[80] (cf. Zajączkowski 1965, 1968 and 1969).

The lexical material in *Durra* is divided into 24 chapters (the terms *qism* 'part', *bāb* 'chapter', *faṣl* 'section' seem to reflect the same level) that are arranged according to semantic subject, traditionally beginning with the names of God and Heaven and ending with a chapter containing colloquial language material (19ʳ5-24ʳ17). *Durra* contains no grammatical information.

As is clear from its title, the Turkic language in *Durra* is called *turkiyya* but no further information is given.

---

[77] The text in the MS is clearly vocalised *Dāwad* and not *Dāwud* as might be expected.

[78] Karatay (1961 II, 19 no. 2036; MS Hazine 1088). Although the introduction is the same as in the Florentine MS, it deals with Anatolian Turkic, rather than Qipčaq.

[79] Zajączkowski based this assumption on the Arabic language material in *Bulġa* (Cf. Zajączkowski 1954 II: 67).

[80] Although Ṣāliḥ attributes *Durra* to ʾAbū Ḥayyān al-ʾAndalusī (cf. *Šuḍūr* 1.7), in *Durra* itself I could not find any indications with regard to its author (cf. Section 1.6).

1.10 *An anonymous work in the margins of Veli ed-Dīn MS 2896*

The existence of yet another Turkic grammar written in the margins of
the Veli ed-Dīn Ms. of *Kitāb al-Idrāk* has been known since the thir-
ties of this century. Thusfar neither its exact title nor the name of the
author have been retraced yet and the text itself, too, still remains to
be thoroughly analysed.

This Margin Grammar (henceforth in references as MG) consists of
a long word list and a large grammatical part. The word list is scribbled
in the margins and between the lines of the lexicographical part of
'Idrāk (2^r-32^r); indeed, there are so many glosses that the original text
has become barely legible. The glosses were compiled and edited in
Roman script by İzbudak and Rifʿat in 1936. In addition to the
glosses, there is a grammatical text in the margins of the sections on
morphology and syntax (32^v-65^r) of *'Idrāk*, which seems to stand
independently of the latter. The handwriting on these pages is much
more legible than that in the word list, and there are only a few glosses
between the lines of the main text. The manuscript has yet to be
carefully examined in order to determine which fragments are in the
same hand as the body text of *'Idrāk* and which ones are not, but there
are enough data to make the following observations. [81]

First, one would expect that this grammar, being written in the
margins of *'Idrāk*, shows many resemblances to the latter. This is only
partially true, for only one chapter heading starts with *al-Qawl fī...*
(41^{v}bm), which is typical of *'Idrāk* and can be translated by 'Chapter
on...'. Some passages are identical with some in *'Idrāk*, albeit they are
not always found on the same folio. [82]

Second, most chapter and section headings are introduced by *Bāb*
("Chapter") or *Faṣl* ("Section"), *Tanbīh* ("Remark"), *an-Nawʿ al-
ʾawwal* ("The first type is..."), *aḍ-Ḍarb al-ʾawwal* ("The first kind
is...") and *al-Qāʿida fī...* ("The rule with regard to..."). These
headings are also typically found in *Qawānīn*. Indeed, there are more
points of resemblance with *Qawānīn*. Apart from an explicit reference
and a quotation (33^r, cf. *Qawānīn* 4,5-15) there are also numerous
other passages that—in some cases with only minor changes—are also
found in *Qawānīn*. [83]

---

[81] The last pages of the volume (65^v-66^r), after the manuscript of *Idrāk*, are dedicated
solely to this, or perhaps yet another grammar.

[82] E.g., 34^{v}rt = *Idrāk* 105,12; 35^{v}bm/rt = *Idrāk* 107,15; 56^r top = *Idrāk* 135,18f

[83] Cf. 37^{v}rt = *Qaw.* 12,9-10; 38^{v}top = *Qaw.* 11,7-8; 38^{r}lt=*Qaw.* 10, 21-3; 41^{v}rt =
Qaw. 23,5-15; 43^{v}top = *Qaw.* 49-50; 44^{v}lt/ult = *Qaw.* 8,4-5; 59^{v}rt/ult. = *Qaw.* 44,5-9
(with minor change); 64^{r}top = *Qaw.* 21,14-7; 64^{v}lt/ult = *Qaw.* 21,18.

Thirdly, other fragments are identical with passages found in
Tarǧumān, e.g. *lafẓa ʾidī at-turkiyya* ("the Turkic word ʾ**idī**", MG 46ᵣlt
= *Tarǧumān* 50,11). This typical style of *Tarǧumān* is found further in
passages like *lafẓa **dakūl** at-turkiyya* ("the Turkic word **dakūl**," MG
47ʳ) which, however (with other examples in the same style on 39ᵣlt
and 64ᵛrt; 50ᵣlt = *Targ.* 50,6-10), is not found in the latter. From this
it is possible to infer that the author of the Margin Grammar had a dif-
ferent copy of *Tarǧumān* at his disposal than the one Houtsma used
for his famous edition.

In the fourth place, some passages are literally copied from *Dīwān*,
e.g., the fragment on the consonants of Turkic (33ᵣlt) is exactly the
same as *Dīwān* (6,6-7,6), albeit without reference. There are more pas-
sages copied from *Dīwān*, for example the fragment on (40ᵛrt) is a lit-
eral copy of *Dīwān* (538,5-8); MG (45ᵣlt) resembles to a large extent
Dīwān (301,13-16), and MG (39ᵛrt) is identical with *Dīwān* 536,13ff.
The resemblance always includes the Arabic versions of the examples,
although the Turkic language is in most cases probably another one.

Five, the fragments are generally not related to the subjects dealt
with in the corresponding pages of main text of *ʾIdrāk*. In most cases
the same issue is dealt with in different ways in several passages on a
given page (e.g. particles of the genitive *(ḥurūf al-ǧarr)* 58ᵣlt; 58ʳtop;
58ʳrt; 58ᵛrt). Fragments from different manuscripts are in a few in-
stances introduced with the phrase *wa-fī nusḥa* (occasionally *nusaḥin*)
ʾuḥrā ("and in another manuscript[s]", cf. 50ʳbm; 64ᵣlt/ml; 50ᵣlt;
51ᵛrt/bm).

Finally, apart from *Qawānīn*, *Tarǧumān*, *ʾIdrāk* and *Dīwān*, there
must have been at least one, and perhaps more than one, other main
source that the copyist/author used for his compilation. This is, for ex-
ample, evident from the approach in MG to linguistic matters, such as
the analysis of the possesssive and the accusative case endings which, as
will be shown in Chapters Four and Five, differs considerably from the
way these issues are dealt with in the other sources. There is also an-
other reference to a name or a title, i.e. *al-ʿAllāma baqiyyat as-salaf
wa-qudrat al-ḥalaf tarǧumān al-ʾadab... (?) ...lisān al-ʿarab* (1ᵛ).

The author of some of the fragments may have been of Turkic de-
scent himself, since he refers to the Turks as "we" (e.g., *ʿindanā* "with
us", 34ᵛtop) and to Turkic as "our language" (*luǧatunā*, 34ᵛtop). In the
body of the text Oğuz forms are given quite often, indicated as such

with the term *Turkmān* or *Turkmānī,* as opposed to *Turk* or *Turkī.*[84] No other ethnonyms are found.

Summarising, the Margin Grammar in both style, set up and contents has more in common with, and contains passages from *Dīwān,* *Qawānīn* and *Tarğumān,* rather than *'Idrāk.* Furthermore, many passages display an entirely new, as yet unknown source. Therefore, the Margin Grammar is best characterised as a compilation of several sources, based on other and, perhaps, more extended manuscripts of *Qawānīn, Tarğumān* and *Dīwān.* In this sense the whole grammar cannot be ascribed to one single author.

2. THE SOURCES AND THEIR PRIMARY MATERIAL

The references in the sources first of all allow us to draw conclusions in regard to the Turkic and Arabic material on which they were based. In this respect I use the term primary sources. In most cases this means an evaluation of the importance of texts that are no longer available to us. Second, under the heading 'indirect evidence', in which I continue the discussion of the Arabic primary sources, I present my findings with regard to the respective internal structures of the sources.

2.1 *Direct evidence: references*

In the first sections of this chapter I pointed out that each of the sources was based on previously compiled material. This primary material can be divided into two categories. The first category comprises Turkic sources that provided the lexical and syntactic material, i.e. previous studies of Turkic in Arabic or in other languages, Turkic texts and, of course, native speakers.[85] The second category of primary material comprises Arabic sources on which the model of the work was based, i.e. studies of Arabic grammar (or lexicography, for that matter) and, of course, the author's teachers. As a result, treatises that describe very different Turkic languages, or even unrelated ones like Mongolian or Persian, may show a high degree of resemblance in regard to the arrangement and interpretation of the lexical and syntactic material.[86]

---

[84] The ethnonym *Turkmānī* - *Turkmān* occurs on: 59ʳrt; 44ᵛlt/ult; 51ʳrt; 51ʳrt; 53ᵛbm/lt; 51ᵛrt; 42ᵛbm; 45ᵛtop; 59ᵛmd; 49ᵛrt; 58ᵛtop/ult; 58ᵛrt.

[85] Indications for a common Turkic basic source might be found after careful analysis of the lexical and syntactic material, but this is beyond the scope of the present study.

[86] I already established this for *Dīwān luğāt at-Turk.*

2.1.1 Turkic primary material

The best indication for a direct dependence on primary material is, of course, a reference to a work or an author, like the references I found for each of the sources. Based on the available data, at least two different groups of works with common primary material can be distinguished. To the first group belong *'Idrāk, Tuḥfa, Bulġa, Durra* and *Šuḏūr*. To the second group belongs *Ḥilya* with its primary material. *Dīwān, Qawānīn* and *Tarǧumān* cannot be incorporated into either group, since none of them mentions any Turkic sources.

This is illustrated in Figure 1 (direct references are indicated with a black line, indirect references with a dotted line):
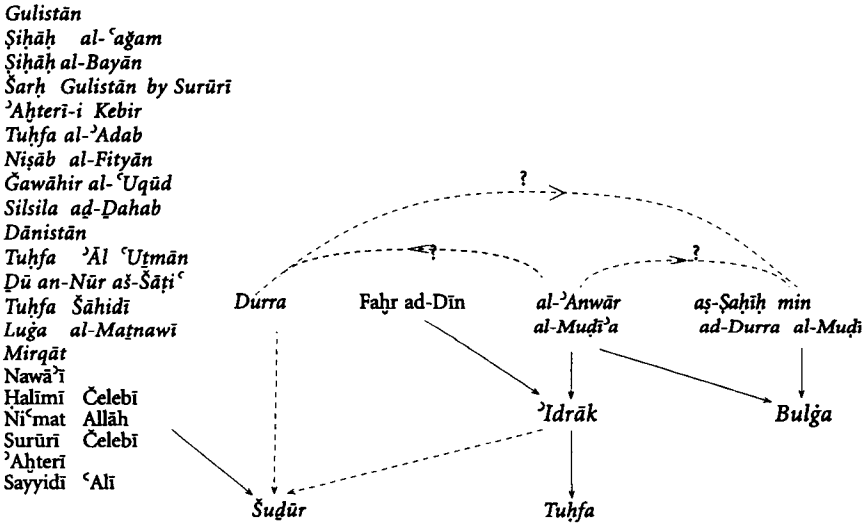
Figure 1 Turkic primary sources of
Durra, Šuḏūr, 'Idrāk, Bulġa and *Tuḥfa*

The author of *Šuḏūr* lists a number of works, most of which do not originate from the Arabic tradition but rather from the Ottoman. To this tradition belong word lists and texts in Arabic, Ottoman Turkish, Persian and Čaġatay. Most of these were compiled in the fourteenth and fifteenth century AD by well-known Ottoman writers, namely Aḥterī, Ḥalīmī, Niʿmat Allāh, Surūrī and Sayyidī ʿAlī. The only connection of *Šuḏūr* with the other sources mentioned in Figure 1 is the author's reference to the work *ad-Durra al-Muḍīʾa fī l-Luġa at-Turkiyya*, at the same time attributing it to 'Abū Ḥayyān, the author of *'Idrāk*. 'Abū Ḥayyān's fame for his Turkic grammars may have become

so widespread that works whose author's name had been forgotten, were attributed to him. I have already pointed out a similar attribution of *Qawānīn* to 'Abū Ḥayyān; below (3.2) it is shown that 'Abū Ḥayyān is not likely to have been its author.[87]

In *'Idrāk* 'Abū Ḥayyān often refers to *šayḫunā* ("our master") Faḫr ad-Dīn; he always does so in the context of a Turkic word. Therefore I assume that his 'master' Faḫr ad-Dīn knew Turkic, although the fact that 'Abū Ḥayyān studied with him the seven readings of the Qur'ān would not indicate this (see Section 1.3). As a consequence, Faḫr ad-Dīn must be taken as a source for Turkic primary material rather than Arabic. Furthermore, 'Abū Ḥayyān mentions the name of Baylik al-Qifǧāqī, who is the author of *al-'Anwār al-muḍī'a*, a work whose contents are not exactly known.

Bulǧa mentions two direct sources: *aṣ-Ṣaḥīḥ min ad-Durra al-muḍī'a* ('Imād ad-Dīn) and *al-'Anwār al-muḍī'a*. Although I do not know the contents of these works, the resemblance of their titles to *ad-Durra al-muḍī'a fī l-luġa at-turkiyya* is in my view no coincidence: the titles are meant to refer to some mutual link; *Durra* was either based on *al-'Anwār al-muḍī'a* or formed the basis for it itself. The work *aṣ-Ṣaḥīḥ min ad-durra al-muḍī'a*, likewise on the basis of its title, may have been based on *Durra*. Further research on both *Durra*'s and *Bulǧa*'s Turkic language material and their structure has to be done in order to show whether any more common features exist.

A nice example of direct reference is *Tuḥfa*'s reference to 'Abū Ḥayyān's *'Idrāk*. Since we find in *Tuḥfa* not only 'Abū Ḥayyān's name but also a literal copy of the passage in which 'Abū Ḥayyān explains the principles for the description of a language, it seems reasonable to assume that *'Idrāk* forms a direct source for *Tuḥfa*. In the scope of the present study, however, we shall not attempt to determine the extent to which Turkic language material has been transmitted from *'Idrāk* to *Tuḥfa*, but concentrate on the way the Arabic model was applied in both works. The findings with regard to *Ḥilya* are summarised in Figure 2 (see page 46). In the first place we find a reference to Muḥammad bn Qays, who may have been the author of one or more of the works mentioned. As a result, the link between the name Muḥammad bn Qays and *Ḥilya* may have to be replaced by any one of the other links. As for the title *Kitāb Yaḥyā al-Malik*, it may be a false reading *Kitāb Ḥilā al-Malik* (see Section 1.2). The findings with regard to the Margin Grammar are summarised in Figure 3 (see page 46).

---

[87] Because of this misunderstanding *Qawānīn* is registered under 'Abū Ḥayyān's name in the library of the Süleymaniye in Istanbul.

Muḥammad bn Qays

Tuḥfa al-Malik 　　　　*Kitāb Nādir ad-Dahr*

　　　　　　　　　　　　　　　Kitāb Yaḥyā
　　　　　　　　　　　　　　　al-Malik
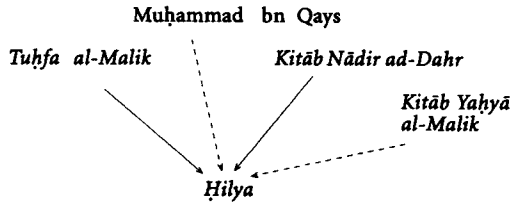
Ḥilya

Figure 2 *Ḥilya* and its Turkic primary sources

Qawānīn　　*Tarǧumān*　　*ʾIdrāk*　　*Dīwān*
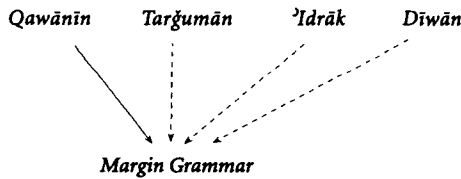
Margin Grammar

Figure 3 The Margin Grammar and its Turkic primary sources.

As is shown in Section 1.10, there is ample direct and indirect evidence for the Margin Grammar's relations to *Qawānīn*, and convincing indirect evidence for relations with *Tarǧumān*, *ʾIdrāk*, and *Dīwān*.

2.1.2 Arabic primary sources

The second category of primary material, the Arabic sources, is not easy to point out for each work separately. The authors of most of the sources do not make reference to their teachers, nor to the Arabic works they used. In most cases they probably saw no need to stress a link between grammars of Arabic and the work on Turkic they were compiling. As a consequence, they considered it superfluous to refer to their Arabic primary material. Nevertheless, some of the authors do

mention some of their teachers and the Arabic primary works they used. The data are summarised in figure 4:

Tāğ ad-Dīn

'Idrāk   at-Taftazānī   Ibn al-Ḥabbāz   *Kitāb al-ʿAyn*
('Abū Ḥayyān)   (al-Ḥalīl)

Tuḥfa   *Šuḍūr*   *Qawānīn*   *Dīwān*
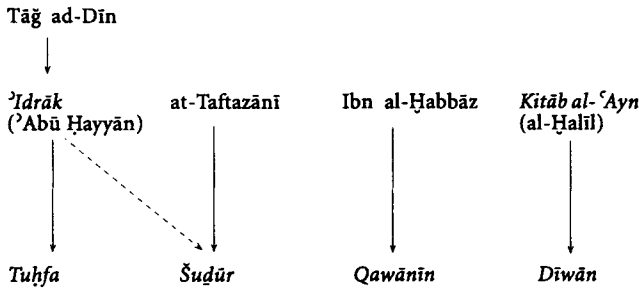
Figure 4 Arabic primary sources.

The few names they mention—*Dīwān* refers to al-Ḥalīl, *'Idrāk* to Tāğ ad-Dīn, *Qawānīn* to Ibn al-Ḥabbāz, *Šuḍūr* to both 'Abū Ḥayyān (implying *Durra*) and at-Taftazānī and, finally, *Tuḥfa* to 'Abū Ḥayyān—show that the authors had had a general education in Arabic linguistics. The direct influence of any of this primary material, how-ever, is impossible to measure. Some relationships are even uncertain, for instance, *Šuḍūr*'s dependence on 'Abū Ḥayyān.

A very sound dependency relation is the link between *Tuḥfa* and 'Abū Ḥayyān's *'Idrāk*. *Tuḥfa*'s dependence on *'Idrāk* is reflected in both the direct reference to 'Abū Ḥayyān's principles and in its adop-tion of the same structure as *'Idrāk*, which I discuss in Section 2.2. In this respect, 'Abū Ḥayyān—or perhaps just *'Idrāk*—may be interpreted as *Tuḥfa*'s main source for the Arabic linguistic model and perhaps also for a part of the Turkic language material.

Tarğumān does not mention any Arabic primary sources, therefore it cannot be incorporated in either one of the figures.

2.2 *Indirect evidence: internal structure*

In Section 2.1. I showed evidence for direct relations between the sources, let us now define the place of each one of the sources within the Arabic linguistic tradition based on the findings with regard to the internal structure of the primary sources (See Appendix to chapter Two, page 55 ff.).

Before proceeding, it is appropriate to say something about the im-portance of the way the grammatical data are structured in the sources.

In the first section (1.1-10) it is shown that some of them (namely
Dīwān, *Ḥilya*, *ʾIdrāk*, *Tuḥfa*, *Qawānīn*, *Tarǧumān*, the Margin Gram-
mar and *Šuḏūr*) contain grammatical information on a Turkic lan-
guage —or even more than one, for that matter. Other works (*Bulġa*
and *Durra*) just consist of a word list. Let us consider again the subject
of this study: the way Arabic grammarians described Turkic lan guages.
In this respect I must obviously concentrate on treatises that contain
grammatical rather than lexical data data. But, apart from the amount
of the data, for this study it is equally important that the grammatical
data upon which it is based are structured coherently. In other words,
it is not only the body of data itself that is of importance, but also its
typological context.

It is possible to divide some of the sources into two 'traditions',
based on a common internal structure that already existed within the
Arabic grammatical tradition, namely one that follows ʾAbū Ḥayyān's
Irtišāf and a second one that follows the traditional pattern of Arabic
grammar works, for which Zamaḫšarī's *Mufaṣṣal* is taken as an exam-
ple. The only work that stands beyond this division is Kāšġarī's *Dīwān*,
which belongs to the lexicographical rather than the grammatical tradi-
tion.

2.2.1 *ʾIdrāk* and *Tuḥfa*

In the first place, as has already been noted, there seem to exist some
connections between *ʾIdrāk* and *Tuḥfa*. I have already referred to the
fact that *Tuḥfa*'s anonymous author quotes ʾAbū Ḥayyān. When com-
paring the structure of both works, the consequences of his adherence
to ʾAbū Ḥayyān's principles are evident. Both works are divided into
four parts: 1. a word list (lexicology); 2. a brief introduction into
phonological matters, 3. morphology and 4. syntax. Even the names of
the chapter headings are in many cases literally the same. Thus, etymo-
logical, morphological and syntactic subjects are treated separately.

ʾIdrāk's organisation is only slightly different from that of another
work of ʾAbū Ḥayyān's, namely *Irtišāf aḍ-ḍarb fī lisān al-ʿArab*. *Irtišāf*
consists of three parts: first, phonology (I 4-12), second, morphology
(*ʾaḥkām al-kalim qabla t-tarkīb*) and third syntax (*ʾaḥkām al-kalim
ḥālata t-tarkīb*). To the second part belong, for example, the patterns
of the word (14-162); the diminutive form (I 169); feminine form (I
179); the broken plural (192); regular plural forms (I 262), etc. The
third part deals with all aspects of nouns, verbs and particles; in the
first place the nouns: declension (I 411); pronouns (I 481); proper
names (I 496); the relative (I 523); topic and predicate (II 3); locatives

and genitives as predicate (II 54); the verb *kāna* 'he was' (II 72); the verbs of the heart (*'afʿāl al-qulūb*; II 118); verb-like particles *'inna* and *'anna* (II 128); words with the ending *-a* (*al-manṣūbāt*; II 201): (a) nouns (II 202), i.e. different kinds of objects; (b) verbs (II 387); words, i.e. nouns, with the ending *-i* (II 426); discussion of the particles of the genitive (II 426) and annexation (II 501); words, i.e. verbs, with zero-ending (II 541). After this follow various issues such as the attribute (*naʿt*, II 579); corroborative (II 613); substitution (II 619) the conjunction (II 629). Volume III starts the discussion of topics related to verbs (III 3); transitive to more than one object (III 56); the verbal noun (III 170); active participle (III 181); and finally the particles (III 255), such as the particle of interrogation (*hal*, III 257).

This division has some points of similarity with that of Sībawayhi's[88] *Kitāb*. The contents of the 574 *'abwāb* (chapters) the *Kitāb* can be roughly divided into three parts, i.e. syntax (1-284), morphology (285-476) and phonology (477-574) (cf. also Versteegh 1987:154).

In *Irtišāf* the division of speech into the three main categories, i.e. noun, verb and particle, plays a role on a different level than in other works of grammar, such as Zamaḫšarī's *Mufaṣṣal*. In works like *Mufaṣṣal*, as is pointed out below, the tripartite division according to word class (nouns, verbs and particles) is the principal basis.

2.2.2 *Ḥilya*, *Qawānīn*, *Tarǧumān*, Margin Grammar and *Šuḏūr*

As has been briefly touched upon in the previous section, the other sources, namely *Ḥilya*, *Qawānīn*, Margin Grammar and *Tarǧumān* (while there are indications for this in *Šuḏūr* as well) show an entirely distinct division. This division is based on the tripartite classification of the elements of speech in Arabic grammar, i.e. 1. noun, 2. verb and 3. particle. Inasfar as they contain word lists, the different items are divided among the three main sections.

This lay-out very much resembles that of the traditional organisation of Arabic grammar that started with Ibn Sarrāǧ's *'Uṣūl* (for which see Owens 1988:28f.). As an illustration I take Zamaḫšarī's (467/1075 – 538/1144) *al-Mufaṣṣal fī n-naḥw*. After an introduction, *Mufaṣṣal* is divided into three main parts, and an additional fourth part: 1. *'asmā'* "nouns" (4-108); 2. *'afʿāl* "verbs" (108-130); 3. *ḥurūf* "particles" (130-

---

[88] 'Abū Bišr ʿAmr (d. 177/793), whose *Kitāb* was the first known grammar of Arabic. Sībawayh's master was al-Ḫalīl Ibn 'Aḥmad (d. 175/791) who is said to have compiled the first dictionary of Arabic: *Kitāb al-ʿayn* (cf. Wild 1965 and Talmon 1997). Bohas et al. (1990) claim that Sībawayhi's *Kitāb* was not intended as a coherent 'book' on Arabic grammar, but rather as a bundle of case studies of different grammatical subjects.

158) and 4. *al-muštarik* "common" (158-197). The fourth part deals
with features that occur in nouns as well as verbs and particles, such as
phonological issues among which assimilation of consonants (cf. also
Versteegh 1987:154).

All this of course does not imply that I believe these works are liter-
ally based on *Mufaṣṣal*, it merely indicates the fact that they are organ-
ised much in a similar way. Indeed, it was the standard organisation of
grammar works. For *Bulġa* and *Durra* it is hardly possible to discern
any overall main division since both are mere wordlists.

2.2.3 *Dīwān luġāt at-turk*

With regard to Kāšġarī's *Dīwān*, we can discern a similar division into
nouns and verbs. However, this is not enough to add it to the second
group. More important is the fact that *Dīwān*'s structure is very much
different, for it was based on Fārābī's lexicon *Dīwān al-ʾadab fī bayān
luġa al-ʿarab*, in spite of the fact that no direct reference to it is found.
I have already given some examples of *Dīwān*'s structure in Section
1.1. In the present section further details about the structure of both
works are provided, and some statements as to the aptness of this
structure for Turkic languages.[89]

To start with, as Muḥtār, the editor of *Dīwān al-ʾAdab*, points out,
its principle of arranging entries is not very practical (18). The user has
to know the structure of a word—and to know whether it is either
sālim, *muḍāʿaf*, *ʾaǧwaf* or *nāqiṣ* etc.—before looking it up. Such a sys-
tem can be of use to people who are looking for rhyming words, like
poets. On the other hand, Muḥtār admits, it is true that the last conso-
nant of a root is subject to very few changes, unlike its first consonant,
which can be preceded or followed by additional consonants for se-
mantic reasons; in this way the original consonant pattern of a word
can become blurred. An example of this is: (/f-ʿ-l/ 'write') /k-t-b/ can
be changed into /k-t-t-b/ 'to make (someone) write, to write intensive-
ly', /ʾ-k-t-b/ "to make (someone) write", /k-"-t-b/ 'to write someone',
/y-k-t-b/ 'he writes' /m-k-t-b/ 'place to write' etc. The last consonant
is affected only when the root gets pronominal suffixes, e.g. /k-t-b-t/
(*katabtu* 'I wrote', *katabta* 'you wrote', *katabti* 'you (f) wrote', *katabat*
'she wrote') or /k-t-b-n-"/ (*katabnā* 'I wrote') etc., and in some pat-

---

[89] Another striking feature which sets *Dīwān* apart from the other sources is its
terminology. Kāšġarī uses an unusual, i.e. non-canonical, set of terms for describing ob-
jects in constructions with causative verbs (Chapter Six, 3.3.2), and some terms which
remind of the Kufan tradition, e.g., *ǧaḥd* 'negation', *ǧābir* 'present tense' and *mubārā*
'reciprocity'.

terns in which the last consonant is doubled, e.g. the pattern /ʾ-f-ʿ-l-l/ e.g. *ibyaḍḍa* 'to become white'.

With this in mind, Kāšġarī's choice of Fārābī's system is quite sur-prising, since in Turkic languages all semantic changes affect the last part of the root, never the first. On the other hand, though, his system allows the reader to have a quick view on roots with the same type of ending (such as, e.g., vowel, voiceless, or voiced consonant) which de-termine the choice of suffixes the root takes.

Let us consider an example: The stem **biti-** means 'to write' (427,9); **bitik biti-t-tī** *ʾanna-hu ʾaktaba l-kitāba* ("he made [someone] write the book", 415,10) under the heading *Bāb aṭ-ṯulāṯī—Bāb faʿaldī muḥar-rak al-ḥašw fī ḥarakāti-hi* ("Chapter of triradical words—Chapter on *faʿaldī*, with the middle radical vocalised", 415,2). The pattern *faʿaldī* is adapted to Turkic verbs, which are conjugated for the imperfect form of the third person singular by adding the ending -**dī** to the stem. The word **biti-k** means "anything written" (193,9), **bitik-či** "scribe", **biti-k-lik** "something prepired to be written on" (254,1), **biti-k-liḵ** "its owner [sc. of writing material]" (254,1-2). Of course, the stem can receive pronominal suffixes too, hence: **biti-dī** (427,9) 'he wrote' and **biti-di-k** 'we wrote'. We see that most of these words are listed in *Dīwān*, but scattered throughout the work.

In other vocabularies, in which the first consonant of the root is the main listing criterion, words of the same stem are neatly put together. An example of this is the word **biti-k** in ʾAbū Ḥayyān's *ʾIdrāk* (28). The main entry is **bitī** while **biti-k** is given as an alternative. The fact that alternatives can be found relatively easily can be even more important for at least one reason: it is not likely that all potential users of the dictionary would spell a word in the same way, since there were no orthographic conventions for Turkic. With the entries rigidly listed according to the last radical, as in *Dīwān*, a user looking for a dialectal or regional variant would not be able to find this word.

3. CONCLUSIONS

In the preceding section I have shown that the data found in the ten sources can be divided into two groups: direct evidence and indirect evidence. With regard to the direct evidence, a distinction was drawn between references to Turkic primary material on the one hand and Arabic primary material on the other. Since direct references to the Arabic primary material are very scarce, the internal structure of the sources was taken as additional 'indirect evidence'. There are two types

of internal structure: the first (A) is similar to the traditional orga-
nisation, illustrated by Zamaḫšarī's *Mufaṣṣal*, whereas the second (B)
to ʾAbū Ḥayyān's *Irtišāf.*

This is illustrated in Figure 5:
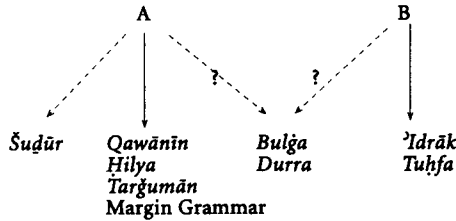
Figure 5 The sources in the context of the Arabic
grammatical tradition

There remains one work which has not been placed in either group,
namely Kāšġarī's *Dīwān Luġāt at-Turk*. The reason for this lies in the
fact that it was based on Fārābī's *Dīwān al-ʾAdab* and, perhaps, to
some extent on al-Ḫalīl's *Kitāb al-ʿAyn*. From the structure of *Dīwān
al-ʾAdab*, *Kitāb al-ʿAyn* and *Dīwān Luġāt at-Turk*, it becomes obvious
that none of them was meant as a grammar, but rather as a lexico-
graphical work.

It is difficult to decide which place *Kitāb al-ʿAyn* occupies with re-
gard to *Dīwān Luġāt at-Turk*, since in the first place it is mentioned as
a direct source—although its influence cannot be measured—and sec-
ond it must have been a basis for *Dīwān al-ʾAdab*, since *Kitāb al-ʿAyn*
was the first lexicographical work in the Arabic linguistic tradition (Cf.
Wild 1965, Haywood 1965). In other words, *Kitāb al-ʿAyn* has served
as a direct primary source for *Dīwān Luġāt at-Turk* and simultaneously
stands at the head of a tradition of lexicographical works of which
Dīwān Luġāt at-Turk also is a part, illustrated in figure 6 (page 53).
Dīwān deserves a separate study with regard to its position within the
Arabic lexicographical tradition—which stands more or less indepen-
dently of the grammatical tradition—and, further, the way Kāšġarī in-
terprets the system Fārābī used to arrange the Arabic roots and the ex-
tent to which *Dīwān* shows the influence of al-Ḫalīl's *Kitāb al-ʿAyn*.

Kitāb al-ʿAyn
(al-Ḫalīl)

Dīwān al-ʾAdab
(Farābī)

Dīwān Luġāt at-Turk
(Kāšġarī)
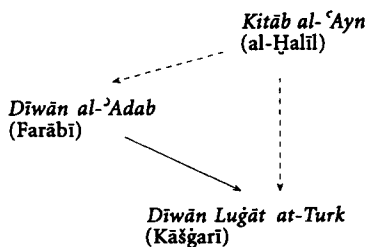
Figure 6 *Dīwān Luġāt at-Turk* and its place
in Arabic lexicography

As to the relations between the sources with regard to the Turkic language material, there is reliable evidence for dependency relations between two of them (*Tuḥfa—ʾIdrāk*). There are also references to a number of primary sources (*ʾIdrāk*—Faḫr ad-Dīn and *ʾAnwār, Bulġa—ʾAnwār* and *Ṣaḥīḥ; Ḥilya—Tuḥfa al-Malik,* Muḥammad bn Qays, *Kitāb Nādir ad-Dahr* and *Ḥilā al-Malik*).[90] In some other cases, in which the evidence was not that strong, some claims were made with regard to possible relations (e.g. regarding *Šuḍūr* vs. both *ʾIdrāk* and *Durra* and *Durra* vs. *ʾAnwār* and *Ṣaḥīḥ*).

As for the place in the Arabic linguistic tradition, there are in the first place direct references to Arabic scholars (al-Ḫalīl, ʾAbū Ḥayyān, at-Taftazānī, Ibn al-Ḥabbāz, Faḫr ad-Dīn, Tāǧ ad-Dīn). In the second place evidence has been found for the existence of at least two 'traditions' within the Arabic - Turkic linguistic grammatical tradition. The first 'tradition', which comprises both *ʾIdrāk* and *Tuḥfa*, largely follows the pattern of Sībawayhi's *Kitāb*. The second 'tradition', to which belong *Ḥilya, Qawānīn, Tarǧumān* and, perhaps, *Šuḍūr*, is the traditional organisation. The fact that *Qawānīn* has a structure which is quite different from that of *ʾIdrāk* proves that it is not likely to have been compiled by ʾAbū Ḥayyān, for the latter generally adhered to his own linguistic principles, even when he wrote his grammar of Turkic. The same goes for *Durra*, although its contents are, admittedly, not sufficiently elaborate as to sustain this assumption in full. As a matter

---

[90] I propose to read *Ḥilā al-Malik* instead of *Yaḥyā al-Malik* (cf. Section 1.2).

of fact, neither *Durra* nor *Bulġa* nor *Šuḍūr* can be incorporated with certainty in either of the two 'traditions'.

Dīwān, in spite of its length and the abundance of information that it presents, belongs to an entirely different class. This can be deduced from both its own structure as well as from the typology of its sources, namely *Kitāb al-ʿAyn* and *Dīwān al-ʾAdab*. Therefore, *Dīwān* belongs to the Arabic lexicographical rather than to the grammatical tradition.

The findings with regard to the different structure of our respective sources have consequences for the way they can be used in an analysis of the way Arabic grammarians described Turkic languages. In my view, the best method to accomplish this is to compare those of the sources that were compiled with a similar goal: to explain grammatical features of Turkic in terms of Arabic grammar. As long as all or the most important grammatical data are given, it is not necessary that all sources have a similar structure. It is obvious, however, that those sources that only contain a word list, and none or very few grammatical data, namely *Bulġa*, *Durra* and to a certain extent *Šuḍūr*, cannot meet this criterion. Similarly, a source that to some extent does contain the desired data but within an entirely different context, namely Kāšġarī's *Dīwān*, does not fit entirely within the frame of this study either. In some cases the data that these sources provide may be useful, e.g. when the structural context is not relevant. For this reason, we shall concentrate on the representatives of the two streams, i.e. *Ḥilya*, *Qawānīn*, *Tarġumān*, the Margin Grammar, *ʾIdrāk* and *Tuḥfa*, and on the other sources as far as they provide grammatical analyses.

APPENDIX TO CHAPTER TWO: CONTENTS OF THE SOURCES

The following gives an overview of the most important sources in this study. These are *Tuḥfa, Ḥilya, Qawānīn, Targumān* and the Margin Grammar. An overview of *ʾIdrāk* is given in Part Two of this book.

Kitāb Ḥilya al-Insān wa-Ḥalaba al-Lisān

al-Qawānīn al-Kulliya li-Ḍabṭ al-luġa at-turkiyya

Tarǧumān

Anonymous Grammar in the Margins of Veli ed-Dīn Ms 2896

(The fragments are given in the sequence in which they occur in the manuscript Since their sequence has not yet been established, the numbers in the ultimate left column do not necessarily indicate separate sections Abbreviations: bm = bottom; md = middle; lt = left; rt = right)

CHAPTER THREE

PERCEPTION AND DESCRIPTION OF PHONOLOGICAL AND PHONETIC FEATURES OF TURKIC

INTRODUCTION

This chapter aims at shedding light on the way the Arabic sources per-
ceived and described the phonetic and phonological features of Turkic.
It deals with matters such as the quantity and quality of Turkic conso-
nants and vowels, vowel harmony and consonant assimilation. The
present survey is a study of concepts and perceptions. It does therefore
not give a description of the phonetics and phonology of 'medieval
Turkic'.

The study of perceptions and concepts is not related to just one is-
sue, but two. The first issue is how the Arab authors perceived Turkic
phonetics and phonology. The second issue is how they formulated
their perceptions by means of the concepts which they had acquired
within the Arabic linguistic tradition. One would expect the descrip-
tions to be fairly similar, if not uniform, since they eventually derive
from the same model. The present study shows that they are not; at
least there are more discrepancies than one would expect, even allow-
ing for differences in the Turkic languages, and it appears that the au-
thors had various phonetic and phonological concepts of Arabic as
well.

Although studies by Cantineau (1960), Fleisch (1961) and Roman
(1983) have shed some light on the interpretation of phonetic descrip-
tions in Arabic source material (al-Ḥalīl, Sībawayh, Ibn Ǧinnī), there
still is no general agreement as to the exact phonetic characteristics of
the classical Arabic language of their times. However, for the purposes
of this study the exact characteristics of classical Arabic are, in fact, of
minor importance. It suffices as a starting point that there indeed was
agreement among the Arab grammarians and phoneticians themselves
on the quantity and quality of consonants and vowels, regardless of
whether we comprehend the descriptions they provided in full detail.

This brings us to a following problem, namely that of accuracy of
the descriptions. Without a full understanding the Arabic descriptions
and knowledge of the phonetic features of 'medieval Turkic', judging

their accuracy would be, indeed, a very complex matter. However, a
number of features can safely be postulated for all Turkic languages.
In this respect one should think of vowel harmony, consonant assimi-
lation and the occurrence of certain consonants and vowels, such as,
e.g., [ö, o] and [ŋ]. It shall be seen that basic principles from Arabic
theory on velarisation and palatalisation play an important role in the
description of Turkic vowels. These principles suffice for the rather
rough distinction between back and front vowels, but they lack much
of the precision which is needed for a clear distinction between
rounded and unrounded vowels. Nevertheless, my intention is to show
that with these instruments it is possible to accurately describe Turkic
vowels.

   The descriptions of Turkic are interesting from yet another point of
view. As is generally know, the Arab grammarians as a rule hardly de-
scribe dialectal features. Such descriptions could be instructive for both
the historical development of dialectal phonology, and for better assess-
ing their concepts of phonetics. The study of how they applied their
phonetic and phonological concepts on foreign languages provides a
good opportunity for this. In this chapter I shall show that the per-
ceived phonetic qualities of dialectal pronunciation is more likely to be
the basis for those concepts than the traditional prescriptions of
Classical Arabic.

   In this chapter, section I gives an outline of Arabic thinking on
phonetics and morphonology, followed by a study of Turkic phonemes
in Section 2. Further, sections 3 provides an extensive discussion of the
Arab grammarians' perception of the phonetical and morphonological
aspects of Turkic.

1. ARABIC THINKING ON PHONETICS AND MORPHOLOGY

Although the main interest of the Arab grammarians was the descrip-
tion of syntactic and morphological issues, they also described the pro-
nunciation of Arabic. They based their discussions on speech examples
of Bedouin tribes living on the Arabian Peninsula. Although the di-
alects of those tribes also differed considerably from each other and
from the Classical Arabic language, their status was very much higher
than that of the vernaculars of the people in the cities. For a long time
after the Islamic conquests, Arab scholars continued to refer to the old
speech examples taken from the speech of so-called unspoilt Bedouin
tribes, which were supposed to have preserved the Arabic language from

external influences. The Bedouins were considered 'native speakers' of the language of the holy Qurʾān (cf. Levin 1994).

Arab grammarians, in their treatises on Arabic grammar, discussed the speech and poetry of a small number of Bedouin tribes living on the Arabian Peninsula. Although they were principally not interested in phonetics, they did not fail to notice the variations in use and pronunciation of the Arabic language, especially among the Bedouin tribes.[1] Studying the speech examples they obtained from selected tribes, the grammarians found themselves confronted with different ways of pronouncing some consonants and vowels, consonant and vowel exchange, and assimilation of consonants. These subjects had to be described and explained, since some of them could affect the form and, hence, the meaning.

The first Arabic work on grammar was Sībawayh's *Kitāb*, compiled in the eighth century AD, which became the ultimate reference source for all grammarians throughout for many centuries to follow. A relatively small number of pages at the end of *Kitāb* are dedicated to phonology and phonetics. Most later grammarians also treated phonetics at the end of their works (Owens 1988:28-9). An exception to this convention is ʾAbū Ḥayyān's *Irtišāf*, in which the first sections are dedicated to phonological matters such as the places of articulation, the substitution of consonants and assimilation.

Before discussing how exactly Arab grammarians perceived and described Turkic phonemes, it is necessary to give a brief outline of Arabic theory on phonology and phonetics, with special attention for the description of consonants and allophones, and those of vowels and vowel harmony.[2]

1.1 *The Arabic consonant system*

In the Arabic tradition the concept of phoneme in the modern sense did not exist. What did exist, though, was a basic list of consonantal sounds, each of which was associated with a (prescribed) pronunciation and a unique grapheme. The grammarians knew that the basic consonants were occasionally realised in a non-standard manner, in some instances those changes could be related to laws of an apparent phonological nature, such as *ṣ* before *d* becomes *z*, *š* before *d* > *ǧ*, etc.

---

[1] For a discussion about the pre-Islamic dialects on the Arabian Peninsula, see Rabin (1951).

[2] Research on Arabic phonological theory started with Schaade (1911), Bravmann (1934) and Cantineau (1960).

In *Kitāb* (II 404-407; also Ibn Ǧinnī, *Sirr* I 436ff), the sounds of speech are 29 'basic consonants' *ḥurūf ʾaṣliyya*[3] and the three vowels, *ḥarakāt*, literally 'movements', *a (fatḥa)*, *u (ḍamma)* and *i (kasra)*.

1 ʾ	8 k	15 r	22 s	29 w[4]
2 ʾalif	9 q	16 n	23 ẓ	
3 h	10 ḍ	17 ṭ	24 ḏ	
4 ʿ	11 ǧ	18 d	25 ṭ	
5 ḥ	12 š	19 t	26 f	
6 ġ	13 y	20 ṣ	27 b	
7 ḫ	14 l	21 z	28 m	

These consonants are produced on 16 'places of articulation' *(maḫraǧ* pl. *mahāriǧ)*,[5] ranging from the back of the throat up to the lips, here summarised in nine main areas:

1   *ḥalqiyya* (laryngal): ʾ h ʾalif[6] ʿ ḥ ġ ḫ
2   *lahwiyya* (uvular): q k
3   *šaǧariyya* ("consonnes de l'ouverture de la bouche" [Cantineau 1960:20]): ǧ š ḍ y
4   *ḏawlaqiyya* (liquids): r l n[7]
5   *ʾasaliyya* (sibilants): ṣ s z
6   *niṭʿiyya* (dental): ṭ d t
7   *liṭawiyya* (alveolar): ẓ ḏ ṯ
8   *šafawiyya* (labial): f b m w
9   *ḫayšūmiyya* (nasal): ŋ ('soft n' *[nūn ḫafīfa]*).

Consonants that share the same place of articulation or whose places of articulation are close may assimilate, or be substituted with each other, and a new (allophonic) sound may arise from the combination. The allophones are discussed below in 1.2.[8]

---

[3] Some later grammarians, however, did not share this opinion. ʾAbū Ḥayyān al-ʾAndalusī mentions some of the opposing grammarians by name and discusses their views. Points of disagreement are e.g. the question whether or not the *hamza* should be considered a phoneme, since according to some Arab grammarians the glottal stop *(hamza)* should not be considered a phoneme (cf. *Irtišāf* I 4). A further point of discussion is the exact number of the places of articulation.

[4] This sequence is according to Sībawayh. Cf. *Kitāb* ed. Būlāq II 404 in: *Bāb al-ʾidġām* 'Chapter about the assimilation', at the end of the second volume. There are indications that *ʾidġām*, usually translated as 'assimilation', in fact means something like 'pronouncing simultaneously'. (Mr Michel Limpens, pc. April 1995).

[5] The exact number of the places of articulation, however, always remained a subject of discussion among the grammarians. In *Irtišāf* it is 16.

[6] *ʾalif* is posited here theoretically, since in reality it cannot have a sound; it is an orthographical device only.

[7] This categorisation does not necessarily agree with the one used in western analysis of consonants (see 2.1 below).

[8] Roman (1983: I 48) 'zone d'émission'.

The consonants are subcategorised according to a total of seventeen criteria based on some characteristics they have in common, such as *šadīda* 'occlusive' vs. *riḫwa* 'spirant', and *mahmūsa* vs. *maġhūra*.[9] The latter pair is generally understood as as 'voiceless' and 'voiced', respectively. The consonants listed as *mahmūsa* are: /h ḥ ḫ k š s t ṣ ṯ f/. *Maġhūra* are /ʾ " ʿ q ǧ y ḍ l n r ṭ d z ẓ ḏ b m w/ (*Irtišāf* I 10,5; for further discussion see Cantineau 1960:22 and Fleisch 1961).[10]

1.1.1 Classification of consonants: velarisation and palatalisation

Apart from a classification according to their inherent qualities, the consonants are divided into several subcategories, in which to their effect on neighbouring vowels is the main criterion.

One such categorisation of the consonants is that in *ḥurūf mustaʿliya* 'elevated consonants' and *mutasaffil* 'depressed'. The former category includes the four 'covered consonants' (*ḥurūf muṭbaqa*), i.e. *ṣ*, *ṭ*, *ḍ* and *ẓ*, also called 'emphatic' (discussed below 1.1.2), along with the guttural consonants, i.e., *q*, *ḫ*, *ġ* and, occasionally, *ḥ* and *ʿayn*. Ibn Ǧinnī describes *istiʿlāʾ* as an 'elevation' to the 'palate' (*Sirr* I 26,4f), which is not identical with 'covering' (*ʾiṭbāq*).

The main effect of the so-called 'elevated consonants' on the neighboring vowels, is that *ʾimāla*, the inclination of *a* toward *i*, does not occur. This distinction between elevated and depressed consonants is important for our study since sources use this as the main instrument for describing front and back vowels in Turkic words. This common feature of all 'elevated' consonants is called *tafḫīm* 'velarisation' (also *taġlīẓ*) With *tafḫīm* *a* is pronounced [a], [å], or even [o]. The result is a neutral, or a 'back' pronunciation of the preceding and following vowels, especially evident in the realisation of *a* as [a], [å] or [o]. For example, the word *kalb* 'dog' is pronounced [kälb], whereas *qalb* 'heart' is [qalb]. Ibn Ǧinnī posits the latter /a/ "between *a* and *u*" (see further descriptions of vowels in 1.4.2).

---

[9] Cantineau (1960:22) mentions the confusion that exists with relation to the meanings of these terms. An adequate interpretation has not yet been found. Roman (1983: I 54) proposes for *mahmūsa* 'chuchoté' (whispered) and for *maġhūra* 'éclatant' (shrill). Interesting in this context is al-Kindī's reference to musical terminology when discussing these terms: al-Kindī's category *naġma* 'melody' almost completely covers the *ḥurūf maġhūra* (cf. Celentano 1979:71).

[10] If *mahmūsa* and *maġhūra* indeed stand for voiceless and voiced, respectively, then the labelling of *hamza* /ʾ/ as *maġhūra* is surprising. The labelling of *q* and *ṭ* as 'voiced' is surprising only within our perception of the Arabic phonetic system, which is based on modern Arabic languages.

In some ancient dialectal variants of Arabic velarisation is reported to occur independently, i.e. without being conditioned by the occurrence of a preceding or following elevated consonant. In the dialect of the Ḥiǧāz, for example, the the word /zaka"t/ 'charity' is pronounced [zako:t]. Ibn Ǧinnī calls the *'alif* in this word which he uses to express *ā*, "the *'alif* of velarisation" (*'alif at-tafḫīm*), which he explains as "the *'alif* inclines to *w*" (*al-'alif mālat naḥwa l-wāw, Sirr* I 50,5). In order to illustrate this, *zakāt* is occasionally written with /w/ instead of *'alif*, viz., /zakawt/; and in fact Ibn Ǧinnī is using here the diphthong /aw/ as a device for indicating [o:].

It is interesting to note that although in the preceding quotations *fatḥa* is compared to *ḍamma*, and *'alif* to *wāw*, *'alif* and *wāw* themselves do not expresses a vowel sound, being orthographic devices only.

1.1.2 The 'covered consonants'

A subcategory of the elevated consonants are the four 'covered consonants' (*ḥurūf muṭbaqa*), i.e. *ṣ, ḍ, ṭ, ẓ* (all other consonants are considered *ḥurūf munfatiḥa*, 'opened consonants').[11] The *ḥurūf muṭbaqa* share a common position of the tongue in the mouth while being pronounced, namely "from their points of articulation until the part of the tongue that is in 'front' of the palate" (*min mawāḍiʿihinna 'ilā mā ḥāḍā l-ḥanak al-'a'lā min al-lisān, Kitāb* II 406).

Three of the four 'covered consonants' have an 'opened' counterpart, i.e., *ṣ - s, ẓ - ḍ*, and *ṭ - d* (cf. also ar-Rāzī *Tafsīr* I 104,6 f). This last pair is surprising at first sight; although *ṭ* is nowadays regarded as a mere velarised *t*, historically it goes back to a velarised *d* (see however, Ibn Sīna on *ṭ - t* 1.1.4; also Cantineau 1960:18).

Of special interest is ض, traditionally transcribed as *ḍ*, which has no counterpart. It was probably lateralised, and Cantineau (1960:55) interprets it as /ḍl/. Roman, on the other hand, prefers /ẓ/ (Roman 1983:162-206). For the purposes of this study it suffices to conclude that whatever its exact historical pronunciation was like, it was certainly not a mere velarised *d*, and, as a result, *ḍ* cannot form a pair with *d* (cf. also *Sirr* I 61,16ff). (The value of *ḍ* is important in regard to its status in Turkic; two sources, i.e., *Qawānīn* and *Durra* use this grapheme for expressing a velarised **d**; cf. 3.1)

---

[11] Al-Ḥalīl (d. 160/776), Sībawayh's teacher, reserves the term *muṭbaq* for the *m*. In *Kitāb al-ʿAyn* "al-Ḥalīl calls the *mīm* 'covered', because it covers the mouth when it is pronounced." (*wa-kāna l-ḥalīl yusammī l-mīm muṭbaqa li-'annahā tuṭbiq al-fam 'iḍā nuṭiqa bihā, Kitāb al-ʿAyn* 58,16.) This statement is related to a different, probably older, concept of the term *muṭbaq* (Michel Limpens, p.c. April 1995).

1.1.3 Effects on other consonants

The covered consonants may also have an effect on neighboring other consonants. This occurs, for example, when /t/ is inserted after the first consonant of the root /faʿala/ in the pattern /iftaʿala/. If the first consonant of the root is a 'covered consonant', the inserted /t/ changes into ṭ, e.g., iṣṭāda 'he hunted' from ṣāda 'he hunted'. In Arabic theory, this is not considered a phonetic phenomenon, but rather an instance of substitution (ʾibdāl) (cf. below 1.3; cf. Zamaḫšarī Mufaṣṣal 172ff; cf. Fleisch 1961:95).

1.1.4 r and l

In addition to the elevated consonants, r and l, too, may cause a conditioned 'back' realisation of accompanying vowels. This back realisation is conditioned by the occurrence of u and a. In Arabic grammar, these allophones are called lām mufaḫḫama and rāʾ mufaḫḫama, velarised l [ḷ] and r [ṛ], respectively. I discuss these the underlying principles in some detail as they are important for the interpretation of r and l in Turkic.

The default position of l and r is palatalisation (tarqīq). In Irtišāf (I 249,5; I 20f) this term is applied to r, when it is not pronounced velarised (also in ar-Rāzī's Tafsīr for the default l, I 104,4). The perception of l and r as velarised in certain conditions plays an important role in the perception of Turkic words with r and l (3.1.2). These conditions are outlined in works by ʾAbū Ḥayyān, ar-Rāzī and Ibn Sīnā. It should be stressed that for the present study the importance of Arabic descriptions lies not in the question whether or not r and l are velarised, but rather in the observation that the Arab grammarians perceived them as such. For this reason I quite extensively deal with the question of how, in their opinion, the velarisation of l and r is conditioned.[12]

1.1.4.1 Velarised r

In Irtišāf (I 248,20ff) ʾAbū Ḥayyān gives a detailed account of the instances in which r has a conditioned velarised pronunciation, i.e. when the following and/or preceding vowels are pronounced 'back'. Velarisation of r is permitted in the following instances:

---

[12] Studies of dialectal Arabic (Grotzfeld 1965:7, Blau 1977:23, Mitchell 1993 and Younes 1994) also point to a velarised realisation of r close to a or u. There seems a similar tendency for z, which in similar instances becomes ẓ e.g., zaffara > ẓaffar "fettig machen"... zunnar > ẓənnār "Gürtel" (Grotzfeld 1965:8).

1　if r is followed by y, e.g., [qaṛya] 'village', [maṛyam] 'Mary'
2　if r is followed (?) by an elevated consonant that is vocalised with i, e.g., [faṛqi] 'difference', [miṛfaqan] 'elbow' (ACC)
3　if r is vocalised with a, e.g., [ǧīṛān] 'neighbours'
4　if r is vocalised with u, while preceded by y, e.g., [ḥayṛun] 'blessing', [qadīṛun] 'powerful'
5　if r is preceded by a basic i (kasra lāzima), e.g., [ḥasira] 'he suffered', [ḥasirū] 'they suffered'
6　if r is followed by another r that is vocalised with i, e.g., [bi-šaṛarin] 'with evil'
7　if r occurs in an Arabic word between i and a (bayna l-maftūḥa wa-l-kasr qablahā), not being followed by another r nor an elevated consonant, e.g., [dikṛ] 'remembrance' (for ḏikr)

Velarisation of r is obligatory,

8　if the preceding i is non-basic ('āriḍ) (i.e. short), e.g., [bi-ṛasūl] 'with a messenger', [bi-ṛūḥ] 'with a spirit'
9　if r follows a consonant vocalised with i, e.g., [siṛāḥīl] 'wolves' (em. for sarāḥīl [sing. sirḥāl])
10　if r is followed by a second r, e.g., [midṛār] 'abundant with rain'
11　if r is followed by an elevated consonant, e.g., ['aʿṛāḍ] 'honour' (pl)
12　if r occurs in foreign words ['ibṛāhīm] 'Ibrāhīm'
13　if r is preceded by an accidental i and followed by an elevated consonant, e.g., ['irṣād] 'to provide'.

'Abū Ḥayyān's account is not very illuminating in the sense that he does not give overall rules; instead, one loses track because of the details and exceptions (discussed by Zamaḫšarī in Mufaṣṣal [159,20ff] in similar terms). The clear tendency, though, is that r is velarised if preceded or followed by an elevated consonant, or a or u. This is true even in 'Abū Ḥayyān's examples in which the velarisation is reportedly associated with the occurrence of i—distinguishing between accidental and basic i—the two other vowels a and u play a role. (cf. also Cantineau 1960:48f; 98; 182).

1.1.4.2 Velarised l
The l is velarised when it is preceded or followed by one of the four 'covered consonants' (ḥurūf muṭbaqa), e.g., [naṭlubu] for naṭlubu 'we demand', ['iṣlāḥ] 'reparation', [yaṣlubu] 'he is firm', [aṣ-ṣalā] 'the prayer', [iḫtalaṭa] 'he mingled' and [istaġlaẓa] 'he became crude' (Irtišāf I 248,10ff). In fact, the examples show that velarisation of l too can

easily be related to either the effect of neighbouring elevated conso-
nants, or the occurrence of *a* or *u*. (cf. Cantineau 1960:50ff).

The most frequent and only non-allophonical—example of velarisa-
tion of the *l*, as far as it is mentioned by the Arab grammarians, occurs
in the word ['alla:h] 'God' (cf. *Kitāb* II 267-270). The double *l* is pro-
nounced velarised when the word *'allāh* is preceded by *u* or *a*, After *i*,
the velarisation is lifted, which results in pronuciation with *'imāla* of
the following *a*, e.g., [bi-smillä:hi] 'In the name of God'. [13]

1.1.4.3 Descriptions of the velarised *l* and its status

Since the velarised l occupies a special position in Arabic phonetics,
which may be enhanced by its occurrence in the word *'allāh*, its status
is more detailedly described than that of other velarised consonants. In
Irtišāf, for example, 'Abū Ḥayyān describes the regular, or neutral, real-
isation of the *l* as "*l* between velarisation and palatalisation" (*al-lām al-
mutawassiṭa bayna t-tafḫīm wa-t-tarqīq*, *Irtišāf* I 8; also I 248,10f). In
ar-Rāzī's (d. 543/1149) *Tafsīr* (I 103,22) the *lām mufaḫḫama* is re-
garded as an 'elevated consonant' (*ḥarf mustaʿlin*).[14] The relation be-
tween the two variants of *l* is described as follows:

> "The palatal *l* is pronounced with the side of the tongue, whereas this ve-
> larised *l* is pronounced with the whole tongue, which requires more ef-
> fort." (*['anna] l-lām ar-raqīqa 'innamā tudkaru bi-ṭaraf al-lisān wa-
> 'ammā hādihi l-lām al-muġallaẓa fa-'innamā tudkar bi-kull al-lisān fa-
> kāna l-ʿamal fīhi 'aktar,* ar-Rāzī *Tafsīr* I 104,2.)[15]

The relation between the palatal *l* (*al-lām ar-raqīqa*) and the velarised *l*
(*al-lām al-ġalīẓa*) is further compared with the relation between *ṣ* and
s, and ar-Rāzī wonders why the two *l*s are, in fact, not considered dis-
tinct phonemes, like *ṣ* and *s*:

> "And likewise the *s* is a consonant and the *ṣ* is another, and the [linguists]
> should say that the palatal *l* is a consonant, and the velarised *l* another, but
> they do not do so, although there is a difference." (*wa-kadālika s-sīn ḥarf
> wa-ṣ-ṣād ḥarf 'āḫar fa-kāna l-wāǧib 'aydan 'an yaqūlū al-lām ar-raqīqa*

---

[13] Because the velarised *l* occurs in the standard language in one word only, Roman
does not consider it a productive allophone: "Le /l/ de /ʔallāh/, s'agissant du nom de Dieu,
s'explique bien par une emphase rhétorique" (Roman 1983: I 321).

[14] Cf. Bravmann (1934:30). It seems, though, that *mustaʿlī* and *mufaḫḫam* were often
interchanged; Sībawayh uses *mufaḫḫam* only once (cf. Troupeau 1976).

[15] Ar-Rāzī continues this passage saying that giving more effort is a good deed, and
compares it with Moses who was told to love his Lord with all his heart. In this case man
has to pronounce the name of God with his entire tongue, which means that he does so
with all his heart (*Tafsīr* I 104,4).

ḥarf wa-l-lām al-ġalīẓa ḥarf ʾāḫar wa-ʾinnahum mā faʿalū ḏālika wa-lā budd min al-farq, ar-Rāzī *Tafsīr* I 104,9.)

All this does not answer the question why the velarised *l* is considered 'elevated' *(mustaʿlin)* rather than 'covered' *(muṭbaq)*, since, according to ar-Rāzī, the difference is the position of the tongue, which is exactly how the covered consonants distinguished from the elevated consonants. Ibn Sīnā (d. 428/1037), indeed, applies the term *muṭbaqa*, 'covered', to the velarised *l* whose effect he compares with *ṭ* (incidentally, he relates the *ṭ* to *t*, rather than *d*, as prescribed in the other manuals):

> "There is a covered *l* which is related to the regular *l* in the same way the *ṭ* is related to the *t*. It occurs frequently in the language of the Turks." *(wa-hāhunā lām muṭbaqa nisbatuhā ʾilā l-lām al-maʿrūfa nisba aṭ-ṭāʾ ʾilā t-tāʾ wa-takṯuru fī luġa at-Turk,* Ibn Sīnā, *ʾAsbāb* 16,9.)

This statement confirms that the original phonetic interpretation as it was given by al-Ḫalīl (and perhaps later phoneticians and grammarians)[16] no longer reflected the phonetic reality anymore. It shall be seen in our discussion in 3.1.1 that *ṭ* represents velarised *t* in the Turkic sources too.

1.2 *Allophones*

This section deals with the description of consonantal allophones that have no effect on the preceding and following vowels. It is important to note that the Arab grammarians themselves had no concept comparable to our notions of 'phoneme' and 'allophone'. However, they did make a distinction between *ḥurūf ʾaṣliyya* (basic consonants)—or *ḫāliṣa* (*Irtišāf* I 11), or *al-ḥurūf aṣ-ṣiḥāḥ* (*Irtišāf* 12,12; *K. al-ʿAyn* 51,14)—from which the *ḥurūf farʿiyya* (secondary consonants) are derived, a concept which is similar. It shall be seen in 3.6.4 that the same or very similar principles and descriptions are used for Turkic consonantal phonemes.

1.2.1 Basic and secondary consonants

Along with the base consonants *(ḥurūf ʾaṣliyya)*, a number of allophonic or secondary consonants *(ḥurūf farʿiyya)* is taken into account.

---

[16] In this respect it is possible that phoneticians after al-Ḫalīl maintained his definitions, although they could not verify them with the speech examples at their disposal. As an example of this I refer again to *m*, which was considered 'covered' by al-Ḫalīl, but not by Ibn Ǧinnī.

The fact that they are considered branches, or derivations of the base consonants is clear from their name. A large number of the allophones mentioned above occur conditioned by a given phonetic environment which is for each case discussed in detail by Fleisch (1961). For the present study it is not necessary to determine in detail the phonetic value of all allophones mentioned in *Kitāb*, Ibn Ǧinnī's works (*Sirr* and *Ḥaṣā'iṣ*), and *Irtišāf*. Instead, I shall briefly discuss some of them in order to give an impression of the way the terminology is applied.

The *ḥurūf far'iyya* are either 'approved', *(mustaḥsan)*, or 'inelegant' *(mustaqbaḥ)*, and they are described in detail in order to distinguish them from the basic consonants. It was necessary to discern between approved and non-aproved consonants, in order to determine which ones could be used while reciting the Qur'ān (*Sirr* I 46,8). However, 'Abū Ḥayyān and Ibn Ǧinnī do not agree on all points on which allophones to approved for the recitation.

Ibn Ǧinnī mentions six approved allophones:

1 [ŋ] 'the light *n*' *(an-nūn al-ḥafīfa)*. In 'Abū Ḥayyān's *Irtišāf* this is the 'silent alleviated *n*' *(nūn sākina ḥafīfa)*, "the resonance is a derivation from the *n*" (*al-ǧunna far' 'an an-nūn, Irtišāf* I 8,4). This allophone of *n* especially occurs when the *n* precedes *k* and *q* (and some 13 other consonants; cf. Cantineau 1960: 38). According to Ibn Ǧinnī (*Sirr* I 48,3) the *nūn ḥafīfa* is produced on the *ḥayāšīm* (sing. *ḥayšūm*) which is usually interpreted as referring to the nasal cavity.

2 'the alleviated *hamza*' *(al-hamza al-muḥaffafa)*.[17]

3 'the velarised *'alif* *('alif at-tafḫīm)*

4 'the palatalised *'alif* *('alif al-'imāla)*[18]

5 [č] 'the *š* that is as *ǧ*' *(aš-šīn allatī ka-l-ǧīm)*

6 [ẓ] velarised *z*—'the *ṣ* that is as *z*' *(aṣ-ṣād allatī ka-l-zāy)*. By 'Abū Ḥayyān described as '*ṣ* between *ṣ* and *z*' *(ṣād bayna ṣād wa-zāy)*.

'Abū Ḥayyān mentions three additional approved allophones:

7 [ž] 'the *ǧ* that is as *z*' *(ǧīm ka-zāy)*

8 'the *s* that is as *z*' *(sīn ka-zāy)*

9 [l] 'the velarised *l* *(lām mufaḫḫama)*.

---

[17] In *Sirr* (I 48,8ff) Ibn Ǧinnī equates this allophone of the *hamza* with *hamza bayna bayna* literally '*hamza* between between' by which was meant a soft pronunciation of *hamza* (Cantineau 1960:77). In *Irtišāf* also a *hamza musahhala* 'softened *hamza*' is mentioned, although its phonetic value is difficult to interpret.

[18] It shoult be stressed that, in spite its association with a vowel, in Arabic theory *'alif* is basically a consonant, and its velarisation and palatalisation should be understood as applied to a con sonant, not a vowel.

Ibn Ǧinnī lists further eight disapproved variations which cannot be used during the recitation of the Qurʾān and in poetry, viz.,

10 [č] 'the *k* that is between *ǧ* and *k̓* (*al-kāf allatī bayna al-ǧīm wa-l-kāf*), such as *ičtamaʿū* for *iǧtamaʿū* 'they gathered'.

11 [g] 'the *ǧ* that is as *k̓* (*al-ǧīm allatī ka-l-kāf*), which ʾAbū Ḥayyān describes as "they pronounce the [ǧ] close to *k*" (*yuqarribūna [l-ǧīm] min al-kāf, Irtišāf* I 8,13), e.g., in *rakul* for *raǧul* 'man'.

12 [ž] [ž] 'the *ǧ* that is as *š̓* (*al-ǧīm allatī ka-š-šīn*)

13 a velarised *d* 'the weak *ḍ*' (*aḍ-ḍād aḍ-ḍaʿīfa*); 'approved' with Sībawayh.

14 'the *ṣ* that is as *s̓* (*aṣ-ṣād allatī ka-s-sīn*)

15 'the *ṭ* that is as *t̓* (*aṭ-ṭāʾ allatī ka-t-tāʾ*)

16 'the *ẓ* that is as *ṭ̓* (*aẓ-ẓāʾ allatī ka-ṭ-ṭāʾ*)

17 'the *b* that is as *m̓* (*al-bāʾ allatī ka-l-mīm*).

ʾAbū Ḥayyān (*Irtišāf* I 8-9) gives four additional disapproved allophones, viz.,

18 [p] 'the *b* that is as *f* (*al-bāʾ allatī ka-l-fāʾ, Irtišāf* I 9,5; also in *Kitāb* II 404), which is noted in the speech of the people of Furs, i.e. the Persians in, e.g., *ʾiṣbahān ~ ʾiṣpahān* 'Isfahan'.

19 [t] *ḍād ḍaʿīfa*, i.e. velarised *t*—'they pronounce the *t* close to the *ḍ*' (*fa-yuqarribūna aṭ-ṭāʾ min aḍ-ḍād, Irtišāf* I 9,11).

20 [g] The 'knotted *q*' (*qāf maʿqūda*), 'the *q* between *q* and *k̓* (*qāf bayna qāf wa-kāf*), or 'the *q* as *k̓* (*al-qāf ka-l-kāf, Irtišāf* I 10,4). Also described by Sībawayh (*Kitāb* II 342) where he says that Persian [g] is as 'the consonant between *k* and *ǧ*' (*al-ḥarf bayna l-kāf wa-l-ǧīm*). A few lines further he adds "the [Persians] sometimes replace [*k*] with *q* because of its closeness to it" (*wa-rubbamā ʾabdalū al-qāf li-ʾannahā qarība minhā*).

Ibn Sīnā calls the latter allophone *kāf ʿarabiyya*, referring by *ʿarabiyya* to the Bedouin dialects of the Arabian Peninsula (cf. Ibn Sīnā *ʾAsbāb* 10; 14; Bravmann 1934:121; 127). Note that none of the aforementioned allophones is marked in orthography.[19]

1.2.2 Analysis of descriptive terminology used for allophones

From the descriptions given above, it follows that the grammarians had three ways to describe allophones. In the first place they use words like

---

[19] Roman draws the conclusion that *q* had different realisations varying from [γ] with a variant [g] in Sībawayh's time, via [g], with a variant [q] with Ibn al-ʿArabī (d. 638/1240), down to the modern realisation [q], with a variant [g] (cf. Roman 1983:140).

ka 'like', secondly, *bayna* 'between'. The third possibility is the use of adjectives.

With *ka*, it would appears at first sight that a resemblance, or even a simple replacement of consonants is intended. For example, the expression *kāf ka-ǧīm* could simply mean that instead of [k] a [ǧ] is pronounced, the result of which, strictly speaking, does not have to be a sound. However, the fact that a new sound are intended is in some instances evidenced by alternative, more explicit phonetic descriptions, e.g., "The *ǧ* as *š* is a branch of the pure *ǧ*" (*wa-ǧīm ka-šīn far ͨ ͨan al-ǧīm al-ḫālisa*, *Irtišāf* I 9,1).

The second term, *bayna*, is more precise in the sense that a new allophone is described with features of two basic sounds. For example, the allophone defined as ' *ṣ* between *ṣ* and *z*' (*ṣād bayna ṣād wa-zāy*, ʾAbū Ḥayyān) can in this sense be interpreted as some kind of voiced *ṣ*, or a velarised *z*, but not as an interchange of *z* for *ṣ*. Ibn al-Qāṣiḥ in his commentary on the *Šāṭibiyya* uses the term *ʾišmām*, viz.,

> "With this *ʾišmām* is meant the mingling of the sound of ṣ with the sound of z, they are combined and a sound arises that is neither ṣ nor z." (*wa-l-murād bi-hāḏā l-ʾišmām ḫalaṭ ṣawt aṣ-ṣād bi-ṣawt az-zāy fa-yamtaz-iǧāni fa-yatawalladu minhumā ḥarf laysa bi-ṣād wa-lā bi-zāy*, Ibn al-Qāṣiḥ [d.?] in his Commentary on the *Šāṭibiyya apud* Grünert 1912:236).

Apart from some non-technical terms, such as *ḫalaṭ* 'blend', and *yamtaziǧu* 'it is combined', Ibn al-Qāṣiḥ uses the term *ʾišmām* with this meaning, although its basic use seems to be with vowels (see 1.4.2 and 1.5.1.2). Another detail worth mentioning is that the term *ḥarf* is used here in a sense that comes close to 'sound'.

The third way used for the description of allophones is the use of adjectives. An example of this is *maͨqūda* 'knotted' in [g], or *mufaḫḫam* for 'velarised'.[20]

---

[20] ʾAbū Ḥayyān quotes Sīrāfī as saying: "As for the knotted *qāf*, Sīrāfī said 'we saw people who pronounce *q* between *q* and *k* (end of quote). [The 'knotted *q*'] is nowadays frequently used by the ͨArabs that live outside the cities, to such an extent that almost all ͨArabs speak with a *q*, rather than with the pure *q* which is described in the books of the linguists...." (*wa-ʾammā l-qāf al-maͨqūda fa-qāla s-Sīrāfī raʾaynā man yatakallamu bi-l-qāf baynahā wa-l-kāf* [end of quote] *wa-hiya l-ʾān ǧāliba fī lisān man yūǧadu fī l-bawādī min al-ͨarab ḥattā lā yakādu ͨarabī yanṭuqu ʾillā bi-l-qāf al-maͨqūda lā bi-l-qāf al-ḫālisa al-mawṣūfa fī kutub an-naḥwiyyīn...*, *Irtišāf* I 9,15.) Interesting to note in this respect is that, according to the biographer Ibn ʾAybak aṣ-Ṣafadī, ʾAbū Ḥayyān himself had this pronunciation of *q*: "His pronunciation is correct, the dialect of the Andalus, he knots the *q* close to the *k* but pronounces it correctly when reciting the Qurʾān." (*ͨibāratuhu faṣīḥa, luġa al-ʾandalus, yaͨqudu al-qāf qarīban min al-kāf ͨalā ʾannahu yanṭuqu bihā fī l-qurʾān faṣīḥatan, al-Wāfī*, 268, 12-13.)

1.3 *Substitution, transformation and assimilation*

In Arabic linguistic theory, the principle of substitution *(badal)* plays
an important role in syntax, where elements are assigned a function
because they can substitute others (Owens 1990:58ff), and in mor-
pho(no)logy, in which seemingly non-regular forms are explained by
means of substitution of consonants (Bohas 1982:337ff). In Arabic
phonology substitution is applied in instances where in western theory
phonetic rules are used. In section 3.6. it shall be seen that in the de -
scription of Turkic similar concepts are applied.

1.3.1 Substitution *(ʾibdāl)*

In Arabic phonology a given consonant is substituted by another when
certain conditions are met. Although the grammarians recognise the
conditions in which a given shift takes place, they explain the phe-
nomenon in phonological terms, not in terms of phonetic laws.

In *Mufaṣṣal* (172ff), for example, az-Zamaḫšarī mentions a number
of consonantal phonemes that can substitute others. For example, *s* is
substituted by *ṣ* when it is preceded (or followed) by *ġ*, *ḫ*, *q*, *ṭ*, e.g.,
ṣalaḥa for *salaḥa* 'he changed [his] skin' (said of a snake) (*Sirr* I
211,16ff). Similarly, the *t* is substituted by *ṭ* when it is preceded by *ṭ* or
ṣ, e.g., *faḥaṣṭu* instead of *faḥaṣtu bi-riǧlī* 'I examined my leg', *iṣṭabara*
instead of *iṣtabara* 'he was patient' and *ḥabaṭṭu* for *ḥabaṭtu* 'I hit'
(*Sirr* I 219,11ff).

Here we see an apparent contradition in the prescribed value of a
phoneme and its perceived value in actual speech. The neutral *s* is the
counterpart of the emphatic *ṣ*, and the fact that they interchange in
certain conditions makes sense. On the other hand, however, nowhere
is ﺕ (*t*) regarded as related to ﻁ (*ṭ*) on the contrary, the communis
opinio dictates that ﻁ represent the emphatic counterpart of ﺩ (*d*).
Therefore, the shift ﺕ *t* - ﻁ *ṭ* cannot be described in the same terms as
s-ṣ. Conversely, one could say, what would be needed here is a velarised
t, which, it appears, has no official reflection in Arabic orthography.
These examples and statements point to a discrepancy between the
prescribed value of a given phoneme/grapheme and its actual perceived
phonetic value. We shall see that in the descriptions of Turkic ﻁ is used
for a velarised t; cf. section 3.

1.3.2 Transformation *(qalb)*

Apart from its meaning of 'metathesis' of consonants, the term *qalb*
has an an overlap in meaning with *badal*. Like *badal*, it refers to the

change of consonants. More specifically, *qalb* is used for changes in the root of the word, in which often, not always, one of the glides is involved (cf. Robakidze 1986). One such proces is the *qalb* of a medium glide into *hamza* in the active participle, viz., */qa"wil/ > /qa"ʾil/ 'speaker', and */ba"yiʿ/ > /ba"ʾiʿ/ 'seller'. In all other instances, it appears to be synonymous with *badal*.

Another example of *qalb* (alluded to in az-Zaǧǧāǧī's *ʿIlal*) is, e.g., the change of /y/ via /"/ into /w/, or elision of /y/ via /"/, viz.,

/muṣṭafay-u-n/	The declensional ending /y-un/yin/ is too heavy *(istiṭqāl)*, hence elision of /u/ (and /i/)
/muṣṭafa-y-n/	change *(qalb)* of /y/ into /"/.
/muṣṭafa-"-n/	Two vowelless consonants cannot stand in a sequence *(iltiqāʾ)*, therefore elision of /"/.
/muṣṭafa-n/	Result, re-spelt as /muṣṭafa"/, pronounced [muṣṭafā] (also in Versteegh 1995:228).

The origin of the discrepancy between *qalb* and *badal*, is perhaps that for *qalb* the interchanging consonants do not necessarily have the same point of articulation, whereas this argument often recurs for *badal*.[21] In the Turkic sources *qalb* is used as a synonym of *badal*.

1.3.3 Assimilation *(ʾidǧām)*

Another important feature is assimilation *(idǧām)*. Two similar consonants assimilate for alleviation *(ḫiffa)*:

> "The co-occurrence of two similar consonants is [considered] heavy in their language, and therefore they aim at a kind of alleviation by means of assimilation." (*taqula iltiqāʾ al-mutaǧānisayni ʿalā ʾalsinatihim fa-ʿamadū bi-l-ʾidǧām ʾilā ḍarb min al-ḫiffa, Mufaṣṣal* 188,3.)

The goal of assimilation is *ḫiffa* 'lightness of speech', as opposed to *tiql* 'heaviness'. It occurs in a sequence of two similar *(mutaǧānisāni)*, or resembling consonants within a word. Resembling consonants are consonants whose places of articulation are close *(mutaqāribāni)*, for example, each one of *ṭ, d, t, ẓ, ḍ* and *t* can assimilate with any other of these consonants (cf. *Mufaṣṣal* 194,18), e.g., ['aratta] for *ʾaradta* 'you wanted', and *iẓẓahara* for */iẓtahara/* 'he became manifest'. In some instances, one could say a phenomenon can be covered by both *ʾibdāl*

---

[21] Cf. Heǧazi (1971:84) "Der Terminus qalb wird von Sīrāfī verwended solange der entstandene Laut in der Schrift des arabischen ein Zeichen hat, wie es bei der Emphatisierung der Fall ist." However, when the result cannot be reflected in orthography, as-Sīrāfī uses *muḍāraʿa* 'likeness'. Heǧazi may be right as far as as-Sīrāfī (280/893-368/979) is concerned, but not in a general sense, since we have seen that other authors use *badal* in those cases.

and *'idġām*, e.g., the form [ḥabaṭṭu], although it is not always repre-
sented in orthography. Assimilation also occurs between two words,
when the first consonant in the sequence is unvocalised, and the sec-
ond is vocalised, e.g., [yaġfillakum] for ❀ *yaġfir lakum* ❀ 'he is forgiv-
ing to you' (Qur'ān 46/30).

1.3.4 Orthographical representation of conditioned changes of conso-
nants

Although the Arab grammarians were not interested in the ortho-
graphic reflection of allophonic sounds, and the Arabic script is essen-
tially phonological, phonetic and phonological changes are reflected in
orthography in various ways. In some instances phonologically condi-
tioned changes are visible in orthography, e.g., the change of *t* into *d*
after *z* in *izdahara* <*/'iztahara/ 'he flourished', or into *ṭ* after *ṣ* in
iṣṭabara <*/'iṣtabara/ 'he was patient'. On the other hand, though,
other changes, such as -*ṭu* for *tu* in [faḥasṭu] 'I examined', assimilation
of -*dtu* to -*ttu* in ['arattu] 'I wanted', and [ṣirāṭ] for *sirāṭ*, are not. In
the same way, the velarised pronunciation of *l* and *r* is often not indi-
cated in orthography.

   In 3.3 it is shown that these and similar principles of substitution
and assimilation also play a role in the way Turkic consonant assimila-
tion is perceived and reflected in orthography.

1.4 *The description of vowels*

Since Turkic languages have an elaborate system with eight vowels that
are subject to the principle of vowel harmony, it is important to de-
scribe the prevalent theories in Arabic linguistics with regard to vowels
other than the three standard ones and vowel harmony. In this section,
we shall try to determine the extent to which the grammarians per-
ceived vowels other than the three basic ones of Arabic, *a*, *i*, and *u*, and
how these are described. In this section I shall also pay attention to an-
other type of vowels, i.e. [ü].

   As already discussed in some detail above in 1.1.1, in Arabic theory
the quality of the vowels depends to a great extent on the surrounding
consonants. Velar consonants cause a 'back' realisation of vowels,
whereas palatal consonants have a 'fronting' effect.

1.4.1 Basic vowels

Arabic phonology recognises only three vowels (called *ḥarakāt*, literally
'movements'): *a* (*fatḥa*), *i* (*kasra*) and *u* (*ḍamma*), along with their

long realisations, *ā, ī* and *ū*. In orthography, the short vowels are expressed by means of vowels signs which do not have the status of a consonant, and which are often omitted. The vowels have no independent status, instead, they are tied to a consonant, which, then, is *mutaḥarrik* 'vocalised', literally 'moving'. Vocalised consonants are assigned adjectives reflecting the name of the vowel they take, viz., *maftūḥ* 'vocalised with *a*', *maksūr* 'vocalised with *ī*' and *maḍmūm* 'vocalised with *u*'.

In orthography, the lengthening of a vowel is indicated by means of one of the weak consonants, i.e. *ʾalif* /"/, /y/ or /w/, respectively, preceded by the appropriate vowel sign. In this way, *ā* is expressed as /a"/, *ī* as /iy/, and *ū* as /uw/. In the word *kātib* /ka"tib/ 'scribe', for example, *k* is vocalised with *a* (*maftūḥ*), whereas *ʾalif* /"/ is not. *ʾAlif, w* and *y* retain their status of consonant, they merely serve as an orthographic device to indicate lengthening (cf. also Ermers and Limpens 1995).

1.4.2 Allophonical realisations of vowels

The phonetic realisation of the three vowels depends on the preceding (and following) consonants. According to Ibn Ğinnī, three additional vowels should be posited with each pair of long and short vowels, depending on the way they are coloured by the surrounding consonants. With neutral consonants, the vowels, especially *a*, are pronounced with 'inclination' (*ʾimāla*).

This means that *a* in varying degrees inclines to [i], resulting in [ä], in other words, a 'fronting' of *a* towards [i] (cf. *Kitāb* II II 407; and ar-Rāzī *Tafsīr* I 103ff; *Irtišāf* I 238ff).[22] In Ibn Ğinnī's terms *a* in /ka"tib/ [kä:tib] 'writer' is pronounced "between *a* and *ī*", alternatively described as "*a* mixed with *ī*" (*al-fatḥa al-mašūba bi-l-kasra, Sirr* I 52,4).

ʾAbū Ḥayyān describes two types of *ʾimāla*; one is *ʾimāla šadīda* 'strong inclination' (also called 'pure' [*maḥḍa*] with Sībawayhi), "as if [the *ʾalif*] has become another consonant, close to the *y*" (*kaʾannahā ḥarf ʾāḥar qarīb min al-yāʾ*, Sībawayhi *apud* ʾAbū Ḥayyān, *Irtišāf* I 8,5), i.e. toward [e] or [i]. The second type is *ʾimāla mutawassiṭa* 'medium inclination', i.e. toward [ä].

Incidentally, if the vowel sound is long, the *ʾalif*, which is a consonant, is considered to take over the quality of the *fatḥa*, because the "*ʾalif* follows *a*" (*tābiʿa li-l-fatḥa, Sirr* I 52,8).

---

[22] In *Irtišāf*, however, ʾAbū Ḥayyān uses this term to indicate a pronunciation of *u* as [ü], in the case of /maḍʿuwr/ 'frightened' [maḍʿū:r].

Although the term *ʾimāla* is used typically for the inclination of *a* to
[i], in *Irtišāf* it also occurs for indicating a certain pronunciation of
other vowels. An example is the realisation of *u* as [ü], e.g., /maḍ⁽uwr/
[maḍ⁽ü:r] 'frightened'. This effect is said to be caused by the *⁽ayn*,
which is an 'elevated' consonant. ʾAbū Ḥayyān quotes various opinions
according to which "both the *w* and the preceding *u* have an inclina-
tion" *(tumīl al-wāw wa-ḍ-ḍamma qablahā)*, or just "the *u* has an in-
clination, whereas the *w* has not" *(tumīl aḍ-ḍamma lā al-wāw, Irtišāf* I
248,2). For this same phenomenon Ibn Ǧinnī uses the term *ʾišmām*,
although he gives other examples, e.g., /siyra/ 'it was gone' [sü:ra], in
which "the *i* is flavoured with an *u*" *(fa-hāḍihi l-kasra al-mušamma
ḍamman, Ḥaṣāʾiṣ* III 121,2), or "the *i* is flavoured with *w*" *(tušammu l-
kasra fī l-wāw, Irtišāf* I 248,2). Conversely, the *u* can be flavoured with
i too, e.g., [bü⁽tu] 'I was sold' for *bu⁽tu* (< *buyi⁽tu), [ṣüntum] 'you
were defended' for *ṣuntum* (< *ṣuwintum), and [šüdda] 'he was
strengthened' for *šudda* (< *šudida). In these cases [ü] for [u] is ac-
counted for by the influence of the phonologically reconstructed *i*.[23]

   In addition to *ʾišmām*, Ibn Ǧinnī applies *mašūb* too. For example,
in /qiyla/ 'it was said' (*Ḥaṣāʾiṣ* III 120,9ff; also *Sirr* I 52,13), the *i* is
"mixed with something from the *u*" *(mašūba bi-šayʾ min aḍ-ḍamma)*,
hence [qü:la]. In this way, a mix of *i* with *a* yields [ä], *u* mixed with *i*
results in [ü], whereas *i* nor *u* cannot be mixed with *a*.

   The concept of mixing of vowels recurs in the descriptions of Turkic
which are dealt with in section 3.4.4.

1.5 *Long and short vowels*

Another important issue for this study is the status of the 'consonants
of softness and prolongation' *ḥurūf al-līn wa-l-madd* (cf. Bravmann
1934:13-14). To this category belong the glides i.e. /w/ *(wāw)*, /y/ *(yāʾ)*,
/ʾ/ *(ʾalif)*, all of which are considered consonants in Arabic theory. The
w and *y* can be vocalised like all other consonants, albeit not in all
positions, whereas the *ʾalif* is never vocalised (cf. also Troupeau
1989:34). In orthography, the signs representing the glides often de-
note lengthening of the three short vowels *(ḥarakāt)*, *a*, *i*, and *u*, re-
spectively.

---

[23] Cf. Grünert (1912:234-5); Wright (1986 [1896]: i 71, 84 and 89); *Kitāb* II:260,
361, 282; Bravmann (1934:89-90); Cantineau (1960:101). An alternative term for
ʾišmām is *ʾišrāb* 'absorption': "the *i* has absorbed an *u*" *(ʾašrabat al-kasra ḍamman)*.
The *a* cannot be absorbed by either one of the other two vowels.

In the description of glides in Turkic (3.5) it shall be shown that they may have altogether different functions. Apart from non-morphological lengthening, they indicate velarisation, or serve as plene spelling of vowels.

1.5.1 Lengthening and reduction: the status of the glides

The distribution of long and short vowels is an important feature in Arabic morphology. In verbs, for instance, the quantity and quality of vowels are used to indicate the valency or the 'mood'. In some instances a specific meaning is expressed by a long vowel. In Arabic phonology this is, however, described in terms of the insertion of a consonant, mor specifically a glide. Examples are, e.g., *ḍāraba* /ḍa"raba/ 'he fought (someone)', passive form *ḍūriba* /ḍuwriba/ 'he was fought', derived from the root /ḍ-r-b/ 'to beat'. The active form of the verb has the morphological pattern, /fa"ᶜala/, and the latter is set up according to /fu"ᶜila/, whereas both are CVCCVCV.

The Arab grammarians argue that in the case of /ḍa"raba/ an *ʾalif*—which is considered a consonant—is inserted between the radicals of the root, and, subsequently, becomes part of the verb itself. This procedure is equal to the insertion of other consonants, such as *t* in, e.g., /iftaᶜala/. In these cases both /t/ and /"/ are considered *zāʾid* 'additional', and as such a part of the word (cf. Bohas 1982:168).[24] After the insertion of *ʾalif*, the preceding consonant is vocalised with *a*, because *a* is of the same kind (*ǧins*) as *ʾalif*. Likewise, /uw/ becomes /uw/, for *u* is of the same kind as *w*, and *i* matches *y* (*ḥaraka min ǧinsihā*, *Ḥaṣāʾiṣ* III 121,9).

In Arabic morphology *ʾalif*—unlike *y* and *w*—is never considered part of the root. Words that contain an *ʾalif* are usually related to radical patterns with *w* or *y*, for which *ʾalif* is substituted in certain circumstances. For instance, in the word *bāb* /ba"bun/ 'door', the *ʾalif* /"/ substitutes /w/ because for some morphonological reasons /w/ cannot stand between two vowels, which is the case in */bawabun/. After substitution with *ʾalif* and subsequent metathesis, then, the result is /ba"bun/. In these instances, *ʾalif* may replace another glide in the root; it is never authentic itself.

---

[24] Bohas elsewhere (1982:393) remarks that the process of derivation is determined for each word separately, rather than in a general way according to patterns.

1.5.1.1 Lengthening
In Arabic vowels can be lengthened for prosodic reasons, i.e. without
morpho(no)logical necessity. This occurs, for example, in poetical con-
texts, in which lengthening may be necessary in order to match the
metre (cf. Wright 1986 [1898]: ii 382). In such a case especially the
vowels of declension are lengthened.

This type of prosodic lengthening is called *ʾišbāʿ*, lit. 'saturation', or
alternatively, *maṭl* (*Ḥaṣāʾiṣ* 121,8ff). For example, *i* (*kasra*) is *mušbaʿa*
'saturated' or 'lengthened', which may be expressed orthographically by
means of insertion of *yāʾ* (cf. *Irtišāf* I 9; also *Sirr* I 338,7; II 630,8).[25]

Thus, the *ʾalif*, *w* and *y* "...are the consonants of lengthening that
arise from the preceding vowels" (...*ḥurūf al-ʾišbāʿ tawalladat ʿan al-
ḥarakāt allatī qablahā*, *Irtišāf* I 423,10). In the case of morphonologi-
cal lengthening, the glide is inserted for a meaning, whereas with *ʾišbāʿ*
it arises from the preceding vowel. In the latter case, therefore, it has
no morphological function, and is not represented in the patterns.
Below, in section 3.5, I shall show that in the grammars of Turkic
ʾišbāʿ is not only used for indicating prosodic lengthening, but also ve-
larisation.

1.5.1.2 Reduction
The term *ʾišmām*, literally 'flavouring', is applied by Sībawayh in the
sense of the reduction of a long vowel or, in the case of a short vowel,
of leaving a slight fragrance of it (cf. *Kitāb* II 283; also *Irtišāf* I 247,16f;
Bravmann 1934: 82-89). Zamaḫšarī describes *ʾišmām* as "joining of the
lips after silencing" (*ḍamm aš-šafatayni baʿd al-ʾiskān*, *Mufaṣṣal*
160,16), and it occurs especially with the nominative case (in *u*). Like
ʾišbāʿ, *ʾišmām* occurs for prosodic reasons only and is not indicated in
orthography.

Although *ʾišmām* is basically applied to vowels (discussed earlier in
1.4.2), as noted in section 1.2.1, it is used for an allophonic pronunci-
ation of consonants too. In the Turkic sources *ʾišmām* is applied to
both vowels and consonants with different meanings (see 3.1.2.2 and
3.5).

---

[25] In *Irtišāf*, ʾAbū Ḥayyān applies the term *ʾišbāʿ* to a consonant other than a glide, i.e.
h. "And some of the [Arabs] lengthen the *h* when it is connected, and thus *y* originates,
e.g., *hidihī* [for *hāḏihi*] 'this' [f.]." (*wa-minhum man yušbiʿ al-hāʾa fī l-waṣl fa-
tatawalladu l-yāʾ, fa-taqūl hidihī*, *Irtišāf* I 407,16.) In these instances a type of vowel
harmony is involved too.

1.5.2 A concept of vowel harmony

In Arabic linguistic theory some concept of vowel harmony existed, even though in Arabic it only occurs as a minor prosodic feature. A term used in this respect is *'itbā*ʿ, literally 'making follow'. *'Itbā*ʿ basically means that one vowel is substituted with another in order to better match the following vowel. This type of vowel harmony occurs in, e.g., *al-ḥamd-i li-llāhi* or *al-ḥamd-u lu-llāhi* instead of *al-ḥamdu li-llāhi* 'praise be to God', and *'imṛi²-in* 'man', instead of *'imṛu²-in* (man-GEN). The replacing vowel is called 'vowel of agreement' *(ḥaraka al-'itbāʿ)*. The term *'itbā*ʿ is of course semantically related to the term *tābi*ʿ, 'following'. This term is used for the (consonant) *'alif* when it 'follows' the colouring of the *a*, which is also a type of vowel harmony (see discussion above 1.4.2).[26]

Another term, *tawālī l-mutaḥarrikāt*, which occurs in the descriptions of Turkic too (3.2.3), is used for a sequence of vocalised consonants. In Arabic morphology some sequences are permitted, whereas others, such as CVCVCVCV, are not. A non-aproved sequence occurs, for instance, with verbs of the past tense, such as *ḍarabtu* 'I beat' and *inṭalaqnā* 'we left'. According to the Arab grammarians, these verbs consist of two parts, /ḍaraba/ and /inṭalaqa/, and their suffixed agents, the pronouns /tu/ 'I', and /na"/ 'we', respectively. Normally this would result in */ḍarabatu/ and */inṭalaqana"/, which, however, contain too many vowels in a row. In order to prevent this, the grammarians say, it is necessary to delete the vowel before the pronoun, viz.,

> "With regard to *ḍarabtu* and *intalaqnā*... the deletion [of the vowel] is incidental because of their dislike of a sequence of four vocalised consonants in a [construction] which resembles one word." *('ammā naḥwa ḍarabtu wa-intalaqnā... fa-s-sukūn fihi ʿāriḍ 'awǧabahu karāhatuhum tawālī 'arbaʿ mutaḥarrikāt fimā huwa ka-l-kalima al-wāḥida,* 'Ašmūnī apud Bohas 1982:44; cf. also *Ḥaṣā'iṣ* III 115,2; *Sirr* I 220,14.)

Note, however, that this does not apply to words as *ḍarabaka* 'he hit you'. In these verbs a a similar sequence of four vowels occurs, this time fully accepted by the grammarians. The reason for this discrepancy is is that *ḍarabtu* is regarded as one word, in which the mentioned sequence is prohibited, whereas *ḍarabaka* is considered two

---

[26] *'Itbā*ʿ is also applied to consonants that are substituted as a sort of alliteration, e.g., *ra'aytu l-walīd bn al-yazīd* 'I saw al-Walīd son of al-Yazīd', in which the definite article, *al*, is added to *yazīd* in order to make it rhyme with *al-walīd* (*'Ašbāh* I 9,20 - 14; cf. also *Ḥaṣā'iṣ* II 333,9; II 335,3/5; II 336,5; *Irtišāf* I 400,5).

words, i.e. *ḍaraba+ka* 'he hit' + 'you' (for further explanation, cf. Bohas 1982: 133).

2. TURKIC PHONEMES

It is all but impossible to give a general picture of 'Turkic phonetics' of the Turkic languages that are described in the sources. They cover a broad range of languages in different periods of time. Furthermore, the historical phonetic characteristics of the various Turkic languages are not known in detail. Nevertheless, it is possible to make some general statements on Turkic phonology, based on the data provided in the sources themselves.

In linguistics often the term 'archiphoneme' is used to refer to all different realisations of a certain phoneme, vowels and consonants alike. Archiphonemes are usually written in capitals. In front words, for example, the archiphoneme G (not to be taken for the symbol for velarised g [ɢ]) is realised as [g] whereas in back words its realisation may be close to [ġ]. In the same way, the archiphoneme I is realised as [i]/[ü] in front words, or as [ï]/[u] in back words.

The present section provides much of the data on Turkic languages which is needed for the discussion of Turkic phonemes in section 3. The main issues are an assessment of the number and quality of the consonants and vowels, the principles of Turkic vowel harmony and consonant assimilation. In modern Turkic linguistics the vowels are usually regarded as primary to consonants, in the sense that they affect the pronunciation of the surrounding consonantal phonemes. This point of departure makes it unnecessary to posit separate velar or velarised consonants, which play such an important role in Arabic theory.

2.1 *The consonants*

In Turkic the following consonantal phonemes occur:

Labial	Dental	Denti-palatal	Palatal
1 b	7 d [27]	13 ǧ [28]	15 ğ
2 p	8 t	14 č	16 ž
3 v	9 ḍ		17 š
4 (f) [29]	10 n		18 y
5 m	11 z		
6 w	12 s		

Post-Palatal	Velar	Liquids	
19 g	7 q [30]	26 l	
20 k	8 ġ	27 r	
21 ŋ	9 (ḫ)	Aspirant	
	10 ŋ	28 h	(After Clauson 1972:viii)

Consonant assimilation is a very normal feature in many languages. It occurs in Turkic, too, and is especially evident in suffixes. The expression 'progressive consonant assimilation' is used for the conditioning of the first consonant of the suffix in order to match the final consonant of the word it is attached to. After voiced consonants the first consonant of the suffix is voiced.

A good example is the marker for the past tense DI. Both D and I are archiphonemes, i.e. both are subject to conditioned changes; here only the changes of D- concern us. For example, in the verbs **kir-dī** [kir-di] 'he entered' and **kal-dī** [kel-di] 'he came', D is realised as [d] because the last consonant of the verb is voiced. Similarly, it is voiceless, i.e. [t], when the last consonant of the word is voiceless, e.g., **tik-tī** [tik-ti] 'he sewed', **qāč-tī** [qač-tï] 'he fled'.

---

[27] The phoneme ḍ occurs only in so-called aḍaq-languages. Turkic languages are divided into two large groups. In the first group the word aḍaq 'foot' is realised as aḍaq (with variations haḍaq and azaq), whereas in the second group intervocalic [ḍ] has merged with [y], hence ayaq (cf. *Dīwān* 27,12). The Turkic language in *Dīwān* belongs to the group of aḍaq, whereas all other sources describe ayaq-languages (although in *ʾIdrāk* there are occasional references to an aḍaq language).

[28] ǧ occurs only after long vowels in Oġuz languages as an allophonic realisation of č

[29] f and ž occur in loan words only.

[30] ḫ occurs as an allophone of q, e.g., [yaḫšï] 'good' for [yaqšï].

2.2 The vowel system

Turkic languages have nine vowels /o, ö, a, e, é[31] i, ï, u, ü/.[32] A typical phenomenon in Turkic is that all originally Turkic words are either front or back. In 'front' words only the vowels /ü, ö, e, i and é/ occur; in back words their counterparts /u, o, a, and ï/. The ninth vowel, /é/, has no real counterpart (although one could say that it shares a position with /e/), and for this reason it stands outside the vowel system. Since it is not of great importance for the discussion of vowel harmony, I shall not refer to it here, but discuss it in some detail in 3.1.6.5 and 3.5.4.2).

The distribution of vowels in Turkic follows a regular pattern; it is subject to the principles of vowel harmony. For Turkic, vowel harmony is based on two parameters. The first parameter is the distinction between front and back vowels, the second is labial harmony, i.e. vowel harmony of rounded vowels.

2.2.1 Front and back vowels

The primary principle of vowel harmony is the distinction between front and back vowels. 'Front' vowels are articulated relatively in the front part of the mouth, whereas the back vowels are articulated in the back part of the mouth. The front vowels /i, ü, e, ö/ have back counterparts in /ï, u, a and o/.

In this way, a corresponds with [e], [o] with [ö], [ï] with [i], and [u] with [ü], respectively. The vowels [i, ï, u, ü] are pronounced relatively higher in the mouth than [e, a, o, ö] (see scheme 1 on page 92). In a genuine Turkic word all vowels are either back or front, e.g., [yüzük]

---

[31] This closed é is posited in a number of publications (e.g. Nemeth, J. 'Zur Kenntniss des geschlossenen e im Türkischen' in *Körösi Csoma Archivum* 1939:515-31 and Clauson 1962:163), but it can also be explained as an allophonic realisation of [ɛ] and [ɛ:] after [y] (A.H. Nauta, p.c. 1995). Whatever the exact nature of the closed é, there are clear references to it in the Turkic sources and I postulate it here for that reason.

[32] Long vowels [o:, ö:, a:, e:, i:, ï :, u:, ü:] occur only in Turcomanian and Yakut. Turkic vowels are open, and should be phonetically represented as [ɔ, ɔ̈, ɑ, ɛ, ι, ῑ, ʋ, ʋ̈]. These comments are intended as very general indications with respect to the phonological system of Turkic, since, again, each Turkic language has its own characteristic phonetic realisation of the respective phonemes. For example, Turkish ö stands for a phonetic value that is very much different from ö in Kazakh (cf. Vajda 1994). For this and other reasons mentioned above, it is not possible to determine the exact phonetic qualities of the vowels occurring in the languages which are described in the sources; for the purposes of this study a rough transcription suffices. A note on transcription: in turcology, front [e] is also often transcribed as [ä]. Here I shall only use [e] for sake of convenience.

'ring', [bešik] 'cradle', and [balïq] 'fish', [boyun] 'neck', whereas the lower vowels typically only occur in non-final syllables.

The relative places of articulation of the eight vowels of Turkic are reflected in the following scheme :

back vowels

front vowels
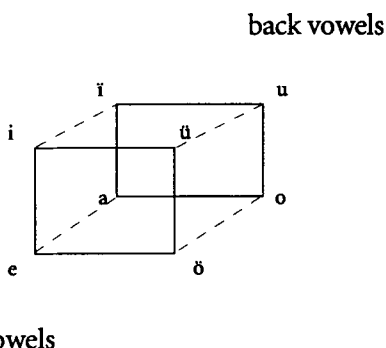
This same principle regarding vowel harmony also holds for suffixes, which implies that all suffixes have at least two forms: one which is used with front words and one for back words.

2.2.3 Labial harmony

The secondary principle of Turkic vowel harmony is labial harmony, which is based on the distinction between rounded and unrounded vowels. The terms 'rounded' and 'unrounded' refer to the position of the lips during the pronunciation. Thus, [ï, a, i, e] are unrounded, and [u, o, ü, ö] are rounded vowels. In the schematic representation above, the unrounded vowels are posited at the left side, and the rounded ones at the right.

Labial harmony is a prosodic feature that occurs in suffixes (and particles) that contain a high vowel. Suffixes that are subject to this principle have four variants. In Oġuz languages, for example, the par-ticle of interrogation has the forms /mi, mü, mï, mu/, depending on two parameters, i.e. whether the preceding vowel is front or back, and further whether it is rounded or unrounded. This is summarised as fol-lows:

	unrounded	rounded
back	mï	mu
front	mi	mü

Examples of this are [keldi mi] 'has he come?', [küldü mü] 'has he laughed?', [qaldï mï] 'has he stayed?', and [urdu mu] 'has he beaten?'.

In suffixes containing a low vowel, e.g., [a, e], only the principal distinction between back and front words is indicated, e.g. in the plural suffix lAr, e.g., [balïq-lar] 'fishes', [gül-ler] 'roses', while the principle of labial harmony is not observed.[33] In phonological terms the labial harmony is inhibited because the counterparts of [a] and [e], i.e. [o] and [ö], as a rule do not occur in non-first syllables, and therefore never occur in suffixes.

The diversity of the vowels in suffixes is indicated by means of archiphonemes too. For example, mI (or mI[4]) implies that the particle of interrogation has four forms, whereas lAr (or lAr[2]) means that the plural suffix has two forms. In this way, a sufix that, combining both parameters, is subject to the rules of both vowel harmony and those of consonant assimilation, has eight different forms, viz.,

	voiced rounded/unrounded	voiceless rounded/unrounded
back	du / dï	tu / tï
front	dü / di	tü / ti

3. TURKIC PHONEMES AS DESCRIBED BY THE SOURCES

In Section 1 of this chapter we dealt with the descriptions of consonants and vowels in Arabic phonetic theory, and believe it has become evident that the Arab grammarians did have instruments to describe allophones that occur in Classical Arabic, as well in some dialects. This section shows how they applied these instruments to Turkic.

This section in addition discusses how regular features of Turkic, such as vowel harmony and consonant assimilation, are perceived by the Arab grammarians. Since in Arabic grammar a discussion on vowels involves one of consonants too, especially inasmuch they are related to matters of velarisation and palatalisation, the section on vowels deals in some detail with those issues as well.

In the present section I have chosen for the following approach. Section 3.1 deals with descriptions between back and front words, 3.2 with vowel harmony. Section 3.3 discusses with the question of how the distribution of velar and palatal suffixes is related to velar(ised) and palatal(ised) words. Section 3.4 shows the phonetic qualities of vowels

---

[33] As the saying goes, there is no rule without exception. In Qirġiz, for example, [o] and [ö] also occur in non-initial syllables, e.g., [ölkö] 'country' (whereas other languages have [ülke]). In addition, the plural suffix has developed rounded vowels, hence [qoy-lor] 'sheep' >*qoy-lar, and [ölkö-lör] 'countries' >*ölkö-ler.

that cannot be described in terms of palatalisation and velarisation alone. Section 3.5 discusses the various functions of the glides. Finally, section 3.6 is dedicated to consonants and consonant assimilation.

3.1 Vowels: front and back

As shown in section 2, Turkic features a vowel harmony between front and back words. Oppositions between front and back words can in Arabic orthography be conveniently indicated by means of velar consonants in back words, and with their palatal (or neutral) counterparts in front words. In instances in which a given consonant does not form a pair with another one, new solutions must be found in order to indicate that it is pronounced in a non-standard way.

Roughly, there are two methods for indicating a non-standard quality of neutral consonants. The first is the use of certain labels which stand for 'velarisation' or 'palatalisation'. The second method is 'to add' a feature of a second consonant that lacks in the basic consonant. Both methods are applied to Arabic consonants in order to indicate a non-standard pronunciation (cf. section 1.2).

3.1.1 The opposition front vs. back

The opposition of front and back words is reflected in Arabic script by means of the appropriate consonants. The Arabic consonants in as far as they are used for Turkic may be divided into four groups. In the first group are q, ġ, ṭ or ṣ, which occur only with back vowels. The second comprises one consonant, k, which typically occurs with front vowels. The third group contains the consonants ', b, č, d, f, ǧ, m, n, ŋ, w, y, z, which, bascially neutral, depending on source, can occur with both front and back vowels. Finally, the fourth group contains r and l which occur with back vowels when preceded or followed by a/u (and there are indications that this hold for z too; cf. 3.4.3).

The so-called elevated consonants typically indicate that the word must be pronounced with back vowels. The sources have different opinions as to which of those occur in Turkic, and, as we shall see, they disagree especially on the covered consonants. Most (except Ḥilya and Dīwān) though, accept ṭ and ṣ, excluding ḍ, ʿ, ḥ and ẓ.

A number of neutral/palatal consonants, i.e. ḵ, k, s, t, d, and z, is related to related to velar counterparts, represented by elevated consonants, in the following manner, viz.,

	Velar (with back vowels)		Palatal (with front vowels)	
1	خ	ġ	گ	ḵ [g]
2	ق	q	ك	k
3	ص	ṣ	س	s (not in *Dīwān*; occ. in *Ḥilya*)
4	ط	ṭ	ت	t (not in *Dīwān*; occ. in *Ḥilya*)
5	ض	ḍ	د	d (in *Qawānīn* and *Durra*)
6	ظ	ẓ	ز	z (in *Durra* only).

The six pairs mentioned here, however, cannot all be based on the Arabic tradition. In fact there is only one, i.e. s - ṣ, that is related to Arabic concepts according to which, ṣ without ʾiṭbāq becomes s (see above 1.1.2). In Arabic theory, q is not regarded as a pendant of k Nevertheless, they may be linked as a matter of convenience or a new perception inasfar as Turkic is concerned. The same can be said for ġ - ḵ, with the important difference that ḵ is considered an allophone in Arabic (and sometimes for Turkic too).

The other pairs (i.e. 4, 5, 6) are more surprising in the light of Arabic linguistic thinking, and as such they convey much about the Arab grammarians' perception of certain Arabic emphatic phonemes. In Arabic theory, as we recall from section 1.1.2, d reportedly forms a pair with ṭ, and ḍ with ẓ. The stand-alone ḍ, has no neutral counterpart, and, conversely therefore, d has no emphatic variant. Although in Arabic a velarised z is recognised and described as an allophone, but never associated with the grapheme ẓ.

It appears therefore that much of what the phoneticians wrote in their Arabic treatises was of a mere prescriptive nature which did not necessarily reflect everyday practice. One example occurs in their descriptions of Arabic itself. If we recall *faḥaṣtu* for *faḥastu*, and *iṣṭabara* for *iṣtabara* (cf. 1.3.1), in which phonological t is conditioned by ṣ and changes into ط . For this ط is to be interpreted as a voiceless emphatic variant of t rather than voiced (if indeed ص is both voiceless and velarised, that is). In the descriptions of Turkic velarised [t] is the only sensible interpretation of ط . The remaining two interpretations of ẓ and ḍ are real innovations, and disclose features of everyday pronunciation.

The sources display four different levels of acceptance regarding emphatic consonants as representatives of velarised consonants. The first level is Kāšġarī who does not accept any of the four emphatic conso-

nants.[34] The second is represented by ʾAbū Ḥayyān, Ibn Muḥammad
Ṣāliḥ (along with, perhaps, Ibn al-Muhannā), and *Bulġa* and *Tuḥfa*'s
respective anonymous authors who use ṣ and ṭ. On the third level is
Qawānīn who in addition ḍ accepts a velarised d. The fourth and final
level is represented by *Durra*, in which all emphatic consonants are
used, including ẓ as velarised z.

Thus, whatever the prescribed pronunciation may have been, the
emphatic consonants are used by the sources as presented in the
scheme in order to mark a back pronunciation of the vowels, e.g.,
ʾalṭun [altun] 'gold', qul [qul] 'slave', ṭur- 'to stand up'. Only in *Durra*
and *Qawānīn* ḍ stands for a velarised d, e.g., ḍūn 'cloth' (*Durra* 12ᵛ2);
ḍuḍaġ 'lip' (*Qawānīn* 60,14); in *Durra* ẓ represents a velarised z, e.g.,
qūẓ [qoz/quz] 'nut' (6ᵛ15). Words with palatal consonants and, there-
fore, a front pronunciation are, e.g., kal- [kel] 'to come', kuk [kök]
'blue'. In all of the examples mentioned here, the choice between a
back or a front pronunciation can easily be inferred from the ortho-
graphic representation, with due consequences for the assignment of
suffixes containing velar or palatal consonants (for which see 3.3).

3.1.2 Instruments for indicating velarisation and palatalisation

Thus in Arabic orthography a number of velar and palatal consonants
form pairs, which can be used to indicate the distinction between a
back and front pronunciation. Fourteen consonants, however, do not
have a counterpart to form a pair with. These are ʾ, b, d, ǧ, l, m, n, r, s,
š, w, y, z (d and z having velarised counterparts in two sources only).

Nevertheless, they do occur in back words, and therefore, their
'velarised' pronunciation has to be indicated in some manner. A logi-
cal means is assigning a label for indicating that the consonant in ques-
tion does not have the default Arabic front pronunciation. In
Qawānīn, *ʾIdrāk*, *Tuḥfa*, *Tarǧumān*, *Šudūr* and the Margin Grammar
the term *tafḫīm* 'velarisation' (or *mufaḫḫam* 'velarised') is applied,
whereas *Dīwān* and *Ḥilya* use *ʾišbāʿ*; in *Tarǧumān* *mašmūm* is occa-
sionally found too.[35] In *Durra* a clear terminology lacks; the velarised
consonants instead being marked with small signs whose exact mean-

---

[34] Kāšġarī's categorical rejection of all four elevated consonants can be understood in
two ways. In the first place, he wanted to adhere to a close transcription of the Uygur
alphabet in which the emphatics do not occur. Second, according to his perception of
Arabic emphatic conso nants he did not think that they could serve as velarised counter-
parts of neutral consonants. In other words, the Arabaic emphatic consonans possessed
certain qualities that made them unsuitable for this purpose. Whatever Kāšġarī's exact
motivation, it is unlikely that he regarded the vowels ss primary.

[35] In the Velieddin MS of *ʾIdrāk*, *mufaḫḫam* 'velarised' is abbreviated as ڂ (ḫ).

ing remains obscure, or with doubling of the consonant, e.g., yyīl [yïl] 'year' (*Durra* 24ʳ14). The latter method is also used to indicate voice-lessness (see 3.6.2.2). In these sources the term 'velarisation' stands opposed to *tarqīq* 'palatalisation' (or *muraqqaq* 'palatalised'), respectively. *Tarqīq* is, of course, used much less often; it is applied in oppositional pairs and on r and l.

In the manuscripts both terms are often abbreviated as mere markers scribbled above the words; a small م (m) usually stands for *mufahham* whereas ق (q) represents *muraqqaq* (see further discussion on palatalisation in Section 3.1.6). The use of the respective terms regarding velarisation and palatalisation are summarised as follows:

	Velarisation	Palatalisation
Qaw, Idr, Tarǧ, Tuḥ, Šuḏ, MG	tafḫīm	tarqīq
Dīwān, Ḥilya	ʾišbāʿ	ʾišmām, ʾimāla
Occasionally: Tarǧ., Qaw.	ʾišmām	

Examples of words labelled for velarisation are, e.g., ʾaš [aš] 'food' (*ʾIdrāk* 14; MS 3ᵛ20; marker above š); ʾant bir [ant] 'swear' (lit. 'give an oath'), "velarisation of the *hamza*" (*bi-tafḫīm al-hamza*, *Qawānīn* 77,8; also 74,8; in *ʾIdrāk* 24 ʾanṭ); būš [boš] 'divorced' (*ʾIdrāk* 32); bul- [bol-] 'to be' (*ʾIdrāk* 126,9); yab for [yap] 'build!' (velarisation [of b?], *Qawānīn* 75,11) and yay- [yay] 'to spread out' "velarised" (MG 45ᵛrt/bm).[36]

3.1.2.1 Velarised *l* [ḷ]
The velarised *l* occupies a special position among the non-standard Turkic phonemes, since it occurs as a well-accepted allophone in Classical Arabic; one could say that its position is close to that of a phoneme in the Arabic tradition. The usual term is 'velarised *l*' (*lām mufahhama*, *Tarǧumān*, *Qawānīn*, *Šuḏūr*, *Ḥilya*), which serves as the terminological base for coining other terms for velarised phonemes. The velarised *l* occurs, e.g., in ʾalṭūn-liġ 'with gold' and ʾat-lïġ 'with a horse' (*Ḥilya* 74,8).[37] Its pronunciation is described as follows:

"The Turk pronounces the *l*s mentioned here like the Arab pronounces the *l*s in the word for the Divine" [i.e. Allāh 'God']." (*at-turk yanṭuqu bi-*

---

[36] *Qawānīn* describes the Turkic negation suffix mA as "a velarised m vocalised with a" (*mīm maftūḥa mufahhama*, *Qawānīn* 13,9), in which an interpretation as 'velarised' is probably not correct, for it obviously occurs in front words too. Here *mufahham* should perhaps be interpreted as 'lengthed' or 'stressed' (see discussion in 3.5.1)

[37] In Rifʿat's edition these words are imperfectly vocalised.

hāḏihi l-lāmāt al-maḏkūra kamā yanṭuqu l- ʿarabī bi-l-lām min lafẓa al-ǧalāla ta ʿālā musammāhā wa-ǧalla, Ḥilya 77,5.)

In *Šuḏūr* the velarised *l* is systematically indicated in orthography by means of three dots underneath the word, e.g., 'al̤ [al] 'take' - [el] 'hand' (although in the MS here both ls have three dots underneath, 5ʳ5f).

3.1.2.2 Mixing and flavouring of consonants

Apart from the above mentioned terminology, the sources often use the concept of a consonant that is 'mixed' *(mašūb)* or 'flavoured of' *(mušamm)* others for indicating velarisation (see discussions 1.1.2). Both concepts are also used for Turkic in order to describe voiced and voiceless phonemes based on basic Arabic phonemes. For example, the velarised *z* is described as " *z* mixed with *ṣ*" (*zāy mašūba ṣādan*, 8,7)[38] or elliptically "mixed *z*" (*zāy mašūba*, 10,4), and occasionally also *bi-l-ʾišmām*, literally 'with flavouring' (75,9).

This description of a velarised *z* is identical with the one used in Arabic for the same or a similar allophone of *ṣ* (see 1.2.1). Further, in *Tarǧumān* and *Durra* the velarised *d* is described as " *d* flavoured with *ṭ*", e.g., 'adim [adïm] 'step' (*ad-dāl mašmūma bi-ṭ-ṭāʾ, Tarǧumān* 21,9).

In the manuscripts the concept of 'mixture' is orthographically reflected by means of a small velar consonant scribbled above the basic one. In *Tuḥfa*, for example, the initial *š* in *šāš-* [čaš] 'to come unexpectedly' (7ᵛ12) is marked with a small *ṣ*, in order to indicate a velarised *č*. In *Qawānīn* velarised *z* is in one instance indicated with the combination of *zṣ*, e.g., buzṣ 'destroy!' (75,9; MS 81ʳ), but usually a full description is given. Further examples are qizḍir 'heat!', and buǧazla 'kill!' (by cutting the throat) (*Qawānīn* 75,4)[39], duwār [duwaːr] 'wall' with a *ḍ* scribbled underneath d (*Durra* 3ʳ8). In *Bulǧa* the orthographic reflection has become the base for the description itself, viz., " *d* with a superposed *ṭ*, or a *ṭ* with a superposed *d*; it is pronounced between [d and ṭ]" (*ad-dāl fawqahā ṭā wa-ṭ-ṭā fawqahā dāl yunṭaq bihā baynahumā*, MS 1ʳ8).

---

[38] Em. for *mašūma* in MS 8ᵛ.

[39] Throughout the edition of *Qawānīn*, though, *ẓ* is erroneously described as " *z* mixed with *ḍ*" (*zāy mašūba ḍādan*, 29,11), in print it is indicated in numerous instances with zḍ, e.g., 'azḍ qal-ḍī 'it was near' (literally 'little was left') (*qāraba*, 29,18). These must be attributed to misreadings by a copyist or the printer.

3.1.3 Labelling of neutral consonants

The use of velar consonants, for those sources that accept them as con-
sonants/phonemes, should make application of the labels mentioned
in the previous section superfluous. Nevertheless, in some instances the
basic consonant with a label is used instead. In these cases it is difficult
to interpret the perceived qualitative difference between the velar and
'velarised' variants—if any difference is intended. In *'Idrāk mufaḫḫam*
is used in the following instances where an emphatic consonant could
have been used, viz.,

Neutral consonant with label	Velar consonant
ʾatā [ata] 'father' (marker above t)	ʾaṭā [ata] (*'Idrāk* 8; MS 2ᵛ3)
yat [yat] 'power' (*'Idrāk* 91)	yaṭ [yat] 'stranger' (*'Idrāk* 94)
sibā [sïpa] 'young donkey' (*'Idr.* 51).	

An interesting case in point is the grapheme /ʾat/. In *'Idrāk* it has three
meanings, depending on the label. When marked for velarisation, i.e.
[at], it means 'throw!', whereas when it is palatalised, i.e., [et] its
meaning is 'meat' (*'Idrāk* 7; MS 2ᵛ2; t marked in both instances). The
marker for palatalisation is, in fact superfluous, but may be put here in
order to stress the opposition. In *Dīwān* it is labelled, probably for the
same reason. But there is more to it. The back variant of this grapheme
has a second meaning, i.e. 'horse'. For the former, however, ʾAbū
Ḥayyān (14) uses the consonant one would expect in a back word all
along, i.e. ṭ. *Qawānīn* (57,16), *Dīwān* (29,10; 12; 52,6; 95,1) and
Ḥilya (87,8; 93,12) have yet other solutions.[40] The overall scheme is as
follows:

[at / et]	*Idrāk*	*Ḥilya*	*Dīwān*	*Qawānīn*
'shoot!'/'throw!'	velarised		no label	ṭ
'horse'	ṭ	velarised / ṭ	velarised	ṭ
'meat'	palatalised		palatalised	t

Qawānīn shows the expected orthographical signs, i.e., ṭ/t in order to
express the distinction. The use of both velarised t and ṭ in *'Idrāk* is
difficult to explain; the phonetic difference between [at] 'horse' and
[at] 'throw' is not clear, if there is any. Perhaps ʾAbū Ḥayyān could not
decide between the two alternatives he had, or, indeed, he heard differ-

---

[40] In some Turkic languages there is a third meaning of this grapheme, 'name'; ʾAbū
Ḥayyān (9) gives the Oġuz variant of this word which ends in -d. *Qawānīn* uses both -ṭ
and -ḍ.

ent pronunciations. A second option is that the words without ṣ and ṭ were copied in this form from an older source that does not use velar consonants.

For historical reasons a back pronunciation must be assumed for a number of words even though this is, or cannot be reflected orthographically, and markers are not used either. Examples are, e.g., ʾač [ač] 'open' (ʾIdrāk 8); čatir [čatïr] 'sunshade' (ʾIdrāk 41); sina-dī [sïna] 'he tested' (ʾIdrāk 54; elsewhere 112,13 ṣ-); mūm [mum] 'candle' (Qawānīn 64,8); siŋ-dī [sïŋ] 'he hid in the earth...' (ʾIdrāk 54); ʾuč[uč] 'fly' (ʾIdrāk 7; ʾuš Qawānīn 78,8) and ʾut [ot] 'grass' (ʾIdrāk 7).

There are very few examples of words that are marked for velarisation whose only vowel is i, and the phonetic value it represents is then difficult to interpret. In some instances, no doubt, [ï] must be meant, whereas in others probably the stand-alone vowel [é] is intended. In ʾIdrāk there are a few examples of this, e.g.,

   did-dī [dïd] 'to pluck (wool and cotton)' (ʾIdrāk 48; MS 14ᵛ8, initial d labelled; EDT 450; cf. Targumān 36,20; Tuḥfa 37ʳ1)

   ʾil [él] 'town' (ʾIdrāk 20; MS 5ᵛ2; marker above l; EDT 122 "él")

   ʾin [én] 'breadth' (opposite of length) (ʾIdrāk 23; MS 6ᵛ9; with hamza marked).

In these instances the sources apparently posit [é] as a secondary realisation of i. Other descriptions of [é] are given in 3.1.6.5.

3.1.4 The status of velarised consonants vis-à-vis their unlabelled counterparts

Each of the sources used in this study gives a listing of the consonants of Turkic which are discussed in 3.6.1; with an appropiate modern term, the consonants in the listings could be described as 'phonemes'. Other phonemes than the basic ones that are included in the lists are typically indicated by means of labels. For example, in the instances mentioned above the term *mufaḫḫam* is used to describe 'allophonical variants', rather than phonemes.[41]

The status of these non-regular consonants in relation to the 'regular' velar consonants is elucidated in *Qawānīn* as follows:

> "The consonants of elevation are seven [in number] and they are gathered in the [mnemonic sequence] qẓ ḥṣ ḍġṭ. [However,] there is no ẓ in this language and therefore it has only six [consonants of elevation]. The ve-

---

[41] In *Ḥilya* (72,4) a bāʾ and zāy mufaḫḫama (79,4) are mentioned, but in there they must be interpreted as [p] and [ž], respectively, rather than velarised (see discussion in Section 3.6.2.1).

larised *b*, *n* and *l* resemble them; like this is every velarised consonant and so is the *z* that is mixed with *ṣ*." (*wa-ḥurūf al-istiʿlāʾ sabʿa yuǧmiʿuhā qawluka qẓ ḥṣ ḍġt wa-hāḏihi l-luġa laysa fīhā ẓāʾ muʿǧama fa-takūnu fīhā sitta wa-šibhuhā al-bāʾ wa-n-nūn wa-l-lām al-mufaḥḥama wa-kaḏālika kull ḥarf mufaḥḥam wa-kaḏalika az-zāy al-mašūba ṣādan,* Qawānīn 8,5-7.)

Four comments to this statement are in place here. First, although the author compares the velarised **b**, **n**, **l** and the '*z* that is mixed with *ṣ*' to the elevated consonants of Arabic, I am not quite sure that default conditioned velarisation is indeed a general principle for **n** to the same extent it is for **r** and **l**. There is, however, some supporting evidence. *Qawānīn* itself specifies elsewhere that **n** in **nādir bū** [nedir bu] 'what is this' is palatalised ("with inclination of **n**" *bi-ʾimāla an-nūn*, 16,9), which only makes sense if back is the default pronunciation. Further-more, **n** is mentioned in *Dīwān* and the Margin Grammar too as a velarised consonant (see 3.1.6.3).

The second remark, in addition to the velar and velarised conso-nants he mentions, the author of *Qawānīn* refers "to every velarised consonant" *(kull ḥarf mufaḥḥam)*, which I interpret here as an indica-tion that basically all neutral consonants can be labelled as velarised.

In the third place, *Qawānīn* is the only source to explicitly mention *b*, which, in fact, makes very little sense. Velarisation is, according to this quotation (see previous remark), a fairly productive feature that can be applied to all neutral consonants; in other words, there is no reason to particularly indicate this for **b**. On this ground, I tend to interpret *bāʾ* here as a copyist's or printer's error for *rāʾ*. Together with with *n* and *l*, *r* forms the category of the liquids *(ḥurūf dawlaqiyya)* (see also 1.1.1) Indeed, *Qawānīn* is not the only source to refer to this allegedly common feature of all three liquids, for similar descriptions are found in *Dīwān* and the Margin Grammar (cf. 3.1.6.3).

In the fourth place, according to other data in *Qawānīn*, **ḥ** is not a basic Turkic phoneme; it has a velarising effect only inasmuch it is an allophone of **q**, such as in **ʾaḥšam** 'evening' and is described in terms of mixing (see discussion of Turkic allophones below in 3.6.4.1 and no. 21b in the appendix to this chapter).

Another interesting point is that ظ is apparently not considered a mere velarised *z*; there seems more to it. In spite of his reservations for *ẓ*, though, the author does accept the grapheme ض, i.e. **ḍ**, for velarised **d**.

3.1.5 Velarisation in *Ḥilya* and *Dīwān*

Ḥilya and *Dīwān* occupy a special place among the sources, in the sense that they reject the existence of all, or a number of Arabic cov-ered consonants, respectively. The opinions reflected in these works with regard to velarisation and palatalisation too differ considerably from the ones found in the other sources.

3.1.5.1 Velar consonants in *Ḥilya*

According to Ibn al-Muhannā, t and s can be substituted *('ibdāl)* with ṭ and ṣ, respectively, viz., s/ṣiǧir 'cattle'; s/ṣūl 'īliŋ (lit.) 'your left hand') t/ṭūrum 'camel colt'; t/ṭāy 'colt' (*Ḥilya* 79,8); 'urt/ṭā 'middle' (90,12). [42] In the rest of his work, though, Ibn al-Muhannā displays a preference for non-velar consonants in these same words (172,14; 90,14; 172,4), which means that he does not use ṭ and ṣ in a regular way for indicating a back pronunciation. The other sources (except *Dīwān*, which rejects all four 'covered consonants', cf. 3.1.5.3 and 3.6.1.2) typically do use ṭ and ṣ (cf. *Qawānīn* 37,16; *'Idrāk* 111,8; MG 52ʳlt; *Tarǧumān* 39,17; *Tuḥfa* 45ʳ8). [43]

In a number of occasions in *Ḥilya* s and t co-occur with ṭ, e.g., ṭūt ǧil 'grasp!' (129,6), sāṭin 'buy' (126,12), whereas in the other sources velar and neutral consonants typically do not occur together in the same stem.

3.1.5.2 Ibn al-Muhannā's concept of velarisation

Ibn al-Muhannā applies the term *mufaḫḫam* to a small number of consonants only. He understands *tafḫīm* as a general term for 'stress' or 'emphasis', of which velarisation is one possibility, viz.,

> "The consonants that are emphasised like the *l*s that resemble the pro-nunciation of the Arabs in the utterance 'the almighty God' are marked with *f* (ڢ)." (*wa-mā kāna min al-ḥurūf mufaḫḫaman ka-l-lāmāt al-mušābiha li-nuṭq al-ʿarab bi-lafẓa ʾallāh taʿālā fa-ʿalāmatuhu f*, *Ḥilya* 74,8.)

The abbreviation *f* ڢ recurs with other consonants that are labelled as *mufaḫḫam*, e.g., tūy /tuɣ/ (with marker above y) [toy] and būy /buɣ/ [boy] (see for other use of *mufaḫḫam* in *Ḥilya* 3.6.4.1). In *Ḥilya* velari-sation is one of the features, apart from lengthening, that can be ex-pressed by means of insertion of a glide, called *tafḫīm al-ḥurūf*

---

[42] Some of these examples from *Ḥilya* are not fully vocalised.

[43] Interestingly, most of these words contain r or l, so that there may be a relation with the default back interpretation of r and l as discussed in section 3.1.6.2).

'velarisation of the consonants'. Here /w/ serves as a marker of velarisation, viz., **būy** /bu<u>w</u>y/ [boy] 'fenugreek' and **tūy** /tu<u>w</u>y/ [toy] 'a large gathering'.[44] Evidently, Ibn al-Muhannā here uses an orthographic device that is usually associated with lengthening for indicating velarisation.[45]

3.1.5.3 Velarisation in *Dīwān*

In *Dīwān*, Kāšġarī, too, basically distinguishes between back and front words by means of the appropriate consonants (cf. examples in 3.6.2). Unlike the other sources, though, Kāšġarī categorically rejects all four 'covered' consonants (along with the Arabic consonants *ḫ, h* and *ʿ*) for Turkic:

> "In none of the Turkic languages do you find *ṭ*, nor any of the 'covered consonants' *ṭ, ẓ, ṣ, ḍ* or the guttural consonants *ḫ, h* and *ʿ*." (*wa-lā tūġad fī luġāt at-Turk bi-ʾasrihā aṭ-ṭāʾ wa-kaḏālika min ḥurūf al-ʾiṭbāq aṭ-ṭāʾ wa-ẓ-ẓāʾ wa-ṣ-ṣād wa-ḍ-ḍād wa-kaḏālika min ḥurūf al-ḥalq al-ḫāʾ wa-l-hāʾ wa-l-ʿayn, Dīwān 7,4-6.*)

The only velar consonants Kāšġarī does accept are **q** and **ġ**.

An important consequence of the rejection of all 'covered' consonants is that more instances of back vowels must be indicated for a back pronunciation than is the case in the other sources. For this purpose Kāšġarī has a set of labels at his disposal, which is also described in detail by Dankoff and Kelly (1982: 56ff) and Kelly (1972). For velarised words Kāšġarī uses the terms *ʾišbāʿ* 'saturation' and *mušbaʿ*, literally 'saturated' (cf. 3.1.5.5). These terms stand opposed to *ʾišmām* 'palatalisation' (cf. 3.1.6). In *Dīwān ʾišbāʿ* is used 16 times, and *ʾišmām* 15. As a general reference, Kāšġarī used "the word that contains a **k**" (*al-kalima allatī fīhā kāf*), 'fine' (*rakīka*)[46] ,"inclining to the

---

[44] In two other instances the term *tafḫīm* does not relate to velarisation but to other qualities of the consonants. For example, with *tafḫīm* **b** is pronounced as [p], and **z** as [ž], respectively. In *Tarğumān* *tafḫīm* is used for both velarisation and descriptions of voicelessness (3.6.2.1).

[45] Three other examples are in fact already interpreted as back because of the occurrence of **ġ**, and any marker would be superfluous. In those cases, though, a marker might serve for indicating that **u** is pronounced [o] or the like, e.g., **būġdāy** /buwġda"y/ [boġda:y] 'wheat'; **būġrā** /bu<u>w</u>ġra"/ [boġra] 'camel stallion'; **tūġrā** /tu<u>w</u>ġra"/ [toġra] 'signature'. In modern Turkic languages, however, there are no indications for [o] in these words. (In *Ḥilya* these same words recur as examples of 'dead **ġ**' *[ġayn mayyita]*, see Section 3.6.4.1).

[46] In *Dīwān* the term *rakīk* is further applied to the distinction between voiced, 'fine' *K̲ (kāf rakīka)* [g] and voiceless 'firm' *K̲ (kāf ṣulba)* [k] (also for **p**, **t**, **č** and **q**, see discussion in Section 4.8). The same term is used to describe the way of speaking of certain tribes who speak with *rikka* (by Dankoff and Kelly translated as 'slurring').

three vowels" (*mumāla ʾilā l-ḥarakāt aṭ-ṭalāṭa*, or simply *mumāla* (302,15; *mumāl* recurs in MG, cf. 3.4.1). In a similar way he differentiates between velar "the word which contains q or ġ" *(al-kalima allatī fīhā qāf ʾaw ġayn)* on the one hand, and "velarised" words *(muṣbaʿa fī l-lafẓ)* on the other.

3.1.5.4 Back and front pairs in *Dīwān*
'Velarised' and 'palatalised' basically serve to indicate the opposition between back and front words. In *Dīwān* all graphemes that are labelled indeed form a so-called minimal pair of which at least one member is marked (although the words themselves do not necessarily occur close to each other in the text.) In some instances marking serves to distinguish between [u - o] or [ü - ö].

There are three possible pairs. The first is that the default pronunication is front, therefore the word must be labelled if a back pronunciation is intended; secondly, the opposite, i.e. the default is back, the front pronunciation being labelled. The third is that both members are labelled for the respective pronunciations in order to indicate the opposition, regardless of the default pronunciation, viz.,

Unlabelled	Velarisation
ʾaba [epe] 'mother'	[apa] 'bear' (55,13; *hamza* velarised)
ʾuǧ [üč] 'three' (29,17)	[uč] 'well-known town' (30,1; *ʾalif* velarised; EDT "Uç Turfan").
taŋ /tank/ [teŋ] 'peer' (599,15)	[taŋ] 'sieve' (600,2).

Unlabelled	Palatalisation
turā qalqan [tora] 'shield' (542,10)	[töre] 'front side of the house' (542,11)

Velarisation	Palatalisation
ʾat [at] 'horse' (29,10)	[et] 'meat' (29,12).
tūš [tuš] 'buckle' (497,11)	[tü:š] 'breast' (497,10).

However, the number of words Kāšġarī labels in this manner is much smaller than one would deem necessary to cover all instances where confusion might occur, especially the number of verbs is very small. A plausible explanation for the relatively small number of labelled verbs is that in *Dīwān* all verbs are already marked in another way. Throughout *Dīwān* Kāšġarī provides for each verbal entry its appropriate infi

nitive form, which is, obviously, **māk** with front, and **māq** with back
verbs, viz.,[47]
butur-māk [bütür-] 'to heal'   **batur-māq** 'to hide' (306,1)
(305,15)
tutuš-māk [tütüš-] 'to quarrel'   **tutuš-māq** 'to hold one another'
(313,8)   (313,6).
According to principles he does not disclose—although in some in-
stances they are obviously related to the occurrence of certain conso-
nants in the stem—Kāšġarī gives the appropriate suffix for each verbal
stem. In further references the velar stems are called *qāfiyya* i.e.
'ending in **q**' (occasionally *ġayniyya*, 'ending in **ġ**'), and palatal verbs
are called *kāfiyya*, i.e. 'ending in **k**'. These denominations either derive
from the infinitive suffixes Kāšġarī assignes them, or from the general
principle that words containing a **q** take velar suffixes, whereas those
that have **k** get palatal suffixes (see further discussion in 3.3.2).

In this way, it is left to the readership to infer that deverbal nouns
and adjectives follow the verb in the type of suffix, for no deverbal
word appears to be labelled for either palatalisation or velarisation (see
listing in Dankoff and Kelly 1982:56-8).[48]

3.1.5.5 Kāšġarī's use of the term ʾišbāʿ

As pointed out above in section 1.5.1.1, in Arabic morpho(no)logical
terminology ʾišbāʿ means 'lengthening' for poetical reasons, such as
the metre or rhyme. In other sources, such as ʾIdrāk, it is used in this
sense exclusively (see below Section 3.5.1). In *Dīwān* ʾišbāʿ is applied
in the sense of lengthening only once, viz., *mušbaʿa fī l-lafẓ*, literally
"lengthened in the pronunciation" (284,12). This concerns a passage
in which Kāšġarī explains that biradical words, such as /bar-/ 'to go',
and /tur-/ 'to stand' can be considered triradical if one also counts the
weak consonant, *y*, *w* and ʾalif, which are inserted when these verbs are
pronounced with lengthening (ʾišbāʿ), /ba"r-/ **bār-** and /tuwr/ **tūr-** (bi-
n-nuṭq ʿinda l-ʾišbāʿ, *Dīwān* 284,4). (The opposite, reduction is possi-
ble too; cf. above 1.5.1.2).[49] It is therefore evident that Kāšġarī was
aware of the signification the term ʾišbāʿ had in Arabic linguistics, and

---

[47] Sometimes even a form **mākq** (with both k and q) is given, interpreted by Dankoff
and Kelly (1982:64) as an indication for velar-palatal alternation in the root.

[48] In Dankoff and Kelly's list of occurrences of ʾišbāʿ and ʾišmām (1982:56ff) only
one verb is labelled. This does not concern the opposition back - front, but a description
of the vowel [o] instead of [u] in a back word.

[49] This passage also shows that Kāšġarī did not regard ʾalif as a vowel, but as a con-
sonant that can be a part of the root of a word (although in practice ʾalif is never a part of
the root in Arabic).

that he for some reason preferred to choose it for indicating 'velarisation', rather than another term he could have applied, such as *tafḫīm*. (In *Dīwān tafḫīm* occurs once, 91,14).

The word *lafẓ*, 'pronunciation', immediately calls to mind its opposite, *taqdīr*, 'underlying structure'. The phrase in question means that the word is lengthened in pronunciation, for prosodic, rather than morpho(no)logical reasons. Thus Kāšġarī quite correctly understands *ʾišbāʿ* and its opposite *ʾišmām* (for which see 3.1.2 and 3.1.6.2) in their broadest sense as features that are reflected in pronunciation (or the surface structure) only, and which have no representation in the underlying structure *(taqdīr)*. In this argumentation, velarisation and palatalisation alike, may be understood as mere prosodic, secondary features.

There is an obvious parallel with Ibn al-Muhannā's use of the glides in *Ḥilya* (discussed above in 3.1.5.2). Among other features, Ibn al-Muhannā uses marking with glides as an orthographic instrument to indicate velarisation.[50] As we have seen, the use of glides for non-morpho(no)logical reasons is called *ʾišbāʿ*. It appears that Ibn al-Muhannā uses this principle for indicating velarisation, although he does not use the term *ʾišbāʿ* in this respect. In Kāšġarī's *Dīwān*, however, it is only the term that is used, without involvement of the glides themselves.[51]

3.1.6 Palatality and palatalisation

This section discusses the instances in which 'palatalisation' is used in the description of Turkic. As a rule, words without a velar consonant have a default pronunciation with front vowels, based on concepts derived from Arabic. For Arabic a palatal pronunciation implies a standard inclination of *a* towards [ä/e]. In the same way, Arabic *u* in neutral words is not always palatal, i.e. [ü], it may be neutral, [u], but never [o] (see discussion in 1.1.1 and 1.4.2). With regard to *i*, things are less evident, since the Arab grammarians do not explicitly describe a change in the quality of *i* with velar or velarised words.

---

[50] Ibn al-Muhannā once uses the term *mušbaʿ* in the sense of 'emphasised' when describing *ġ*, where he writes that the people of Turkastān pronounce the *ġ* "emphasised as it is" *(mušbaʿa bi-ḥālihā)* in words like **tūġrā** /tuwġra"/ 'signature' [tuġra], **būġdāy** /buwġda"y/ 'wheat' [buġday], **būġrā** /buwġra"/ 'camel stallion' [buġra]—as opposed to other Turks who pronounce it as w (*Ḥilya* 78,2; cf. Section 3.6.2 'dead *ġ*').

[51] Unless, indeed, one considers the two examples given above /ba"r/ and /tuwr/ as examples of velarisation by means of insertion of a glide.

Nevertheless, it seems reasonable to assume for Turkic a certain colouring in velarised words, i.e. [ï]. However, in Turkic vowels can never be neutral, they have to be either front or back in view of the suffixes the stem is to take.[52] In this respect, in front words, **u** is gen-erally interpreted as [ü], **a** as [e], and **i** as [i]. To mention some exam-ples with **a**, **'amkak** [emkek] 'hardship' (*Dīwān* 68,1); **dak-dī** [dek] 'he reached' (*'Idrāk* 49). In front words **u** is generally read [ü], e.g., **'uǎun** [üǎün] 'for' (*Tarǧumān* 54,20); **duš-tī** [düš-] 'he fell' (*'Idrāk* 48) and **'iǎtur-dī** [iǎtür-] 'he made drink' (*Dīwān* 116,16).

3.1.6.1 Velar consonants labelled palatalised
In two instances 'Abū Ḥayyān indicates that a velar consonant must be pronounced palatalised *(muraqqaq)*, which appears contradictory, since he could have used the non-velar counterpart instead. In these cases the 'velarising' effect on the following vowel is probably perceived as too strong, viz., **ṭuy-dī** [tuy] 'he felt'; **ṭuy-sū** [tuysu] 'feeling' (*'Idrāk* 67; MS 21ᵛ1). The default pronunciation of these words then, may be [o]. (Note that here **s** is used rather than **ṣ**).[53]

3.1.6.2 Palatalisation of **r** and **l**: back is default
The labels for palatalisation in question are especially often applied to words with **r** and **l**. This is a general tendency in at least four sources (*Dīwān*, *'Idrāk*, the Margin Grammar and *Qawānīn*), which can easily be explained with the aid of rules from Arabic phonology as pointed out above in section 1.1.4.[54]

There is abundant evidence for this in all sources mentioned above. Words with a default back pronunciation with out being marked are, e.g., **'arī** [arï] 'clean' (*'Idrāk* 10); **bul-dī** [bul] 'he found' (Oġuz) (*'Idrāk* 36); **bar** 'existing' (*'Idrāk* 29); **nār** 'pomegranate' (*Qawānīn* 63,8);

---

[52] That is, within the system of a given Turkic language the opposition must exist. In absolute phonetic terms it may well be the case that in certain Turkic languages the opposition could more aptly be described as neutral-front, or neutral-back.

[53] There are two other explanations. The first is that **ṭ** must not be pronounced palatalised, but rather like velarised *d* If this is the case, then **ṭ** only serves to indicate velarisation of the initial consonant. Interpreted in this second way, the description yields [duy] and [duysu], which is close to Oġuz forms such as **duy-** in Turkish and modern Türkmen (cf. Hanser 1977:239). The second option is that the marker here is used for indicating an oppositional pair with [toy-] 'to be saturated', which is marked for velarisation (cf. 3.4.1).

[54] As stated already in 3.1.4 and below in 3.1.6.2, there is as yet no convincing evidence that the same rule applies to the third liquid, **n**, although some sources make statements to this effect., There are indications that it holds to some extent for Turkic words with **z** preceded by **u** as well, although this principle is, to my knowledge, not described in Arabic grammar itself; see discussion in section 3.4.3.

'ulu-du [ulu] 'he howled' (said of dogs) (*'Idrāk* 20); **yarat-tï** 'he creat-
ed' (*'Idrāk* 93); **yar-dï** 'to split' (*'Idrāk* 93); **yïlān** [yïlan] 'snake and
viper' (*Qawānīn* 62,9).

Naturally, if a back pronunciation of **a/u** close to **r** and **l** is default,
it is necessary to specifically indicate a front pronunciation with a la-
bel. This too can be exemplified with many examples, e.g., **'al** [el]
'hand' (*'Idrāk* 20; MS 5ʳ21; *Šuḏūr* 5ʳ11 palatalised l); **'ala-dï** [ele] 'he
sieved' (*'Idrāk* 20; MS 5ᵛ1, l palatalised); **yabar** [yeber] 'send!'
(*Qawānīn* 74,12); **tabra-t** [tebret] 'move!' (*Qawānīn* 77,10); (in the
last two words "**r** [is] palatalised" *[bi-rā' muraqqaqa]*). For *Dīwān* this
can be illustrated with two minimal pairs in which the default form has
back vowels, whereas the front version has a label, namely

Palatalised	Unlabelled
sar- [ser] 'be patient' (285,7)	[sar] 'to rebuke' (267,14 **sār-**) (285,8).
turā [töre] 'front(side) of the house' (542,11)	[tura qalqan] 'shield' (542,10)

3.1.6.3 Discussions of l and r in the sources

Apart from rather indirect clues such as the marking of consonants
and words with abreviated labels, as has been shown above, the rule
with regard to l and r is referred to quite explicitly in *Dīwān* and in the
Margin Grammar. Kāšġarī notes that a given suffix has **ġ** when it is at-
tached to 'velarised' *(muṣbaʿa)* 'liquids' *(ḥurūf aḏ-ḏalāqa)* (i.e. l, r and
n), and to words containing **ġ**, whereas the suffix contains **k** when it is
attached to their palatal counterparts. He writes:

> "The **ġ** is attached in the case of the liquids that are velarised or [words]
> that contain a **ġ**, and the k in the opposite case." *(tuzād al-ġayn fī ḥurūf
> aḏ-ḏalāqa al-muṣbaʿa 'aw al-ġayniyya wa-l-kāf fī ḏiddihā...,* Dīwān
> 582,5; Dankoff and Kelly 1984: 311.)

In the Margin Grammar it is explained in quite similar terms that
words with either velar consonants, or with l, r or n receive a velar suf-
fix:

> "If the verb contains a q, ṭ, ḫ, ṣ, ḍ, ʿ, ẓ, then it is correct to use **qil**, and if it
> contains ..., l, r or n, it is correct to use **ġil**, and for [verbs that] contain
> any other consonant you use **kil**. The first category of consonants are
> called consonants of elevation" *(al-fiʿl 'in kāna fīhi min al-qāf wa-ṭ-ṭā'
> wa-l-ḫā' wa-ṣ-ṣād wa-ḍ-ḍād wa-l-ʿayn wa-ẓ-ẓā' ḥasuna fīhi qil wa-'in
> kāna fīhi al-...ā' [?] wa-l-lām wa-r-rā' wa-n-nūn ḥasuna fīhi ġil wa-*

mā ḥalā min hāḏihi l-ḥurūf kullihā yaḥsun fīhi **kil** *wa-summiya l-ʾūlā*
ḥurūf al-istiʿlāʾ, MG 44ʳlt in text.)

I am not sure whether the rules that apply to l and r equally holds for
n, but there are at least three authors that assume that all three liquids
share this conditioned pronunciation.[55] Note however, that neither
source mentions any conditioning by surrounding vowels. The inclu-
sion of a number of consonants that never occurs in Turkic (such as ʿ
and ẓ) shows that this rule is basically an Arabic principle that is trans-
ferred to Turkic. Incidentally, it could be interesting to know which
fourth consonant the anonymous author mentioned; unfortunately
the manuscript is illegible at this point. A fair guess would be z (as *zāʾ*),
since there are indications that the rule of l and r applies to z too.

Furthermore, the application of the term *ʾimāla* to a consonant
other than *ʾalif* or *a*, is interesting too (for which see 1.4.2).

3.1.6.4 Counterexamples of l and r
There are some counterexamples too. A few words are marked ve-
larised whereas this label seems redundant, considering the principle
explained in the previous section, viz., **ʾāl** /ʾaˈl/ [al] 'trick, and blond
horse' (*ʾIdrāk* 20; MS 5ʳ21, velarisation indicated above middle *ʾalif*);
ʾāl̤ [al] 'take' (*Qawānīn* 77,12; "velarisation of *hamza* and l"; also
Šuḏūr 5ʳ11, velarised l); **bašar-dī** [bašar] 'to complete' (*ʾIdrāk* 32; MS
9ᵛ8, š marked velarised). In the case of [el] 'hand' and [al] 'trick' the
markers may have been put there in order to stress the opposition be-
tween the two ways of pronunciation (see also Kelly 1972:183), espe-
cially when the words occur near to each other in the text, although,
strictly speaking, the label for velarisation is redundant.

In other instances, words known to be front remain unlabelled
even though a label would be expected here, viz., **buldī** [böl] 'to re-
move' (*ʾIdrāk* 36; MS 10ᵛ18); **durā** [döre] 'a code of law and conduct'
(*ʾIdrāk* 48) and **tar** [ter] 'perspiration' (*ʾIdrāk* 38).

3.1.6.5 r and l preceded by i
In accordance with the general rule in Arabic (see 1.1.4), the unla-
belled word with r or l preceded by i has default front vowels, which is
also the general principle in all sources, viz., **ʾir** 'make disappear!'
(*ʾIdrāk* 10); **birdī** 'he gave' (*ʾIdrāk* 29; *Ḥilya* 73ff; however EDT 354
"**bér-**") and **bir** 'one' (*ʾIdrāk* 29).

---

[55] In section 3.1.4 I discuss in more detail a statement in *Qawānīn*, with similar ref-
erences to this allegedly common feature of all three liquids. There is a similar reference
in ʿAlī bn Sulṭān al-Qārī"s *al-Minaḥ al-fikriyya* (22,37).

It is not quite clear how one should interpret **i** near an **r** and/or **l** when they are marked for palatalisation or velarisation, of which there are a few examples in *'Idrāk*. Possible interpretations are [é], or explicit back i,[ï], viz., e.g.,

Palatalised:
'ir   [ir] 'feel bored!'
       [ér] 'reach!' (*'Idrāk* 10; MS 3ʳ2, **r** palatalised; EDT 194)
'ir   [ïr] 'song' (*'Idrāk* 10; MS 2ᵛ22; **r** palatalised; EDT 192)

Velarised:
'il   [él] 'town, peace' (*'Idrāk* 20; MS 5ᵛ2, **l** velarised; cf. 3.1.3).

3.1.7 Summary

In some instances words with neutral consonants must be interpreted as having a back pronunciation. There are clear indications that this holds mainly for nouns. The grammarians considered the marking of verbal stems more important than nouns. As is briefly discussed in 3.1.5.4, *Dīwān* verbs are already marked for the opposition velar - palatal by means of the attachment of the infinitive suffix, and therefore in *Dīwān* only nouns need to be labelled for this opposition. Furthermore, the labelling mainly serves to distinguish between words that have the same orthographic representation (cf. Dankoff and Kelly 1982:56-8).

A similar situation holds for *Tarğumān*, which gives a list of verbal stems to which either a velar or palatal suffix is attached. In this way it is not necessary to specially indicate verbs. In both *Dīwān* and *Tarğumān*, therefore, only a relatively small number of nouns (in *Tarğumān* only those with **ẓ** and **l**) are labelled for the opposition velar - palatal. In *'Idrāk* and *Qawānīn* there is no such division; in *'Idrāk* 98 entries of the word list are labelled, only 28 of which are verbs, whereas in *Qawānīn* almost all back words are marked by the occurrence of a velar consonant, making an additional label superfluous. In *Ḥilya* marking is very scarce, which is probably related to Ibn al-Muhannā's peculiar insights in the distribution of velar and palatal suffixes (see discussion in Section 4.3).

The principles on which the sources base the distribution of terminology for indicating palatalisation and velarisation are summarised below in Table I:

neighbouring consonant	vowel sign	no label	label 'palatalised'	label 'velarised'
		Pronunciation of vowel when neighbouring consonant of the word has		

	neighbouring consonant	vowel sign	no label	label 'palatalised'	label 'velarised'
1	front: k, ḵ				
		a	e	—	—
		i	i	—	—
		u	ü	—	ö
2a	neutral ', b, m, n, w, y, š, č				
		a	e	é	a
		i	i	é	ï / e
		u	ü	—	u/o
2b	s, t (always neutral in Dīwān, occasionally neutral in Ḥilya; palatal in other sources; between brackets Dīwān/Ḥilya/ʾIdrāk)				
		a	e (a)	(e)	a
		i	i (ï)	—	ï / e
		u	ü/ö (u)	(ü)	u (o)
3	r, l (perhaps also z), with u or a				
		a	a	e	—
		i	i	é	ï/é
		u	u	ü/ö	o
4a	back: ṣ ṭ (not in Dīwān, not structural in Ḥilya; also ḍ in Qawanīn, Durra; ẓ in Durra)				
		a	a	?	o
		i	ï	—	—
		u	u	?	o
4b	q, ġ				
		a	a	?	—
		i	ï	—	—
		u	u	?	o

Table I: The labels assigned to Turkic consonants
and their implication for the pronunciation of vowels

3.2. Vowel harmony

As briefly discussed above in 2.2, Turkic suffixes are subject to two types of vowel harmony. In the first place, there is the so-called twofold vowel harmony, according to which, for example, the archiphoneme A is realised as either [a] or [e]. This occurs, for instance, in the locative case DA [da - de], the dative case GA [ġa - ge], and in the plural suffix lAr [lar - ler]. In the second place, there is the fourfold vowel harmony, in which the vowel in the suffix has four forms,

depending on whether the stem is back or front, and rounded or unrounded, e.g. the imperative suffix GIl, the past tense suffix DI and the interrogative suffix mI.

The twofold vowel harmony can conveniently be described in terms of velarisation and palatalisation, whereas for the fourfold vowel harmony more detailed descriptions are needed. We shall see on which principles the distribution of back and front suffixes is based.

3.2.1 Twofold vowel harmony

According to Arabic theory the palatal or velar character of a given suffix evolves as a prosodic feature after elevated or depressed consonants. Vowel harmony thus is often correlated with a change of the consonant, and I am inclined to believe that this change, especially if it can be represented in Arabic script, was the main, or perhaps the sole reason, for the Arab authors to pay attention to this phenomenon. All sources (except *Hilya*) describe the distribution of the velar and palatal forms of the suffixes based on these principles. (For ample discussion of the attachment of suffixes that contain the twofold vowel harmony, he reader is referred to sections 3.3).

In Turkic languages a change in the quality of the vowel must also be inferred when the suffix does not undergo a change that can be made explicit in orthography. Usually, though, the fact that **a** in velar suffixes is pronounced [a] and [e] in palatal or palatalised words is not explicitly referred to in the sources, since they consider it a secondary feature. In this the following paragraphs my intentions is to show that velarity and palatality are nevertheless assumed even where this is not visible in orthography. Cases in point are lAr, sA and DA.[56]

3.2.1.1 The plural suffix lAr

Although not explicitly stated in the sources, there is evidence that the plural suffix lAr follows the basic word in palatalisation and velarisation. The evidence consists of consonantal changes in the suffix, e.g., **kiši-lar-ḳā** person-PLUR-DAT 'to people' (*ʾIdrāk* 132,11). In view of the palatal dative suffix, the pronunciation of the plural suffix must here be interpreted as front, i.e. [ler]. In a second example the plural

---

[56] Another instance of a suffix whose twofold vowel harmony is, as a rule, not recognised, is the negation suffix **mā**, which is attached to verbal stems. However, the author of *Qawānīn* states that the negation suffix is "a velarised **m** which is vocalised with **a**", ie. **ma** (or **mā**), (*mīm maftūḥa mufaḥḥama, Qawānīn* 13,9). This apparently holds for both velarised and palatal verbs indiscriminately, since it is illustrated with a palatal verb, i.e. **kal-ma-dum** 'I did not come'.

suffix is realised back, viz., **bu-n-lar-ǧuq** this- n -PLUR- DIM 'those little [ones]' (*'Idrāk* 105,14), which is inferred from the attachment of the diminutive suffix **ǧuq** (*'Idrāk* 105,10). The suffix **lar** itself, however, is never explicitly described as either palatal or velar.

3.2.1.2 The conditional suffix sA

Another instance of a suffix that does not change is the conditional suffix sA. Normally, the quality of the vowel varies according to the twofold vowel harmony, with front verbs it is [se], e.g., **kal-sä** [kelse] 'if he comes' (*'Idrāk* 152,22); **kat-ar-sä** [keterse] 'if he goes' (*Qawānīn* 20,6); **kal-ur sä** [kelürse] 'if he comes' (MG 39ʳlt), and with back verbs [sa]. In the sources, though, it is invariably spelled **sä**, never *-**ṣä**, with back and front verbs indiscriminately, even in the sources that in other cases express a velarised pronunciation by means of **ṣ**, viz., **'al-ḍu-m sä** [aldumsa] 'if I have taken' (*Qawānīn* 19,14); **'al-ur-sä** [alursa] 'if he takes' (*Qawānīn* 20,6); **ṭur-di-sä** [turdïsa] 'if he has stood up' (MG 63ʳrt); **ṭur-ur-sä** [turursa] 'whenever he should stand up' (MG 64ʳtop/lt).

It is difficult to account for this discrepancy; there are two clues. First, for some reasons the distinction [ṣ] - [s] is apparently perceived as less significant than the opposition [q] - [k], as it is posited in **maq** - **mak**. Second, there may have been certain spelling conventions in which **ṣ** is never, or rarely, used in suffixes.

3.2.1.3 The locative suffix DA

In the same way the locative case DA is added to both velar and palatal words indiscriminately, e.g., **ṭaŋ-dä** [taŋda] morning-LOC 'tomorrow' (*'Idrāk* 126,23) **'aw-dä** [evde] house-LOC 'in the house' (*'Idrāk* 137,21). There are no explicit references to a perceived difference in pronunciation.

Only in *Qawānīn* the quality of the **d** in back words is made explicit by means of the grapheme **ḍ**, e.g., **šar-ḍä** town-LOC 'in the town' (*Qawānīn* 41,8). In this way, the quality of **a** in the suffix can easily be inferred. In *Qawānīn* suffixes with -D- (e.g., locative **dä**, ablative **dan**, and past tense **dī**) have their own conditioned distribution (**d-, t-, ḍ-** and **ṭ-**). This is not the case in the other sources, in which **ḍ** lacks alto - gether (see discussion below in 3.6).

3.2.2 Fourfold vowel harmony

In general, the sources do not refer in a systematic way to complete vowel harmony, which also includes labial vowel harmony. The present

section discusses instances of suffixes, i.e. DI, GIl and mI, for which most Turkic languages observe a complete, fourfold vowel harmony of the type sketched above in 2.2.3.

3.2.2.1 The past tense suffix DI

The first example is the past tense suffix DI:

> "The consonant of the past tense is a d vocalised with i, if the consonant before it is vocalised with i or a. It is vocalised with u if the consonant before it is vocalised with an u, e.g. ṭur-dū ['he stood up'], 'ulṭur-dū ['he sat down'], bar-dī ['he went'], and bir-dī ['he gave']. In Oġuz d is always vocalised with i." (*ḥarf al-māḍī huwa dāl maksūra 'in kāna mā qablahā maksūran 'aw maftūḥan wa-maḍmūma 'in kāna mā qablahā maḍmū- man naḥwa ṭurdū wa-'ulṭurdū wa-bardī wa-birdī wa-fī t-turk- māniyya tuksar ad-dāl muṭlaqan,* 'Idrāk 111,11.)

'Abū Ḥayyān thus relates the quality of the vowel to the one in the fi- nal syllable. This principle is the same as the one generally held in western studies of Turkic. 'Abū Ḥayyān's comments can be formalised as follows: Ci(C)/Ca(C) → dī, Cu(C) → dū. Note however, that he does not give any clues in regard to front/back harmony.

An interesting remark is that, apparently, the labial harmony is not applied in Oġuz (Kāšġarī makes a similar statement in *Dīwān* 539). It is difficult to interpret this information. In most cases, 'Abū Ḥayyān's remarks on Oġuz are confirmed by evidence from modern languages, which proves that they should be taken seriously. To my knowledge, however, the complete vowel harmony subsists in the past tense suffix of most Turkic languages, including Oġuz languages such as Turkish and Turkmen (e.g., Hanser 1977:85).[57]

3.2.2.2 The pausality suffix of the imperative, GIl

Further descriptions of the fourfold vowel harmony are found with the forms of the pausality suffix in the imperative form:

> "If the first consonant of the imperative form is vocalised with an u, then the penultimate consonant [sc. ġ / ḵ in the suffix] is vocalised with an u too, except when there is an a in the verb. Examples of this are ṭur-ġul ['stand up'] kul ḵul ['laugh'] kustar-ḵil ['show'], 'ur-ġul ['beat']. If [the first consonant of the verb] is vocalised with an a or an i, then the penul- timate consonant is vocalised with an i, except when there is an u in the

---

[57] In some languages groups, such as Kazakh (Qipčaq), the complete vowel harmony (i.e. including labial harmony) in the interrogative suffix has become twofold: keldi me? 'has he come?' qaradï ma? 'has he looked?', for geldi mi? and baktï mï?, respectively, in Turkish (Oġuz).

verb. Examples of this are **bar-ġil** and ʾišit-ḳil."[58] (*wa-fiʾl al-ʾamr ʾin kāna ʾawwaluhu maḍmūman fa-mā qabla ʾāḥirihi maḍmūm ʾillā ʾin wuǧidat fī l-fiʾl fatḥa, miṯāl ḏālika ṭurġul kul ḳul kustar ḳil ʾurġul. wa-ʾin kānat maftūḥan ʾaw maksūran fa-mā qabla ʾāḥirihi maksūr ʾillā ʾin wuǧidat fī l-fiʾl ḍamma miṯāl ḏālika **barġil**, ʾišitḳil, ʾIdrāk* 120,19ff; cf. also *Qawānīn* 8,14; further descriptions in *ʾIdrāk* 106,19.)

Here ʾAbū Ḥayyān gives a complicated set of rules, which are schematised as follows:[59]

$C_1 u C_2$ or $C_1 a C_2 u(C_3)$ or $C_1 i C_2 u(C_3)$ → **ġul/ḳul**

$C_1 u C_2 a(C_3)$ or $C_1 a(C_2)$ or $C_1 i(C_2)$ → **ġil/ḳil**

In ʾAbū Ḥayyān's reasoning not only the quality of vowel in the final syllable of the verb plays a role, but also that of other vowels in the verb, which, one is lead to think, complicates the rules more than necessary.

3.2.2.3 The interrogative particle mI

The third fourfold suffix I discuss here is the interrogative particle mI. In the sources the complete vowel harmony is not always reflected in orthography, viz., **sangar kal-di-mü** Sangar come-3sg/PAST-INT 'Has Sangar come?' (*ʾIdrāk* 131,8). Nevertheless, in one instance vowel harmony is referred to as "The base form is **u**, whereas i is [used] for vowel harmony." (*wa-l-ʾaṣl aḍ-ḍamm wa-l-kasr li-l-ʾitbāʿ, ʾIdrāk* 131,8; also 155,4; and MG 62ʳtop.) This statement serves for the reconstruction of an alternative, viz., ***sangar kal-di-mī.

Apparently, the form with **u** is regarded as basic, and, therefore, the alternation with i is considered a prosodic phenomenon.

3.2.3 Terms for indicating vowel harmony

As briefly pointed out in the previous section, the term ʾAbū Ḥayyān applies is *ʾitbāʿ*, 'making follow' (also MG 60ʳbm; 50ᵛbm), which is also used in this sense in Arabic phonology (cf. 1.5.2).[60] The term *ʾitbāʿ* implies that the optional vowel i has no morphological significance, it merely follows as a replacement of a basic vowel. ʾAbū Ḥayyān uses in addition a second term, i.e. *tawālī l-ḥarakāt* 'sequence of the vowels'.

---

[58] The k in ḳil is explicitly described as 'Bedouin *k*' (*kāf badawiyya, ʾIdrāk* 120,17).

[59] According to *Qawānīn* the form of the causative suffix is also determined by the preceding vowel: "The verb contains no **u**, therefore as a rule the d has an i" (*wa-laysa fīhi ḍamma fa-li-ḏālika buniyat ad-dāl ʿalā l-kasr, Qawānīn* 68,13; also 67,14).

[60] *ʾItbāʿ* is further used to indicate a jingle, repetition of a word with change of the first consonant, like *kaṯīr baṯīr*. Kašġarī uses *ʾitbāʿ* (or *tabʿ*) also in this sense, e.g. **äwürdi täwürdi** 'he turned and turned' (*Dīwān* 310,2; also 217,12ff), thus using a meaningless word as a complement (cf. Dankoff and Kelly 1982:52).

He does so when he describes the insertion of a vowel, i.e. **a**, in the Oġuz verb **kal-mi-yā-san** 'you will not come', because of "the sequence of the vowels" (*li-tawālī l-ḥarakāt, ʾIdrāk* 130,21). In *Qawānīn* a third term is used, i.e. *murā'ā* 'agreement', to account for the **u** in the suffix **luq** in **yuq-luq** 'poverty', e.g., "the l is here vocalised with an **u** for the agreement." (*wa-l-lām maḍmūma hunā li-murā'ā ḍamma al-yāʾ, Qawānīn* 23,13). *Qawānīn* does not specify what type of agreement is exactly meant.[61]

3.2.4 Summary

In this section I have deliberately selected three suffixes in which the sources for different reasons do not, or cannot reflect the twofold vowel harmony. It appears that the twofold vowel harmony, which is covered by the distinction back - front, is only signalled by the sources when it can be related to a change in the consonantal form of the suffix. The fourfold, or complete vowel harmony, which includes labial harmony, is only occasionally referred to as a prosodic feature, and appears not to be considered a structural phenomenon.

3.3 *The distribution of palatal and velar suffixes as a phenomenon*

The preceding sections have shown how the sources distinguish between front words on the one hand, and back words on the other, and that they only occasionally indicate vowel harmony. Words that are not evidently either front or back are labelled in order to make a classification possible. The question arises as to why exactly the grammarians found it important to distinguish between these two types of words, and how do they relate the type of suffix to a given stem.

In my opinion, this basic division between front and back words is related to the interest the grammarians had in indicating certain principles to which they could relate the distribution of the various forms, rather than a genuine wish of describing the phonetic features of Turkic. This lack of interest in describing the quality of all vowels throughout has been evidenced in the preceding section with a suffix in which—in their view—morphological criteria do not play a role; i.e.

---

[61] Further, the short vowel preceding each one of the glides, u with w, i.e. /uw/, a with ʾalif, i.e. /aʾ/ and i with y, i.e. /iy/, is put because of *munāsaba* a 'relationship' with the glide in question; it is a *ḥaraka tuġānisuhā* 'a vowel that is of the same kind' (*Qawānīn* 44,8; 44,4). In the latter case, the vowels are said to be homogenic with the glides. In this sense these terms cannot be interpreted as expressing vowel harmony. It is used in the same sense in *Irtišāf* I 130,14f.

when the only difference is the pronunciation of the vowels (or a per-
ceived velarisation of the consonants), rather than a change of conso-
nants. These change are considered mere prosodic features which do
not affect any consonants. In this subsection I intend to show to
which principles the grammarians relate the distribution of front and
back suffixes. The approaches in *Dīwān* and *Tarǧumān* are very similar,
whereas it appears that Ibn al-Muhannā did not regard it as a coherent
feature at all.

3.3.1 Palatal and velar suffixes

The basic division between palatal and velar suffices to account for the
twofold forms of some suffixes as a prosodic feature, which can rela-
tively easily be reflected in Arabic script. Palatal words, or words
marked as 'palatalised', typically have suffixes with palatal consonants,
e.g., **kant-kā** [kentke] 'to the town', and of course, velar or velarised
words get velar suffixes, and **'arslan-ġā** [arslanġa] 'to the lion'. This
principle is reflected in *Qawānīn*, *'Idrāk* and the Margin Grammar in
quite similar terms. The following quotation is from the Margin
Grammar:

> "To the particles of the genitive belongs the word ġā, which has the
> meaning of 'towards' that governs the genitive, and whose form varies ac-
> cording to [the principles of] palatalisation and velarisation in the nouns,
> because the noun affects it." (*wa-min ḥurūf al-ǧarr lafẓa ġā bi-maʿnā 'ilā
> l-ǧarra wa-taḥtalifu ṣūratuhā ʿalā ḥasab at-tarqīq wa-t-tafḥīm fī l-
> 'asmā' li-'anna l-ism yaʿmalu fīhā*[62] ... MG 58 top; also MG 36ᵛ in text;
> 45ᵛrt/bm.)

In other words, suffixes differ in form depending on whether they are
attached to a velar(ised) word, or to a palatal(ised) word. With each dis-
cussion of a suffix, its forms both for velar and palatal words are given.

As a general reference, *'Idrāk*, too, uses the terms 'the velarised
word' *(al-kalima al-mufaḥḥama)* and 'the palatalised word' *(al-kalima
al-muraqqaqa)*, respectively. For example, ḳā is suffixed instead of ġā,
when a consonant that is pronounced palatalised *(ḥarf tarqīq)* precedes
(*'Idrāk* 139,9). These terms stand for all types of velarity and palatality,
regardless of whether the consonant is velar, such as **q** and **ġ**, or ve-
larised, such as **ḷ** (105,6). The form of the infinitive suffix varies accord-
ing to this principle too:

---

[62] In syntax *ʿamila fī* is the usual term to express governance. See further discussion
in Chapters Four, Five and Six.

> "The marker [for the verbal noun] is the attachment of **maq** to the im-
> perative form of the verb if the preceding consonant is velarised, and **mak**
> if the preceding consonant is palatalised." (ʿalāmatuhu ʾan yalḥaqa ʾāḫir
> fī ʾl al-ʾamr **maq** ʾin kāna qablahu mufaḫḫam, wa-**mak** ʾin kāna qablahu
> muraqqaq, ʾIdrāk 108,13.)

These statements agree with the distributional rules for the suffix that
indicates the noun of place (ğak or ğaq) as explained in Qawānīn:

> "The marker of the noun of place consists of adding the word ğak or ğaq
> to the stem of [a word]. The [word with] q is attached to words that con-
> tains velarised consonants, and the [form with] k to other [words] than
> these." (ism al-makān ʿalāmatuhu ʾan tuzīd ʿalā l-ʾuss[63] lafẓa ğak ʾaw
> ğaq fa-l-qāf li-mā fīhi al-ḥurūf al-mufaḫḫama wa-l-kāf li-mā siwā ḏā-
> lika, Qawānīn 25,11.)

In Qawānīn, the category of velarised consonants (ḥurūf mufaḫḫama)
is divided into three groups, viz., "a consonant of elevation, a velarised
consonant, or a consonant that is flavoured" (ḥarf istiʿlāʾ ʾaw ḥarf mu-
faḫḫam ʾaw mašmūm,[64] Qawānīn 22,13). The latter term refers to ẓ,
the velarised z (see discussion in 3.1.2.2). The verbs that contains any
of these consonants get a velar suffix, all other verbs take a palatal one.

The respective authors of Qawānīn and the Margin Grammar refer
to the distribution of these suffixes with the term "position" (mawḍiʿ),
viz., "each of them on its position" (kull fī mawḍiʿihā, 21,18; 41,21;
42,14; 50,19; also MG 61ᵛbm; 59ᵛrt; 58ᵛlt). Palatal suffixes, for exam-
ple, occur in the "position of palatalisation" (mawḍiʿ at-tarqīq, MG
45ᵛlt).

3.3.2 Dīwān

Kāšġarī's principles for the attachment of the velar and palatal variants
of suffixes are based on similar principles as those mentioned in the
previous section, viz.,

> "The verbal noun is [construed] in one way in all verbs. To the [last] con-
> sonant of the [verbal] stem an **m**, an ʾalif and a **q** [i.e. **māq**] are added in
> the case of words that contain q or ġ, or which have ʾišbāʿ in pronuncia-
> tion. The **k** [is used] instead of q in words [i.e. **māk**] which contain k or
> which are palatal or which have ʾimāla to the three vowels." (wa-l-maṣ-
> dar yakūn ʿalā minwāl wāḥid fī ğamīʿ al-ʾafʿāl wa-huwa ʾan yuzād ʿalā

---

[63] The expression ʾuss is quite unusual; it must be understood as 'stem', unlike the
Arabic term ʾaṣl which stands for the abstract notion 'underlying form'.

[64] Note that mašmūm is the participle passive of šamma 'he flavoured', the first form
of the verb, whereas ʾišmām 'make flavour' represents the fourth form, a causative.

*ḥarf al-ʾaṣl al-mīm wa-l-ʾalif wa-l-qāf fī l-kalima allatī fīhā qāf ʾaw ġayn
ʾaw kānat muṣbaʿa fī l-lafẓ. wa-l-kāf makān al-qāf fī l-kalima allatī fīhā
kāf ʾaw kānat rakīka ʾaw mumāla ʾilā l-ḥarakāt aṭ-ṭalāṭa, Dīwān*
284,10ff.)

The most basic note Kāšġarī makes is that the velarised form of the
suffix is used when the word contains a **q** (and/or **ġ**), and the
palatalised form when it has a **k**. By extension, velarised and palatalised
words are called *al-kalima al-qāfiyya* (e.g., 482,6)—also occasionally
al-kalima al-ġayniyya—and *al-kāfiyya* (303,17), respectively.[65]

These terms are often used to indicate the opposition between velar
and palatal. A good example is the pausal suffix GIl (**ġil~qil/kil**) which
is added to the imperative of the second person singular, viz.,

> "[Most Turks] use... the **ġ** with the **l** [i.e. **ġil**] in the *qāf*-word or the ve-
> larised word, and the **q** with the **l** [i.e. **qil**] in the word whose last conso-
> nant is a **ġ**; and [they use] the **k** with the **l** [i.e. **kil**] in the *kāf*-word or the
> palatalised or the light word." (*ʾakṭar at-turk tadkuru... al-ġayn wa-l-
> lām fī l-kalima al-qāfiyya ʾaw al-muṣbaʿa wa-l-qāf wa-l-lām fī l-kalima
> allatī fī ʾāḫirihā ġayn; wa-l-kāf maʿa l-lām fī l-kalima al-kāfiyya ʾaw al-
> mumāla ʾaw ar-rakīka, Dīwān* 288,2ff.)

The expressions *al-kalima al-qāfiyya* and *al-kalima al-kāfiyya* are most
likely related to the infinitive suffixes that are attached to the verbal
stem; back words get the infinitive suffix **māq**, front words **māk**. In
Arabic poetry the poem, especially the *qaṣīda*, is often named after the
consonant that causes end rhyme, called *ar-rawī*. A poem ending in *l*,
then, is called *lāmiyya*, or *tāʾiyya* when it ends in *t*. Kāšġarī names the
back and front verbs after the last, i.e. rhyming, consonant of the in-
finitive suffix that is attached, rather than the stem itself (cf. 3.1.5).

Still, the distribution of the suffixes is, to my opinion, not satisfac-
torily accounted for in *Dīwān*. According to Kāšġarī, the suffix **ġil** is
attached to velarised verbs in general, and **qil** to words that end in **ġ**.
This last statement does not agree with the principles of consonant as-
similation, according to which the voiced final -**ġ** should be followed
by **ġil** (see 3.6.3).

3.3.3 Targumān

Much like Kāšġarī, *Targumān* bases the distribution of verbal suffixes
on his own extensive list of verbal stems (33,8 - 44,10). In this list, he
assigns each verbal stem one out of three forms of the pausality suffix

---

[65] The terms *kāfiyya, rakīka* and *mumāla* recur in MG (40ʳrt) which is a literal copy
of *Dīwān* (538,5).

GII (i.e. ǧil, qil and ḳil). The form of other verbal suffixes also depends on the form of the suffix that is attached to the stem in the list. The infinitive suffix **maǧ**, for example, occurs with verbal stems that are assigned ǧil:

> "You consider any imperative and if the [assigned] pausal form is like ǧil, you omit the word ǧil, then you attach the form **maǧ** to what precedes it [i.e. the stem], and thus [the combination of the stem and suffix] becomes a verbal noun." (*fa-ʾinnaka taʿtabir ʾayy ʾamr šiʾta fa-ʾin kānat lafẓa al-istirāḥa min ḏawāt ǧil fa-tusqiṭu lafẓa ǧil wa-tulḥiqu bi-mā taqaddamahā min al-ʾamr lafẓa maǧ fa-yakūn maṣdaran, Tarǧumān* 45,2ff; similar statement on 49,1.)[66]

Similarly, if the pausal suffix is qil, the infinitive is **maq**, and if it is **ḳil**, **mak** is attached. Needless to say that **maǧ** is, in fact, an allophonical variant of **maq**, and it is not likely to be conditioned in the same way as GII. In fact, -ǧ is very unlikely to be conditioned at all, but rather an allophone of **q** which is used throughout by certain speakers of Turkic. Indeed, the attachment of all verbal suffixes is described in terms of the data in the same list. In this way, the choice of a given suffix is not to be inferred from the features of the verbal stem itself, but depends on a prescription by the author of *Tarǧumān*.

3.3.4 Ḥilya

In *Ḥilya*, Ibn al-Muhannā links the distribution of the velar and palatal suffixes to the first consonant of the word. With regard to the ending of the first person plural in -**diq** he states:

> "You add a **q** to the base of the verb [i.e. stem+di], and vocalise the consonant that precedes it [i.e. d/t, marker of the past tense] with **i**, if the first consonant of the verb is **q** [i.e. diq/tiq], or else **k** if the first consonant of the verb is **k** [i.e. dik/tik]." (*ʾannaka tuzīd baʿda ʾaṣl al-fiʿl qāfan maksūran mā qablahā ʾin kāna ʾawwal ḥurūf al-fiʿl qāfan, ʾaw kāfan ʾin kāna ʾawwal ḥurūf al-fiʿl kāfan, Ḥilya* 81,9f.)

Ibn al-Muhannā attaches a number of examples to this statement, one of which, as a matter of fact, has neither **q**- nor **k**- in its stem, e.g.,

---

[66] A similar reference to the pausality suffix is also found in MG 36ʳlt/bm, and on MG 35ʳlt with regard to the attachment of **qān**, **ǧān** and **kān** for the active participle. These fragments no doubt derive from another copy of *Tarǧumān*. I have not been able to verify whether the word list of the Margin Grammar (2ʳ-32ᵛ) is set up in the same way, because it is extremely difficult to read.

qāz-diq 'we dug'; 'īn-diq[67] 'we descended'; kas-tik 'we cut'. Some of these verbs share the same back marker for the second person plural, i.e., -ngiz (in which ng is to be read as a velarised ŋ) e.g., qāz-di-ng-iz [qazdïŋïz] 'you dug', and, surprisingly, kas-ti-ng-iz [kestiŋiz] 'you cut'.

In another instance (101,9) he uses the same criterion, i.e. q or ġ for words or verbs whose first consonant is a q, and k for words or verbs beginning with k. In his listings, however, he gives many instances of words that do not match this criterion. One is inclined to conclude that this should not be taken too literally, in the sense that it also in-cludes words which have a k or q at another place, or, perhaps, share other features with q and k. It appears that the mention of q and k is to be interpreted more generally as reference to front and back words, because a word containing q can only be back, and a word with k is per se front, much like the reference to q and k is used in *Dīwān* (see 3.3.2).

In general, though, Ibn al-Muhannā does not regard the distribution of the respective suffixes as a coherent phenomenon that is subject to specific rules. For example, the suffix for the active participle *(ism al-fāʿil)*, Ibn al-Muhannā writes, is ġān which can be substituted by k͓ān,[68] because of the closeness of their place of articulation (*li-taqārub maḥraǧayhimā*, *Ḥilya* 116,17ff; similar statement on 127,3). The closeness of the place of articulation regards ġ vs. k͓ rather than either of them on the one hand, and the final consonant of the verbal stem on the other, as one would expect. Ibn al-Muhannā himself too seems to have been puzzled by the distribution of the suffixes. He quotes one of his sources who had asked the same question of his informants:

> "The compiler of *Tuḥfa al-Malik* said: 'I asked the Turkic scholars why they in some instances use a ġ and an l [i.e. ġil] after the imperative form, and k͓ and l [i.e. k͓il] in others, e.g., 'al-ġil 'take' bīr k͓il 'give'... and [why] they do not use a k͓ on the place of the ġ nor a ġ on the place of a k͓... but they did not answer." (*qāla ṣāḥib kitāb tuḥfa al-malik saʾaltu ʿulamāʾ at-turk lima ḥaṣṣū mawḍiʿan min al-ʾamr bi-l-ġayn wa-l-lām fī ʾāḫirihi wa-mawḍiʿan bi-l-kāf wa-l-lām miṯl ḥud ʾal-ġil ʾaʿṭi bīr k͓il... wa-lā yastaʿmilūna mawḍiʿ al-ġayn kāfan wa-lā mawḍiʿ al-kāf ġaynan... fa-lam yuǧībū bi-šayʾ, *Ḥilya* 193,3ff.)

---

[67] In both instances q is marked with a small *k* on top of it, probably indicating a voiced g. According to Ibn al-Muhannā's description of this verb, it can only be inter-preted as 'back'. In modern languages, however, the verb 'in- 'to descend' is usually 'front'.

[68] Rifʿat emends *mutaqqala* 'heavy' to *mutaffā*.

The apparent misunderstanding and mistaken application of the prin-
ciples of the distribution of the suffixes, raises the question whether
Ibn al-Muhannā indeed was a native speaker of Turkic.

3.3.5 Summary

The conclusion of this exposé is that all sources except *Ḥilya* regard the
distribution of the twofold forms of a suffix as a phenomenon which is
subject to certain rules. In most cases these rules are related to the
principles of velarisation and palatalisation, and in this sense one could
say that they are conditioned by palatality or velarity of the stem. In
two instances, *Tarǧumān* and *Dīwān*, the reader is referred to intuitive,
or non-argumented knowledge of the author which, however, results in
a similar distribution. The sources, as a rule, describe an incomplete
distribution of three suffixes rather than four, and it seems that their
interest concerned in the first place with describing instances of a con-
sonantal change. More about this will be said below in section 3.6.3

3.4 *Phonetic descriptions of rounded vowels*

In the preceding sections it has been shown that the grammarians had
instruments to describe velarisation and palatalisation. They use these
tools for determining which type of suffix a given stem takes. With the
same instruments the can describe six of the eight Turkic vowels.
Nevertheless, they scarcely dedicate any words to the description of
rounded vowels. Therefore, in front words it is often difficult to decide
whether the grapheme **u** corresponds to [ü] or [ö], whereas in back
words the choice between [u] and [o] is equally complex. However, the
Arab grammarians did perceive these distinctions and occasionally in-
dicated them in orthography.

3.4.1 [o]

In some instances the velar consonants alone apparently are not suffi-
cient enough to indicate the sound of the vowel. In those cases the la-
bel 'velarised' is applied. With words containing **u**, the label 'velarised',
then, indicates a pronunciation as [o], viz.,

ṭuy-dī [toy] 'he was saturated' (*'Idrāk* 67; marker above the word;
MS 21ʳ24), "with ṭ inclining to **u**, with velarisation" (*mumāla aṭ-ṭāʾ
ʾilā ḍamm maʿa t-tafḫīm*, MG 42ʳlt; cf. *mumāla* in *Dīwān* 3.1.5.3 and
ṭuysu palatalisation marker in 3.1.6.1)

qūz [qoz] 'walnut' (with marker above q) (*'Idrāk* 71)

ʾūṭ /ʾuwṭ/ [ot] 'fire' (*ʾIdrāk* 15; MS 4ʳ14, marker up between **w** and
ṭ).
The same can especially be illustrated with words containing **l** and **r**
from all relevant sources. For example, the grapheme **tūl** /tuwl/ may
have different phonetic realisations, depending on the label. When la-
belled for velarisation, an interpretation as [tol] 'widow' is possible, al-
though not attested in modern Turkic languages (*Dīwān* 501,11;
Dankoff and Kelly 1984:216 "**tūl**"). When marked for palatalisation
(here *ʾimāla*) the pronunciation must be [töl] 'litter' (*Dīwān* 501,13;
Dankoff and Kelly 1982:58/1984:216).

Marked for velarisation are, e.g.,
ʾul [ol] 'he' (*ʾIdrāk* 20; MS 5ᵛ3, marker above **l**)
būl [bol] 'much' (*ʾIdrāk* 36; MS 10ᵛ15; marker above /w/)
yul [yol] 'road' (*ʾIdrāk* 97; MS 31ᵛ1, marker above the word).
The following minimal pairs of [u - o] and [ü - ö] occur:

Unlabelled	Velarised
qul [qul] 'slave' (*ʾIdrāk* 73)	[qol] 'arm' (*ʾIdrāk* 73; *Tarǧ.* 19)[69]
turum [turum] 'act of' standing (*Dīwān* 200,1).	[torum] 'a camel colt' (200,2)

Unlabelled	Unlabelled
ʾulu- [ulu] 'to howl' (said of dogs) (*ʾIdrāk* 20)	[olu]? 'to grow' (*ʾIdrāk* 20; MS 5ᵛ7 (but see related word **ʾulū** [ulu] 'great' (MS 5ᵛ7; marker above **l**).

Velarised	Palatalised
ʾūr ʾat [or] '(a horse) be- tween sorrel and bay' (34,12)	[ör] 'the waist' (of a cloak) (34,13)

Unlabelled	Palatalised
tur [tor] 'snare or net to hunt birds or fish' (*Dīwān* 495,13).	[tör] 'front side of the house' (495,12)

3.4.2 [ö]

The label *mufaḫḫam* is sometimes applied in instances in which a back
pronunciation cannot be intended, because the marked consonant or

---

[69] In *ʾIdrāk*, the label 'velarised' is also applied to a ʾaq [aq] 'white' (*al-ʾabyaḍ, ʾIdrāk*
17; MS 4ᵛ17; with marker above q; EDT 75 "aːk"), which is difficult to account for.

other consonants in the word can only be palatal, especially **k**. There are two explicit examples of this in *ʾIdrāk*, viz.,

   tušak [töšek] 'carpet' (velarised, marker above **t**; *ʾIdrāk* 39; MS 11ᵛ18)
   kun [kön] 'hide' (*ʾIdrāk* İÜ 39ᵛ8)
   ʾulkan [öl-gen] 'name of the absent', i.e. 'deceased' *(ism al-ġāʾib, Ḥilya* 74,13ff).

Other examples of this type must also be understood in this sense, although the marker *mufaḫḫam* is applied to a neutral consonant (but see 3.4.3 for a discussion of **z**), viz.,

   dūz- [dö:z] 'to endure' (*ʾIdrāk* 48; MS 14ᵛ19; marker above /w/)
   suz [söz] 'speech' (*ʾIdrāk* 52; MS 16ʳ21; marker above word)
   suz-la- [sözle] 'to speak' (*ʾIdrāk* 52; MS 16ʳ21; maker above /s/)
   ʾup [öp] 'to kiss' (*ʾIdrāk* 7; MS 2ʳ19, velarised and voiceless).

In view of evidence from modern Turkic languages, in these instances the label is unlikely to refer to back vowels, and [ö] then is the only reasonable alternative.

Because the default pronunciation of **u** preceding **l** is [u], rather than [ü], the word must be labelled if a front pronunciation is intended, viz.,

   ʾul- [öl] 'to die' (*ʾIdrāk* 20; MS 5ᵛ7, marker above /l/)
   ʾula- [üle] 'to divide' (*ʾIdrāk* 20; MS 5ᵛ1, markers above /l/ and /d/).

The label 'palatalised' *(muraqqaq)* is applied in two instances to words containing **u** that have a default front pronunciation. First, **suz-** 'to filter' which is simply pronounced as [süz] (*ʾIdrāk* 52). Here, the marker may be put because of the default back pronunciation of **u** and **a** when they occur close to **z** (see discussion in section 3.4.2). For the second word, **ʾub**, there are indications from modern Turkic languages (for which cf. EDT 4) that it must be read [ö], viz., [öp] 'swallow, gulp down' (*ʾIdrāk* 7; MS 2ʳ19, badly legible marker for palatality above *hamza*).

3.4.3 Excursus: confusing labels—the case of **z**

The sources sometimes assign contradictory labels. In one source a given grapheme word is marked palatalised or velarised, whereas it has no, or an contradictory label, in another. In Section 3.1.6 I show that **a** and **u** have a default back pronunciation when they occur close to **l** and **r**. There are some indications that the same applies to some extent to words containing **z** (see also [süz-], discussed in 3.4.3.) A case in point would be the grapheme بوز (**būz** /buwz/), which in *ʾIdrāk* to ex-

presses three different meanings, i.e. 'ice', 'grey' and 'cloth', depending on a front or back pronunciation, and the quality of the vowel.[70]

1 'Ice'
In *Idrāk* **būz** (with palatalised **b**; MS 9ʳ2) means 'ice' (*Idrāk* 30), and according to the label, the pronunciation would have to be something like [büz]. However, [buz] is a very common word in several Turkic languages, and never found pronounced with front vowels (EDT 389 too, could not interpret the label).

   Here, we might interpret the default interpretation of u before z as [o] or, in any case, as [u], and ʾAbū Ḥayyān may have used the label 'palatalised' to express the opposition with other instances in which he intends a different phonetical realisation (discusssed below). In *Dīwān*, Kāšġarī labels **būz** 'ice' as "with velarisation"(*bi-ʾišbāʿ*, *Dīwān* 496,8), which would points to a default front pronunciation.

2 'Grey'
Another meaning of the same grapheme is 'grey' (said of a horse) (*Dīwān* 496,5). In *Dīwān* the word remains unmarked for this meaning, whereas in *Idrāk* it is labeled velarised (*Idrāk* 30; MS 9ʳ2). A third opinion complicates the matter even further; in *Qawānīn* both words have merged in one single, unmarked pronunciation, i.e., "'cold', and 'white' for cattle, mules and donkeys" (*Qawānīn*, 58,5). Based on ʾAbū Ḥayyān's system of labelling I would suggest here [boz], which is confirmed by data from modern Turkic languages.

3 'Cotton cloth'
There is a third word which is written with the same grapheme, i.e. **būz** 'cloth' [böz]. In *Dīwān* it is marked for palatalisation (*Dīwān* 496,7), while it does not occur in *Idrāk* and *Qawānīn*.

3.4.3.1 Summary of z
Ideally, of course, the labels in all sources should indicate the same or a similar phonetic interpretation for one word. I admit, however, that this point of view entails the danger of assigning a known or desired phonetic interpretation to a language in which the pronunciation of the word in question is doubtful. Nevertheless, it seems that the in *Idrāk* the labels are used merely for expressing the opposition between the two words. In *Dīwān* on the other hand they indicate a default

---

[70] Reference to the meaning 'to destroy' is of no help in this discussion since it is realised as [boz-] in Oġuz languages, and as [buz-] in others (EDT 388-9).

back interpretation. The confusion evolving from these data does not permit drawing more precise conclusions as to the application of the labels applied to this grapheme. There are supporting (but also some contradicting) examples in *Dīwān* (discussed in Dankoff and Kelly 1982:56). All these words recur on different places in *Ḥilya* where they remain unlabelled (184,5, 168,15, 167,9).

The following scheme summarises the various opinions:

/buwz/		*Dīwān*	*ʾIdrāk*	*Qawānīn*	*Ḥilya*	Intended phonetic value
1 'ice'		'velarised'	'palatalised'	unlabelled	unlabelled	[buz]
2 'grey'		unlabelled	'velarised'	unlabelled	unlabelled	[boz]
3 'cloth'		'palatalised'	—	—	unlabelled	[böz]

3.4.4 Ibn al-Muhannā on rounded vowels

Ibn al-Muhannā is the only source to compare and discuss in detail five different pronunciations of one grapheme, اُوت (ʾūt /ʾuwt/) which he seems to have chosen as a specific illustrative example of several possible meanings depending on the pronunciations.[71] In this section I give Ibn al-Muhannā's descriptions which are followed by a phonetic interpretation and argumentation. As a point of departure I suggest in-terpreting the grapheme as neutral or back, rather than front.

1 'to win in a game'

> "If you pronounce it as it is, its meaning is the imperative for the singular of winning in a game, like you say 'beat him'." *(ʾin talaffaẓta bihā ʿalā ʾiṭlāqihā kāna maʿnāhā al-ʾamr li-l-wāḥid al-ḥāḍir bi-l-ġalab fī l-qumār kamā taqūlu iqmarhu.)*

This description speaks for itself: the vowel, reflected by /uw/, is to be pronounced 'as it is', or, in other words, as it would be in Arabic. This implies a lengthened, neutral vowel, hence [u:t] (EDT 38 "ut-").

2 'grass'

> "Without clear pronunciation of the w, its meaning is the singular of 'grass' or 'medicine' or *qayd*." *(wa-ʾin lam tatalaffaẓ bi-l-wāw talaffuẓan bayyinan kāna maʿnāhā mufrada al-ḥašīš ʾaw ad-dawāʾ ʾaw al-qayd.)*

---

[71] He further refers to one of his sources, *Kitāb nādir ad-dahr ʿalā luġa mulk al-ʿaṣr*, in which the author discusses all instances of homophonous words (*wa-qad ḏakara ṣāḥib kitāb nādir ad-dahr ʿalā luġa mulk al-ʿaṣr šayʾan tāmman min iḫtilāf al-maʿānī wa-ttifāq al-lafẓ, Ḥilya* 78,13).

Here the pronunciation of /ʾuw/ is not the regular one; /w/ must not
be pronounced very clearly, which most likely implies that it does not
indicate lengthening. Perhaps /w/ merely serves as an orthographical
device to indicate velarisation, to which Ibn al-Muhannā nevertheless
assigns a phonetic value. Such would not be a unique case, since in
Arabic theory *ʾalif* too is associated with a phonetic value, even though
it is a mere orthographic device for indicating lengthening (cf. 1.1
above). The result, then, would be an interpretation as [ot] (cf. EDT
34).

3 'hole'

> "If the *hamza* is joined to the w with a kind of velarisation, then its
> meaning is 'the piercing of an ear' or 'the eye of a needle' or the like." *(wa-*
> *in wuṣilat al-hamza maʿa l-wāw bi-nawʿin min tafḫīm kāna maʿnāhā*
> *ṭaqb al-ʾuḏun ʾaw ṭaqb al-ʾibra wa-šibhuhu.)*

In this case, it seems, /w/ is pronounced long, since the hamza is
joined to it, while being velarised *(tafḫīm)* at the same time. However,
the velarisation is not regular but of 'a certain kind' *(nawʿ min)*, or
perhaps it is just 'a bit of velarisation'. Therefore, I propose to interpret
this as [ö:t].

There are two arguments against the reading [ü:t] as proposed in
EDT (36), and an additional one that supports the interpretation I sug-
gest here. Firstly, the use of the word *tafḫīm* excludes a full palatalisa-
tion (although, admittedly, it may stand for a general 'emphasis'). The
second argument is that if it were [ü:t], the description could have
been much simpler. The argument in favor of my suggestion is that for
at least one language [öt] is attested for this meaning (cf. EDT 36).

4 'to burn the skin of an animal'

> "If you do not pronounce w and join the *hamza* to the t and slightly
> lengthen [the u ?], then its meaning is the imperative for the second per-
> son singular 'to burn animals'." *(wa-ʾin lam tatalaffaẓ bi-l-wāw wa-*
> *wuṣilat al-hamza bi-t-tāʾ bi-madd ḫafīf kāna maʿnāhā al-ʾamr li-l-*
> *wāḥid al-ḥāḍir bi-ʾiḥrāq al-ġanam.)*

In other words, the /w/ does not stand for a long vowel, for the vowel
between *hamza* and t is only somewhat lengthened *(madd ḫafīf).*[72]

---

[72] In Arabic phonology the term *waṣl* is associated with the *hamza al-waṣl*, the
prothetic *hamza* in, for example, the definite article /ʾal/. The *hamza* and the vowel *a*
disappear, or merge, when /ʾal/ is preceded by a word that ends in a vowel: /ʾal-kitaⁿbu/ -
wa-l-kitābu. When /ʾal/ is not preceded by a vowel, the *hamza* remains.

Here, it seems, /w/ again indicates velarisation, and I would interpret this as referring to a back vowel, probably [ot] (or [o:t]).

In spite of this description, the original verb is most likely [üt], although it is often interpreted as [ot] because of a false association with [o:t] 'fire'. This false association is, it seems, also made by Ibn al-Muhannā (cf. EDT 40).

5 'gall'

> "If the u of the *hamza* gets a medium inclination towards i, then its meaning is 'gall'." (*'in 'amalta ḍamma al-hamza 'ilā l-kasra 'imāla mutawaṣṣita kāna ma 'nāhā al-marāra.*)

Assuming that a 'full inclination' would result in [ü:t], 'a medium inclination' is probably best reflected with [ö]. Furthermore, Ibn al-Muhannā does not mention lengthening, so I tend to interpret this as [öt]. (EDT 35; see for *'imāla* applied to *u* above section 1.4.2).

In summary, the interpretations which I have suggested in the previous comments are the following: 1 'to win' [u:t], 2 'grass' [ot], 3 'hole' [ö:t], 4 'to burn' [ot], 5 'gall' [öt]. Thus Ibn al-Muhannā's descriptions show that it is possible to describe quite detailedly the phonetic qualities of vowels with the use of Arabic phonetic terminology.[73] (In *Dīwān* too rounded vowels are described, for which I refer to Dankoff and Kelly 1982:56).

3.4.5 Summary

This subsection has dealt with two issues. The first is the description of rounded vowels. In this respect. The second issue is the question whether the rules for l and r also apply to z; I think the supporting evidence is not entirely convincing as yet.

In regard to the first point, I would like to make the following comments. In the first place it is important to note that the Arabic authors indeed perceived—to a certain extent—differences between the rounded vowels. Second, it would have been strange indeed if they had had same concept of rounded vowels as ours. Considering their concepts of vowels and consonants, they are not likely to have thought in terms of oppositional pairs, and high versus low vowels. Instead, they

---

[73] The various meanings of this grapheme are also discussed by Nawā'ī in *Muḥākama al-Luġatayni* (cf. Devereux 1964:283). From Devereux's translation, however, it is difficult to get a clear view regarding Nawā'ī's terminology used to describe [ü] and [ö].

applied their concepts of terms of velarisation and palatalisation of consonants in these cases too. The data indicate that they posit the rounded vowels on a scale of velarisation and palatalisation, which runs from [o] to [ü].[74] The appropriate vowel is produced when a certain measure of palatalisation or velarisation is assigned to the consonant.[75] The continuum, in its ideal shape, can be envisualised as follows:

consonant	o	u	ö	ü
palatal/neutral	—	velarisation	velarisation (?/palatal)	no label (/palatalisation)
velarised	velarisation	no label	—	—

For example, velarisation of a velarised consonant results in [o], whereas a similar measure of velarisation to a palatal consonant yields [ö]. Conversely, palatalisation of a palatal consonant gives [ü]. Of course, the scheme does not account for labels that merely express an opposition.

3.5 The status of the glides

The status of the glides, i.e. ʾalif /"/, y and w in the sources is not always easy to interpret. The glides are pronounced in various ways, and in some cases not at all. In a number of sources any reference lacks and one has guess as to whether the glide indicates vowel length or stands for plene writing only. When they are added for lengthening (ʾišbāʿ) of a vowel, such as in poetry, the glides are not considered a basic part of the word, and therefore, as a rule not, or only occasionally expressed in orthography (see 1.5.1). This section does investigate vowel length in Turkic languages of which there are traces in modern Turkmen and Yakut; it merely discusses in a broad way the functions the sources assign to the glides in their descriptions of Turkic.

As said before, the sources indeed display various opinions on the status of the glides. In some (Ḥilya, Dīwān, Tuḥfa and—to some ex-tent—ʾIdrāk) principles of the Uygur alphabet shine through. In the Uygur alphabet there were no signs for short vowels, long vowel signs being used both for indicating long vowels and as plene writing of short vowels.

---

[74] It is important, though, not to be misguided by the labelling of oppositional pairs.

[75] The idea of a scale is enhanced by data from Kāšġarī who remarks that a given word (which is known as having [ö]) requires 'less šamma' than another word which has [ü] (Dankoff and Kelly 1982:56).

3.5.1 'Idrāk, Ḥilya, Qawānīn and MG: glides indicate prosodic lengthening

In 'Idrāk 'Abū Ḥayyān considers all occurrences in which the glides are used to indicate a lengthened vowel, prosodict lengthening, i.e. 'išbā ͨ, rather than a part of the root:

> "None of the three weak consonants [sc. 'alif, wāw and yā'] are ever part of the root in this language, they arise from the lengthening of the vow-els." (ǧamī ͨ ḥurūf al-madd wa-l-līn aṭ-ṭalāṭa lā yakūnu šay' minhā 'aṣlan fī hāḏihi l-luġa 'innamā hiya nawāši' ͨan 'išbā ͨ al-ḥarakāt, 'Idrāk 101,10; compare Irtišāf I 423)[76]

In other words, 'Abū Ḥayyān means that unlike in Arabic, in Turkic the glides are never part of the root; instead they always serve to indi-cate secondary lengthening ('išbā ͨ). In Qawānīn and the Margin Grammar, too, the term 'išbā ͨ is used in the sense of lengthening, viz., ('išbā ͨ fatḥa [al-mīm]). The lengthening is shown by means of inser-tion of a glide, as in kalmās /kalma"s/ 'he will not come' (Qawānīn 14,2; similar statement in MG 62ᵛtop; 46ᵣlt/ult). A similar opinion is reflected in Ḥilya:

> "An 'alif that evolves from the lengthening of a..." (wa-'alif tanša'u ͨan 'išbā ͨ fatḥa, Ḥilya 75.)

Thus, it appears that most grammarians accept both long and short vowels for Turkic, long vowels occurring basically as a secondary, per-haps marginal effect, i.e. similar to the status of the lengthening of vowels in Arabic poetry.

It is interesting to note that, although in Arabic linguistics the term 'lengthening' ('išbā ͨ) it is sometimes refers to the consonant instead. For example, in the word 'aqǧā 'white' (used as the name of a coin) ǧ is both vocalised with a (maftūḥa) and lengthened (mušba ͨa, M G 34ᵛlt): "the noun whose last consonant is vocalised with a and length-ened" (al-ism al-maftūḥ al-'āḫir al-mušba ͨ, MG 59ᵛrt). Similarily, the m (in the negation suffix) is "lengthened with an 'alif" (al-mīm al-mušba ͨa bi-l-'alif, MG 62ᵛtop). The last quote calls to mind Qawānīn's description of the same suffix, but there the author uses the term mu-faḫḫam (mīm maftūḥa mufaḫḫama, Qawānīn 13,9), which is to be in-terpreted as 'lengthened', or—perhaps, 'emphasised'—rather than 'vel-arised' (see also 3.1.2).

---

[76] It will be shown in Chapters Five and Six that this point of view has consequences for the interpretation of some Turkic morphemes that are expressed by means of a glide.

3.5.2 *Tuḥfa*: plene spelling as default

Traces from the Uygur orthographical tradition are evident in *Tuḥfa* too. The anonymous author formulates his concept of the glides as fol-lows:

> "Know that there are no signs for short vowels and by the lack of it much confusion occurs. Thus, they set up a rule for all nouns, verbs and parti-cles alike in order to write without vocalisation and without confusion. [This rule] is that a is followed by *ʾalif*, i by y, and u by w." (*iʿlam ʾanna l-luġa t-turkiyya laysa fīhā šakl wa-bi-ʿadamihi yaqaʿu l-ʾiltibās fa-ġaʿalū li-kull min al-ʾasmāʾ wa-l-ʾafʿāl wa-l-ḥurūf ḍābiṭan li-yaktubū [?] bi-ġayr šakl wa-lā iltibās, wa-hiya [sic] ʾanna l-fatḥa talīhā ʾalif, wa-l-kasra yāʾ wa-ḍ-ḍamma wāw, Tuḥfa 2ᵛ9.*)

Indeed, in *Tuḥfa* all vowels are indicated by means of *ʾalif*, w and y, re-spectively, while his plene orthography is observed in cases where the other sources have a short vowel, e.g., ṣāqāl /ṣa"qa"l/ 'beard', qānāṭ /qa"na"ṭ/ 'wing'. Initial *hamza*, for example, is followed by *ʾalif* to in-dicate a, e.g., ʾāṭā /ʾa"ṭa"/ 'father' *(ʾab)*; ʾānā /ʾa"na"/ 'mother'; ʾāq /ʾa"q/ 'white' (*Tuḥfa* 3ʳ11ff). This plene writing may be related to the Uygur script.

Nevertheless, in a few instances a vowel is represented by a vowel sign only, e.g., in ʾawrāt /ʾawra"t/ 'woman', ṣuwūq /ṣuwuwq/ 'cold' (6ᵛ10). This is probably related to the fact that otherwise these words would have contained a sequence of two *ʾalifs* and a /w/, i.e. *أارات /ʾa"wra"t/, or one of three /w/s, i.e., *صوووق /ṣuwwuwq/, respectively, which are unacceptable in Arabic phonology. [77]

3.5.3 The use of glides in *Dīwān*

Kāšġarī's principles with regard to the use of the glides are described in detail in Dankoff and Kelly (1982:67ff) (and Kelly [1973]). They are right when they say that Kāšġarī at times is inconsistent in his use of a plene orthography.

According to them, the instances in which Kāšġarī uses plene or-thography are summarised in four categories which I repeat here for sake of convenience: (1) anomalous plene orthography in the intro-ductory pages, probably due to errors during copying; (2) inconsistent use of the glides for the vowel sign as in Uygur script; (3) secondary lengthening, or 'pausal lengthening' in proverbs and verses, and in the rhyme position; (4) secondary lengthening which is to be interpreted as

---

[77] This plene writing of all vowels in *Tuḥfa* is an important difference with *ʾIdrāk*, in spite of the structural resemblances of the two works.

indicating stress, especially occurring in two-word phrases, "where it points to stress on the root as against the general rule." Kelly (1973) gives a detailed account of the instances in which Kāšġarī makes statements about vowel length and how these should be interpreted. One of the conclusions Kelly reaches is that Kāšġarī understood that it is possible to express vowel length more accurately in Arabic script, as opposed to Uygur script, in which this is not possible. In other instances the glide indicates stress in the first syllable. At the same time, though, Kāšġarī's attitude with regard to the status of the glides could be called ambivalent. In some instances he accepts the glides as part of the root, in others he considers them additional and suggests deletion because a short pronunciation is "the most eloquent" (al-'afṣaḥ).

In Arabic grammar deletion of a glide as a basic consonant is only possible under certain conditions (cf. Bohas 1982:91ff). On 515,16-18 (also in Kelly 1973:156) Kāšġarī says that it is possible in Turkic to delete the glides just as this is possible in Arabic. He exemplifies this with the verb zāna 'he decorated' yazīnu 'he decorates', "then one forms the imperative by saying zin 'decorate!'…" (ṯumma yu'mar minhu fa-yuqāl zin…, Dīwān 516,6). This argumentation only makes sense if we regard yazīnu as the phonetic realisation of /yaziynu/. In the case of the imperative form the medium /y/, which is a part of the root /z-y-n/, is omitted, leaving /i/, hence /zin/. In other words, in Arabic theory the shift of yazīnu—zin is not the shortening of a long vowel, since in this theory long vowels do not exist, but the elision of a consonant. In the same way, the pattern of qūl 'arm' is /quwl/; after deletion of the glide /w/, the pattern becomes /qul/, realised phonetically with a short vowel. Depending on which form is regarded as primary, /w/ is either deleted from the root, or inserted.

Kāšġarī makes a yet another statement in regard to words whose middle consonant is a glide. In these words, he writes,

> "the glides disappear when they are pronounced fast… [T]he glide is elided in speech, not in writing." (wa-ḥurūf al-līn yasquṭu minhā ʿinda surʿat an-nuṭq bihā… fa-…yunqaṣu minhā ḥarf al-līn fī n-nuṭq dūna l-kitāba, Dīwān 493,13).

Here Kāšġarī says that even though a given word contains a glide this does not necessarily indicate lengthening, for the glide can be elided in speech. In other words, shortening is the result of a prosodic process which is not always indicated by means of a short vowel; thus the glides sometimes stand for plene writing of vowels rather than actual lengthening.

3.5.4 The functions of the glides in Ḥilya

Ibn al-Muhannā developed a rather complicated system of abbrevia-
tions and markers which he used to indicate aspects of the vowels that
cannot be reflected with the regular script. In this system, described on
pp. 73ff., the glides play an important role.

3.5.4.1 Labels of the glides, and the functions of the glides as markers

As discussed in section 3.1.5.2 above, in Ibn al-Muhannā's orthogra-
phy the glides serve more functions than.lengthening alone; they occa-
sionally serve as markers for velarisation and as plene writing of vowels.
When they are used in a non-regular way, Ibn al-Muhannā marks
them with *m* (ـ). For example, in some words the occurrence of /w/
stand for plene writing of a vowel, not for vowel length. Examples are,
e.g., **būrġū** /bu̲w̲ṟ̲ġuw/ 'trumpet', in which the first /w/ is marked,
hence [burġu] and **ʾūš** /ʾuw̲š̲/ [uš] 'the one that' *(alladī)*.[78]

Similarly, *ʾalif* is marked when it does not stand for vowel length,
but indicates a vowel, much like 'the pausal *h*' *(hāʾ al-ʾistirāḥa)*, Ibn al-
Muhannā writes, in Arabic words like *farsiyyah* 'Persian', in which fi-
nal *h* indicates a preceding short vowel *a*, rather than a real *h*.[79] For
Turkic this applies to **yā** /ya"/, [ya], 'bow'; **mundā** /munda"/ [munda]
'here', **ʾalmā** /ʾalma"/ [alma] 'apple', in all of which the *ʾalif* is
marked.[80]

At best, Ibn al-Muhannā's system of marking and labelling can be
characterised as quite confusing; it may well be the case that he has
mixed up the writing conventions he found in his own sources.

3.5.4.2 *ʾImāla* applied to y

In Ḥilya the terms *ʾimāla* and *ʾišmām* are applied to words whose de-
fault interpretation is already front. The labels, then, perhaps must be
interpreted as indications for a stronger inclination to [e] (or [é]) than
usual. For this type of *ʾimāla*, Ibn al-Muhannā uses the abbreviation
mk—which stands for *māla ʾilā l-kasr*, i.e. "inclining to *ī*". Examples of

---

[78] Interestingly, /w/ remains unmarked when it is pronounced 'soft' *(ḥafīfa)*, i.e. "not
extended" *(lā mušbaʿa)*, which I tend to interpret as plene writing too. This is ex-
emplified with the very same word, i.e. **burġū** /burġuw/ 'trumpet' [burġu].

[79] Another regular use in orthography is when *ʾalif* is added after ū /uw/ at the end of
a word. In Arabic this occurs for instance in verbal forms conjugated for the third person
plural past tense, e.g., *ḍarabū* (ضربوا) 'they beat'. Ibn al-Muhannā does the same in
Turkic too: **bū** 'this' (بُو) (cf. also *Durra* Ms 4ʳ ṣuw 'water /ṣ-w-"/).

[80] The *ʾalif* and w remain unmarked when they are basic consonants of the word
(ʾaṣlī). This is, for example, the case in **qām** /qa"m/ 'soothsayer'. They remain, paradox-
ically, also unmarked when they indeed do indicate vowel length *(mušbaʿa)*, e.g., **kūrār**
'he sees', **yarār** 'he matches'; **ʾūz** /ʾuwz/ 'master' [u:z] and **ʾūr** /ʾuwr/ 'beat!', [u:r].

this are, e.g., **man** [men] 'I', **'at** [et] 'meat', **bak** [beg] 'chief' (*Ḥilya* 78,3).

The same label is applied to **y**, which here indicates a long vowel (*mušbaʿa* 76,11), furthermore, **y** is "flavoured of **i**" (*mušamma ʾilā l-kasr*). At this particular point Ibn al-Muhannā does not mention the quality of the vowel that precedes **y**, but a few lines further (76,14) he refers to it as **i** (see examples in listing below).

Elsewhere (77,7) some of these words are mentioned again, albeit with yet another marker, i.e. *m-l-ḥ* (ملح), which also indicates *ʾimāla*. Here the quality of the vowel represented by **y** is compared to that of *y* in four Persian words:

1 "lengthened *y*", as in e.g., *šīr* 'milk' and *tīr* 'arrow'.

2 "a light *y* with inclination to *i*" (*yāʾ ḥafīfa bi-ʾimāla al-kasr*), as in e.g., *šīr* 'lion' and *sīr* 'garlic'.[81]

In contrast to these instances of marked **y**, Ibn al-Muhannā (*Ḥilya* 76) gives some words that remain unmarked, in which the pronunciation of **y** is 'clear' (*bayyina ẓāhira*). When combined in one scheme, the following pairs become evident, viz.,

Labelled (*ʾImāla*)	Unlabelled
ʾīš /ʾiyš/ [éːš][82] 'companion'	[iš] 'work'
bīr /biyr/ [béːr] 'give'	[bir] 'one'
kīš /kiyš/ [kéš] 'quiver'[83]	[kiš] 'sable'
bīz /biyz/ [béːz] 'swelling'[84]	[biz] 'cloth' (*al-ʾišfā*[?]), EDT "béːz")
	('we'; see also 77,2)

---

[81] These data are not entirely confirmed by Western interpretations of Persian. Steingass gives 'lion' as *sher*, the three other nouns with *ī*. Meier (1981:86-7) notes that in Arabic sources Kurdish *ē* is marked with a special marker ("haken") above *y*, and *š* with this marker above *w*. Further, "zur wiedergabe eines *ē* im sudanesischen arabisch schreibt Yusuf Faḍl Ḥasan in seiner kritischen ausgabe von Muḥammad an-Nūr b. Ḍayfallāh's *Ṭabaqāt fī ḥuṣūṣ al-ʾawliyāʾ wa-ṣ-ṣāliḥīn wa-l-ʿulamāʾ wa-š-šuʿarāʾ fī s-Sūdān*, Chartum 1974, einen bogen unter den buchstaben vor dem y." And: "In manchen arabischen dialektaufnahmen wird das ē mit den worten *bi-ʾimālat al-yāʾ* angedeutet."

[82] With no further vocalisation of *hamza*. From an earlier statement (74,3) on this subject, however, it is evident that *hamza* must be vocalised with *a*. Other sources do not point at this specific pronunciation, e.g., *ʾIdrāk* 14 ʾiš.

[83] In *Dīwān*, too, two instances of [e] and [é] are described. The term used in this respect is *bi-l-ʾimāla*, e.g., kīš 'quiver' (*Dīwān* 498,6; EDT 752 "kéːṣ"; Dankoff and Kelly "kéš"); sayš 'surety' (*Dīwān* 497,17; EDT 856 "sɪːṣ").

[84] In another instances [éː] is represented by *ʾalif*, viz., bāšik /baˈšik/ 'cradle' (*Ḥilya* 179,3; EDT 390).

3.6 *Consonants*

The present section deals with two main issues. First, it shows which phonemes each source posits for Turkic. Secondly, it discusses in which terms they describe those phonemes that are not regular phonemes in Arabic are described.

In the present section I shall not deal in great detail with velarisation and palatalisation, which are already amply discussed in the previous sections (3.1-3.5). Instead, I shall describe which consonantal phon-emes are described in the sources, and which terminological in-struments are used in those descriptions. Furthermore, I shall attempt to assess the status of the phonemes the sources posit for Turkic. It shall be seen that there are three types of Arabic descriptions of Turkic phonemes. First a number of basic phonemes that also exist as such in Arabic. In a sense these could, therefore, be called phonemes. Second, there is a number of phonemes that are described as allophones of Arabic, sometimes for dialectal variants. The third type are phonemes that are entirely new.

In regard to the front-back opposition, there is an important basic difference in approach between the Western analysis on the one hand, and the Arabic tradition in the other. In Western analysis the vowels are primary elements and thus affect the surrounding consonants, whereas in the Arabic tradition this is the other way around. Western scholars, for example, apart from only four velar consonants does not posit any velar or velarised consonants, such as [ṣ], [ṭ], [ẓ], [ḷ] and [ṛ] whereas these are so important in the Arabic analysis.

It shall be seen in this section that all sources have a list of those Turkic consonantal phonemes they consider the basic ones, in addition to a number of secondary sounds, whose status comes close to what in western theory are called allophones. The description of the con-sonants falls into two parts, based on the approach the sources choose in their respective descriptions of Turkic phonemes. It shall be seen that the first group, which consists of *Ḥilya* and *Dīwān*, takes the Uygur alphabet as its basic list of Turkic phonemes, whereas the second group bases itself on Arabic phonemes instead.

3.6.1 The number and status of consonantal phonemes

The sources display various opinions on the number of Turkic phonemes. *Dīwān* gives 18 basic consonants, *Šuḏūr* counts as many as 29, and even though the listing of basic phonemes in a particular source may be small initially, this number is usually extended with a

number of additional phonemes. The diffences between the sources in this respect are, as we shall see, related to their respective concepts of both Turkic and Arabic phonemes.

The question arises what exactly the status of the basic and additional phonemes is. As a rule of thumb, I suggest considering all phonemes that are explicitly mentioned in the basic listing as 'phonemes', and all those that are mentioned in passim as secondary or, with a modern term, 'allophones'. But, again, each source has different concepts in this matter.

The scheme in the appendix to this chapter (pp. 160-162) brings together all descriptions of Turkic phonemes and allophones that are mentioned in the sources, and serves as a reference.

3.6.1.1 The numbers of phonemes

Five of our sources each give an elaborate listing of the basic consonantal phonemes that occur in Turkic. The sources that provide such a list are *Ḥilya* (82,4f.), *Dīwān* (6,7ff), *Qawānīn* (4,7ff), *ʾIdrāk* (101,2ff) and *Tuḥfa* (2ʳ). In spite of this identical point of departure, there are considerable discrepancies in the numbers of consonants the sources posits for Turkish. The lowest number is 23, the highest 29. Four sources classify the consonants into 'basic' and 'secondary', while showing considerable differences as to the respective numbers of these subcategories. The secondary consonants are in all but one source (i.e. *Dīwān*) phonemes that do not occur in Arabic. Accounts of secondary consonants are provided in *Tuḥfa*, *Ḥilya*, *Dīwān* and *Tarğumān* (3,4ff), although the latter does not give an overall inventory (nor does *Šuḏūr*[85] 4ᵛ12;20).

The consonants in the basic list are presented as such in various ways. Kāšġarī and *Tuḥfa*'s anonymous author calls them simply 'basic' (ʾaṣliyya, and ʾaṣlan, respectively). The secondary consonants are called 'derivational' *(farʿiyya)* in *Dīwān*, and 'repeated' *(mukarrara)*[86] in *Tuḥfa*. *Qawānīn*, however, writes that the basic consonants occur in 'orthography' *(fī ṣ-ṣūra)*, and, along with six secondary ones, they are 'in speech' *(fī l-lafẓ)*. The following scheme summarises the data:

---

[85] In *Šuḏūr*, Ibn Muḥammad Ṣāliḥ counts 29 Turkic consonants (4ᵛ12;20), but he refrains from giving a detailed list, contenting himself with a description of the non-Arabic phonemes instead. He further specifies (without giving examples) that some of these are 'used' *(mustaʿmal)* whereas others are 'avoided' *(mahğūr fīhi)*.

[86] Doubling of consonants as a means to indicate a non-standard pronunciation also a method applied in *Durra* (4.2.7).

	Basic		Secondary		Total		
Tuḥfa	19	(*ʾaṣlan*)	4	(*mukarrara*)	23[87]		
Idrāk					23		
Ḥilya	19		4		23		
Dīwān	18	(*ʾaṣliyya*)	7	(*farʿiyya*)	25		
Qawānīn	21	(*fī ṣ-ṣūra*)	6		27 (28)		(*fī l-lafẓ*)
Šuḏūr					29		

In the overall scheme of Turkic phonemes in the appendix to this chapter, I have distinguished the basic (B) consonants from the secondary ones (S) by means of appropriate labels.

In addition to these two categories of consonantal phonemes, all but two sources (i.e. *Tuḥfa*, *Šuḏūr*) list yet other consonants in passim. Those are either Arabic consonants which had not been included in the list of basic consonants, or full-fledged phonetic descriptions of different sounds. A number of these are additional or alternative definitions for some of the secondary phonemes. I have marked these P (for passim) in the overall scheme, in as far as they can be identified in the respective sources. The secondary phonemes are the following: *Idrāk* [d̞]; *Ḥilya* dead ġ, pausal *h* (orthographic device), velarised t, [l̩] and [ṇ]; *Dīwān* [g] and [ʔ]; *Qawānīn* [ġ], [ʔ,] [r], and [ḫ] (as allophone of [q]); *Durra* [g], and *Tarǧumān* velarised *d* and velarised *z*, respectively.

3.6.1.2 Lists of absent phonemes
Six of the sources give a list of Arabic phonemes of which they explicitly state that they do not occur in Turkic. Such lists are found in *Qawānīn* (4,13), *Tarǧumān* (2,12), *Dīwān* (6,4), *Bulġa* (MS 1ʳ5), *Ḥilya* (82-83) and *Šuḏūr* (4ʳ25), whereas they lack in *ʾIdrāk*, *Durra* and *Tuḥfa*. As expected, the sources are not unanimous as to which Arabic phonemes do not occur in Turkic, but there are enough points of agreement too.

The listings display the following consonants/phonemes:

---

[87] I counted only 22 in *Tuḥfa*; the *dāl* must have been omitted erroneously.

	Qawānīn	Tarǧumān	Šudūr	Ḥilya	Dīwān	Bulġa
1	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ
2	ḥ	ḥ		ḥ	ḥ	
3		ḫ				
4	ḏ		ḏ			
5				ṣ	ṣ	
6		ḍ	ḍ	ḍ	ḍ	ḍ
7					ṭ	
8	ẓ	ẓ	ẓ	ẓ	ẓ	ẓ
9		ʿ		ʿ	ʿ	ʿ
10	f	f				
11	h		h			

Table I Absent consonants in Turkic according to the sources

In *ʾIdrāk* ʾAbū Ḥayyān writes that any Arabic consonant other than the ones mentioned in his listing of Turkic consonants cannot be genuinely Turkic. He exemplifies this statement with some examples of Persian and Sogian loan words, which contain the non-Turkic *f* and *ḫ*, e.g., **farman** 'decree', from Persian *farmān*, **firišti-lār** 'angels', from *färište*, and **kulaf** 'rose', from *gülab*, and Sogdian *ʾaḫšam* 'evening' (EDT 96:).

In *ʾIdrāk* (9; 26) **ḍ** is mentioned in passim as occuring in *Bulġar*. Although, as a rule, *ʿayn* does not occur in Turkic languages, including loan words, it occurs in two instances in *Qawānīn*, viz., *ʿaḍam* 'man', (59,11; 53,10) and *ʿawrat* 'woman' (59,14), in which it is probably to be interpreted as a kind of velarised, or stressed *hamza*. The pair *ẓ* - *z* in *Durra* is surprising, because according to the Arabic prescriptions, *ẓ* forms a pair with **ḍ**. *Durra*'s concept must be regarded as the reflection of a dialectal pronunciation.

An interesting point is the question of the contradictive statements regarding *ṣ* and *ṭ* in *Ḥilya*. Although explicitly given in the listing of Turkic phonemes, *ṣ* is here mentioned again as a non-Turkic consonant. There is no mention to this effect in regard of *ṭ* (cf. 3.1).

3.6.1.3 The Uygur alphabet as a base
The respective lists of consonantal phonemes in *Dīwān* and *Ḥilya* are based on the Uygur alphabet (cf. for example Kelly 1973). Kāšġarī indeed displays the Uygur alphabet and gives examples of how Turkic words are written in it, along with a transcription. His listing of eighteen basic phonemes, therefore, is essentially transcription from the Uygur alphabet, which has consequences for the interpretation of some

signs. (This point is further discussed in 3.6.1.3). *Ḥilya*'s listing bears a clear relationship to the Uygur alphabet too, although it is apparently a more indirect one, and his listing of 23 consonants is quite different from Kāšġarī's. *Dīwān* (6,7ff) and *Ḥilya* list the basic consonants of Turkic as follows (the sequence all but identical):

Dīwān	/ *Ḥilya*	*Dīwān*	*Dīwān*	/ *Ḥilya*	*Ḥilya*
1 ا ʾalif -*hamza*		7 ى y	13 ب b		19 ص ṣ
2 ڤ v	/ f	8 ك k	14 ج č	/ ǧ	20 د d
3 خ ḫ	/ غ ġ	9 ذ ḏ	15 ر r		21 ه h
4 و w		10 م m	16 ش š		22 خ ḫ
5 ز z		11 ن n	17 ت t	/ ط ṭ	23 ت t
6 ق q		12 س s	18 ل l		

In his listing Kāšġarī does not mention **d**, although he uses it quite often in the course of his text. Apart from being used in the verbal ending -**dī**, **d** also occurs in verbal and nominal stems. There are indications that in at least some of these **d** is not interchangable with **ḏ**. Fore instance, there is evidence for this assumption lies in the fact that intervocalic -**ḏ**- in *Dīwān* coincides with -**y**- in the so-called ayaq-branch of modern Turkic languages (viz. ʾayaq ~ ʾaḏaq 'foot', 45,1), whereas words with -**d**- apparently never change (viz. ʾidiš 'cup', 42,12).[88]

Incidentally, there is also confusion in regard to **d** in Ibn al-Muhannā's list, because of its occurrence on two places, i.e. on (9) and (20), and I tend to interpret the former as **ḏ**.[89]

Finally, Kāšġarī's use of the term 'basic consonants' (*ḥurūf ʾaṣliyya*) presupposes a second category, namely one of 'derived consonants' (*ḥurūf farʿiyya*). Indeed, Kāšġarī describes seven additional phonemes which he in concord with his earlier preferences derives from the Uygur alphabet. In his view, then, the 'Arabic *ǧ* (*ǧīm ʿarabiyya*), the regular [ǧ] and the 'Arabic *f* (*al-fāʾ al-ʿarabiyya*) [f] are derivatives of [č] and [v], respectively. There is one Arabic phoneme, i.e. [ġ], which has no representation in the Uygur script, and another two that cannot be

---

[88] Mr Hans Nugteren, p.c. November 3rd, 1997. // A copy of this description of basic and secondary consonants is given in the Margin Grammar (33ʳlt). Dankoff and Kelly (1982: 50f and 53ff) summarise the consonants occurring in the Turkic language of *Dīwān* as follows: /b, p, m; d, ḏ, t; n, ŋ; l, r; g, k, q; s, z, š, j, č, ž; f, h, x; y, w, v/.

[89] In another instance, indeed, **d** merges with **ḏ** (although this may be due to a copyist's error), where Ibn al-Muhannā discusses substitution of **d** by **t**. The editor Rifʿat gives the etymologically correct forms with **ḏ** in square brackets after each word (79,2ff).

expressed in either script, i.e. [ž] and [g] (further discussion in section 3.6.3.1).

3.6.3.1 Transcription from the Uygur alphabet

Kāšġarī's basic approach to reflection the sound system of Turkic in Arabic script is that of a transcription from the Uygur alphabet, which he apparently considers quite suitable for Turkic. His transcription, in its turn, has to be reflected in the Latin alphabet for our purposes.

There are two possibilities for the transcription of a given phoneme from the Uygur into the Arabic script. The first is that the phoneme can be reflected in both, i.e. each script has an unambiguous sign for it, e.g., [s]. The second is that a non-Arabic is reflected in the Arabic script with a sign that in Arabic has a different phonetic value. This is, for example, the case for the phonemes [č] and [v], which Kāšġarī tran-scribes with the regular Arabic signs for ğ and f, respectively. Problems occur if these Arabic signs are needed in the value they have in Arabic. Such a need in fact did occur.

There are some foreign loan words in Turkic, Kāšġarī writes, which contain [f] and [ğ], for which the Uygur alphabet, as explained above, has no signs. In these instances the appropriate Uygur signs are marked by means of additional diacritical dots. Kāšġarī subsequently transfers this method to the Arabic script, inasfar as he uses it for Turkic. In this way, both ف and ج are marked with extra dots, which results in ڤ and چ, respectively. As a consequence, then, ڤ and چ should be interpreted as respectively [v], and [č]. However, I have to add immediately that in practice for [v] it is the other way around, probably because a copyist, or even Kāšġarī himself found the system too confusing to apply throughout, and therefore, ڤ seems to represent [v], rather than [f].

Dankoff and Kelly (1982:55) discuss in some detail the transcrip-tion of these graphemes. Their conclusion in regard to [f/v] is that ڤ is to be interpreted as [w], and و as [v]. Their arguments are twofold. First, the Oġuz (and New Persian?) pronunciation of و is [v]. Second, the Sogdian grapheme for F was used to signify [w] or [β]. Etymolo-gically speaking, Dankoff and Kelly are probably right. However, what concerns us here is Kāšġarī's perception of Turkic phonemes and the way he described them in terms of Arabic phonetics, and therefore we should be careful in positing features of Turkic Kāšġarī probably did not, or could not notice. In this respect, I believe there are strong argu-ments in favour of interpreting ڤ as [v], and و as regular [w], all of them based on Kāšġarī's text. The first is that ڤ is clearly described as 'between Arabic f and b', which point at [v], rather than [w] (cf.

3.6.2.1). The second argument is that Kāšġarī nowhere says that he uses و for a pronunciation that is different from the standard Arabic one, while in other instances he is very accurate on this. (The standard Arabic pronunciation is, no doubt, [w].) The third argument combines the two first ones, and I put it in the form of a rhetorical question. Why would Kāšġarī complicate matters and introduce a new symbol (ڤ) for an existing phoneme, i.e. [w], for which he had an appropriate and non-ambiguous grapheme at his disposal?

The problem of transcription plays no role in *Hilya*; Ibn al-Muhannā lists 23 basic consonants, five more than Kāšġarī. The respective basic lists differ on a few more points; Ibn al-Muhannā includes the following consonants which are absent from Kāšġarī's: f, ġ t, ṣ, d and h. Furthermore, there are differences in transcription from the Uygur, viz., (*Dīwān-Hilya*) ḫ - ġ, d - ḍ, t - ṭ, v - f and č - ǧ, whereas in *Hilya* [v] and [č] are described elsewhere (cf. 3.6.2.1).

The different transcriptions are of interest from two more general points of view too. First, they indicate that Ibn al-Muhannā had no first-hand knowledge of the Uygur script. If he had, he might have found out, as Kāšġarī did before him, that there are only 18 basic signs. Secondly, the two variants show that there was no uniform transcription for the basic signs of the Uygur script. For example, some may have transcribed the phoneme on no. 3 as خ ḫ, whereas others preferred خ ġ, and no. 17 as ṭ, rather than t. It appears that, as in the course of time scholarly knowledge of the Uygur script dwindled, transcriptions that initially had the status of alternatives (ġ, ṣ, ṭ) were erroneously added to the basic list as phonemes in their own right.

In *Hilya* too, the non-standard realisations have to be marked, and in Arabic script this is reflected by means of additional diacritical dots on the basic signs:

> "The marker of these consonants which distinguishes them in writing from the pure Arabic consonants consists of three dots under the consonant concerned… which indicate their emphasis." (*wa-ʿalāma hādihi l-ḥurūf farqan baynahā wa-bayna ṣarīḥ al-ḥarf al-ʿarabī fī l-ḥaṭṭ ʾanna taḥta l-ḥarf alladī hādā šaʾnuhu ṭalāṭa nuqaṭ muṭaffā… tadullu ʿalā tafḫīmihi, Hilya* 73,2.)

Indeed, for two non-Arabic phonemes, i.e., [p] and [č], respectively Ibn al-Muhannā occasionally uses the signs پ and چ. A similar system, in which velarised *l* is marked with dots too, is used more extensively in *Šuḏūr* (see 3.1.2.1).

3.6.1.4 The Arabic alphabet as a base

As pointed out above in section 3.6.1.1, five sources give a listing of consonantal phonemes. Two of those, i.e. *Dīwān* and *Ḥilya*, take the Uygur alphabet as a starting point. The present subsection discusses the sources that take the Arabic alphabet as the point of departure without any reference to other scripts. These are *Qawānīn* (4,7ff), *'Idrāk* (101,2ff) and *Tuḥfa* (2ʳ). The remaining sources, i.e., *Šuḏūr*, *Tarǧumān*, *Bulġa* and *Durra*, although they do not give any listings, provide descriptions of a number of Turkic consonants. These phonetic descriptions are given below in 3.6.2.1. *Qawānīn* (4,7ff) counts 28 consonants (but actually gives 27; š is erroneously left out) that are used 'in speech' *(fī l-lafẓ)*, 21 of which are reflected 'in writing' *(fī ṣ-ṣūra)*, by which the author means that seven consonants do not have a unique grapheme in the Arabic script. By interpretation, then, the ones 'in writing' must be [ʾ, b, t, ǧ, ḥ, d, r, z, s, š, ṣ, ḍ, ṭ, ġ, q, k, l, m, n, w, y], whereas the seven consonant that occur in speech only, are, viz., [p, č, z̠, G, g, ḷ, ŋ]. *Tuḥfa* distinguishes between nineteen 'basic' *(ʾaṣlan)* and four 'repeated' *(mukarrara)* consonants. The basic consonants are [ʾ, b, t, ǧ, d, r, z, s, š, ṣ, ṭ, ġ, q, k, l, m, n, w, y], the derived ones are [p, č, g, ŋ]. The term 'repeated' too refers to a means of marking them in the Arabic script as non-standard (see 3.6.2.2).

One would, again, assume that the use of the same alphabet, i.e. the same basic concept leads to very similar perceptions of Turkic phonemes. This is, as we shall see in the following subsections, only partially true.

3.6.2 Descriptions of non-Arabic phonemes

Having sketched a general context in the preceding subsections, we now turn to the actual phonetic descriptions of Turkic consonantal phonemes as they are given in the sources. Descriptions of velarised consonants are not given here, since they have been amply dealt with in section 3.1.2. After the respective descriptions, this subsection contains a brief excursus on how doubling is used as a marker of a non-standard pronunciation, followed by a summary.

Standard or regular Arabic consonants are, of course, not discussed, although they are indicated too, opposing them to the non-regular pronunciation. The sources describe them as *ḫāliṣ, ṣarīḥ* (*Ḥilya* 73,2), *ṣaḥīḥ* (*Dīwān* 26,17) all of them meaning 'pure', which are regular terms for describing the Arabic basic phonemes as opposed to the allophones (cf. 1.2). Kāšġarī uses 'Arabic' (*ʿarabiyya*, 26,17), when referring to standard *f* and *ǧ*, and 'firm' *(ṣulba)* for the Arabic *k* (as opposed

to the 'fine' *rakīka* for [v] and [g], respectively). This makes sense, since his point of reference is the Uygur script and the pronunciations he associates with that, rather than the Arabic (cf. 3.1.5.3).

3.6.2.1 Non-Arabic consonantal phonemes in detail

The present subsection deals with the actual phonetic descriptions of eleven specifically Turkic phonemes as they are given in the sources. In addition there is [ḷ] which is discussed above in 3.1.2.1, which makes a total of twelve. As we shall see, seven of the twelve phonemes, i.e., [p, č, g, ḷ ž, ŋ, z̤], are described in very similar terms as allophones that oc-cur in Arabic. The remaining five phonemes, [ṇ, v], the velarised g [ɢ], the mixed h, and the 'dead ǧ' are alien to the Arabic consonantal sys-tem and are described in new terms.

1 [p]

The phoneme [p] is typically described as 'a blend' of two Arabic phonemes, [b] and [f], viz., 'the *b* that is mixed with the *f* (*al-bāʾ al-mašūba bi-l-fāʾ, Tuḥfa*), or elliptically, 'the mixed *bʾ* (*al-bāʾ al-mašūba, ʾIdrāk*). The same relation of this phoneme with *b* and *f* is posited in *Qawānīn*, viz., 'the *b* that is mixed with the crackling *f* (*al-bāʾ al-mašūba bi-l-fāʾ al-mušanšana*). The term *mušanšana* is also used in *Qawānīn*'s description of [č] (see below). The concept of mixing, or flavouring recurs without any explication in *Durra*, where a small *f* is scribbled above **b**, e.g., 'arba [arpa] 'barley' (*Durra* 6ʳ9). The few oc-currences of **f** in *Qawānīn* therefore are to be interpreted as [p], for [f] is not listed as a Turkic consonant, viz., **'arfa** [arpa] 'barley' (63,12); **šaftalū** [šaptalu] 'peach' (63,8), and **šifšaq** [šipšaq] 'sparrow' (62,14). Finally, *Šuḏūr* compares Turkic p to that in Persian, whose place of ar-ticulation is "between the Arabic *b* and *f*" e.g., *pādišāh*.

Targumān and Ḥilya use in this respect *tafḥīm*, viz., 'emphasised *b* between *b* and *f* (*al-bāʾ al-mufaḥḥama bayn al-bāʾ wa-l-fāʾ*, *Targumān* 3,5). This 'emphasis of *bʾ* is in *Ḥilya* further specified as "as if you pronounce [b] outside your lips" (*ka-ʾannaka tanfaḥu bihā min ḥāriǧ aš-šafatayn*). In **tap-tī** 'he found' (99,17) voicelessness is indi-cated with doubling (viz., تَبِّتِي). (For further use of *tafḥīm* for in a voiceless phoneme, see below [č].)

The phoneme [p] is described in *Dīwān* with an adjective, i.e. 'firm *bʾ* (*bāʾ ṣulba*, *Dīwān* 6,15)—which is, along with [t, č, k], and [q] col-lectively referred to as the 'consonants of firmness' (*ḥurūf aṣ-ṣalāba*, 281,14; 266,4)—much like the equally voiceless [k] (*kāf ṣulba*), a regu-

lar Arabic phoneme. There is, however, no expression * *bāʾrakīka* for the regular, voiced *b* to complete match the pair.

2 [v]

Dīwān and *Ḥilya* are the only sources to posit a Turkic phoneme [v]. In *Dīwān* it is named 'fine *f* (*fāʾrakīka*, *Dīwān* 26,17), analoguous to 'fine **k**' (see [g]), described more precisely as being "articulated between the points of articulation of the Arabic *f* and the correct *b*" (*ḫāriǧa bayna maḫraǧay al-fāʾ al-ʿarabiyya wa-l-bāʾ aṣ-ṣaḥīḥa*, 26,17), or simply "*f* between the two points of articulation" (*bayn al-maḫraǧayn*, 54,10; 544,12). Ibn al-Muhannā describes it as 'emphasised *f* (*fāʾ mufaḫḫama*), or, more detailedly, "*f* flavoured of *w*" (*fāʾ muṣamma ʾilā l-wāw*). This phoneme, in *Ḥilya* expressed by means of ڤ, occurs, for example, in ʾ**āv** [av] 'hunt', **čāv** [čav] 'fame', **savinğ** [sevinč] 'joy' (*Ḥilya* 79,15). Based on our knowledge of Arabic phonetics in regard to [b] and [f], which both Kāšǧarī and Ibn al-Muhannā here use as their reference, this description exactly matches [v].[90]

According to Ibn al-Muhannā in these words [v] alternates with [w], viz., ʾ**āw**, **čāw**, **sawinğ**, a phenomenon which Kāšǧarī mentions as an option too: "every **f** that is articulated between the points of articulation may be replaced by a **w**" (*kull fāʾ bayna l-maḫraǧayn fa-ʾinnahu yaǧūz mubādalatuhā bi-l-wāw*, 54,10), whereas for the Oġuz [w] is the standard pronunciation (27,1).

3 'Dead *ġ*'

Ibn al-Muhannā mentions a 'dead *ġ*' (*ġayn mayyita* 74,10; 77,9f), which he also calls 'melting' (*maḏāba*). The 'dead *ġ*' is in fact an allophone of *ġ*, and realised as [w], and therefore in essence not a new phoneme, viz.,

> "The pronunciation of [the 'dead'] *ġ* in the languages of the Turks in our realms is like the *w*'s, and this is because the *ġ* is a guttural sound whereas the *w* is a labial sound. The place of articulation of both is inclining toward the lips." (*fa-n-nuṭq bi-hāḏihi l-ġayn fī ʾalsina t-turk fī ʾarḍinā ka-ʾannamā huwa bi-l-wāw wa-ḏālika li-ʾanna l-ġayn ḥarf ḥalqī wa-l-wāw ḥarf šafawī, fa-l-maḫraǧ baynahumā māʾil ʾilā š-šafa, Ḥilya* 77,17ff.)

The 'dead *ġ*' occurs in words like, e.g., **tūġrā** /tuwġra"/ 'signature' [tuwra], **būġdāy** /buwġda"y/ 'wheat' [buwday], and **būġrā** /buwġra"/ 'he-camel' [buwra].

---

[90] Rather than [w] as proposed by Dankoff and Kelly; for which see discussion in 3.6.3.1.

The 'dead *ġ*' is described as a rather soft allophone of *ġ*. In fact, the description Ibn al-Muhannā gives calls to mind the realisation of the socalled "yumuşak ge" in modern Turkish. This, indeed, may not be a coincidence, for there is other evidence too that points to a relation of the language in *Ḥilya* with Azeri, or even Anatolian Turkish (see discussion in Chapter 2). Although the 'dead *ġ*' forms a terminological pair with the 'dead *k*' (*kāf mayyita*) [g], in pronunciation there is no obvious relationship.

In the spirit of the Arabic tradition, in which *badal* typically is associated with a common point of articulation of the consonants in question, Ibn al-Muhannā postulates that the 'dead *ġ*' is articulated between *ġ* and *w*.

4 [č]

The sources typically describe the phoneme [č] in terms of a mixture of Arabic [š] and [ǧ], 'the *ǧ* that is mixed with the *š* (*al-ǧīm al-mašūba bi-š-šīn, Tuḥfa, Qawānīn*)—also elliptically 'the mixed *ǧ* (*al-ǧīm al-mašūba, ʾIdrāk*)[91]—"*ǧ* flavoured of *š*" (*ǧīm mušamma ʾilā šīn, Ḥilya*)— or simply as "*ǧ* between *š* and *ǧ*" (*al-ǧīm bayna š-šīn wa-l-ǧīm, Bulġa* MS 1ʳ9). The concept of mixing, or flavouring is evident in *Durra* two ways; the first of which is a small *š* scribbled above *ǧ*, **nāčuk** [nečük] 'how', and the second a small *ǧ* written underneath *š*, e.g., **bišāq** [bičaq] 'knife' (*Durra* 10ᵛ6). See further discussion of methods of transcription in *Durra* 4.2.7.[92]

In both *Ḥilya* and *Tarǧumān* the term *mufaḫḫam* is applied. In the the former it is simply an 'emphasised *ǧ*' (*ǧīm mufaḫḫama*), whereas in the latter [č] is an 'emphasised *ǧ* between *ǧ* and *š*' (*al-ǧīm al-mu-faḫḫama bayn al-ǧīm wa-š-šīn*), reflecting the concept of the blend of two consonants. According to *Tarǧumān*, this phoneme "resembles the pronunciation of *ǧ* by the Nabateans of Baʿalbak" (*tušbih talaffuẓ nabaṭ bilād baʿalbakk bi-l-ǧīm, Tarǧumān* 3,8). The phoneme [č] occurs, e.g., in **bičaq** 'knife' (86,6) and **ʾāč** 'open' (*Ḥilya* 99,3). In *Qawānīn* [č] is further described as a "crackling *ǧ*" (*ǧīm mušanšana*,

---

[91] In the manuscript p and č are indicated with *ǧ* and b, each marked with a superposed *š* (ش), which stands for *mašūb* 'mixed'. To avoid confusion because of the defective Arabic script, ʾAbū Ḥayyān states "In all [words] whose first consonant is a *ǧ*, [the *ǧ*] is mixed, except in *ǧiǧā* ['mother' (42)] and *ǧumart* ['generous' (46)]." (*wa-ǧamīʿ mā ʾawwaluhu ǧīm hiya fīhi mašūba ʾillā ǧiǧā wa-ǧumart faqaṭ, ʾIdrāk* 47.) This means that in all instances, except in the two words mentioned, the grapheme *ǧ* on word-initial position stands for [č].

[92] In *Ḥilya* and *Durra* occasionally the grapheme ج is used too, e.g., **bičqač** [bičqač] 'scissors' (*Durra* 9ᵛ13).

Qawānīn 55,18). The term *mušanšana* may be a general reference to voicelessness, since it recurs in *Qawānīn*'s description of [p]. Finally, *Šuḏūr* compares [č] to Persian *č* (see similar comparisons to Persian in below in 4.3.1.3).

5 [g]

The phoneme [g] is described as a 'knotted *k*' (*al-kāf al-maʿqūda*, *Tuḥfa*), a 'bedouin *k*' (*kāf badawiyya*, '*Idrāk* 106,18; *Qawānīn* 50,19) or as a combination of both, i.e. 'the knotted bedouin *k*' (*al-kāf al-badawiyya al-maʿqūda*, *Qawānīn*, *Ḥilya*). *Tarğumān* writes "the tied *q* of the Arabs" (*qāf al-ʿarab al-maʿqūda*[93], 40,18; also 21,16), which calls to mind its regular description in Arabic theory (cf. 1.2.1). *Dīwān* specifies it as "k that is produced between the places of articulation of q and k" (*al-kāf al-mutawallida bayna mahraġay al-qāf wa-l-kāf*, 6,17), and, in short reference 'thin k' (*kāf rakīka*) (as opposed to 'hard k' *[kāf ṣulba]*, i.e. voiceless Arabic *k*).

In *Ḥilya* (76-77) in addition the terms 'dead *k*' *kāf mayyita*, or 'melting *k*', *kāf ḏāʾiba* are used. Examples of *Ḥilya*'s 'dead *k*' are tīrik̂ [tirig] 'living', and sīlīk̂ [silig] 'clean' *an-naẓīf*, 77); bak̂ [beg] 'chief' (75,12), k̂un [gün] 'sun'. Finally, *Šuḏūr* compares it to 'Persian *k*' (*kāf ʿaǧamiyya*, *Šuḏūr* 19ʳ20).

6 Velarised g [ɢ]

Two sources posit a velarised voiced phoneme, which I would interpret as [ɢ]. It is described as an " ʿayn that is [pronounced] between ʿand ġ" (*al-ʿayn allatī bayna l-ʿayn wa-l-ġayn al-muʿǧama*, *Qawānīn*, 4,9), or, in *Tarğumān*, "the g is resonant between ġ and k; its place of articulation is from the nasal cavity" (*fa-l-k̂āf... bayna al-ġayn wa-l-kāf muġannanatan wa-mahraġuhā min al-ḥayšūm*, 3,9)

7 Velarised z [z̧]

Two sources, *Qawānīn* and *Tarğumān*, posit a velarised z, 'the z that is mixed with the ṣ' (*az-zāy al-mašūba bi-ṣ-ṣād*, *Qawānīn* 4,9 and various places), also found as an Arabic allophone. Examples are, e.g., yaz- to write', buġazla- 'to kill' (*Qawānīn*). In *Tarğumān* it is described in terms of an 'emphasised z' (*zāʾ mufaḫḫama*), which "resembles the pronunciation of *ḍ* by the Nabateans among the people of the Ṣaʿīd" (*tušbih talaffuẓ nabaṭ ʾahl aṣ-ṣaʿīd bi-ḍ-ḍād*, *Tarğumān* 3,9). This very same comparison with a dialectal pronunciation of *ḍ* is also made in *Ḥilya*, there representing ẓ.

---

[93] Em. for *al-manqūṭa*

8 [ž]

Dīwān and *Ḥilya* posit a Turkic phoneme *ž*, which only occurs in loan words. It is pronounced "between the place of articulation of *ǧ* and *š*" (*bayna maḥraǧ al-ǧīm wa-š-šīn, Ḥilya* 77), or, more precisely, 'the *z* that is produced between the places of articulation of *z* and *š*' (*az-zāy al-mutawallida bayna maḥraǧay az-zāy wa-š-šīn, Dīwān* 6,17). *Ḥilya* further describes it as an 'emphasised *z*' (*zāy mufaḫḫama,* 77), which "resembles the pronunciation of *ḍ* by the Nabateans among the people of the Ṣaʿīd" *(tušbih talaffuẓ nabaṭ ʾahl aṣ-ṣaʿīd bi-ḍ-ḍād)* (This same description recurs in *Targumān*, there signifying velarised z). The *ž* occurs, for instance, in ʾ**ažūn** 'world'; ʾ**aržū** 'jackal' and **kužāk** 'a lock of hair' (*Ḥilya* 79,5),[94] generally considered Sogdian or Iranian borrowings (cf. EDT). Although not specifically described in *Targumān*, it occurs there once (32,1 ṭāy ʾ**ažā** 'maternal aunt' (with ژ).

9 Palatal [ŋ]

The palatal *ŋ* is usually described as ' *n* of the nasal cavity' *(an-nūn al-ḫayšūmiyya, Idrāk, Qawānīn, Tuḥfa; al-ḥarf al-ḫayšūmī, Dīwān* 8,11) *ḫayšūm* probably referring to the nasal cavity. A second term is 'resonant *n*' *(nūn maǧnūna)*; e.g., 'move!' **tabraŋ** (*Qawānīn* 77,11). In *Targumān* it is described as "there is some resonance in [the n]; its place of articulation (i.e. the *ḫayšūm*) [ranges] from the ceiling of the throat up till the end of the nose" *(fīhā šayʾ min al-ǧunna wa-maḥraǧuhā min saqf al-ḥalq ʾilā raʾs al-ʾanf, Targumān* 46,6), The *ŋ* occurs in, eg., **čiz-mā-ŋ** [čïzmaŋ] 'do not write' (pl). In *Durra* (2ʳ6) we find the expression "soft *k*" *(kāf ḫafīfa)*, which reminds of the *nūn ḫafīfa* in the descriptions in the Arabic sources (1.1).

In the text of *Qawānīn* [ŋ] is indicated by means of an *n* and a superposed *k* (*Qawānīn* 27,14), while in *Tuḥfa, Durra* and *Ḥilya*'s it is represented by the digraph /nk/, e.g., **kāl-dī-ŋ** /kaʺldiynk/ 'you came' (*Tuḥfa* 57ʳ6), or a single n, viz., /tanriy/ (*Durra* 2ʳ6). In *Ḥilya* the k and the n reportedly assimilate with eachother in speech, even though they remain fixed in orthography *(al-ḫaṭṭ)*. [ŋ] is described as both a k̂ assimilating with n in **tankiz**, i.e., [teŋiz] 'sea', and, conversely, as n assimilating with k̂, e.g., **sinkir** [siŋir] 'sinew' (75,1; cf. Cantineau 1966:206).

---

[94] According to Ibn al-Muhannā *ž* in these words alternates with *š* (ʾaršū ~ ʾaržū), s (kusāk ~ kužāk), and *ǧ* (ʾaǧūn ~ ʾažūn), respectively.

10 Velarised ŋ [n̠]

Two sources posit a velar variant of the ŋ, here transcribed as n̠. They probably did so because the palatal version is closely associated with k which is the palatal consonant *par excellence* and therefore cannot refer to anything velar. Not surprisingly, therefore, they describe the velarised ŋ in terms of ġ, viz., n̠ 'nasalised ġ' (*ġayn maġnūna, Qawānīn* 29,6), or as "k of resonance that is produced between ġ and q and [between] n and q" (*kāf al-ġunna al-mutawallida bayna l-ġayn wa-l-qāf wa-n-nūn wa-l-qāf, Dīwān* 7,1). Examples are, e.g., 'an̠la-dī /ʾaġla-dī/ 'he understood' (*Qawānīn* 29,6), and tan̠rī /taġrī/ 'God' (*Qawānīn* 58,1; 74,5).

'Abū Ḥayyān describes this velarised n̠ in phonological terms, rather than in the phonetic qualities of the consonants:

> "As far as ṣaġā ['to you'] is concerned, its base form is ṣan-ġā. The n was elided because of the frequency of its use; the base form is used less often." (*wa-ṣaġā ʾaṣluhu ṣanġā fa-ḥuḏifat an-nūn li-katra al-istiʿmāl wa-yaqillu stiʿmāl al-ʾaṣl, ʾIdrāk* 127,2.)

In *ʾIdrāk* this rare instance of velarised ŋ is understood as a crasis of n and the dative suffix ġā in the personal pronouns man 'I' and san 'you' (sg), producing [maŋa] and [saŋa], respectively. In *Idrāk* it is in these two instances reflected by means of a single ġ (127,2; 142,2).

In *Ḥilya*, surprisingly, the 'dead' k (*kāf mayyita*), apart from referring to palatal [g], can refer to a velar sound too, i.e., a 'k flavoured of ġ' (*kāf mušamma ʾilā l-ġayn*, 77,4). Ibn al-Muhannā gives the following description, viz.,

> "When there is a n before a 'dead' k of this kind, the k is pronounced flavoured of the ġ, and the sound becomes more nasalised." (*wa-matā kāna qabla mā hāḏā sabīluhu min al-kāfāt al-mayyita nūn nuṭiqat bi-hāḏihi l-kāfāt mušamma ʾilā l-ġayn fa-yakūn aṣ-ṣawt bihā ʾaġann, Ḥilya* 77,14.)

Examples are, e.g., 'ankā [aŋa] 'to him', munk̬ā [muŋa] 'to this one', sank̬ā [saŋa] 'to you', mank̬ā [maŋa] 'to me'. All these words are originally front, but shift to back in the dative case.

11 "Mixed ḫ"

This phoneme, described in *Qawānīn* only, is a fricative velar phoneme, similar to but apparently not identical with Arabic ḫ. It occurs in Turkic as an allophone of q only, within one language, or as a replacement of q throughout. In orthography occasionally reflected with ḫ (as in 'aḫšā 'dirham', 57,21), this sound is described in various ways,

viz., ʿq mixed with a ǧ (*qāf mašūba bi-ġayn muʿǧama*, 76,3), ʿq be-
tween ġ and q, but which is closer to q' (*qāf bayna l-ġayn wa-l-qāf hiya
'aqrab 'ilā l-qāf*, 64,1). Two alternative descriptions are 'a q between q
and ḫ' (*qāf bayna l-qāf wa-l-ḫā'*, 64,7), and ʿh which they pronounce
mixed with ḫ' (*yanṭiqūna bi-[l-ḫā] mašūba bi-ḫā*[95], 4,14). This sound
occurs in, e.g., **qalqān** [ḫalḫan] 'shield', **bardaq** [bardaḫ] 'cup' and
'aḫšam 'evening' (latter spelled with ḫ).

Although this phoneme is listed as a regular Turkic phoneme, the
author evidently regards it as an allophone of q, which is evidenced by
his comment that wherever ḫ occurs, "the pure Turks pronounce [the
word] with q" (*wa-t-turk al-ḫullaṣ yaqūlūnahu bi-l-qāf*, *Qawānīn*
62,20). In *Tarǧumān* too ḫ interchanges in one occasion with q, but
still occurs in some loan words (cf. Houtsma 73).

3.6.2.2 Doubling as a marker for a non-standard pronunciation

In *Durra* no listing of consonants is provided, and the quality of the
phonemes can only be determined by interpretation of the markers
that are put below and above some consonants. Doubling is the main
instrument to indicate a non-standard pronunciation. A voiceless
pronunciation of otherewise voiced basic phonemes is indicated by
doubling, e.g., **ǧǧuwmǧǧa** [čümče] 'ladle' (10ᵛ6) and **'ūbb** [öp] 'kiss'
(20ʳ9). Doubling as a means to indicate voicelessness may go back to
an older tradition. Ibn Sīnā, for example, uses the term 'doubled b̄'
(*bā' mušaddada*) for Persian p ('*Asbāb* 16,15). In certain instances
doubling indicates velarisation (see discussion in Section 3.6.1). Doub-
ling also serves to indicate a fricative pronunciation, e.g., **kāššī** [keči]
'goat' (7ʳ15); **ššūb** [čöp] 'straw' (6ʳ15; also 5ᵛ15) (*Tuḥfa* too refers to
'doubling' for indicating non-standard phonemes, see 3.6.1.3).

Doubling of voiceless consonants, on the other hand, indicates a
voiced pronunciation, e.g., **kkāl** [gel] 'come' (20ᵛ10; 21ʳ12), **kkāz** [gez]
'walk' (21ʳ12) and **kkāt** [get] 'go' (20ʳ10). (These words may be Oġuz
which typically has voiced initial consonans).

A second method used in *Durra* to mark non-Arabic phonemes is
scribbling a small consonant above or underneath a regular Arabic
phoneme, instances of which are discussed in [č] and [p].

3.6.2.4 Summary

Out of the total of twelve phonemes whose descriptions are given
above, seven [p, č, g, ḷ, ž, ŋ, ẓ] are also described as allophones in

---

[95] Em. for ḫā'.

Arabic.[96] For these phonemes the sources give, as expected, much the same, or in any case very similar descriptions (cf. 1.2.1). The remaining five, i.e. [v], 'dead *ǧ*', velarised [ŋ], 'mixed [ḥ]' and [ɢ], are foreign in the sense that they do not occur as Arabic allophones. Of these, 'dead *ǧ*' and 'mixed *ḥ*' are recognised as allophones that can also expressed by means of a standard Arabic grapheme, or even more than one. The velarised ŋ, [ɢ] and [v] are entirely new. One possible way to analyse the two former ones is to regard them as variants of [ŋ] and [g]. The sources, nevertheless, instead relate them to basic Arabic consonants.

For the Arab authors, there is a set of basic (*ʾaṣlī*) Arabic consonants each of which is associated with a given phonetic realisation. This phonetic realisation, much like our own definitions, is described in terms of the position of the tongue in the mouth, vis-à-vis the tooth and the palate. (In fact, there are two definitions; first, a prescriptive one, and second the ones that are based on the actual dialectal pronunciation, such as *ḍ* and *ẓ*; see discussion in 3.1.1.) The choice of the Uygur alphabet does not lead to different descriptions of non-Arabic phonemes. Kāšġarī describes all foreign phonemes in terms of Arabic phonemes.

As has been pointed out, in their descriptions of non-Arabic consonantal phonemes, the sources take one basic Arabic phoneme to which a particular characteristic of a second (and, occasionally a third) basic phoneme is added. This can be done in terms of a blend, or by positing the new phoneme between two (or more) basic ones. The characteristic of the added or blended second phoneme is often not named; the only terms used are *tafḫīm* (and *tarqīq*), *tašanšun* and *ǧunna*. But in the majority of cases it is possible to do reconstruct the argumentation. For example, [p] consists of a (basic) [b] which is mixed with *f*. The aspect needed to derive [p] from [b] is 'voicelessness' (*tašanšun*, lit. 'crackling'), which is, in this case, supplied by *f*. An interpretation of the data is summarised below in table II:

---

[96] Although velarisation of r is often implied, it is nowhere described as such.

no.	phonetic value	basic con- sonant	velari- sation (tafḫīm)	voice- lessness (tašanšun)	voiced- ness	nasality, reson- ance (ġunna)	sibil- ant	friction
2	p	b		f				
3	v	f			w			
7b	velar. d 1	d	ṭ					
	2	ṭ			d			
11b	velar. z	z	ṣ					
14	č	ǧ		š				
15	ž	z					š	
18a	g 1	k				+		
	2	q			+			
18b	[ɢ] 1	k			ġ			
	2	k	q		ġ	n		
	3	ʕ						ġ
20a	ŋ 1	k				n		
	2	n				+		
20b	velar. ŋ 1	ġ				n		
	2	n	ġ					
21b	velar. ḫ 1	q						ġ
	2	q						ḥ
	3	h						ḫ

Table II Schematic description of mixing of consonants

3.6.3 Consonant assimilation

Consonant assimilation is a progressive feature in Turkic languages. This implies that the shape of the suffix changes depending on whether the final consonant of the verb or noun is voiced [+voiced] or voiceless [-voiced]. If it is voiceless, the initial consonant of the suffix is too. Similarly, if the word ends in a vowel or a voiced consonant, the first consonant of the suffix is voiced.

In Section 3.3 it is shown that most sources regard the distribution of palatal and velar suffixes as a phenomenon. Because in the opinion of the Arab grammarians (palatal - velar) vowel harmony is explicitly correlated with the qualities of the consonants, the form of the suffixes is influenced by both consonant assimilation and vowel harmony. One could say that to express both principles, ideally a type of extended consonant assimilation that matches two parameters is required. Palatal(ised) words ending in a voiceless consonant would get a suffix with an initial voiceless palatal consonant, whereas the suffix attached to a velar(ised) words ending in a voiced consonant would be both voiced and velar(ised). In this subsection it is pointed out for the past

tense suffix DI that these principles are only partially observed by the Arab grammarians. (3.6.3)

An important principle with regard to consonant assimilation in Arabic is *'ibdāl* 'substitution' (see Section 2.2). Although the term *'ibdāl* stands for a qualitative change in the pronunciation of a consonant, e.g., [ṣ] as [s], or [d] as [t] (which may occur phonetically conditioned, or else as a dialectal variant), in Arabic grammar it is defined in phonological terms, and interpreted as substitution of one consonant by another.

In regard to Turkic the grammarians hold similar principles, and it shall be seen that in the discussion of the consonant assimilation the notions 'substitution' *('ibdāl)* or 'transformation' *(qalb)* of one consonant into the other play an important role. Research on Arabic linguistic theory traditionally relates voiceless and voiceled to the terms *mahmūsa* and *maǧhūra*, respectively. Contrary to what one would expect, though, the sources do not use these terms at all in relation to consonant assimilation in Turkic.

3.6.3.2 An instance of consonant assimilation: the past tense suffix -DI

A very frequently occurring suffix which is subject to consonant assimilation is the one used for the past tense DI (which in modern Turkic languages stands for di, dï, dü, du and ti, tï, tu and tü, respectively). After vowels and voiced consonants, DI is realised as [di], e.g., **suzla-dī** 'he spoke', **kir-dī** 'he entered', while

> "this **d** is substituted by **t** because the place of articulation is close. You say **tāp-tī** ['he found'], **qāč-tī** ['he flied'], **sūk-tī** ['he tore apart'], **sīq-tī** ['he squeezed'] and it changes into **t** in order to protection against heaviness in speech." (*tubdal hāḏihi d-dāl tā'an li-qurb al-maḥraǧ taqūl ṭābtī wa-qāǧtī, wa-sūktī wa-sīqtī wa-tudǧam fī t-tā' ittiqā'an li-ṯiql al-kalām, Tuḥfa* 52ʳ13.)[97]

Although the rule appears to be sound and familiar, it is not entirely what one would have expected. The *maḥraǧ* 'place of articulation' is this context does not signify the places of articulation of **p, č, q** and **k** on the one hand, and **t** on the other, are meant, since they cannot be considered quite close, but most likely those of **d** and **t**. The phenomenon, thus, is not brought in direct relation with the preceding consonant.

---

[97] In these verbs **b** is interpreted as **p**.

Kāšġarī's account, though, is quite to the point, and I quote him in full:

> "Know that the past tense is bound [to be indicated] by means of **d** and **y** [i.e. **dī**] in all words. There is no exeption to [this rule] except that the **d** changes into **t** [i.e. **tī**] when it is accompanied by the consonants of firmness because of the tightness of their place of articulation. The consonants of firmness in this language are four, the firm **p, t,** the firm **č** and the firm **k**... The **d** is changed into **t** because of the firmness of the [respective] places of articulation of these consonants... even though the base [of the suffix] is **d**, it is preferable to pronounce it as **t**." (*wa-ʿlam bi-ʾanna l-māḍī muqayyad bi-d-dāl wa-l-yāʾ fī ǧamīʿ l-ʾafʿāl faqaṭ lā yataġayyar ʿanhā ʾillā ʾanna d-dāl ʾiḏā ṣaḥibat ḥurūf aṣ-ṣalāba tanqalibu tāʾan li-šiddat al-maḥraǧ. wa-ḥurūf aṣ-ṣalāba fī hāḏihi l-luġa ʾarbaʿa wa-hiya l-pāʾ aṣ-ṣulba wa-t-tāʾ wa-l-ǧīm aṣ-ṣulba, wa-l-kāf aṣ-ṣulba... ʾinnamā ǧuʿilat ad-dāl tāʾan li-ṣalābat mahāriǧ hāḏihi l-ḥurūf... wa-l-ʾaṣl fīhā d-dāl wa-n-nuṭq bi-t-tāʾ afṣaḥ,* Dīwān 281,14-16 and 282,4; similar statement to this effect on 266,4.)

Examples with which he illustrates his statement are, e.g., **tap-tī** [tep] 'to kick', **tut-tī** 'to hold', **qač-tī** 'to flee', **baq-tī** 'to look at'.[98] It is of great significance that Kāšġarī here introduces a new category, i.e. the *ḥurūf aṣ-ṣalāba*, the 'voiceless consonants', rather than using the term *maḥmūsa* which in Western studies is traditionally translated as 'voiceless'. Kāšġarī however does not mention [s, š] and [q].

Further three sources give no clues in regard to the principles of consonant assimilation, in spite of mentioning it as a phenomenon. In *ʾIdrāk* (114,6), for example, ʾAbū Ḥayyān regards the different forms of DI as substitution (*ʾibdāl*) of consonants, i.e. change of **d** into **t** and **ṭ**, respectively, i.e. as a secondary prosodic feature, without explaining the underlying principles. Some examples given elsewhere in *ʾIdrāk* are ʾurt-tī 'he covered' (*ʾIdrāk* 110,18); **kas-tī** and **bič-tī** both 'he cut' (*ʾIdrāk* 109,10) (Examples with **ṭ** are discussed below 3.6.3.3).[99] In

---

[98] In another instance (*Dīwān* 354,4ff) the alternation of the causative suffix DIr is noted, **tal-tur-** ~ **tal-dur-** 'make (someone) pierce (a wall)'. This instance of alternation is compared to the non-conditioned change of *t~ d* in the Arabic verb *qadara* he decreed' - *qatara* 'he was miserly' (Qurʾān 65/7 and 89/16). In regard to DIr, it seems that Kāšġarī considers **tur** as the basic form of the causative suffix, and that he does not relate the change to phonetic conditions. Another likely explanation is that here is a difference between dialects; it may well be that Kāšġarī's Ḥāqānī Turkic did not observe consonant assimilation in this suffix.

[99] A similar alternation between the voiced and voiceless shapes of the suffix occurs in the causative suffix DIr, which in *ʾIdrāk* is also considered an instance of substitution (*ʾibdāl*) of a base from **-dur** with a secondary form **-tur**: ʾišit-dur-dī vs ʾišit-tur-dī 'he made listen' (*ʾIdrāk* 116, 19).

Šuḏūr consonant assimilation is only very marginally observed; in some instances, when the verbal stem ends in **t**, consonant assimila-tion is indicated with with a *šadda* (ˑ), which indicates that the **d** of the suffix assimilates with the preceding **t**, e.g., **kat-dī** [ketti] 'he went'; **ʾišit-dī** [išitti] 'he heard' (10ᵛ-11ʳ). In all other instances, the default -**dī** is used. The same can be said of *Ḥilya* in which the principles of consonant assimilation are observed only occasionally, e.g., **ʾāč-dī** - **ʾāč-tī** 'he opened'; **baq-dī** - **baq-tī** 'he looked' (79,14).

3.6.3.3 Extended consonant assimilation

The extended consonant assimilation requires that a stem ending in a voiceless consonant gets a voiceless suffix, while also agreeing in front and back quality (cf. 3.3.1). These two parameters yield four positions for the suffixes DI and GIl, respectively

	voiced	voiceless
back	ḍī / ġil	ṭī / qil
front	dī / k̯il	tī / k̯il[100]

As far as GIl is concerned, none of the sources are all four of these ex-plicitly described at the same time. *ʾIdrāk*, for example, only has a basic set of two, i.e. a palatal and a velar form. In *Qawānīn* three are de-scribed, two of which are attached to velar words. However, the form of the suffix is related to the type of velar or velarised consonant that occurs *in* the word:

> "The corroboration [of the imperative] consists of the addition of the word qil, i.e. a q vocalised with i and a silent l, if the verb contains one of the consonants of elevation. If it contains a consonant that resembles those, then change the q into a ġ [i.e. ġil] and if there is nothing of this [type] in the verb, then change the q into a k̯[101] [i.e. k̯il]." (*wa-tawkīduhu bi-ʾan yuzād [sic] ʿalayhi lafẓa qil wa-hiya qāf maksūra wa-lām sākina ʾin kāna fī l-fiʿl ḥarf min ḥurūf al-istiʿlāʾ. wa-ʾin kāna fīhi ḥarf min šibhihā fa-ʾabdil al-qāf ġaynan. wa-ʾin lam yakun fī l-fiʿl šayʾ minhā fa-ʾabdil al-qāf kāfan*, *Qawānīn* 8,9f; similar statements on 11,21; 12,9; MG 44ʳlt; 44ᵛbm; cf. 3.1.6.3.)

---

[100] In fact, the ideal suffix should have eight forms, depending on a third parameter, i.e. [+/- rounded vowel]—which is dealt with in section 3.3.2.2—viz.,

	voiced rounded/unrounded	voiceless rounded/unrounded
back	ġul / ġil	qul / qil
front	k̯ul / k̯il	kul /k̯il

[101] Elsewhere in *Qawānīn* described as [g].

It is not clear to what extent the principle explained in the quotation really holds. One would rather say, as shown in the scheme above, that the attachment of ġil or qil depends on whether the last consonant of the stem is voiced or voiceless, rather than on the occurrence of a ve-larised consonant *in* the stem.

But let us see where the text leads us to if we nevertheless interpret this statement as referring to the final consonant of the stem. Then, the author of *Qawānīn* in the first place says that all elevated conso-nants—which for the Turkic in his concept cover ṣ, ṭ, ḍ, q, ḥ and ġ—require qil. In the second place, ġil follows after "the consonants that resemble those"—which refers to at least, velarised r, l, and z, and, in practice any velarised neutral consonant, including velarised *ʾalif* /ˌ/. The third rule collects the rest; palatal words simply take kil. In regard to the first rule: includes ḍ and ġ that are both voiced and not likely to take a voiceless suffix. For the latter there is no mention of a possibility to choose between ḳil [ġil] and kil, and about conditions that could be phrased similar to those of ġ (discussion continued below). (A similar, incomplete tripartite division recurs in *Tuḥfa* [53ᵛ8] for the future tense suffix qāy/ġāy versus kāy.) In *ʾIdrāk*, there are only two basic forms, i.e. ġil and kil. These incomplete representations occur in spite of the fact that [g] is well described and listed as a basic consonant. Apparently, when it comes to suffixes, the shift [k - g] is a mere prosodic feature, which cannot be reflected in orthography and there-fore is not worthy of explicit mention.

In regard to DI, depending on which set of emphatic consonants the respective sources take, most sources should be able to show at least three different suffixes in orthography. Only one, i.e. the voiced ve-larised position (which is only used in *Qawānīn*; cf. 3.6.3.4) should then be indicated with a label.

On this position, indeed, generally (neutral) d is used, viz., ʾal-dī [al] 'he took' (134,3); ʾur-ul-dī [uruḷ-] 'he was beaten' (134,9) and ṭur-dī 'he stood up' (*ʾIdrāk* 121,17; with vowel harmony 111,13). In one instance, however, *ʾIdrāk* has ṭ, i.e., ṭur-ṭu-q 'we stood up' (153,20). As far as I could check, in no instance d is labelled as 'velarised'. Only one source, i.e. *Qawānīn* apparently observes the extended consonant as-similation (discussed below in next subsection).

In regard to the voiceless velarised position, ʾAbū Ḥayyān writes,

> "The t is substituted by a ṭ; they say ṭuṭ-ṭī ['he held']; its base form is ṭut-tī." (*tubdal [at-tāʾ] ṭāʾan qālū ṭuṭ-ṭī ʾaṣluhu ṭut-tī, ʾIdrāk* 117,2.)

Further examples are **baq-ṭi-m** 'I looked' (139,11) and **ṣaṭ-ṭim** 'I sold' (*'Idrāk* 117,19). This form of consonant assimilation is also observed occasionally in the ablative case ending DAn, where the base form (*'aṣl*) is **d-**; "**d** is substituted for **ṭ**, and subsequently assimilates with [preceding] **ṭ**." *'ubdilat ad-dāl ṭā'an wa-'udġimat fī ṭ-ṭā', 'Idrāk* 143,6.)

Nevertheless, although we have seen that the appropriate instruments are available, the principles of the extended consonant assimilation are far from being systematically observed in the sources, even where this is possible in regard to the available orthographic tools. After the final voiceless consonant of a velarised verb sometimes **-t** occurs on the voiceless velarised position, e.g., **ṣangiš-tī** 'he battled (with someone)', but more often **-d**, **ṣang-dī** [sanč-] 'he stabbed' (111,3) and **'aq-dī** [aq-] 'he flowed' (*'Idrāk* 17), which is the 'base form' of the suffix.

3.6.3.4 A closer look at consonant assimilation in *Qawānīn*

Only in *Qawānīn* the extended consonant assimilation is apparently structurally and fully maintained in orthography. However, here the regular equivalents of the terms 'voiceless' and 'voiced' are not used either. The distribution of the four markers of the past tense, i.e., **tī**, **dī**, **ḍī** and **ṭī**, runs according to four rules that are summarised as follows, viz.,

1 If the final consonant of the stem is an unvocalised **b** [p] **t**, **ǧ** [č], **s**, **k**, **š** or **q**, then the marker of the past tense is **tī**. Several examples of each instance are given, for example, **tap-tī** [tep-ti] 'he kicked', **'up-tī** 'he kissed', **baq-tī** [baqti̊] 'he looked', **buk-tī** 'he bent'.
   Note that here **b** and **ǧ** are to be interpreted as voiceless, since words cannot end in **-b/-ǧ**. Furthermore, the consonants mentioned here are exactly those which Kāšġarī calls *ḥurūf aṣ-ṣalāba*, i.e. the consonants after which the suffix has **t-** (see above 3.6.3.2).

2 If the final consonant is **d, r, z** or **l** (non-velarised), **n, y**, or if the last consonant is vocalised, the marker of the past tense is **dī**.[102] Examples, e.g., **kir-dī** 'he entered', **kal-dī** 'he came', **min-dī** 'to mount'.

---

[102] Here **m** is not listed, which may be an error. On the other hand, no examples with a verbal stem ending in **-m** are given, and the author of *Qawānīn* may have assumed that there are no Turkic verbs ending in **-m**. In fact, of the few verbs ending in **-m**, *Qawānīn* gives only one, i.e., **kum** 'bury!' (76,1).

Some of the verbs given, one with -d, e.g., ʾizda-dī 'he sought', and -y, viz., ya-dī 'he ate' end in a vowel; for this reason they can hardly serve as appropriate examples. Furthermore, **bar-dī** 'he went' and ʾur-dī 'he beat' are given here, although they are characterised as 'velarised' elsewhere (22,15; 22,16; 8,16). In these instances one would therefore expect -ḍī instead.

3 If the final consonant is ẓ, ṛ, ḷ (velarised), or ġ, then the marker of the past tense is ḍī. Some of the examples given are, e.g., **yaẓ-ḍī** 'he wrote', ʾal-ḍī 'he took'. To these he adds w, exemplifying with **yū-ḍī** /yuw-ḍī/ 'he washed'.

4 If the final consonant of the verb is ṣ or ṭ, the marker of the past tense is ṭī, e.g., **yuṭ-ṭī** 'he swallowed', ʾaṣ-ṭī 'he hung', and **quṣ-ṭī** 'he vomited'.[103]

Close examination of the rules themselves shows that they do not account for the occurrence of ṭī after neutral consonants, such as [p], [č] and [š] in back words, such as -ṭ in ṭap-ṭi [tap-tï] 'he found' (10,10).[104] This is, however, acknowledged by the author, and he gives a partial solution:

"The marker of the past tense in this word is ṭ, even though the final consonant of the verb is not one of [the consonants] I mentioned; [I refer to the first] ṭ which is meant as if it were the last consonant [of the stem]. [The ṭ] is a consonant of elevation, the b is a single consonant, not vocalised; a single unvocalised consonant is not considered an insuperable border. Therefore a marker of the same kind [i.e. a velarised consonant] is used according to the principle of the last [i.e. 4], with the intention of agreement and alleviation." (fa-ʾinna ʿalāma l-maḍiy fīhā ṭāʾ wa-laysa ʾāḫir al-fiʿl šayʾan mimmā ḍakartu qultu [?] aṭ-ṭāʾ hunā bi-maʿnā al-ʾāḫir wa-hiya min ḥurūf al-ʾistiʿlāʾ wa-l-bāʾ ḥarf wāḥid sākin wa-laysa l-ḥarf al-wāḥid as-sākin bi-ḥāǧiz ḥaṣin fa-ǧīʾa bi-l-ʿalāma min ǧinsihā li-kawnihā fī ḥukm al-ʾāḫir li-qaṣd al-munāsaba wa-t-taḫfīf, Qawānīn 10,10ff.)

In other words, the effect of the initial ṭ reaches to the end of the word, without being hindered by b [p] and makes -ṭi acceptable. Even though he alludes to the velarising effect of the initial ṭ, he does not refer to the voicelessness of the final [p]. Note, furthermore, that the au-

---

[103] In neither one of these cases, the labial vowel harmony, [o/u ~ u] and [ö/ü ~ ü], is observed.
[104] The rules do not predict a voiced velarised suffix in back verbs ending in n or y either.

thor does not really amend his rules in order to give a more general and fuller coverage of all possible suffixes; he merely gives an addition, which has the appearance of an improvisation for a specific case.

The four rules still do not account for the ending of ʾaṇla-ḍī 'he understood' (29,6) and other verbal stems that end in a vowel. To complicate matters further, a number of verbs that end in a vowel are separately inserted between the rules, e.g., ʾuqu-dī 'he read' (9,14)[105], tuša-dī [töše] 'he spread out' (9,11), two of them are actually presented as regular examples of verbs ending in -d, viz., ʾizda-dī 'he sought' and ʾanda-dī 'he called' (9,17), and one as a verb ending in -y, i.e. ya-dī 'he ate (10,1).

In sum, even though the four rules constitute a fair starting point for the determination of the appropriate suffixes, they do not give a real insight into the underlying distributional principles; i.e. they have no full explicative and predicative value. The inclusion of verbs ending in vowels indicate that even *Qawānīn*'s author in fact does not in all instances relate the form of the suffix to the final consonant of the verbal stem, nor does he its velarity or palatality take into account.

3.6.3.5 Summary

My conclusion with regard to consonant assimilation is that in the source not all rules of the ideal extended consonant assimilation are observed. In one source, i.e. *Qawānīn*, consonant assimilation is based on a set of four potential markers whose distribution does not—as one would expect—depend on general Arabic principles that are interpreted as equivalent to voiced - voiceless and palatal(ised) - velar(ised). Except *Qawānīn*, the sources take dī as the base form from which the others are derived.

The sources in general deviate from the Arabic practice in the sense that, even though they are in most cases incomplete, their reflections of Turkic are probably closer to the phonetic reality than written Arabic usually is. They take pains to express a prosodic feature which as a rule is only marginally observed in Arabic orthography.

4. CONCLUSIONS

The differences between the sources in their enumeration of basic, secondary and tertiary consonants, the way they described them depends on two factors. The first is their factual perception of a given

---

[105] But cf. ṭ in ʾuqu-ṭ 'make read!' (69,10).

sound, and each author's subsequent choice of relating the new sound
to his own concepts of the Arabic phonemes and allophones. The
speech sounds one perceives as meaningful are, of course, related to
one's concept of what type of sounds are likely to be meaningful. The
second is the importance or the status the author assigns to the
phoneme. In other words, is a given sound considered a phoneme, i.e.
a meaningful sound that has to be indicated throughout, or is it re-
garded it as a mere allophone that occurs as a prosodic feature, perhaps
worthy of signalising and describing as an isolated phenomenon, but
nothing more.

In relation to these three factors we can say in the first place that
the Arab authors did indeed perceive all meaningful phonemes of
Turkic, vowels and consonants alike, and even describe sounds that are
considered allophones in Western descriptions of Turkic, e.g., [ḫ]. In
this respect there are three types of Turkic phonemes. The first type are
those that are identical with existing Arabic sounds, e.g., [k, q, z, m,
etc.]. Second, there are phonemes that can be described in terms of a
non-standard , e.g., [p, č, g, z, ḷ, r], or dialectal pronunciation of
Arabic sounds, e.g. [ḍ, ẓ]. This class includes neutral consonants that
are labelled in passing as 'velarised', e.g., *hamza*, d, t, etc.. The third
type are foreign phonemes for which new descriptions have to be in-
vented, in terms of the basic Arabic phonemes, not as elaborations of
the allophones. These phonemes and allophones are not necessarily
identical with those posited for Turkic in Western analysis, such as [ɢ]
In relation to the second factor we can state that the status attributed
to a given phoneme is determined by whether it can be correlated with
an existing Arabic phoneme. This is evident in the only partial observa-
tion of most non-standard phonemes in the texts, e.g., [p, č, g], and
the incomplete reflection of vowel harmony and consonant harmony
in suffixes.

Vowel harmony is reflected only in asmuchas it involves a change
that can be expressed by choosing another consonant. Furthermore, it
appears that the sources did not have a real concept of voiced vs voice-
less to which they could relate the distribution of the suffixes. In regard
to the non-Arabic vowel phonemes, such as [ö, o], it is evident that al-
though they were occasionally indicated by means of labels, they are
not systematically described.

The differences that arise as a result of taking the Uygur alphabet as
basic (Kāšġarī) on the one hand, and the Arabic alphabet on the other
are obvious. However, the choice for the Arabic alphabet does not au-
tomatically lead to the same solutions for expressing certain Turkic

sounds. The authors apparently had various concepts of a number of Arabic phonemes, e.g., [ḍ, ẓ] that do not match the prescriptions in the Arabic sources. Here they most likely base themselves on dialectal pronunciations instead.

phon.value	Hilya	Diwān	ʾIdrāk	Qawānīn	Tarǧumān	Šuḏūr // Bulġa	Tuḥfa
1 [b]	B bāʾ	B bāʾ (sahīha)	bāʾ hālisa	B bāʾ hālisa			bāʾ
2 [p]	S bāʾ mufaḫḫama kaʾannaka tanfaḫu bi-hā min ḫāriǧ aš-šafatayn	S bāʾ šulba	bāʾ mašūba	B/S bāʾ mašūba bi-fāʾ mušānsana	S bāʾ mufaḫḫama bayn al-bāʾ wa-l-fāʾ	pāʾ fārisiyya bayna bāʾ wa-fāʾ	S bāʾ mašūba bi-l-fāʾ
3 [v]	S 1 fāʾ mušamma ʾilā l-wāw / S 2. fāʾ mufaḫḫama	B 1. fāʾ (rakika) / 2. bayna maḫraǧ al-fāʾ al-ʿarabiyya wal-bāʾ as-sahīha					
4 [f]	S fāʾ	S fāʾ ʿarabiyya					
5 [m]	B mīm	B mīm	mīm	B mīm			mīm
6 [w]	wāw	B wāw	wāw	B wāw			wāw
7a [d]	dāl		dāl	B dāl			
7b velar. d				B ḍād	P dāl mušamma bi-t-tāʾ	// B dāl fawqahā tāʾ yuntaq baynahumā	
8a [t]	tāʾ	B tāʾ	tāʾ	B tāʾ			tāʾ
8b velar. t	B 1 tāʾ / P 2 tāʾ mufaḫḫama		tāʾ	B ṭāʾ			ṭāʾ
9 [d]	dāl	B dāl	P dāl in Bulġar (9)				
11a [z]	zāy	B zāy	zāy	B zāy hālisa			zāy
11b velar. z				B zāy mašūba bi-ṣ-ṣād	S zāʾ mufaḫḫama, tušbih talaffuz nabat ʾahl aṣ-ṣaʿīd bi-ḍ-ḍād		
12a [s]	sīn	B sīn	sīn	B sīn			sīn
12b velar. [s]	ṣād		ṣād	B ṣād			ṣād
13 [ǧ]	ǧīm	S ǧīm ʿarabiyya	ǧīm hālisa	B ǧīm hālisa			ǧīm

phon.value	Hilya	Dīwān	ʾIdrāk	Qawānīn	Tarǧumān	Ṣudūr // Bulġa	Tuḥfa
14 [č]	1 ǧīm mušamma ʾilā šīn / S 2 ǧīm mufaḫḫama	čīm	ǧīm mašūba	S 1 ǧīm mašūba bi-š-šīn / B 2 ǧīm mušanšana	S ǧīm mufaḫḫama bayn aš-šīn wa-l-ǧīm, tušbih talaffuẓ bilād ba ʿibakk bi-l-ǧīm	čīm fārisiyya bayna l-ǧīm wa-š-šīn	ǧīm mašūba bi-š-šīn
15 [ž]	S 1 žāʾ ... bayna maḥ-raǧ al-ǧīm wa-š-šīn, / S 2 zāy mufaḫḫama, tušbih talaffuz nabaṭ ʾahl as-saʿīd bi-ḍ-ḍād	S zāy mutawallida bayna mahraǧay az-zāy wa-š-šīn					
16 [š]	šīn	B šīn	šīn	(B šīn)			šīn
17 [y]	yāʾ	B yāʾ	yāʾ	B yāʾ			yāʾ
18a [g]	S 1 kāf kamā yantuqu l-badawī bi-ʾaqdi l-qāf / P 2 kāf mayyita (ḍāʾiba)	S 1 kāf rakika / 2 kāf mutawallida bayn al-kāf wa-l-qāf	kāf badawiyya	B kāf badawiyya (maʿ qūda)	S 1 kāf mutawallida bayn mahraǧay l-qāf wa-l-kāf / P 2 wa-kāf ka-qāf al-ʿarab al-ma qūda	B kāf ʿaǧamiyya	S kāf ma qūda
18b [ɡ]	P kāf mušamma ʾilā l-ǧayn	kāf al-ǧunna mutawallida bayna l-ǧayn wa-l-qāf wa-n-nūn wa-l-qāf		B ʿayn bayna l-ʿayn wa-l-ǧayn al-mu ǧama	P kāf bayna l-ǧayn wa-l-kāf muǧannana, mahraǧuhā min al-ḥayšūm		
19 [k]	kāf	B 1 kāf ṣaḥīḥa / B 2 kāf ṣulba	kāf ḫāliṣa	B kāf ḫāliṣa			kāf
20a [ŋ]	P kāf mayyita mudaġġama bi-n-nūn	P al-ḥarf al-ḥayšūmī	nūn ḥayšūmiyya	B 1 nūn ḥayšūmiyya / S 2 nūn maġnūna wa-hiya llatī taḫruǧ min al-ḥayšūm	P nūn fīhā šayʾ min al-ġunna, maḫ-raǧuhā min saaf al-ḥalq ʾilā raʾs al-ʾanf	S nūn ḥayšūmiyya	
20b velar [ŋ]	P nūn nuṭiqat... mušamma ʾilā l-ǧayn (77)			P ǧayn maġnūna			

phon.value	Ḥilya	Dīwān	ʾIdrāk	Qawānīn	Tarǧumān	Ṣuḍūr // Bulġa	Tuḥfa
21a [q]	B -qāf-	B qāf	qāf	B qāf			qāf
21b [b] allophone of [q]				P 1 qāf maṣūba bi-ǧayn mu ǧama / P 2 qāf bayna l-ǧayn wa-l-qāf wa-hiya ʾaqrab ilā l-qāf / P 3 qāf bayna l-qāf wa-l-ḥāʾ / P 4 ḥāʾ maṣūba bi-ḥāʾ			
22a [ǧ]	B ǧayn	S ǧayn	ǧayn				ǧayn
22b [w] all. of [ǧ]	P ǧayn mayyita (ḍāʾiba)						
23 [h]	B ḥāʾ	B ḥāʾ		B ḥāʾ			
24 [n]	B nūn	B nūn	nūn ḫāliṣa	B nūn ḫāliṣa			nūn
25a [l]	B lām	B lām	lām	B lām ḫāliṣa			lām
25b velar l	P lām mufaḫḫama			B lām mufaḫḫama	S lām mufaḫḫama	lām mufaḫḫama	
26a [r]	B rāʾ	B rāʾ	rāʾ	B rāʾ			rāʾ
26b velar. r				P rāʾ mufaḫḫama			
27 [h]	B hāʾ	B hāʾ		B hāʾ			hāʾ
28a [ʔ]	B ʾalif (=hamza)	B hamza-alif	hamza	B hamza			ʾalif - hamza
28b velar. [ʔ]		P tafḫīm al-hamza		P hamza mufaḫḫama			
Total	31	28	24	32			22

Appendix to Chapter Three: an overview of the descriptions of consonantal phonemes

CHAPTER FOUR

CASES AND MARKERS

INTRODUCTION

This chapter is meant as an introduction to Chapters Five and Six and deals with two main subjects. In Section A it discusses the concept of case in Arabic linguistic theory, and, second, some typological characteristics of case in Arabic and Turkic which may have formed the base for this concept. In addition, some preliminary statements are made in regard to the suitability of Arab grammarians' approach to language, especially as far as their their method of morphological segmentation concerned.

In Section B this chapter gives an account of the way the sources apply the Arabic terms *kalima* and *lafẓa* to Turkic, and the various means for expressing synonymy between Turkic and Arabic.

PART A 1. *ʿALĀMA* AND *ʾIʿRĀB*

In *Tuḥfa* the author proclaims: "They [sc. the Turks] do not have anything like *ʾiʿrāb* in their speech" (*laysa fī kalāmihim šayʾ min al-ʾiʿrāb*, 50ᵛ13; another remark to this effect is found in Ibn Fāris' *Ṣāḥibī*, 161,12). If we accept *ʾiʿrāb* as the Arabic term for 'declension' or 'case', this statement seems to contradict the communis opinio, for Turkic languages do have a system of case endings, at least in the Western understanding of case. How, then, can we explain this remark?

In this section we intend to show in detail how cases are defined by the Arab grammarians themselves, and point to the differences and similarities with the Western traditional conception of case. We shall elucidate some of the terms that play an important role in their discussion of the Turkic case system and discuss especially the terms *ʿalāma* 'marker' and *ʾiʿrāb* 'declension'. These terms are closely connected, for, as we shall see, *ʾiʿrāb* is regarded as a coherent set of *ʿalāmāt*.

In Arabic treatises the term *ʿalāma* 'marker' (pl. *ʿalāmāt*) is applied to various morphemes, whose common characteristic is the fact that

they do not govern. I wish to stress that in this context the notion of governor should not be interpreted as being identical with that in Generative Linguistics, but rather as a similar notion (see Section 1.2 for further discussion of *ʿāmil*).

Although the Arab grammarians themselves do not specify this as such, it is possible to discern two categories of *ʿalāmāt*. In the first place there is a category of markers that indicate a change in the meaning of the word, without the presence of a governor *(ʿāmil)* in either the surface structure *(lafẓ)* or the underlying structure *(taqdīr)*. The second category comprises markers that are caused by a governor *(ʿāmil)*. This second category of markers coincides with the Arabic concept of *ʾiʿrāb*. This is worked out in more detail below.

1.1 *Markers without a governor*

The first category of markers comprises elements that occur to indicate a certain meaning, without a governor *(ʿāmil)* being implied. In other words, an *ʿalāma* of this type is a marker for a certain meaning; it has an inherent semantic load.

To this group, for example, belong the so-called *tāʾ marbūṭa* 'the connected *t*' that serves to differentiate between feminine and masculine adjectives and nouns. This *t*, or rather the suffix *at*, is added to a masculine adjective or a noun. The case ending of the word follows after the *t*. Thus, *malik-un* 'king' is changed into *malik-at-un* 'queen', *ǧamīl-un* 'beautiful' (m) *ǧamīl-at-un* 'beautiful' (f) and *ḍārib-un* '(the) beating one (m)' vs. *ḍārib-at-un* '(the) beating one (f)'. The Arab grammarians call this *t* more specifically *ʿalāma at-taʾnīṯ* 'marker of the feminine' *(Irtišāf I 293)*.

Another *ʿalāma at-taʾnīṯ*, with the same status, is the *t* in the perfect verb, such as in (1). The verb *ḍaraba-t*, therefore, contains an agent, albeit not the *t*, but one hidden in the verb itself:

1   *ḍaraba-t*   *zayd-an*
     beat/PAST-f/3sg   zayd-ACC
     'She beat Zayd.'

Contrary to what one would expect, in Arabic theory the final *t* is not the agent *(fāʿil)* of the verb *qāma-t* in (1). The reason for this is the fact that an agent, e.g., *hind* can be inserted, viz.,

2   *ḍaraba-t*   *hind-un*   *zayd-an*
     beat/PAST-f/3sg   hind-NOM   zayd-ACC
     'Hind beat Zayd.'

If the *t* represented the agent too, the verb would have two agents, i.e. *hind* and *t*, which is impossible. As a consequence, the *t* can merely have the status of an *ᶜalāma*. Because this *t* remains vowelless, it is called *tāʾ sākina* 'silent *t*'. Širbīnī gives the following explanation:

> "The *t* is the marker of the feminine; the pronoun hidden in it [sc. the verb] with the implicit meaning of *hiya* ['she'] is an agent on the syntactic position of the *rafᶜ* because of the preceding verb and it does not exhibit declension" (*at-tāʾ ᶜalāma li-t-taʾnīṯ wa-ḍ-ḍamīr al-mustatir fīhi al-muqaddar fīhi bi-hiya fāᶜil fī maḥall ar-rafᶜ bi-l-fiᶜl allaḏī qablahu lā yaẓharu fīhi ʾiᶜrāb*, Širbīnī 162).

In this case the agent of the verb *ḍaraba-t* is 'hidden', just like the pronoun of the agent is 'hidden' in *ḍaraba zayd-an* 'he hit Zayd'. In other words, the *t* in the verb *ḍaraba-t* 'she beat' is the *ᶜalāma at-taʾnīṯ* 'marker of the feminine', but it is not the agent, for the agent of the sentence is *mustatir*, 'hidden', in the verb. For the Arab grammarians the fact that the *tāʾ marbūṭa* is added to nouns and the *tāʾ at-taʾnīṯ* to verbs is not significant. For them it is basically a single morpheme that is applied to both verbs and nouns.

The *w* in the verbal form *katabū* /katabuw/, on the other hand, is not an *ᶜalāma* but the agent *(fāᶜil)* of the verb in sentences like *ḍarabū zayd-an* 'they beat Zayd':

> "...*ḍaraba* 'struck', spelled with *a* after the *ḍ*, is a past tense verb, and *ū* [/uw/] 'they' (m) is an agent on the syntactic position of *rafᶜ* through the preceding verb and does not exhibit declension." (*fa-ḍaraba bi-fatḥ aḍ-ḍād fīᶜl māḍin wa-l-wāw fāᶜil fī maḥall rafᶜ bi-l-fiᶜl allaḏī qablahu lā yaẓharu fīhi ʾiᶜrāb*, Širbīnī 162.)

The basic reason for this argumentation is that Arabic sentences typically start with a verb. If the agent is a plural noun the verb agrees in gender, but not in number, e.g., *ḍaraba l-muᶜallimūna zaydan* 'the teachers beat Zayd', but not **ḍarabū l-muᶜallimūna zaydan*. If the agent is a pronoun, the number of the agent is expressed in the verb, e.g., *ḍarabū zaydan* 'they beat Zayd'. According to the Arabic theory, the final *w* in /ḍarabuw/ *ḍarabū* is the agent on the position of the nominative *(rafᶜ)*, although it does not take the appropriate case ending.

Apart from the function of agent in some verbal forms, in nouns the *wāw* serves as the marker for the plural. In this case it is called *ᶜalāma ǧamᶜ al-muḏakkar as-sālim* 'marker of the sound masculine plural'. An example is *muᶜallim-ūna* /muᶜallimuwna/ 'teachers', plural of *muᶜallim-un*. In other syntactic functions, such as *naṣb* and *ǧarr*,

the /w/ is replaced by a /y/, i.e. /iy/, viz., *muʿallimīna* /muʿallimiyna/.
In the dual the *ʾalif* /"/ has a similar function (*Irtišāf* I 252):
muʿallimāni /muʿallima"ni/. In these forms, the *wāw*, *ʾalif* and *yāʾ*, re-
spectively, apparently combine two functions simultaneously, namely
plural and declension (see Section 1.2 for further discussion).

1.2 *Markers with a governor:* ʾIʿrāb

The second category of *ʿalāmāt* comprises the endings that are caused
by a governor. In Arabic grammar this system of endings is called
ʾiʿrāb.

1.2.1 The assignment of declensional endings

The endings of *ʾiʿrāb* indicate the fact that the word is governed by an
ʿāmil 'governor'. *ʾIʿrāb* becomes explicit in four endings: *u*, *i*, *a* and
zero (*ǧazm*), the lack of any of these. Nouns and adjectives may get the
endings *u*, *i* and *a*, and verbs may end in *u*, *a* or zero.[1]
    The form and use of the endings of nouns is as follows:

3   *zayd-un*   *karīm-u-n*
        zayd-NOM   kind-NOM -INDEF
        'Zayd is kind.'

4   *kitāb-u*   *zayd-in*
        book-NOM /DEF   zayd-GEN
        'Zayd's book.'

5   *ḍarab-tu*   *zayd-an*
        beat/PAST -I   zayd-ACC
        'I beat Zayd.'

According to Western grammar the endings in *zayd* are equivalent
with 'nominative', 'genitive', 'accusative', respectively.
    Unlike the concept of case in western theory, *ʾIʿrāb* is not limited
to nouns only; it also occurs in verbs in a specific verbal form
(*muḍāriʿ*), which may end in *u*, *a* or *ǧazm* 'zero':

6   *ya-ḍrib-u*   *zayd-un*
        he-beat/IMPF -IND   zayd-NOM
        'Zayd beats.'

---

[1] In the analysis of Arab grammarians, the nunation, i.e. final *n* in indefinite nouns,
e.g. *kitāb-u-n* (book-NOM -INDEF), is no part of *ʾiʿrāb* itself.

7	kay	ya-ḍrib-a	zayd-un
	so that	he-beat/IMPF-SUBJ	zayd-NOM
	'so that Zayd beats.'		

8	lam	ya-ḍrib-ø	zayd-un
	NEG	he-beat/IMPF-APOC	zayd-NOM
	'Zayd did not beat.'		

It is therefore difficult to find an English equivalent for *ʾiʿrāb* that covers this broad range. With regard to nouns *ʾiʿrāb* is usually translated in Western studies of Arabic with 'declension', and in the case of verbs 'mood' is often used.

A definition of *ʾiʿrāb* is given by ʾAbū Ḥayyān:

"*ʾIʿrāb* is a change at the last consonant of the word [because of a governor] which affects it, and the vowels are markers of *ʾiʿrāb* and an indication of it." (*al-ʾiʿrāb huwa taġyīr fī ʾāḫir al-kalima... [li-ʿāmil] daḫala ʿalayhā nafsihā wa-l-ḥarakāt ʿalāmāt al-ʾiʿrāb wa-dalāʾil ʿalayhi, Irtišāf* I 413, 11.)

Western scholars equated these verbal endings with the Latin/Greek concepts of 'mood': 'indicative', 'subjunctive' and 'apocopate', respectively.

A further main difference between the two conceptions is that in Western linguistic theory each of these forms has a different meaning, whereas the Arab grammarians concentrate on the morphological changes instead, the meaning residing in the particle that acts as a governor, thus causing the change.

In spite of their importance for indicating governance, the endings of *ʾiʿrāb* may be absent due to morphonological rules or in certain prosodic conditions. The word is supposed to show the effect of a governor, but does not do so. This does not imply a change in the relations of governance, which can be shown in the *taqdīr*, the underlying level of the phrase. In the first place this includes nouns whose stem ends in one of the so-called 'weak' consonants, i.e. *w* and *y*. An example of this is the noun *qāḍin* 'judge', having an ending in *-in* in both nominative and genitive (the accusative is 'regular': *qāḍiyan*). According to Arabic grammar its morphological structure is /qaʺḍiyun/ in the nominative, and /qaʺḍiyin/ in genitive, respectively, showing a regular ending in both cases (for an elaborate analysis of the Arabic theory of morphology and phonology, see Bohas 1982).

In the second place endings of *ʾiʿrāb* may be omitted under certain prosodic conditions. The Arab grammarians were well aware of this phenomenon which often occurs in normal speech, but is especially

frequent in poetry and readings of the Qur'ān. Ibn Ǧinnī, for example, gives two occasions in which such a deletion may occur; as a pausal form *(waqf)* and in connected speech *(waṣl)*:

> "With regard to the pausal form, one may pause on every fully inflected noun which has nunation in the case of both *rafʿ* and *ǧarr*, its *ʾiʿrāb* and nunation are deleted... if it is put in *naṣb* the nunation is substituted by an *ʾalif* [= ā]." *(fa-ʾammā l-waqf, fa-kull ism mutamakkin munawwan waqafta ʿalayhi fī rafʿihi ʾaw ǧarrihi ḥaḏafta ʾiʿrābahu wa-tanwīna-hu... fa-ʾin naṣabta ʾabdalta min tanwīnihi ʾalifan, Sirr Ṣināʿa* 518, 10ff; on this subject, see also Širbīnī in Carter 1981:37.)

In a pausal form, *zayd-un* and *zayd-in* change to *zayd*, whereas *zayd-an* changes to *zayd-ā* (cf. *Mufaṣṣal* 162; Wright 1986 [1898] ii 368).

This type of deletion, however, is no indication of *ǧazm*, for it concerns the deletion of an element from the surface structure; the deleted element subsists in the underlying structure *(taqdīr)*. Moreover, unlike *ǧazm* in verbs, in these cases the absence of the ending of *ʾiʿrāb* is not caused by a governor.

When *ʾiʿrāb* is omitted from the surface, or when words have an ending that does not show *ʾiʿrāb*, word order may take over its function of distinguishing between, for example, agent and direct object. Ibn Ǧinnī describes this phenomenon:

> "...one may say 'Yaḥyā hit Bušrā', but a distinguishing declension is not found in it, and likewise in similar expressions. [The answer is] when something like this happens, viz., that the case is not visible in the surface realisation, speech is subjected to what takes over the function of declension, namely preposing the agent and postponing the object." *(...fa-qad taqūlu ḍaraba yaḥyā bušrā, fa-lā taǧidu hunāka ʾiʿrāban fāṣilan wa-kaḏālika naḥwahu. qīla: iḏā ttafaqa mā hāḏihi sabīluhu mimmā yaḫfī fī l-lafẓ ḥāluhu, ʾalzama l-kalāma min taqdīm al-fāʿil wa-taʾḫīr al-mafʿūl mā yaqūmu maqām bayān al-ʾiʿrāb, Ḫaṣāʾiṣ* I 35,7ff.)

Ibn Ǧinnī says here that the nouns *yaḥyā* and *bušrā* both end in *ā*, and do not show the markers of declension, i.e., nominative for *yaḥyā* and accusative for *bušrā*, respectively. For phonological reasons the case endings of /yaḥyayun/ and /bušrayan/ are omitted, and they end in *ā.*

The fact that the endings of *ʾiʿrāb* may be omitted or fuse for phonological reasons does not lead to a change in the hierarchy of elements within the phrase; the respective functions of the elements remain the same, even though the relations of governance *(ʿamal)* are not visible in the surface structure.

1.2.2 Governance

The concepts of governance ('*amal*) and governor ('*āmil*) are indeed crucial for the understanding of *ʾiʿrāb*. An '*āmil* is any element, especially a verb (*fiʿl*) or a particle (*ḥarf*, pl. *ḥurūf*), that causes a morphological effect on another word. In terms of Arabic theory, though, the resulting ending is not part of the word's morphological structure but is part of the domain of syntax (*naḥw*). Some particles govern nouns only (e.g. *bi* 'with': *bi-s-sayf-i* 'with the sword', and *ʾinna*, a particle of topicalisation), whereas others only govern verbs (e.g., *kay* 'so that'). An example of a *ḥarf* governing a noun and causing it to take the ending *-a* is *ʾinna*:

9   *ʾinna*   *zayd-an*   *karīm-u-n*
        TOP   zayd-ACC   kind-NOM-INDEF
        '[topicalisation] Zayd is kind.'

After *bi* the noun gets the ending *-i*, and after *ʾinna* it takes *-a*. A small number of *ḥurūf* may govern both verbs and nouns, causing verbs (in the *muḍāriʿ* form) to take the ending *-a* instead of *-u*. An example of a verb governed by a *ḥarf* is:

10   [*daʿaw-tu*   *zayd-a-n*]   *kay*   *ya-ǧīʾ-a*
        [called/PAST-1sg   zayd-ACC]   so that   he-come/
                                                                   IMPF-SUBJ

        'I called Zayd to come.'

11   *lam*   *ya-ḍrib-ø*   *zayd-un* (= 8)
        NEG   he-beat/IMPF-APOC   zayd-NOM
        'Zayd did not beat.'

The governor in question may be absent from the surface structure and reconstructed in the underlying structure (*taqdīr*) only. An example of such a governor in *taqdīr* occurs in the case of the genitive constructions. In early Arabic grammar (Sībawayh) it was accepted that nouns could govern other nouns. Thus the ending *-in* in *zayd-in* (12a) was explained in terms of governance by the preceding noun, *bayt-u* 'house':

12a   *bayt-u*   *zayd-in*
        house-NOM/DEF   zayd-GEN
        'Zayd's house'

Later grammarians, however, had developed a different analysis of this construction. Rejecting the possibility of a noun governed by another noun, they posited a virtual particle in the underlying structure that acts as a governor on *zayd*. In this way they created a new syntactical

position of the genitive *(ǧarr)*. Thus, phrases of type (12a) are recon-
structed as follows:

12b	*baytun*	*li-*	*zayd-in*
	house	for	zayd-GEN

The effect of the virtual particle *li*, namely the ending *-in*, is visible in
surface structure. The reconstructed phrase *(taqdīr)* does not have the
status of a good or correct sentence; its only goal is to elucidate the
governance relations within the sentence (for a brief sketch of the his-
torical development of this theory, see Owens 1990: 17).[2]

An instance of an equally invisible and even abstract governor is the
notion of *ibtidāʾ*, 'the act of beginning'. This notion serves to account
for the nominative case of the topic in nominal sentences, such as (3)
zaydun karīmun 'zayd is kind', and instances such as, e.g., *zayd-un
ḍaraba ʿamr-an*, in which *zaydun* is regarded as the topic *(mubtadaʾ)*
and *ḍaraba ʿamran* as its predicate *(ḫabar)* (cf., e.g., Ayoub 1991; cf.
also Mosel 1975 on this term in Sībawayhi's *Kitāb*).

Verbs may govern too. In the first place they govern the agent and,
secondly a possible object. In this respect the fact that the verb pre-
cedes agent and object is of importance, since governance can be ex-
ercised in one direction only. In (13) the verb governs two elements,
namely *zayd* and *ʿamr.*

13	*ḍaraba*	*zayd-un*	*ʿamr-an*
	beat/PAST/3sg	zayd-NOM	ʿamr-ACC
	'Zayd beat ʿAmr'		

The verb causes *zayd-* to take *-u(n)* and *ʿamr-* to take *-a(n)*, in agree-
ment with their syntactic functions: *zayd-un* is the agent *(fāʿil)* and
ʿamr-an the direct object *(mafʿūl bihi)* to the verb, although the end-
ings cannot be taken as direct references to their respective syntactic
functions.

1.2.3 ʾIʿrāb as an abstract notion

In *Irtišāf* ʾAbū Ḥayyān refers to a discussion among the Arab grammar-
ians as to whether ʾiʿrāb is identical with the endings themselves, i.e. *u*,
i, *a* and *ḥaḏf*, or whether it is a more abstract notion for which the
endings serve as markers (*Irtišāf* I 413,7ff). In the former case the
ʾiʿrāb is *lafẓī* (formal) and in the latter case he regards it as *maʿnawī*
(abstract). In this view, the endings are mere indications of the abstract
notions of ʾiʿrāb, i.e. of *rafʿ*, *ǧarr*, *naṣb* and *ǧazm*.

---

[2] In the light of the opposition of semantic versus syntactic case, the Arabic genitive
in (12a) could well be interpreted as a semantic. See discussion in section 2.3.

Versteegh (1985) analyses the views of Arab grammarians on this is-
sue based on the accounts of the grammarian ʿUkbarī (d. 616/1219).
ʿUkbarī lists some arguments in favour of the opinion that *ʾiʿrāb* is an
abstract, rather than a formal notion. The most important of these ar-
guments is the fact that "declension uses different means—vowels,
consonants—to express identical meanings" (Versteegh 1985:153). In
other words, since a particular 'meaning'—or rather function—may be
conveyed in different ways, there is no correlation between an ending
and a specific function (see discussion below on whether a consonant
can be an indicator of *ʾiʿrāb*).

The grammarians in favour of the formal approach *(lafẓī)* hold that
if declension is not visible in surface structure, there is no reference to
the differences in the respective functions of the elements. This opin-
ion is exemplified with the sentence *ḍaraba mūsā ʿīsā* 'Mūsā hit ʿĪsā'.
Here the agent and the object have no distinctive markers, which may
lead to confusion. Therefore, the argument goes, the obvious conclu-
sion is that the endings are crucial for indicating the syntactic func-
tions of each element.

Whatever the opinion of the grammarians with regard to the exact
nature of *ʾiʿrāb*, i.e. whether the term refers to the vowels themselves
or to an abstract meaning, they agree that *ʾiʿrāb* is caused by gover-
nance, and that the effect on the endings caused by governance de-
pends on the type of governor (see discussion in Section 2).

With regard to *ʾiʿrāb* in plural and dual forms, the Arab grammari-
ans hold that the *wāw* in *muʿallimūna* /muʿallim-uw-na/ is the marker
of the sound plural (cf. Section 1.2) when the word occupies the syn-
tactic function of *rafʿ* (cf. *Irtišāf* I 262 ff.). In the dual form, e.g.
muʿallimāni /muʿallimaˮni/ 'two teachers', the *ʾalif* /ˮ/ has an equiva-
lent status. In other syntactic positions, sc. *naṣb* and *ǧarr*, the *wāw*
and *ʾalif* are replaced with a *yāʾ*: *muʿallimīna* /muʿallimiyna/ in the
plural, and *muʿallimayni* /muʿallim-ay-ni/ 'two teachers' in the dual,
respectively. Thus it seems that *wāw*, *yāʾ* and *ʾalif* have two functions:
they serve as markers of the plural, respectively the dual, and indicate
the case of the word in question.

Nevertheless, the Arab grammarians could not accept the possibility
of an element with that serves to convey more than one meaning. To
put it in Versteegh's (1985:157) terms, they rejected the "concept of a
multifunctional linguistic element". In their works, therefore, the
grammarians discuss the status of *w, y* and *ʾalif*, and the question of
whether a consonant, instead of a vowel (i.e. *w, y, ʾalif*), may serve as a

'marker of *ʾiʿrāb*'. ʾAbū Ḥayyān summarises the two opposing opin-
ions:

> "Al-Ḥalīl and Sībawayh hold the opinion that the vowels of declension
> were implied in the *ʾalif,* the *w* and the *y*... The Kufans, Quṭrub and
> Zaǧǧāǧī and a group of later [grammarians] maintain that these conso-
> nants are identical with declension" (*wa-ḏahaba l-ḥalīl wa-sībawayh ʾilā
> ʾanna ḥarakāt al-ʾiʿrāb muqaddara fī l-ʾalif wa-l-wāw wa-l-yāʾ... wa-
> ḏahaba l-kūfiyūn wa-quṭrub wa-z-zaǧǧāǧī wa-ṭāʾifa min al-
> mutaʾaḫḫirīn ʾilā ʾanna hāḏihi l-ḥurūf hiya l-ʾiʿrāb nafsuhu, Irtišāf* I
> 264).

In *Manhaǧ* (9,20-21) ʾAbū Ḥayyān takes sides with Sībawayh by saying
that these consonants (i.e. *w, y* and *ʾalif*) are 'the consonants of declen-
sion' (*ḥurūf al-ʾiʿrāb*) like the *d* in *zayd.* This implies that he regards
the glides in question as equivalent to the last consonant of the word
that 'carries' the *ʾiʿrāb,* and, hence, that he does not consider them
ʾiʿrāb themselves (*Manhaǧ* 7,21-22).[3] However, the question as to
whether there is *ʾiʿrāb* or not in these cases is not directly answered by
ʾAbū Ḥayyān. Zamaḫšarī (*Mufaṣṣal* 9ff), on the other hand, holds a
different opinion. He accepts the possibility of a *ḥarf* (i.e. *w, ʾalif* or *y*)
taking the place of one of the vowels of *ʾiʿrāb.*

   In this section we have briefly summarised two discussions between
Arab grammarians. The first discussion is about the exact nature of
ʾiʿrāb; whether it should be interpreted as a formal notion (*lafẓī*) or as
an abstract (*maʿnawī*) one. The second one deals with the question
whether consonants can be interpreted as markers of *ʾiʿrāb,* next to
vowels. We also saw that, in spite of their disagreement in these mat-
ters, the Arab grammarians agree that the basic characteristic of *ʾiʿrāb*
is that it is the result of governance of a word by another element. In
the following section it is pointed out which functions are indicated by
ʾiʿrāb.

2. THE FUNCTIONS OF *ʾIʿRĀB*

This section discusses the functions of *ʾiʿrāb.* The endings of *ʾiʿrāb* as-
signed to the nouns and verbs themselves are no direct indications of

---

[3] ʾAbū Ḥayyān attributes the remark about the *ḥarakāt muqaddara* 'virtual vowels'
to al-ʾAʿlam (= aš-Šantamarī, d. 476/1083; *Manhaǧ* 9,20-21; cf. also Versteegh
1985:160).

their respective syntactic functions; rather they mark the position the word occupies in the syntactic structure of the sentence.

2.1 *ʾIʿrāb and the markers of syntactic position*

The function of *ʾiʿrāb* is to indicate the position of the word in the syntactic structure; i.e. to serve as an indication of the type of governor that acts upon the word. A syntactic position *(mawḍiʿ)* involves a certain kind of government. In other words, *u (ḍamma)*, for example, is not an indication of the fact that a certain noun is an agent in the sentence, it merely refers to the fact that the word in question occupies the syntactic position of *rafʿ*. Because the word occupies this position, a *ḍamma* is assigned to it. To put it in more general terms, the syntactic position is assigned by governors, and any word that occupies it is obliged to take the appropriate marker. In this respect, *ḍamma (u)* is an *ʿalāma li-r-rafʿ* 'marker for the *rafʿ*' (Širbīnī 46), *fatḥa (a)* is an *ʿalāma li-n-naṣb* 'marker of the *naṣb*' (Širbīnī 60), *al-kasra ʿalāma li-l-ḫafḍ ʿkasra (i)* is a marker for the *ḫafḍ* [= *ǧarr*]' (Širbīnī 72), and *al-ḥadf yakūnu ʿalāma li-l-ǧazm* 'deletion is a marker for the apocopate' (Širbīnī 80). The terms *rafʿ* and *naṣb* are applied to verbs and nouns alike.

Of course, not all words are capable of taking every position. Verbs, for instance, may not occupy the *mawḍiʿ* of *ǧarr*, which is usually marked by the ending *i*.[4] In order to receive the *i*, other than as a result of phonological processes, the word must be in the position of *ǧarr*, which is assigned to the *muḍāf ʾilayhi* 'possessor' in a construction of *ʾiḍāfa* 'annexation' only, and a verb cannot be a possessor (cf. Zaǧǧāǧī, *ʾĪḍāḥ* 108ff.). In the same way, the position of *ǧazm* 'apocopate' cannot be assigned to nouns (cf. Zaǧǧāǧī, *ʾĪḍāḥ* 102ff.).

According to the Arab grammarians, *ʾiʿrāb* is basically *(ʾaṣl)* a characteristic of nouns, and a verb may have *ʾiʿrāb* insofar as it resembles a noun.[5] The type of nouns which a verb may resemble is more specifically the active participle *(ism al-fāʿil)*.

---

[4] Some verbs do end in *i*, though, albeit for phonological reasons only, e.g. in connected speech *lam yaktub-i l-muʿallim-u* 'the teacher has not written', in which *yaktub* is originally an apocopate.

[5] This resemblance concerns some functional resemblances, such as the ability of being specified by particles, the internal pattern of consonants and vowels (CaCCiC = the pattern of both the verbal form /yafʿil/ and the active participle /fāʿil/), and the fact that the active participle in some cases may alternate places with the *muḍāriʿ* form of the verb (cf. Owens 1988:208 for more details).

...the imperfective verb [i.e. *muḍāriᶜ* form] has a number of resem-
blances to the noun, which is basically inflectable[6], and because of these
resemblances it becomes inflectable as well. (Owens 1988:207; cf. 66-8;
Zaǧǧāǧī, *Ῑḍāḥ* 77,3.)

The property of having *ʾiᶜrāb* is thus derived from nouns, and not in-
herent in verbs.

In verbs the addition of *ʾiᶜrāb* is limited to the *muḍāriᶜ* form, ex-
cluding the *māḍī* (past tense). The *muḍāriᶜ* ending in *u* expresses what
in Western grammar is called the 'indicative mood' (cf. Wright 1986
[1898] i 51 and 57), even though the term *muḍāriᶜ* itself does not ex-
press a tense but rather an alleged morphological resemblance with
nouns. The third person singular of this form follows the pattern *ya-
fᶜal-u*. The first two vowels may vary in connection with changes in
meaning or with certain verbs, e.g. *ya-ktub-u* 'he writes' (for complete
paradigms, see, e.g., Wright 1986 [1898]). For Arab grammarians the
final vowel, *u*, is the ending of *ʾiᶜrāb*, and it may be substituted with *a*
(*fatḥa*), or disappear (*ǧazm*).

In the way *ḍamma* is a marker for *rafᶜ*—in both verbs and nouns—
rafᶜ, in its turn, is an indication of a certain function:

"each of the *rafᶜ*, the *naṣb* and the *ǧarr* is a marker of a [specific] mean-
ing." (*ar-rafᶜ wa-n-naṣb wa-l-ǧarr wa-kullu wāḥid minhā ᶜalam ᶜalā
maᶜnan*, *Mufaṣṣal* 10,19.)

Note that Zamaḥšarī does not refer to the realisation *rafᶜ*, *naṣb* and
ǧarr may have in the surface structure of the sentence. Due to certain
phonological rules and processes the original endings (*u*, *i* and *a*) may
have undergone changes and may have practically disappeared from the
surface structure. Therefore Zamaḥšarī assigns functions to the abstract
notions of *ʾiᶜrāb*, rather than to their respective realisations on the sur-
face level. For nouns, according to Zamaḥšarī, *rafᶜ* is the *ᶜalam al-
fāᶜiliyya* 'marker of agency', *naṣb* is the *ᶜalam al-mafᶜūliyya* 'marker of
objectivity', and *ǧarr* is the *ᶜalam al-ʾiḍāfa* 'marker of annexation'
(*Mufaṣṣal* 10,19ff).

No such functional definitions are elaborated for verbs:

"Mention of the types of *ʾiᶜrāb* of the *muḍāriᶜ*, they are *rafᶜ*, *naṣb* and
ǧazm; these types are no markers of meanings like the types of *ʾiᶜrāb* of
the noun..." (*ḍikr wuǧūh ʾiᶜrāb al-muḍāriᶜ hiya r-rafᶜ wa-n-naṣb wa-l-
ǧazm wa-laysat hāḍihi l-wuǧūh bi-ʾaᶜlām ᶜalā maᶜānin ka-wuǧūh ʾiᶜrāb
al-ism...*, *Mufaṣṣal* 109,8ff.)

---

[6] As a translation of the term *ʾiᶜrāb*, in this book 'declension' is preferred to
'inflection', as opposed to Owens.

Zamaḫšarī's statements are summarised in the following schemes:

nouns

	indicates		indicates	
ḍamma	→	rafʿ	→	fāʿiliyya
kasra	→	ǧarr	→	ʾiḍāfa
fatḥa	→	naṣb	→	mafʿūliyya

verbs

	indicates	
ḍamma	→	rafʿ
fatḥa	→	naṣb
ḥaḏf	→	ǧazm

Thus the category of rafʿ, for example, serves as a theoretical interme-diate between ḍamma and fāʿiliyya. The advantages of this reasoning are obvious. The ḍamma itself is not the exclusive indication of the agent—one of the main reasons, of course, is the fact that on surface level verbs, too, may end in ḍamma—and, hence, ḍamma itself cannot serve as a marker for 'agency' (fāʿiliyya). Simultaneously, in this way the broader notion of ḍamma as an indication of rafʿ in both nouns and verbs is preserved. In the case of nouns in verbal sentences, rafʿ is an indication of fāʿiliyya. In the same way ǧarr indicates ʾiḍāfa, i.e. the fact of being possessor, and naṣb is an indication of mafʿūliyya, i.e. 'objectivity'.

As can be inferred from the schemes above, no similar set of func-tions has been worked out for the declension of verbs. Zamaḫšarī also gives a reason: "...because the verb originally has no ʾiʿrāb, for it [sc. ʾiʿrāb] is [derived] from the noun" (...li-ʾanna l-fiʿl fī l-ʾiʿrāb ġayr ʾaṣīl bal huwa min al-ism, Mufaṣṣal 109,9).

Since the Arab grammarians did not have a concept of 'mood', a further elaboration would have been complicated indeed. In fact, such a concept is not needed in their system. The so-called 'moods' of the verb, rafʿ for 'indicative', naṣb for 'subjunctive' and ḥaḏf for the 'apocopate', are in their views just different versions of one verbal form: muḍāriʿ. The respective endings change only when the verb is governed by a governor of a different type.[7]

---

[7] The Western concept of 'mood' presupposes a variation in semantic content of the respective verbal forms. Such a concept is absent from Arabic theory, which relates dif-ferences to governance only.

This rather indirect relation between the 'vowels of *ʾiʿrāb*' (*ḥarakāt al-ʾiʿrāb*) and the syntactic function of the nouns they occur on does not seem to have been the common opinion of all grammarians. Ibn as-Sarrāǧ (d. 316/928), for example, equates the endings with the ab-stract meaning they represent. This becomes evident from the fact that he applies the terms *rafʿ*, *naṣb* and *ǧarr* to the endings themselves, rather than interpreting the endings as *markers* of *rafʿ*, *naṣb* and *ǧarr*, viz.,

> "*ʾIʿrāb* occurs with three vowels, *u*, *a* and *i*. When the *u* occurs and disap-pears as *ʾiʿrāb* on the last consonants of nouns and verbs, it is called *rafʿ*. When the *a* is like this, it is called *naṣb*, and when the *i* is like this, it is called *ḥafḍ* and *ǧarr*." (*wa-yakūnu [l-ʾiʿrāb] bi-ḥarakāt talāt: ḍamm wa-fatḥ wa-kasr fa-iḍā kānat aḍ-ḍamma ʾiʿrāban tadḥulu fī ʾawāḥir al-ʾasmāʾ wa-l-ʾafʿāl wa-tazūlu ʿanhā summiyat rafʿan, fa-ʾiḍā kānat al-fatḥa kaḍālika summiyat naṣban wa-ʾiḍā kānat al-kasra kaḍālika summiyat ḥafḍan wa-ǧarran...*, Ibn as-Sarrāǧ *ʾUṣūl* I 45,2-5.)

The markers of declension are needed to express the various functions the word can fulfill in the sentence, regardless of its being a noun or a verb. The grammarian az-Zaǧǧāǧī (d. 340/951) who lived in the same period as Ibn as-Sarrāǧ, uses the term *maʿnā* in this context:

> "Meanings alternate on the nouns, since they can be agent and object and possessor and possessed, and there are no indications of these meanings in their form and their structure; they are homonymous [in all these functions]; therefore, the vowels of the declension are used to denote these meanings." (*ʾinna l-ʾasmāʾa lammā kānat taʿtawiruhā l-maʿānī fa-takūnu fāʿila wa-mafʿūla wa-muḍāfa wa-muḍāf ʾilayhā wa-lam takun fī ṣuwarihā wa-ʾabniyatihā ʾadilla ʿalā hāḍihi l-maʿānī bal kānat muštarika, ǧuʿilat ḥarakāt al-ʾiʿrāb fīhā tunabbiʾu ʿan hāḍihi l-maʿānī*, az-Zaǧǧāǧī *ʾĪḍāḥ* 69,6.)

Versteegh (forthcoming) lists among the applications of the term *maʿnā* its use as a technical term meaning 'syntactic function'. When az-Zaǧǧāǧī uses *maʿnā* he does not refer to 'meanings', but rather to the syntactic functions of the noun. For az-Zaǧǧāǧī, it seems, the vow-els of declension are directly related to a syntactic function, such as agent (*fāʿil*), object (*mafʿūl*) and possessor (*muḍāf ʾilayhi*), the same functions mentioned by Zamaḥšarī. This is no coincidence, for these functions are the only ones in which the type of declension is always predictable. With active verbs, an agent always has *rafʿ*, an object takes *naṣb*, and a possessor *ǧarr*.

13	*ḍaraba*	*zayd-un*	*ʿamr-an*
	beat/PAST/3sg	zayd-NOM	ʿamr-ACC
	'Zayd beat ʿAmr.'		

14	*hāḏā*	*bayt-u*	*zayd-in*
	this	house-NOM/DEF	zayd-GEN
	'This is Zayd's house.'		

The agent, *zaydun*, and the object, *ʿamran* are subjected to governance of the verb only.

All other instances in which *rafʿ* and *naṣb* occur are compared to the governance of the verb on its agent and object. For (9), repeated below, the governance of the particle *ʾinna* is compared to that of a verb:

9	*ʾinna*	*zayd-an*	*karīm-u-n*
	TOP	zayd-ACC	kind-NOM-INDEF
	'[topicalisation] Zayd is kind'.		

With regard to sentences of the type in (9), Zamaḫšarī compares the governance of the particle *ʾinna* to the relation between a verb on the one hand and its agent and object on the other, although *ʾinna* is not a verb (cf. *Mufaṣṣal* 14,20ff. and for further discussion cf. Owens 1988:57 and 223).

In this respect the references to agency *(fāʿiliyya)*, objectivity *(mafʿūliyya)* and annexation *(ʾiḍāfa)* must be understood in their broadest sense, rather than as references to semantic agents and objects (cf. Owens 1988:223; also 1990:120). Or, to put it in another way, the concept of syntactic function is correlated with that of syntactic position assigned by governance, regardless of the semantic associations this governance relation brings about. The goal of this correlation is, no doubt, to match the number of syntactic functions with that of syntactic positions.

These findings confirm Owens' (1988:58) assumptions with regard to the question of whether an element is governed by function or by governing item. Although syntactic position and function are related to each other, and certain functions are associated with particular syntactic positions and the appropriate case endings, the latter are basically assigned as a result of governance, not by their functions.

Summarising the discussion of *ʾiʿrāb* in this section, it appears that the functions of the terms *rafʿ*, *naṣb*, *ǧarr* and *ǧazm* are threefold. In the first place they refer to an abstract *(maʿnawī)* governance relation of a certain type: each type of governor has its own effect. Secondly, they denote the endings caused by these respective governors, and are thus equivalent to the declensional endings (*u*, *i*, *a* and *ø*). Further,

they refer to the syntactic function of the word, albeit in a very broad sense.

2.2 Redundancy of ʾiʿrāb?

In the first sections of this chapter we have seen that the Arab grammarians assign a clearly defined function to each of the endings of ʾiʿrāb. However, the main impression one gets is one of a set of endings that do not themselves convey a meaning; the endings of ʾiʿrāb as such do not automatically indicate certain syntactic functions (at least, not in a semantic sense). Furthermore, the respective functions of words could also—and sometimes even better—be derived from other characteristics, such as their place in the sentence or the construction. The fact that the endings of ʾiʿrāb may be omitted for mere prosodic reasons sustains this impression.

One grammarian, Quṭrub (d. 206/821-2) a student of Sībawayh, advanced a theory on declension in which these arguments play an important role. He says:

> "If declension were applied to speech because of the difference between the meanings [= syntactic positions], then it would be necessary to have a declension for each meaning in order to indicate it." (*law kāna l-ʾiʿrāb innamā daḥala l-kalām li-l-farq bayna l-maʿānī la-waǧaba ʾan yakūna li-kulli maʿnan ʾiʿrāb yadullu ʿalayhi...* Quṭrub *apud* Zaǧǧāǧī, *ʾĪḍāḥ* 70,14ff.)

Unfortunately, Quṭrub argues, the same declensional ending serves to express several these meanings, and, moreover, the same meaning is sometimes conveyed by means of different ʾiʿrāb:

> "There are very many similar examples of [words] whose declension is the same but that differ in meaning and [words] whose meaning is different but that have the same declension." (*wa-miṯlu hāḍā kaṯīr ǧiddan mimmā ttafaqa ʾiʿrābuhu wa-ḥtalafa maʿnāhu wa-mimmā ḥtalafa ʾiʿrābuhu wa-ttafaqa maʿnāhu*, Quṭrub *apud* Zaǧǧāǧī, *ʾĪḍāḥ* 70,13.)

Quṭrub illustrates this with examples: *lā māl-a ʿindaka* and *lā māl-un ʿindaka* both meaning 'you do not have money'; in the former phrase *māl* is marked with *-a* (ACC) and in the latter with *-u(n)* (NOM), without any substantial semantic difference. In other instances different meanings are covered by the same ending, in this case *-an*, e.g. *ka-ʾanna zayd-an* 'as if Zayd' and *la ʿalla zayd-an* 'perhaps Zayd'. Quṭrub's conclusion is that no direct relation exists between the case endings

and a certain function, and that, as a consequence, *ʾiʿrāb* itself has no inherent function.

This radical view, of course, is not shared by grammarians after Quṭrub, including Zaǧǧāǧī who discusses his theories at length. Their arguments refer to the fact that the endings indicate *rafʿ*, *ǧarr* and *naṣb*, respectively. They maintain that there is a correlation between *ʾiʿrāb* (i.e. *rafʿ*, *naṣb* and *ǧarr*) and certain functions, such as agent *(fāʿil)* and direct object *(mafʿūl)*; when governed by a verb, the agent would take the ending *-un* and the direct object *-an* (*ʾĪḍāḥ* 71,8ff.), and the possessor *(muḍāf ʾilayhi)* takes *-in* when in a construction of annexation (see Section 2.1 above for a discussion). If Quṭrub were right, they say, the agent would indiscriminately receive the vowels *u*, *i*, or *a*, which is not the case (for a discussion of Quṭrub's views, see Versteegh 1981a).

The discrepancy between the two views is obvious. While Quṭrub draws attention to the fact that *ʾiʿrāb* does not refer to the syntactic function of the word, the other grammarians argue that it indicates certain 'meanings' *(maʿānī)*, i.e. syntactic positions.

2.3 *Summary*

Summarising the data in the preceding sections, it seems that, in spite of the attempts of the later Arab grammarians to prove otherwise, it is only in a limited sense possible to speak of a correlation between verbal and nominal endings and a certain syntactic function.

The reason for this is in the first place the fact that the endings common to both verbs and nouns are difficult to collect under one single heading. *Ḍamma*, for instance, is the reflection of the same 'meaning' *(maʿnā)*, i.e. *rafʿ*, in both verbs and nouns. *Rafʿ*, thus, occurs in both verbs and nouns, although only for nouns does it serve in some instances as a label for a syntactic function, i.e. agency *(fāʿiliyya)*.

In other cases, however, one might assert that a syntactic position is not equivalent to a syntactic function. A topic and a predicate, for instance, may have various endings, especially *rafʿ* and *naṣb*, denoting the governance of an element, such as a particle or a verb of a special type. In this case the construction is compared with the governance relation between a verb and its object and agent complements.

It seems, therefore, that the endings of *ʾiʿrāb* are conceived of as merely a set of markers of governance. Their function is to indicate the fact that a word is governed, referring to the type of governor,

rather than real indicators of a specific syntactic function. The seman-
tic content of the combination of governed word and governor de-
pends on the governor rather than on *ʾiᶜrāb*. It shall be seen that the
Arab grammarians' concepts of declension is probably closely related
to typological characteristics of Arabic case. From this it follows that
ʾiᶜrāb is quite distinct from the first category of markers that convey
an intrinsic meaning and do not indicate governance.

3. TURKIC CASE ENDINGS

The conclusion of the survey in the preceding sections is that,
notwithstanding some instances in which case fuses with other func-
tions, such as plural and gender, Arab grammarians attempted to seg-
ment their speech into separate morphemes to each of which they as-
sign separate meanings. In this section I give a brief sketch of the case
system of Turkic languages, and show whether an analysis in which the
principle of segmentation plays an important role can be suitable for
Turkic.

   Another matter are the typological characteristics of Turkic case. In
order to find out how Turkic cases from a typologically point of view
can be characterised, I carry out some simple tests. The findings are
discussed in Section 4.

3.1 *Turkic case endings and their functions*

Most Turkic languages possess a clear-cut case system. From the earliest
stages, i.e. the 8th century Runic inscriptions, they show a remarkable
consistency in their common morphological structure, lasting up to
modern times. Even between modern Turkic languages the similarities
are still striking, considering their physical and temporal distances.

   As a the paradigm we here consider the declension of the noun
sanğar, based on data found in our sources. Once again it is empha-
sised that all forms are quoted in a direct basic transliteration from the
Arabic texts. The paradigm presented here is the conventional one in
Western analyses of Turkic.

sanğar	NOM	sanğar-nī	ACC
sanğar-niŋ	GEN	sanğar-dā	LOC
sanğar-ğā	DAT	sanğar-dan	ABL

The forms of the case endings of most Turkic languages are subject to the principles of vowel harmony and consonant assimilation (see brief outline below and Chapter Three).

The nominative is typically assigned to the agent and topic/predicate: **sangar kal-dī** 'Sangar came', and **sangar kurklū-dur** 'Sangar is good'. The genitive usually occurs in combination with the possessive suffix:

15   **bī nin**[8]   **'aṭ-ī**[9]
         chief GEN   horse-POSS/3sg
         'the chief's horse' *(Qawānīn)*.

The genitive is also used as a predicate: **bū qul sangar-niŋ-dur** 'this slave is Sangar's' *('Idrāk)*.

The dative, locative and ablative cases convey meanings that in English are expressedby means of prepositional clauses: **sangar-ġā baqtim** (Sangar-DAT) 'I looked at Sangar'; **sangar 'aw-dā** (house-LOC) 'Sangar is in the house' or 'at home', **sangar-dan kaldim** 'I came from Sangar'. The ablative also serves to express the partitive:

16   **turk-tan**   **bir**   **'ar**
         Turks-ABL   one   man-NOM
         'a man from the Turks' or 'a Turk' *('Idrāk)*

After the **n** of possessive endings **ġā** changes into -ā:

17   **bir**   **'ar**   **qaṭ-i-n-ā**
         a   man   vicinity-POSS/3sg/-n-DAT
         'to a man.'

A place may be more specified by means of an extra noun denoting a time or a place:

18   **'a-niŋ**   **ard-i-n-dā**
         he[OBL]-GEN   behind-POSS-n-LOC
         'behind him.'

The place of **dā** may be occupied by **ġā** DAT and **dan** ABL, respectively.

3.2 *Typological characteristics of Turkic case*

An important issue discussed by the Arab grammarians in regard to Arabic case endings is, as I have shown above, the fact that they may be omitted or fuse according to certain morphonological rules, without major consequences for the meaning of the phrase. Things are differ-

---

[8] Although the author of *Qawānīn* had the instruments to distinguish between ŋ and n, he does not do so in the genitive case ending.

[9] After vowels the possessive changes into -sī: **'aṭa-sī** 'his father'.

ent in Turkic. In general, case endings cannot be omitted without dis-
tortion of the meaning, and deletion of dative, locative and ablative
case makes the phrase ungrammatical. In this section it shall be shown
which Turkic cases may be omitted and in which conditions.

For the dative case, for example, deletion yields the following result
(the examples with asterisk are construed; those without are taken
from *Idrāk*):

19a   [ʾaqǧa-nī]   bir-du-m   sangar-ǧā
        [coin-ACC]   give-PAST-1sg   Sangar-DAT
        'I gave the coin to Sangar.'

The result is an ungrammatical phrase:

19b   *[ʾaqǧa-nī]   bir-du-m   sangar
        [coin-ACC]   give-PAST-1sg   Sangar-ø
        *'I gave the coin Sangar.'

Likewise, it is not possible to delete the locative and ablative case end-
ings:

20a   sangar   ʾaw-dā
        Sangar/NOM   house-LOC
        'Sangar is in the house.'

20b   *sangar   ʾaw
        Sangar/NOM   house
        *'Sangar [is] a house.'

21a   ʾaw-dan   kal-di-m
        house-ABL   come-PAST-1sg
        'I came from the house.'

21b   *ʾaw   kal-di-m
        house/ø   come-PAST-1sg
        *'I came house.'

The accusative case is an exception. The deletion of the accusative case
ending probably denotes a change related to the degree of definiteness
of the noun. The ending **nī** is used when a definite object is in-
tended:[10]

22   bāliq-nī   ya-du-m
        fish-ACC   eat-PAST-1sg
        'I ate the fish'.

---

[10] The Arab grammarians do not explicitly explain this, therefore the translations of
the respective phrases may serve as indirect evidence for this assumption.

In (22) **bāliq** refers to a specific fish known by the second person. When the accusative case is omitted, the object refers must be understood as a collective noun:

23   **'atmak**   **yi-yū**
        bread/ACC   eat-KONV
        'eating bread'.

The object **'atmak** refers to 'bread' in the most general sense. An indefinite object is preceded by the indefinite article **bir** 'one' or 'a':

24   **kur-du-m**   **bir**   **uṣlū**   **'ar**
        see-PAST-1sg   one/a   wise   man
        'I saw a wise man.'

It is also possible to combine the indefinite article **bir** and the accusative suffix in one clause. In this case the object is not yet definite, but will be defined in a subsequent clause: [11]

25   **bir   'ar-nī**   **kur-du-m,**   [**'ur-ul-miš**   **qul-ī**]
        one   man-ACC   see-PAST-1sg   [beat-PASS-PART   slave-POSS]
        'I saw a man whose slave was beaten.'

There are two instances in the language material in *'Idrāk* in which the genitive case is omitted. In the first place so-called compound words:

26   **ṣanğiš**   **kun-ī**
        attack/NOM   day-POSS
        'day of the attack' (*'Idrāk* 111,3).

Another example is **yuzum 'ağāš-ī** 'grape vine' (*Qawānīn* 59,3). The combination of the two nouns with the possessive ending, **kun-ī** 'day' and **'ağaš-ī** 'tree' refers to a general notion expressed by the compound, comparable to compound words like 'apple tree' in English.

Secondly, the genitive case is omitted—regularly, it seems—after proper names:[12]

27   **sanğar**   **'aṭa-s-ī**   **ṭur-dī**
        Sanğar-ø   father-s-POSS/3sg   stand up-PAST/3sg
        'Sanğar's father stood up.'

Many sources provide examples in which the genitive case is omitted after proper names, e.g. **sanğar kuz-ī** 'Sanğar's eye' (*'Idrāk* 146,10), and **'alṭinbuğā 'uğl-ī** 'Alṭinbuğā's son' (*Qawānīn* 44,12), **zayd 'aṭ-ī**

---

[11] This explanation is to some extent tentative, for the issue has not yet been thoroughly investigated. The solution presented here was suggested for Turkish during a lively discussion at the IIIrd 'Turcologendag' on February 17th, 1995 (Utrecht, the Netherlands)

[12] See discussion in Chapter Five.

'Zayd's horse' (*Ḥilya* 84,15ff.)[13] but also with nouns perceived as proper names, such as titles **sulṭān qul-ī** 'the sultan's slave' (*Qawānīn* 44,10).[14] The assumption that this is a general feature is further confirmed by information supplied in *Qawānīn* (44,10) and the Margin Grammar (MG 59ʳrt/ult), where the deletion of **niŋ** is described as an option; all examples of omission concern proper names (there is also quite convincing evidence for this in *ʾIdrāk* 120,1-4). On the other hand, especially in *ʾIdrāk* many examples contain proper names, i.e. **sangar** and **sunqur**. In summary, it appears that the languages described in our sources share the same features with regard to the omission of the genitive case.[15]

One could say that the elision of the ending either conveys a meaning in itself (perhaps that of one semantic unit), or is conditioned by the type of noun it is attached to, i.e. proper names. In any case, in spite of the disappearance of the case ending, the meaning of the genitive case itself is preserved by means of the possessive and the sequence, and therefore the loss is never complete.

---

[13] In *Ḥilya*, all occurrences of a noun (i.e. proper names) without genitive marker (84-87) were 'corrected' by the editor Rifʿat, who put a corrected form with the genitive between square brackets after each word: "Wherever I have indicated the words that are in my opinion not correct, I have put the correct form after that particular word with this sign [ ]…" (Yanlış olduğunu hüküm ettiğim kelimenin doğrusunu gösterdimse, doğrusunu o kelimenin yanına [ ] şu işaretle ilave ettim, *Ḥilya* 1). The assumption that **niŋ** is omitted with proper names is further confirmed by numerous data found in Qipčaq texts, for example the manual on archery *Kitāb fī ʿIlm an-Nuššāb* (cf. Öztopçu [1990]). In this manual all omissions of **niŋ** are related to either a so-called short genitive, e.g., oq tämüri 'the arrow's iron' (38a), ṣahadat barmaqı (39a) 'index finger', or, indeed, a proper name, e.g. ʾAbū Hāšim-i Bāvardı qatında (14b) 'with ʾAbū Hāšim-i Bāvardı', Bahrām-i Gūr qavlı (42b) 'Bahrām-i Gur's saying' (transcription by Öztopçu [1990]).

[14] In view of the absence of a coherent grammatical analysis of the Turkic language(s) of these sources, including the remarks of the grammarians, I base this assumption on my own examination of the material in the sources. Furthermore, comparison with similar cases in monolingual manuscripts in XIV-Century Qipčaq seems irrelevant here, at least before the linguistic relations between them are made more evident, since certain features may occur frequently in one language or dialect and be absent altogether in the other.

[15] The possessive is almost never omitted; there is one instance in *Ḥilya* (87,3): *ġulām-u ʾanta* 'slave you', viz.,

san-niŋ	qūl	san-niŋ	qūl
you-GEN	slave	ʾanta	ġulām
'your slave'			

Omission of the possessive in this manner also occurs frequently in colloquial variants of Turkic languages.

3.3 Exception to the rule

The only exception to this pattern of regular changes in Turkic case is the accusative case. In a limited context, i.e. after a possessive, the suffix -nī is reduced to -n:[16]

28	ʾāš-ī
	meal-POSS/3sg
	'his meal'

Instead of ʾāš-ī-nī, as the rules would demand, the accusative is

29	ʾāš-i-n	ya-dī
	water-POSS/3sg-ACC	eat-PAST/3sg
	'He ate his meal.' (MG)	

Even though the Turkic case endings are variable because of the above-mentioned phonological changes, their forms are predictable throughout (loan-words being the only exceptions).

Because the suffixes do not fuse, even complex compound words are easily segmentable into morphemes:

30	kur-kan-im-niŋ	qul-ī
	see-KONV/PAST-POSS/1sg-GEN	slave-POSS/3sg
	'the slave of [the one] I saw.' (Idrāk)	

Even in a sequence of several suffixes, all morphemes are easily recognisable. We shall see that segmentability is an important element in Arabic theory.

3.4 Case markers in front and back words

In Chapter Three have given a survey of the phonemes of Turkic, and the way words are marked for a front or a back pronunciation. Our conclusion was that the Arab grammarians marked the words for either front or back in order to determine which type of suffix must be added to the word in question. Velar consonants indicate a back pronunciation, whereas neutral consonants indicate a front pronunciation. Back words get velar suffixes, and front words palatal ones. ʾAw 'house' is pronounced [ev] (front), and ʾaw 'game, hunt' [av] (back). In the case of the dative this is visible in the shape of the consonant: 'to the house' (house-DAT) is pronounced [evge], and 'to the game' (game-DAT) [avġa]. In this ending the Arabic script represents, e.g., ʾaw-ḵā vs. ʾaw-

---

[16] The sources (ʾIdrāk, Qawānīn, MG) treat the different realisations of the accusative case as variants of the same morpheme. The deletion of final -i in the accusative is a feature of many Qipčaq languages (cf. discussion in Chapter Six. Further Tatar: Poppe 1963; Kazakh: SKJa 1962:166-7; Turkish: Lewis 1984: 41).

ġā. After unvoiced 'back' consonants ġā changes into qā, e.g., 'uġmaq-qā 'into heaven', and after voiceless 'front' consonants it is kā, e.g., kant-kā 'to the town'.

In order to get the cases for the plural, it suffices to add -lar to the stem. Thus 'aw-lar [evler] 'houses' (house-PLUR/NOM), 'aw-lar-niŋ [evlerniŋ (house-PLUR-GEN), 'aw-lar-dā [evlerde] (house-PLUR-LOC), but qul-lar 'slaves' [qullar] (slave-PLUR/NOM) etc.[17]

4. TYPOLOGICAL DIFFERENCES AND SIMILARITIES OF TURKIC AND ARABIC CASE

The present section briefly discusses the findings with regard to the functions of Arabic and Turkic case in the framework of modern lin-guistic typology. It shall be seen that there are considerable differences from a typological point of view. For our argumentation these are im-portant features, since we assume that certain typological features of Arabic case formed the base for their linguistic theory. The two sec-tions that follow discuss semantical and morphological typological matters and give some predictions as to the methods the Arab gram-marians may apply to Turkic.

4.1 *Semantic typology: semantic versus syntactic case*

It is argued above (Section 1) that the basic function of Arabic case endings is to indicate governance; Arabic case in general is an indica-tion to the type of governance, rather than a marker for a specific syn-tactic function. As we have seen, other features such as word order in-dicate the syntactic function of a word when the case is omitted or is not manifest due to morphonological reasons. In brief: it does not have an inherent semantic content.

On the other hand, in Turkic languages there is a fundamental constraint against the ommission of case endings for it leads to serious distortion of the sentence. Only the accusative and genitive case may be omitted under certain restricted conditions. These are strong indi-cations for typological differences between Turkic and Arabic case.

Similar typological distinctions between cases are described for Russian by Babby (1986:198ff). In his view there are two types of cases, syntactic case and semantic case. Syntactic case is caused by gover-

---

[17] In modern Kazakh the suffix -lar/-ler is also subjected to consonant assimilation, changing into -tar/-ter after voiceless consonants, and to -dar/-der after voiced conso-nants, respectively; hence kisi-ler 'men' vs. köz-der 'eyes' and Qazaq-tar 'Kazakhs'.

nance by other syntactic elements, thus making the occurrence of a given case ending obligatory and, as a consequence, predictable (e.g. after prepositions or transitive verbs). In this way "[t]he Syntactic Case makes no contribution to the sentence's semantic interpretation" (Babby 1986:199).[18] The basic characteristic of semantic case, on the other hand, is that it does contribute to the sentence's semantic interpretation and, therefore, cannot be omitted without distortion of the meaning of the sentence. The occurrence of a semantic case is not subjected to principles of governance, and, therefore, it is unpredictable in terms of governance.

Another aspect of syntactic case is that in languages in which it is realised morphologically, the case marker may not appear on surface level in certain conditions, or disappear altogether when the function of the noun is expressed in other ways, such as word order. Examples of this are the accusative case in Turkic which is not always expressed by means of a marker, and, to give a very different example, the now obsolete case endings in older stages of Standard Dutch (cf. Van der Wal 1992:141-2), respectively. Syntactic case can even be posited for languages with no morphological realisation of case, such as English. For those languages case is posited on a theoretical level only.

Semantic case cannot be assigned on a theoretical level only because it "appears to be found only in languages in which case is realized by means of inflectional morphology" (Babby 1986:170).

In the light of these distinction between semantic and syntactic cases, Arabic cases can be characterised as syntactic, Turkic cases—except the accusative, which is basically syntactic, and the nominative, which is difficult to classify—as semantic. However, for our purposes not only the typological characteristics of Arabic case as such are important for the development of a linguistic concept, of equal significance is the way these chacteristics are perceived by the Arab grammarians, which is to be discussed in Chapters Five and Six.

4.2 Morphological typology: segmentability

An important point in Arabic argumentation is, as I have shown above, the segmentability of the elements of speech and the subse-

---

[18] Babby (1991) subdivides syntactic case into lexical and configuational case, and semantic case into argument and adverbial case. It shall become clear in the course of this chapter that for our purposes the rather simple formal-syntactic distinction (as initally proposed in Babby's publications), and the assumption that some cases can be used in either way, is sufficient.

quent assignment of a specific meaning to each morpheme. For the grammatical descriptions of Turkic words a second aspect is of interest, i.e. the degree to which segmentation can be successful in a morphological analysis of Turkic syntactic elements.

In his suggestions for a morphological typology of languages, Comrie (1983:39ff.) discusses two 'indices'. The first, the index of synthesis, refers to the number of morphemes per word. An isolating language has few, whereas an agglutinating language may have many. The second index, which is on issue here, is the index of fusion. For this index Comrie takes agglutination of invariant suffixes to the stem as the norm. The index of fusion, then, is the deviation from this norm in the sense that suffixes may fuse with each other or with the stem. Agglutination stands for a high degree of segmentability and invariance. Invariance refers to the ideal of the same morpheme always representing the same meaning under all circumstances. Phonological rules of the language—such as in Turkic front and back words—that affect the form of a suffix are not considered violations of the principle of invariance, because in practice they do not affect the predictability of the form. Applying this theory to the Turkic examples given in Section 3, it is possible to conclude that, apart from the accusative, Turkic case endings are both segmentable and invariant.[19]

In regard to the Arabic cases it is possible to conclude that although the declensional endings are both predictable and segmentable from the nouns and verbs to which they are attached, they fuse in many other instances. We have seen that the Arab grammarians, nevertheless, treat all case endings as if they were predictable and segmentable. For this they use the instrument of the underlying form.[20]

4.3 Predictions as to the Arab grammarians' interpretation of Turkic case

Based on these observations, it is possible to predict that Arab grammarians, with their inclination to segment the elements of their own speech—for which they for non-regular forms recur to the concept of underlying level—will indeed recognise the Turkic case endings as

---

[19] Comrie came to the same conclusion with regard to Turkish.

[20] It is important to stress that, although the AG clearly define the declensional endings as separate segments, they do not consider them part of the morphological structure of the words, for their occurrence depends entirely on the syntactic position of the word in question in a sentence. If there is no governance, there can be no case ending either. As a consequence, the declensional endings belong to the domain of syntax *(naḥw)* rather than morphology *(taṣrīf)*.

separate morphemes. As we have shown, Turkic languages are quite susceptible to this kind of analysis, perhaps even more than Arabic. In view of the differences between Arabic and Turkic in terms of syntactic case versus semantic case, though, it is not likely that the Arab grammarians will regard Turkic case endings as equivalent to Arabic *ʾiʿrāb*, except the accusative case. The discussion on the application of these principles is continued in Chapter Five.

PART B 5. GENERAL TERMS FOR TURKIC WORDS AND MORPHEMES

In this part of this chapter I give a short survey of the way Arab grammarians applied the terms *kalima* and *lafẓa* to Turkic elements of speech, and which terms they used to express synonymy between Arabic and Turkic linguistic elements. A brief introduction of this kind seems appropriate before engaging in the detailed discussion presented in Chapters Five and Six.

In the sources *kalima* and *lafẓa* are applied to Turkic words and morphemes. In Arabic grammar *kalima* is the general expression for 'word', both noun and verb, whereas *lafẓa* means both 'word' and 'expression'. The sources display different methods in their application of the available terminology to words. In *ʾIdrāk* and *Tarǧumān*, for instance, the terms *kalima* (pl. *kalim*) and *lafẓa* are used, both of which may mean 'word'. In this section I first discuss the instances in which these terms are used in the sources and then compare this with the analyses given by Levin (1986) and Owens (1988).

5.1 Kalima

The term *kalima* seems to be applied exclusively to nouns and verbs. *Tarǧumān*, for example, discussing **kim** 'who' in the sentence **kim bardī** 'who went away?', states "you use it before the word" (*wa-taʾtī bihā fī ʾawwal al-kalima*, 51,13). In this case *kalima* obviously refers to a verbal form. In the same way, other elements are "used at the last consonant of the word" (*fī ʾāḫir kulli kalima*, 51,15).

In *ʾIdrāk*, *kalima* is used exclusively in the sense of 'word', e.g. "with the palatalised word" (*maʿa l-kalima al-muraqqaqa*, 145,11), and "they do not have a word that is synonymous with the conjunction *wa*..." (*lā kalima ʿindahum turādifu wāw al-ʿaṭf*...149,23); "Morphology means knowledge of the rules of a word before its conjunction with another word" (*at-taṣrīf ʿilm bi-ʾaḥkām al-kalima qabla tarkībihā maʿa kalima ʾuḫrā*, 101,2); and "The word consists of the consonants

of the alphabet" (*wa-l-kalima nāšiʾa min ḥurūf al-muʿǧam*, 101,2),
"the essence of the word" (*nafs al-kalima*, 136,1).[21] *Kalima* is also
found in the general introduction to *ʾIdrāk*, again as a general term for
'word': "...the first is the meanings of all simple words, which is called
'lexicology'." (... *ʾaḥaduhā madlūl mufradāt al-kalim wa-yusammā*
ʿilm al-luġa, ʾIdrāk 5,10), "The word is an utterance, or rather what is
intended by the utterance." (*wa-l-kalima qawl ʾaw manwīy maʿahu*
mawḍūʿ li-maʿnan, 101, 9.). The same holds for *Dīwān*, in which
kalima is applied to nouns and verbs, e.g., *al-kalima ar-rakīka* "the
palatalised word" (15,7; further numerous occurrences, e.g., 5,15; 26,1;
204,6; 581,3).[22]

5.2 Lafẓa

In *Tarǧumān* and *ʾIdrāk* the term *lafẓa* is used too, although they show
differences in the way it is applied. *Tarǧumān*, for instance, applies it
to various kinds of elements, such as personal pronominal endings (**siz**
'you (pl.)' [49,12]), pronominal endings in combination with a tem-
poral suffix (**ǧā-biz** FUT-'we' [48,9]), a temporal suffix alone (**ǧāy** FUT
[48,3]), case endings (**nī** ACC [52,16]), the particle of interrogation
(**kim** 'who' [51,13]). Moreover, *lafẓa* is also applied to various Arabic
constructions and adverbs, viz., *li-ʾaǧli* 'for the sake of' (54,20), *ṯumma*
'then' (56,4).

In *ʾIdrāk lafẓa* is used less frequently, and it is applied especially to
elements other than nouns and verbs, nominal suffixes (**liq**, used to
denote the purpose of a given object [108,14]), verbal suffixes (**ǧalī**
'since' [145,9]), and **bar** 'existent' (132,23). *Lafẓa* is also used in the
introduction to *ʾIdrāk*'s Turkic-Arabic word list as a general term for
'word': "and I mention the Turkic word and then its equivalent in the
Arabic language..." (*fa-ʾaḏkuru l-lafẓa at-turkiyya wa-ʾutbiʿuhā bi-*
murādifihā min al-luġa al-ʿarabiyya, ʾIdrāk 5,16).

---

[21] In this respect *nafs al-kalima* probably refers to the notion of 'stem' which is not
identical with the theoretical Arabic notion *ʾaṣl* 'root' (for which see, e.g., Baalbaki
1988). The expression *nafs al-kalima* is also fairly common in *Dīwān* (e.g. 411,13),
whereas *Qawānīn* uses *ʾuss* 'base' (25,11). These expressions and terms seem to refer to
a notion of the shape a word before the attachment of suffixes. Although this in itself is a
very interesting issue, it cannot be dealt with further in the framework of the present
study.
[22] Other syntactic elements are invariably called *ḥarf*, e.g., vocative particle (*ḥarf*
nidāʾ) qī (538,2); diminutive particle (*ḥarf taṣġīr*) qiyā (519,13) interrogative particle
(*ḥarf istifhām*) mū (539,1); particle with the meaning of 'if' (*ḥarf bi-maʿnā law*) sā
(526,2), I have not come across the word *lafẓa* (cf. also Dankoff and Kelly 1985:273).

In *Tarǧumān* and *'Idrāk* alike *lafẓa* is, as far as we could deduce
from the material, not used for nouns and verbs in particular. In
'Idrāk the term *qawl* 'utterance' is used more or less synonymously
with *lafẓa*. It is typically used to introduce Turkic sentences and
phrases, e.g. *wa-qawluhum* 'their utterance' or 'they say' (cf. *'Idrāk*
106,17; 107,15; 136,2). In *Tuḥfa* the term *ʿalāma* is applied indiscrim-
inately to all categories of words and morphemes. In *Qawānīn* (4,5;
5,15; 6,1; 6,2; 53,16; 63,7) and *Ḥilya* the distribution of *kalima* and
lafẓa seems to be less strictly determined.[23] Instead, *lafẓa* is applied to
Turkic particles, nouns and verbs alike. In *Qawānīn*, for example,
lafẓa is applied to verbs, e.g., **yaz-dī** 'he wrote' or 'he failed' (29,14),
bašla-dī 'he started' (30,7) and the combination **'aẓd qal-dī** 'little was
left' (29,14; in context to be translated with 'almost' or 'near to'), the
postpositions **'ušun** 'for' (34,1) and **dakin** 'until' (42,10) and case
endings. In *Ḥilya*, *lafẓa* is used for verbal forms, e.g., **bul-ġay** ('may be'
[be-FUT] 99,10), **'idī** 'he was' (*lafẓa ǧāmida māḍiya* "an indeclinable
marker [for] the past tense", 99,8), postpositions, e.g., **'ayrūq** 'except'
(97,6), but also for Arabic words, such as *ṣalā* 'prayer'.

5.3 *Summary*

Summarising, it appears that in *'Idrāk* and *Tarǧumān*, and also in the
two instances in *Ḥilya*, the term *kalima* is used exclusively for 'word',
especially with regard to noun and verb, but also sometimes for
(Arabic) particle. 'Abū Ḥayyān's criterion for applying the term *kalima*
may have been the question of whether or not the linguistic element
can be used as an independent unit. The use of the term *lafẓa* in *'Idrāk*
comes quite close to the modern notion 'morpheme' (although there is
one instance in which it seems to convey 'word' in a general sense). In
his final summary of the particles *(ḥurūf)*, 'Abū Ḥayyān considers all
Turkic oblique case endings (**dā, dan, ġā, niŋ**) particles of annexation
(except the accusative **nī**), **'uš** (particle of attention), **'awat** ('yes', par-
ticle of answer), and **maqar** (particle of exception). 'Abū Ḥayyān ap-
plies the criterion of independency to Turkic linguistic elements: "All
of these are independent particles of meaning and [independent]
words" (*wa-hāḏihi kulluhā ḥurūf maʿānin mustaqilla [wa]kalimāt*,
'Idrāk 155,9). This statement, however, is in contradiction to what is
stated elsewhere in *'Idrāk*. For example, **ma** (verbal suffix for nega-
tion), **sun** (suffix for the imperative of the third person singular), **sā**

---

[23] Apart from two occurrences in *Ḥilya* (89,15) *'aṣl al-kalima* 'the root of the word',
and (102,1).

(suffix for conditional verbs), which are also included in the listing, are not considered independent elements. Moreover, he continues

> "...As to what is added for one of the meanings that occur as part of the root of another word, e.g. the diminutive, the plural, the transitive, or the construction of a passive form, we do not intend to list it in this chapter, since it has already been mentioned... " *(wa-ʾammā mā zīda li-maʿnan min al-maʿānī wa-hiya fī sinḥ kalima ʾuḥrā kamā dalla ʿalā taṣġīr ʾaw al-ǧamʿ ʾaw at-taʿaddī ʾaw bināʾ li-l-mafʿūl /*[24] *fa-laysa maqṣūdunā ʿaddahu fī hāḍā l-bāb wa-qad taqaddama ḍikru ḍālika...)*

This quotation, too, is in contradiction to other statements, for in Turkic these meanings do not involve a change within the word. Therefore, our conclusion in regard to this passage is that it must be in-terpreted as a reference to Arabic particles, since both statements do apply to Arabic.

In the other sources, *lafẓa* is used for various types of Turkic and Arabic morphemes and words, or combinations of these.

In regard to the term *kalima* this conclusion agrees to a great ex-tent, albeit not fully, with Owens' (1988) findings. According to Owens (1988:110ff.) and Levin (1986) *kalima* refers to two different concepts. In the first place it refers to an "orthographic word consist-ing of more than one morpheme" and, secondly, to a morpheme, al-beit not all morphemes.[25] In the data we have examined, the term *kalima* never conveys the sense of morpheme; instead its use seems to be limited to those instances in which 'word' is intended.[26] The sec ond term, *lafẓa*, is not clearly defined and, as far as we know, little used in general works in Arabic linguistics. In his article, Levin (1986) shows that *lafẓa*—as described by Ibn Yaʿīš (d. 643/1245)—is a general term conveying our notion of 'word', whereas *kalima* refers to mor phemes:

> "[Ibn Yaʿīš] distinguishes between two categories of *lafẓa*: one category denotes 'a simple sense' *(maʿnan mufrad)* such as the sense of *zaydun*— 'Zayd', while the other category of *lafẓa* is according to him *murakkab*— '(denoting a) composite (sense),' as in the examples: *al-raǧulu*—'the man'... *ḍarabā* 'both of them hit and *ḍarabū*—'they hit'... he believes

---

[24] This passage is found in almost literally the same form in *Qawānīn* (73; here it continues as follows: "...has been mentioned in the appropriate chapters *fa-qad taqad-dama ḍikruhā ʾaydan fī ʾabwābihā....)* and MG (63 'top), and may not be an original part of *ʾIdrāk*. The term *sinḥ al-kalima* recurs in *Dīwān* (283,2).

[25] See Owens (1988) and Levin (1986) for further discussion, and Versteegh (1993:102-3) for a survey of the earliest use of *kalima* as a grammatical term in the sense of 'a lexical unit' or 'word'.

[26] In *Irtišāf*, too, ʾAbū Ḥayyān uses *kalima* exclusively in the sense of 'word', rather than morpheme.

that each of these forms is a *lafẓa* consisting of two *kalimas*: the verb is a *kalima*, and each of the suffixes *-ā* [/a"/] and *-ū* [/uw/], contained in *ḍarabā* and *ḍarabū* respectively is a *kalima*, as it denotes the subject" (Levin 1986: 430).

Lafẓa is used in this general sense of 'word' or 'utterance' by Zamaḫšarī in his definition of *kalima*: "A word is the utterance that indicates a simple sense [when put] in a syntactic position" (*al-kalima hiya l-lafẓa ad-dālla ʿalā maʿnan mufrad bi-l-waḍ*, *Mufaṣṣal* 4,9, also quoted by Owens [1988:331]; cf. also *ʾAšbāh* III 3,16). In this sense the meaning of *lafẓa* seems to come close to the meaning of the general term *qawl* 'utterance' in *ʾIdrāk* (see quotation above).

In summary, although the term *kalima* may denote both 'word' and 'morpheme' for later grammarians as far as Arabic is concerned, with regard to Turkic it is used almost exclusively in the sense of 'independent unit'. *Lafẓa* is a term that is applied to both Turkic morphemes, as well as to words consisting of several morphemes. In *ʾIdrāk*, in which both terms are used, *kalima* is applied to words, whereas *lafẓa* is used exclusively for morphemes and not for words.

In the preceding section we have shown that in the sources the terms like *kalima* and *lafẓa* are applied to the Turkic linguistic elements in various ways. In *ʾIdrāk* and *Tarǧumān* there is a tendency to reserve *kalima* for nouns and verbs, and to apply *lafẓa* to other elements that in Western grammar would be considered morphemes and case endings. In the other sources the term *lafẓa* is used indiscriminately for all Turkic—and in some cases also Arabic—elements, i.e. both nouns and verbs on the one hand, and morphemes and combinations of words on the other.

Finally, interesting to note is that in most sources the terms *fiʿl* ('verb') and *ism* ('noun') are applied quite naturally to Turkic verbs and nouns; this does not hold for *Tuḥfa*, in which the term *ʿalāma* 'marker' is used throughout. The Turkic verbs are usually given in the past tense, i.e. with the suffix *-dī* (3sg), e.g. *ʾur-dī* 'he beat', which is the analogous form of the only verbal form in Arabic that shows the basic consonants of the verb, e.g. *ḍaraba* 'he beat' /ḍ-r-b/. Within the Arabic system of dividing speech into three basic categories, one would expect that all other elements that cannot be classified as verbs or nouns are characterised as particles *(ḥurūf)*. The sources (except *Dīwān*) appear very cautious at this point, for they use the rather vague term 'word' *(lafẓa)* for syntactic elements that cannot be understood as verbs or nouns.

6. SYNONYMY BETWEEN ARABIC AND TURKIC: *FI* AND **DĀ**

In the sources the synonymy of Arabic and Turkic elements is ex-
pressed in various ways. In this section I give a brief survey of the terms
they use, especially in connection with the Arabic particle *fī* and the
Turkic locative ending **-dā**.

MAʿNĀ

Most sources contain statements of the type "**dā** has the meaning of
fī" (*bi-maʿnā fī*, *ʾIdrāk* 144,21; 154,2; 136,10; *Ḥilya* 98,6; MG 40ᵛlt;
55ᵛrt/bm), and, more specifically, "**dā** has the meaning of *fī* that gov-
erns in the genitive" (*bi-maʿnā fī al-ǧarra*, MG 58ʳrt).[27] This term must
probably be understood in the sense of having the same lexical
meaning as *fī*, without any reference to any similarity in function (cf.
Versteegh [forthcoming] for further discussion of the term *maʿnā*).

MAQĀM

In *Tarǧumān*, **dā** is said to "take the place of the Arabic *fī*" (*dā taqūmu
maqāma fī l-ʿarabiyya*, 51,13). This expression, *taqūmu maqāma*, is
used in six similar instances (*Tarǧumān* 50,6; 50,9; 50,11; 51,13;
51,19; 52,2) and it is also found quite frequently in the Margin
Grammar (MG 42ʳbm; 44ᵛrt; 47ʳ; 50ʳtop; 56ᵛtop; 58ᵛrt; 59ʳrt; 60ᵛlt/ult)
and *Dīwān* (54,14; 444,10; 540,4).

According to Versteegh (1978:275), *maqām* as a grammatical term
means that "there is a specific context in which a word or an expres-
sion is normally used, but now it is replaced by another word or ex-
pression, which acts, as it were, on its behalf." In Arabic grammar the
term maqām is used in combination with the notion of ittisāʿ, mean-
ing both 'individual choice of the speaker' and 'flexibility of the Arabic
language' (cf. Versteegh 1990:293). In a case of *ittisāʿ*, an element is
deleted from a sentence, while its syntactic function is fulfilled
(*yaqūmu maqāmahu*)—temporarily—by another element. Obviously,
words may only assume the place of other words when they have the
same capacities in terms of governance. Applied to **dā**, the term
maqām would imply that it has a meaning and a function similar to
that of fī, but, at the same time, that it is not its exact equivalent.

---

[27] In this survey I have listed only the terms that are applied to **dā**.

YANŪBU ʿAN

Ḥilya gives the description "the d and ʾalif [sc. **dā**] substitute for the particle *fī*" (*ḥarfu fī wa-yanūbu ʿanhu dāl wa-ʾalif*, 92,16); in *Ḥilya*, the expression *yanūbu ʿan* is used quite frequently for other Arabic words and their equivalents in Turkic (cf. *Ḥilya* 93,15; 93,14; cf. also MG 64ʾlt).

The expression *yanūbu ʿan* is also used when one Turkic element substitutes for another. This is for example the case for the possessive ending ī, such as in ʾāš-ī 'his food' (food-POSS/3sg). According to the Margin Grammar, in certain conditions ī can be abbreviated to i, espe-cially when a consonant follows, such as in the accusative ʾaš-i-nī 'his food' (food-POSS/3sg-ACC), "the i substitutes for the y that has the meaning of the third person in Arabic" (*tanūbu l-kasra ʿan al-yāʾ al-latī bi-maʿnā ḍamīr al-ġāʾib fī l-luġa l-ʿarabiyya*, MG 55ʾbm).

In Arabic, the root /n-w-b/ conveys the sense of 'to substitute', or 'to deputise' for someone who is absent. In Arabic grammar the ex-pression *nāba ʿan* is, for example, applied in a discussion of the fact that some morphemes may fulfill the function of the *ḍamma* (u) as a marker of the nominative without being declensional markers them-selves. In some instances, e.g. the plural *al-muʿallimūna* /al-muʿallimuwna/ 'the teachers', the Arab grammarians argue that the w "represents the u" (*tanūbu ʿan aḍ-ḍamma*, Širbīnī 46,6) in al-muʿallim-u 'the teacher'. Another instance in which the term *yanūbu*—or a derivative form—is used concerns the form of the pas-sive verb whose agent is not mentioned, e.g. *ḍuriba zayd-un* 'Zayd was beaten'. According to the Arab grammarians, *zaydun* is not the agent of the verb, in spite of the nominative case ending, but rather its syn-tactic replacement: *yanūbu al-fāʿil ʿan al-mafʿūl* ("the agent represents the object") or *al-fāʿil nāʾib ʿan al-mafʿūl* ("the agent is the substitute of the object").[28] The essence of this term, then, is that the deputy el-ement fulfils the function of the one represented, albeit without as-suming a full equivalent status (cf. also *Ḫaṣāʾiṣ* II 274ff.).[29]

The meaning of the verb *nāba ʿan*, therefore, is here interpreted as 'assuming, or representing the function of something which cannot stand at that particular position'. 'Position', in this context, can be in-

---

[28] This term was introduced by Ibn Mālik (d. 672-3/1274), the original term being *al-mafʿūl allaḏī lam yusamma fāʿiluhu* ("the object whose agent is not mentioned") (cf. also *Manhağ* 111,4ff. and Carter 1981:169).

[29] In Zamaḫšarī's *Mufaṣṣal*, *yaqūmu maqāma* is more or less equated with *yanūbu ʿan* in the example of passive verb: "the object fulfils its [sc. the agent's] place" (*fa-ʾuqīma l-mafʿūlu maqāmahu*, *Mufaṣṣal* 116,4; also Širbīnī 168,6).

terpreted as either syntactic or morpho(no)logical position. In this
sense, the *wāw* fulfils the function of *ḍamma*, because the *ḍamma*
cannot stand at that particular morphological position in a word.
Likewise, in the case of passive verbs, the agent cannot occupy that
particular syntactic position and, therefore, it is represented by the ob-
ject. The difference with maqām is that in the case of *maqām* the ele-
ment represented could have occupied the position taken by the repre-
sentative.

NAẒĪR

In *Šuḍūr* the term *naẓīr* (lit. 'peer') is used in various contexts that give
little foundation for any conclusions with regard to the status assigned
to the Turkic equivalent: "and thus you apply it in the construction
with *fī* whose peer is **dah**" (*wa-kaḍā taʿmal fī l-maẓhar at-tarkīb bi-fī
naẓīruhu dah*, *Šuḍūr* 26ᵛ1[30]). According to Versteegh (1993:70), *naẓīr*
is used as a technical term in early grammar "to indicate words with
the same syntactic status because of a resemblance in form." In
Ḫaṣāʾiṣ (II 197,8f) *naẓīr* is used in the sense of 'having the same pat-
tern', which does not necessarily imply an identical meaning or status.
In *ʾIdrāk* (142,8) *naẓīr* is used to indicate a semantic similarity be-
tween two Arabic sentences: *dafaʿtu d-dirhama ʾilā sanǧara* 'I handed
the *dirham* to Sanǧar' is the *naẓīr* of *ʾaʿṭaytu sanǧara d-dirhama* 'I gave
Sanǧar the *dirham*'.

MURĀDIF

Most sources agree that **dā** is a *ḥarf*. But the terminology is not always
very specific about the kind of ḥarf. In some cases it is simply regarded
as 'identical' with *fī*, e.g. "the particle *fī*" *ḥarfu fī* (MG 55ᵛrt/bm).
   A more detailed term is "the particle that is synonymous with *fī*"
(*al-ḥarf al-murādif li-fī*, *ʾIdrāk* 146,3; 136,5; 122,8; *Qawānīn* 41,6; MG
56ʳtop). The term *murādif* is also used for a literal interpretation of a
Turkic sentence (*ʾIdrāk* 132,12; 146,18). In one instance the expres-
sion is used conversely, in the sense that the Arabic particle "*fī* is syn-
onymous with **dā**" (*fī al-murādif li-dā*, *Tuḥfa* 72ᵛ9). The term *murādif*
is, with its related form *mutarādif*, the usual term for 'synonym' in
lexicography (cf. Lane; also Versteegh 1993:153). In other works of
ʾAbū Ḥayyān, however, the term *murādif* is limited to 'having the same
meaning', perhaps including the fact of belonging to the same class,
but explicitly excluding full similarity in distribution. In *Irtišāf*, for ex-

---

[30] The term *maẓhar* is also frequently used in *Sirr*, e.g., I 329,9.

ample the locative *(ẓarf)* *ʾiḏ* 'at the time, when' is discussed, and "*ḥīna* ['at the time'] is synonymous with it" *(yurādifuhā ḥīna)*, but its use, and, hence its function, is not equal, e.g. *ʾiḏ* may follow *ḥīna*, *ḥīn-ʾiḏ* 'when', but not the other way around (*Irtišāf* II 234,6). To give another example, in *Manhaǧ* ʾAbū Ḥayyān discusses the locative *ladun* 'at', and states

> "With regard to *ladun*, its equivalent is *ʿinda* and *ʿinda* 'at, with' may be used in its place, but *ladun* cannot be used in [*ʿinda*'s] place." (*ʾammā ladun fa-ʾinnahā turādifu ʿinda wa-taṣluḥu ʿinda makāna ladun wa-lā taṣluḥu ladun makānahā, Manhaǧ* 293,12.).

This incomplete synonymy of particles is confirmed by Gully's (1994) findings. Gully states that "a particle could only replace another if the two meanings amounted to one, and if the meaning of the speech of which they were part was one, or, at least, could be traced back to that meaning." A linguistic element may be partially synonymous with another but not fully; in this way *ʾilā* 'to' is only partially synonymous with *ḥattā* 'until'.

THE GENITIVE CASE (ǦARR)

INTRODUCTION

This chapter deals with the instances in which a genitive case occurs in Arabic and the way the constructions in which it appears are reflected into Turkic.

In Chapter Four we have seen that the Arabic grammarians acknowledged three types of governance relations, expressing the syntactic positions of *rafʿ*, *ǧarr* (or *ḥafḍ*), *naṣb* and *ǧazm*, respectively. The syntactic positions are indicated in the surface structure by means of the declensional markers *u*, *i* and *a*. With respect to nouns, all instances in which these governance relations occur are reduced to three functional categories; *rafʿ* is related to agency *(fāʿiliyya)*, *naṣb* to objectivity *(mafʿūliyya)*, and *ǧarr* to annexation *(ʾiḍāfa)*. The genitive case *(ǧarr)*, the subject of this chapter, thus typically occurs because of annexation *(ʾiḍāfa)*. The annexation construction is used to express a relation between two nouns. In the opinion of the later grammarians this construction of two nouns is possible only if a particle serves as the connective element. A special type of such a relation is the possessive relation in which the particle does not occur in surface structure. In this chapter I shall also deal with the locative *(ẓarf)*. The locative is connected with the genitive case in two ways: in the first place because it is basically considered a relation between two nouns; and secondly, because the *ẓarf* includes the meaning of a particle, i.e. *fī* 'in'.

The present discussion of the genitive case will be a survey of the way the Arabic grammarians conveyed Arabic particles and locatives into Turkic. We shall see that most Turkic case endings are regarded as equivalents of the Arabic particles, and make some assumptions as to the status the Arabic grammarians assigned to these equivalents in terms of government. In the first place, we shall see how the problem of two Arabic particles with the same equivalent in Turkic was solved. The third issue will be the analysis of one Arabic particle which in Turkic is equivalent to a sequence of two elements. The final issue is the Turkic possessive construction and how it fits into Arabic theory.

Furthermore, in section 2.4.1, we shall discuss the Turkic equivalents of some Arabic annexation constructions and investigate how they fit into their theories of governance.

As has been argued in Chapter Four, it is my hypothesis that the principal place *ʾiʿrāb* and government occupy in Arabic linguistic theory can be accounted for with the assumption that these concepts are essentially based on Arabic syntactic case. It shall be seen in the present chapter and in Chapter Six that only in one instance, the accusative, a Turkic case ending is in fact assigned a status similar to that of an Arabic case ending, whereas all other Turkic cases are considered equivalent to Arabic particles. In our line of argumentation this is hardly a coincidence, for the accusative is the only syntactic case of Turkic.

1. ANNEXATION AND PARTICLES IN ARABIC

1.1 *Annexation*

According to ʾAbū Ḥayyān,

"annexation is a connective relation between two nouns that invariably causes a genitive case in the second noun." (*al-ʾiḍāfa nisba bayna smayni taqayyudiyya tūǧibu li-ṯānīhimā al-ǧarr ʾabadan, Irtišāf* II 501,4; the same definition is found in *Manhaǧ* 263,7.)

The first noun of a so-called annexation construction is called 'the annexed' *(muḍāf)* and the second 'the [noun] to which is annexed' *(al-muḍāf ʾilayhi)*. On the surface level the annexation is realised as in (1):

1   *bayt-u*   *zayd-in*
       house-NOM/DEF   zayd-GEN
       'Zayd's house'.

In the theories of the so-called later grammarians—among whom ʾAbū Ḥayyān is to be reckoned—the two nouns are typically connected by means of a particle in the underlying structure, which causes the genitive case in the second word, viz. (2)

2   *bayt-u-n*   *li-*   *zayd-in*
       house-nom-INDEF   to   zayd-GEN
       'A house to Zayd'.

With regard to the fact that a particle connects the two nouns, this type of particles is often called a 'particle of annexation' *(ḥurūf al-*

ʾiḍāfa). In this way a discussion of the genitive case is to a great extent
equivalent to a discussion of the particles of annexation. In terms of
governance relations, the particle *li* in (2) is the *ǧārr*, 'the [element]
that governs in the genitive', and the governed noun (*zaydin*) is called
maǧrūr, 'the [element] governed in the genitive'. This particle accounts
for the genitive case (*ǧarr*), since—in the view of the later grammari-
ans—a noun cannot govern and, hence, cannot cause a case: "The
basis of the governance of the genitive is because of the particle, not
because of the nouns." (*ʾaṣl ʿamal al-ǧarr ʾinnamā huwa li-l-ḥarf lā li-
l-ʾasmāʾ, Manhaǧ* 265,6, see also discussion in Chapter Four). A noun
may only govern insofar as it substitutes for a particle,

> "except that the Arabs limited the use of the particles of the genitive to cer-
> tain syntactic positions and annexed nouns to each other, and the noun sub-
> stitutes for the particle and exerts governance, i.e. the genitive." (*ʾillā ʾanna
> l-ʿarab iḥtaṣarat ḥurūf al-ǧarr fī mawāḍiʿ wa-ʾaḍāfat baʿḍ al-ʾasmāʾ ʾilā
> baʿḍ fa-nāba l-ism al-muḍāf manāb al-ḥarf fa-ʿamila wa-huwa l-ǧarr,
> Manhaǧ* 265,7-8.)

1.2 The locative

A special class of nouns, the *ẓurūf* (sg. *ẓarf*, 'locative'), which can be
annexed to other nouns, implicitly conveys the sense of a particle in
the underlying structure, i.e. *fī* 'in'. The *ẓarf*, also called *mafʿūl fīhi* 'the
object in which', is one of the optional objects of the verb (see Chapter
Six; cf. also Owens 1988:131ff. and Mosel 1975:345ff.). The *ẓarf* de-
pends on a verb which conveys an action and which is visible in the
surface structure (*maḏkūr*), e.g.

3   *mašā*   *ʿamr-un*   *ʾamām-a*   *zayd-in*
       walk/PAST/3sg   ʿamr-NOM   front-ACC/DEF   zayd-GEN
       "ʿAmr walked in front of Zayd.'

In (3), the verb *mašā* governs *ʾamām-a* in the accusative case (and
ʿamr-un in the nominative). The verb may also be absent from the sur-
face structure,

4a   *ʿamr-un*   *ʿind-a*   *zayd-in*
       ʿAmr-NOM   with-ACC/DEF   zayd-GEN
       "ʿAmr is with Zayd' or 'at Zayd's.'

In that case the verb is implied in the underlying structure (*muqad-
dar*), which can be reconstructed as:

4b	*ʿamr-un*	*kāʾin-u-n*	*ʿind-a*	*zayd-in*
	ʿamr-NOM	be/PART -	with-ACC/DEF	zayd-GEN
		NOM -INDEF		

ʾAbū Ḥayyān defines the *ẓarf* as follows:

> "...'the object in which', or the locative, is [a noun] conveying a time or place [e.g. *layl* 'night', *ʾamām* 'front'] that is assigned the accusative by implication of *fī* with respect to an action [occurring] in it." (*al-mafʿūl fīhi wa-huwa ẓ-ẓarf wa-huwa mā ntaṣaba min waqtin ʾaw makānin ʿalā taqdīr fī bi-ṭṭirād li-wāqiʿin fīhi, Irtišāf* II 225,2f.)

In early Arabic linguistic theory, the particle should not appear in the surface structure, since in that case the noun would depend on the particle instead of on the verb, and, as a result, it would stop being an object. The noun *dār* 'house', for instance, although denoting a place, is not considered a *ẓarf*, because it does not contain the meaning of *fī*. For this reason it cannot serve as a direct dependent of the verb in this sense, viz.,

5a	**ǧalasa*	*ʿamr-un*	*dār-a*	*zayd-in*
	sit/PAST/3sg	ʿamr-NOM	house-	zayd-GEN
			ACC/DEF	

**ʿAmr sat Zayd's house.'

Instead, with nouns like *dār*, *fī* must be used in the surface structure:

5b	*ǧalasa*	*ʿamr-un*	*fī*	*dār-i*	*zayd-in*
	sit/PAST/3sg	ʿamr-NOM	in	house-	zayd-GEN
				GEN/DEF	

'ʿAmr sat in Zayd's house.'

Later grammarians, such as az-Zamaḫšarī and ʾAbū Ḥayyān (*Irtišāf* II 253,7) distinguished two categories of locatives. First, a *ẓarf muḫtaṣṣ* 'specified locative' (*ẓurūf muwaqqata* 'temporary locatives' in Zamaḫšarī, *Mufaṣṣal* 25,15, which includes nouns denoting a specified space or time (e.g. *dār* 'house', *masǧid* 'mosque', or *sūq* 'market), and which can be used only with *fī* in the surface structure (see below Section 2.5.2; Owens 1988:136), and secondly, a *ẓarf mubham*, denoting an unspecified space or time (e.g. *ʿinda*, or *ḫalfa*) and directly governed by the verb.

Locatives of time, on the other hand, never require *fī* in the surface structure: *qumtu l-yawma* 'I stood up today' and *sirtu yawman ṭawīlan* 'I travelled a long day' (*Irtišāf* 226). The locatives of time, too, are divided into *muḫtaṣṣ* and *mubham*. To the *muḫtaṣṣ* belongs, for example, the word *šahr* 'month', that can be annexed to the names of the months, e.g. *ǧiʾtu šahr-a ramaḍān-i* 'I came [in] the month [of] Ramaḍān' (cf. *Irtišāf* 226,19). The *ẓarf mubham* comprises nouns that

denote unspecified quantities of time, e.g. *waqt*, *zamān*, and *ḥīn* all meaning 'time', e.g. *ǧi'tu ḥīna qiyāmihi* 'I came [at] the time he stood up'.

Positing a particle in the underlying level, though, may lead to different effects on the surface level. In the locatives, the accusative is considered to 'contain' the meaning of the particle without showing any traces of governance thereof, for the accusative case is assumed to be caused by the verb (or by a verb in the underlying level), only the meaning of the particle being implied, not its governance. In other instances, especially in the possessive construction (see above Section 1.1) the particle is posited in the underlying level because of the apparent lack of a governor in the surface structure. This way of argumentation reveals an interesting paradox in the Arabic argumentation; the accusative in the locatives is assigned a semantic content, namely that of *fī*, which would make it a semantic case. At the same time, however, its appearance is accounted for in terms of governance, i.e. as syntactic case.

2. TURKIC EQUIVALENTS OF ARABIC PARTICLES

The status of a linguistic element is in the first place indicated by its assignment to one of the three main classes: nouns, verbs and particles. In this section it will be seen that the Turkic case endings and postpositions are basically equated to what in Western studies are considered particles. The status and governing capacities the grammarians assign to the Turkic elements varies according to their use in comparison to Arabic.

I have selected five items which are expected to cause varying difficulties in Arabic theory. The first case is *fī* 'in', which finds an equivalent in the Turkic locative case, which can be simply interpreted as a particle. The second is *bi* 'with', whose Turkic equivalent has several meanings, one of which is best expressed with the Arabic noun *maʿa* 'together with'. An important part of our discussion discussion of the Turkic equivalent deals with the question what the status of the Turkic element should be. Thirdly is *'ilā*, which basically means 'to', but can also be used in the sense of 'until'. In the former sense, it is reflected with simply with a Turkic ending; in the second sense, however, it is used in combination with a postposition. It is expected that the grammarians will have difficulties in assigning a status to the postposition. The fourth instance is the particle *li* 'for' and the possessive construction. In the fifth and last place, I discuss the locative *(ẓarf)*. We

shall see that an innovation of the later grammarians to include con-
structions in which the particle *fī* 'in' (or its translation for that mat-
ter) is visible in surface structure is especially convenient for the analy-
sis of Turkic constructions.

2.1 *The case of* fī

In this section we examine how the particle *fī* is translated into Turkic
and which terminology is used to describe the status of the equivalent.
In this framework we will not be able to examine all instances in
which *fī* occurs, so the discussion will be limited to those cases in which
it is equated with **dā**. Further analyses with *fī* are given in the
discussion of the locative (below in Section 2.5.)

The sources basically agree that *fī* is translated into Turkic as **dā**.[1]
This becomes implicitly evident from the translations of phrases that
contain *fī*, e.g. '**aw-dā** is translated as *fī l-bayt-i* 'in the house' (MG
58ᵛrt). Apart from this rather implicit semantic relation with the
Arabic particle *fī*, the sources also describe the status of **dā** in relation
to *fī*. We have seen above that *fī* is a particle that serves to connect two
nouns to each other, and as such it governs nouns and causes them to
take the genitive case. In this respect, the assignment of the term *ǧārr*
or *harf ǧarr* to its Turkic equivalent, **dā**, and *maǧrūr* to the noun it is
connected to are important indications of its status.

In western studies of Turkic, **dā** is generally considered a locative
case ending whose form differs according to consonant assimilation
'**aw-dā** 'in the house' - **kant-tā** 'in the town', while the pronunciation
depends on whether the word is 'back', e.g., [taŋda] 'in the morning'
or 'front', e.g., [evde] 'in the house' (for further details regarding vowel
and consonant assimilation the reader is referred to Chapter Three).

HARF AL-ʾIḌĀFA

In *ʾIdrāk*, **dā**, along with other Turkic equivalents of the Arabic parti-
cles, is classified as a 'particle of annexation' (*harf al-ʾiḍāfa*): "These
are the particles of annexation that come at the end of the word, as we
showed, unlike the particles of annexation in the Arabic languages"
(*wa-hāḏihi hurūf al-ʾiḍāfa taʾtī ʾahīran kamā mattalnā bi-hilāf hurūf
al-ʾiḍāfa fī l-lisān al-ʿarabī*, 145,18). The status of the Turkic particle
as a governor is inferred only implicitly; the term *hurūf al-ʾiḍāfa* refers

---

[1] In CC (80) **da** is translated into Latin as both 'ad' and 'in'.

to a syntactic function rather than to a syntactic position.[2] In Chapter Four I pointed out that *'idāfa* is a syntactic function for which words are marked with the genitive case *(ǧarr)*. In practice, however, *ḥarf al-'idāfa* refers to the same as *ḥarf ǧarr*, since annexation *('idāfa)* always involves governance in the genitive case *(ǧarr)*. Therefore, its application to Turkic *ḥurūf* can be safely interpreted as a reference to their syntactic position.

ḤARF ǦARR

Although several sources (e.g. *Qawānīn* 44-46) have the term *ḥarf ǧarr*, 'particle of the genitive' or a closely related one, such as *maǧrūr* 'noun governed in the genitive' in their chapter and section headings, these terms must be regarded basically as a reference to the Arabic particles rather than their Turkic counterparts. In addition, there are many instances in which the terms *ǧarr* and *ḥarf ǧarr* are applied to the Turkic equivalent of the Arabic particle *fī* or the element it governs.

In *Qawānīn*, for example, **dā** is explicitly termed both a *ḥarf ǧarr* and the synonym of *fī*: "**dā** is the particle of the genitive which is synonymous with *fī*" *(wa-dā ḥarf al-ǧarr al-murādif bi-fī*, 41,6; also: 36,6). Indeed, not only **dā** but nearly all Turkic equivalents of Arabic *ḥurūf al-ǧarr* (as far as they exist in Turkic, cf. *Qawānīn* 41,2) are explicitly categorised as *ḥurūf ǧarr*. In one passage in *'Idrāk*, 'Abū Ḥayyān, too, assigns the status of *ḥarf ǧarr* directly to a Turkic particle: "it is permitted to use another particle of the genitive than **dā**" *(wa-yaǧūzu duḥūl ḥarf ǧarr ǧayr dā ʿalayhā*,136,5). In *Tuḥfa*, too, the term *ḥarf ǧarr* serves as a reference to Turkic 'particles': "...the particle of the genitive which is synonymous with *fī* in Arabic, which expresses the act of containing, and this [particle] is **dā**" *(...ḥarf al-ǧarr al-murādif li-fī ʿarabiyyatan allatī li-l-wiʿāʾ wa-huwa dā*, 66ᵛ3-5; also 72ᵛ9; also MG 58ʳt)

The fact that in Turkic the *ḥurūf al-'idāfa* follow the noun instead of preceding it does not seem to have disturbed the grammarians, although in Arabic linguistics it is assumed that a governor always precedes the element it governs:

---

[2] In *'Idrāk*, the *ḥurūf al-'idāfa* are discussed under the heading *al-qawl fī l-'idāfa* (144,19, which is separated into two subdivisions, i.e. annexation with [144,19] and without particles [146,15]). This approach is quite unlike the one 'Abū Ḥayyān applies in *Irtišāf* (II 426) and *Manhaǧ* (231), in both of which the *ḥurūf al-ǧarr* are treated in a separate chapter followed by a one on the *'idāfa*, although for *Manhaǧ* this may have been determined by the structure of Ibn Mālik's *'Alfiyya*.

"These are the particles of annexation that come at the end [of the word], as
we showed, unlike the particles of annexation in the Arabic language." (*wa-
hāḏihi ḥurūf al-ʾiḏāfa taʾtī ʾāḫiran kamā maṭṭalnā bi-ḫilāf ḥurūf al-ʾiḏāfa
fī l-lisān al-ʿarabī, ʾIdrāk* 145,18.)

This is stated in a similar way in *Qawānīn*:

"It must be known that the governor of the genitive in this language is con-
nected to the last consonant of the governed [noun] in all circumstances, the
opposite of the Arabic." (*wa-l-yuʿlam ʾawwalan ʾanna l-ǧārr fī hāḏihi l-
luġa ʾinnamā yattaṣilu bi-maǧrūrihi ʾāḫiran fī l-ʾaḥwāli kullihā ʿaks al-
ʿarabiyya, Qawānīn* 40,19.)

In other words, the Turkic particle is the *ǧārr* and the noun that pre-
cedes is its *maǧrūr*, whereas in Arabic the word in the genitive follows
its particle. The discussion of the *fī* continues in Section 2.5.

2.2 *The case of* bi

The second of the Arabic *ḥurūf al-ǧarr* I propose to discuss here is the
particle *bi* 'with'. The basic function of *bi* is to denote instrumentality
(*istiʿāna*): *ḍarabtuhu bi-s-sayf-i* 'I beat him with the sword'. Apart
from this basic meaning, which will be the issue here, *bi* is used in vari-
ous other meanings, such as accompaniment (*muṣāḥaba*), e.g.
[*ištaraytu l-farasa*] *bi-sarǧihi* '[I bought the horse] along with its sad-
dle', compensation (*muqābala*), [*ištaraytuhu*] *bi-dirhamin* '[I bought
it] for a dirham', the oath (*qasam*), *bi-llāhi* 'by God', etc. (cf. *Irtišāf* II
426ff.).

The Arabic grammarians found that the Turkic counterpart of *bi* in
its sense of instrumentality is **bilā** which also serves as the translation
of *maʿa* 'together with', which, in its turn, in Arabic is principally used
to indicate accompaniment (*istiṣḥāb*), e.g. *ǧiʾtu maʿa zayd-in* 'I came
[together] with Zayd'.

In western grammars of Turkic, **bilā** (and its older variant **birlā**) is
regarded as a postposition used for conveying the sense of both in-
strumentality and accompaniment.

In *ʾIdrāk*, **bilā** is used in the following sentences:

6	qalam	bilā	yaz-du-m
	pen	with	write-PAST-1sg
	'I wrote with the pen.'		

7	sanğar	bilā	kal-di-m
	sanğar	with	come-PAST-1sg
	'I came together with Sanğar.'		

In combination with pronouns (except 3pl), **bilā** is preceded by the re-spective genitive suffixes:

8a **man-im**   **bilā**
    I-GEN   with
    'with me'

8b **biz-iŋ**[3]   **bilā**
    we-GEN   with
    'with us'

Apart from the clear difference in meaning, in Arabic grammar there is also a categorial distinction between *bi* and *maˤa*. *Bi* is, as we have seen, a *ḥarf ǧarr*, whereas *maˤa* is a noun: "noun for the place or time of accompaniment" (*ism li-makān al-istiṣḥāb ʾaw waqtihi*, *Irtiśāf* II 267,4ff. also II 449). In this sense, *ǧiʾtu maˤa zaydin* 'I came with Zayd' could also be translated as 'in Zayd's company' or 'at the time Zayd came'. In other words, *maˤa* is a noun used as a *ẓarf* (locative). This status of *maˤa* is usually demonstrated by its use in combination with a *ḥarf ǧarr*, especially *min* 'from': *min maˤī* 'by me', and, further, its full inflection (i.e. with nunation) when used adverbially *ǧiʾnā maˤan* 'we came together'. Both governance by a *ḥarf ǧarr* and full in-flection (with nunation) are typical characteristics of nouns.

In this respect the Arabic grammarians are confronted with the problem of assigning a status to **bilā**, which conveys two meanings that are associated with Arabic words belonging to a different category. The sources display two opinions.

The first opinion is that **bilā**, as an equivalent to *bi*, is primarily re-garded as a *ḥarf ǧarr*. Furthermore, by extension, it may convey the meaning of *maˤa*, albeit without assuming its status of a locative noun (*ẓarf*). This opinion is reflected in ʾAbū Ḥayyān's *ʾIdrāk*, where **bilā** conveys "the meaning of both *bi* and *maˤa*" (*bi-maˤnā al-bāʾ wa-bi-maˤnā maˤa*, *ʾIdrāk* 144,21); "I wrote with the pen" **qalam bilā yaz-dum**, "I came with Sangar" **sangar bilā kaldim** (*katabtu bi-l-qalami, ǧiʾtu maˤa sangar*, *ʾIdrāk* 145,17). ʾAbū Ḥayyān's opinion as to the sta-tus of **bilā** becomes explicit from his remark (after listing its use with the pronouns and their possessives): "**bilā** has the meaning of *maˤa* as if it were a locative" (*wa-bilā bi-maˤnā maˤa ka-ʾannahā ẓarf, ʾIdrāk* 146,3). In other words, ʾAbū Ḥayyān appears to be saying, in some instances **bilā** has the meaning of a *ẓarf*, although in reality it is not. In

---

[3] The fact that the sources have different forms, due to differences in languages and dialects, e.g. **man-iŋ** (1sg) or **biz-im, miz-iŋ** (1pl), etc. instead of the ones mentioned here, is not relevant to this discussion.

'Abū Ḥayyān's view, the major criterion for assigning to any Arabic el-
ement the status of a noun with the function of a *ẓarf* is the question
of whether *fī* is implied. This, for example, is the case with *ʿinda* 'at,
with' and *ʾamām-a* 'in front [of]'. For a Turkic *ẓarf* 'Abū Ḥayyān
seems to apply the same criterion, i.e. whether the meaning of *fī* is
implied, which involves its equivalent in Turkic, i.e. **dā**. In practice this
means that the Turkic equivalent of an Arabic *ẓarf* must have **dā** in
the surface structure.[4]

> "...the equivalent of 'in', which is **dā**, is not used with it, unlike **qaṭ** which
> has the meaning 'with'." *(wa-lā yadhulu ʿalayhā murādif fī lladī huwa dā
> bi-ḫilāf qaṭ bi-maʿnā ʿinda, 'Idrāk 146,2-3.)*

Abū Ḥayyān probably relates this to the fact that in the Turkic equiva-
lents of many Arabic *ẓurūf* the 'particle' **dā** is used, e.g. *fawqa* 'above'
'ustun dā, *tahta* **'altindā** 'underneath'. In other words, it seems that
'Abū Ḥayyān's main reason for not regarding **bilā** as a *ẓarf* is the fact
that it cannot be construed with **dā**, as opposed to other Turkic *ẓurūf*.

The second opinion is that **bilā** is equivalent to *maʿa*, and is consid-
ered a noun with the function of a *ẓarf*. This opinion is evident in
Qawānīn. First it is pointed out that **bilā** is used as a translation of *bi*
(instrumental) in sentences like **qiliǧ bilā 'urdum** *ḍarabtu bi-s-sayfi* 'I
beat with the sword'. Its use as an equivalent of *maʿa* (comitative)
seems secondary, which is inferred from the use of the word *ʾayḍan*
'too': "the word **bilā** is also used with the meaning of *maʿa*, 'I came
with the chief' **bī bilā kaldum** *(wa-tustaʿmal lafẓa bilā 'ayḍan bi-
maʿnā maʿa ǧiʾtu maʿa l-'amīri bī bilā kaldum, Qawānīn 41,9).* In
another passage, however, he seems to assign the primary meaning of
maʿa to **bilā** "the word **bilā**... and its meaning is *maʿa*" *(lafẓa bilā...
wa-maʿnāhā maʿa, Qawānīn 34,4).* The most explicit statement with
regard to the status of **bilā** is found in a discussion related to the
equivalents of Arabic pronouns governed by *bi*. According to the au-
thor of *Qawānīn*, in Turkic pronouns are not governed by *bi*:

> "I do not know in this language a pronoun that is governed in the genitive
> by *bi*; instead, they use *maʿa* whose equivalent is **bilā**." *(wa-'ammā l-bāʾ
> fa-lā 'aʿrifu fī hādhihi l-luġa ḍamīran yuǧarr bi-l-bāʾ wa-'innamā
> yaǧurrūna bi-maʿa wa-yurādifuhā bilā, Qawānīn 45,11.)*

(This statement is followed by a listing of the Turkic pronouns in
combination with **bilā**, e.g. *maʿī* **manim bilā** 'with me', and *maʿahu*

---

[4] This holds as far as locatives of place are concerned. According to Arabic grammar
locatives of time cannot be used with *fī* in surface structure. In Turkic many locatives of
time (adverbs of time) do not have **dā**.

'anin bilā 'with him', etc..) Although the author seems to consider bilā as the equivalent of ma ʿa rather than bi, it is not clear whether he also assigns it the status of ẓarf. The use of the general term lafẓa 'word' for bilā gives no clues in this respect. In short, according to Qawānīn the Arabic particle bi has no equivalent in Turkic.

In Tarǧumān, too, the primary meaning of bilā seems to be ma ʿa. The fact that the word (lafẓa) ma ʿa is listed as an entry among other ẓurūf, e.g. al-wasaṭ 'middle' (used in the sense 'between') and ʿinda 'with', may serve as a first piece of evidence (Tarǧumān 54,3-6). A second indication may be the fact that in Tarǧumān bilā, besides the expected combinations—ma ʿaka 'with you' sanin bilā—is also used with dā:

9   'a-nin   bilā-si-n-dā
       he[OBL]-GEN   with-POSS-n[5]-LOC
       'with that one' (ma ʿa ḏāka).

Likewise, ma ʿa 'ulā'ika 'with them' is 'an-lār bilā-lār-i-n-dā, ma ʿī 'with me' bila-m dā and ma ʿanā 'with us' bilā-muz dā.[6] It is not clear whether bilā is assigned the status of ẓarf.

Šuḏūr's author Ibn Muḥammad Ṣāliḥ seems to have a preference for ma ʿa as the main equivalent of bīlah, his version of bilā, as well.[6] He refers to one of the other meanings of bi (i.e. muṣāḥaba 'accompaniment'), and thus first equates bi and ma ʿa in Arabic before translating them into Turkic. In this way "ma ʿa and bi of accompaniment [are] bīlah" (ma ʿa wa-bā' al-muṣāḥaba bīlah, Šuḏūr 26ʳ2) are equivalent in meaning. Ibn Muḥammad Ṣāliḥ uses the term naẓīr:

> "and thus you apply it in the construction with ma ʿa whose peer is [both] bīlah and lah." (wa-kaḏā ta ʿmal fī l-maẓhar at-tarkīb bi-ma ʿa naẓīruhu bīlah wa-lah, Šuḏūr26ʳ7ff.)

The fact that bīlah is also the equivalent of bi seems of secondary importance: "likewise, 'by me' is bīlah" (wa-kaḏālika bī bīlah, Šuḏūr 26ʳ9).

In the other sources neither opinion is obvious. In MG, for instance, both meanings of bi are expressed. In one instance bilā (along with the older variant birlā) is, as expected, mentioned as the equivalent of bi (cf. MG 58ʳrt). The author may have tried to solve the problem by discerning two dialectal variants of bilā. To one of them, lah (which is

---

[5] This -n- is inserted as an intermediate consonant after possessive endings (3sg and 3pl) when they are followed by cases.

[6] The orthographic differences in the reflection of bilā (bīlah, bilā, bīlā) are due to different writing conventions.

probably Ottoman) he assigns the sense of *bi* and to the other **bilah**, that of *ma'a*, viz., **banūm-lah** *bī* 'by me' versus **manūm bilah** *ma'ī* 'with me' (MG 59ʳbm).[7]

In *Ḥilya*, both *ma'a* and *bi* are regarded as a *ḥarf*, and, as a consequence, they may have one single equivalent: "[As for] the particle *ma'a* and the added [particle] *bi*, the word **bīlā** substitutes both of them in Turkic." (*ḥarf ma'a wa-l-bā' az-zā'id wa-humā fī t-turkiyya yanūbu 'anhumā lafẓa bīlā, Ḥilya* 93,1.)

2.3 *The case of* 'ilā

This section discusses the way the meanings of *'ilā* and *ḥattā*, and the notion 'until' are reflected in Turkic, and the analyses the Arabic grammarians give of the respective Turkic equivalents.

2.3.1 *'Īlā* meaning 'to'

The Arabic particle *'ilā* is basically used to convey a motion or direction toward a place or until a point in time, e.g. *ḏahabtu 'ilā makkata*[8] 'I went to Mecca' and *baqītu 'ilā yawmi l-qiyāmati* 'I waited until the day of Judgement'. When used in this sense, *'ilā* is called *ḥarf al-intihā'*. By extension *'ilā* is also used to express the end point of a motion *(al-ġāya)*, e.g. *waṣaltu 'ilā makkata* 'I arrived at Mecca', especially when it is opposed to *min*, 'from' (*Irtišāf* II 449,16; II 567,3; cf. Wright 1986 [1898] II 144ff.; Gully 1994:41). In this last meaning, *'ilā* is synonymous with the particle *ḥattā*, 'until', e.g. *'akaltu s-samakata ḥattā ra'sihā* 'I ate the fish until its head', excluding the end point itself, i.e. *ra'sihā*, 'its head'.

As in all Turkic languages, in the Turkic languages described in our sources, too, the notion 'to' or 'towards' is conveyed by means of the dative case:

9   **bak-imiz**   **ḥurāsān qā**   **bār-dī**
       chief-POSS/1pl   Khorāsān-DAT   go-PAST/3sg
       'Our chief went to Khorāsān'
       (*'amīrunā mašā 'ilā ḥurāsān, Ḥilya* 92,5).

---

[7] MG (59ʳbm) also lists various other meanings of *bi* (accompaniment, compensation, instrumentality, etc.) and their respective realisations in Turkic. In this respect, interestingly, *bi* is equated with the *'alif* /ˀ/ in ǧā /ġa"/ — and other 'elements' ending in -ā, some of which, in fact, consist of the final consonant of a word and a case ending (dative) — **ṣattūm art-ǧā** *bi'tu bi-zā'idin* 'I sold [it] with a profit'.

[8] *Makka* belongs to a class of nouns that do not get the ending *i* in the genitive case.

2.3.2 *'Ilā* meaning 'until'

Turkic case endings usually denote a very general motion or place, and a specific meaning, such as 'until', is conveyed by means of a postposition after the case ending. (It will be seen that this postposition is not the same in all sources.) In *'Idrāk* the dative case is followed by the postposition **daqin**:

10   makka-ḵā   daqin   yuri-di-m
        Mecca-DAT   until, as far as   walk-PAST-1sg
        'I walked as far as Mecca' (*sirtu ḥattā makkata, 'Idrāk* 145,12).

The issue we deal with in this section is how the grammarians coped with the problem of assigning a status to a combination of two elements, i.e. ḵā and **daqin**, which serve to express a notion that in Arabic is conveyed by means of one word.

The Turkic equivalent of *'ilā* is ġā, viz., "to the particles of the genitive belongs the word ġā, with the meaning of 'towards' that governs the genitive" (*wa-min ḥurūf al-ǧarr lafẓa ġā bi-maʿnā 'ilā l-ǧārra*, MG 58 top).[9] In *Qawānīn* (42,7), too, the first meaning of *'ilā* is ḵā or ġā:

11   šār ġā   bar-du-m
        city-DAT   go-PAST-1sg
        'I went to the city' (*ruḥtu 'ilā l-madīnati*).

12   kant ḵā   kat-tu-m
        town-DAT   go-PAST-1sg
        'I went to the town' (*ḏahabtu 'ilā l-baladi, Qawānīn* 42,7).

In *'Idrāk*, though, in the introductory list of particles, it is not ġā but **daqin** that is mentioned as the particle conveying the meaning of both *'ilā* and *ḥattā* (144,21). Furthermore, elsewhere (145,15), in the analysis of (10), it is specified that ġā and ḵā indeed express the meaning of *'ilā*, but that they are used pleonastically together with **daqin** (*maʿnā kullin minhumā 'ilā lākinnahā tustaʿmal murdifatan bi-daqin*). In other words, in 'Abū Ḥayyān's opinion, the basic meaning of *'ilā* is conveyed by means of **daqin**, whereas ġā has a merely complementary function.

In *Qawānīn* the addition of this element is also discussed, but it is assigned a different status:

---

[9] In the Margin Grammar and *Dīwān* the meaning of the dative case ending ḵā is described as "particle-locative with the function of 'to'" (*harf ẓarf bi-manzila 'ilā*, MG 40ʳrt; identical with *Dīwān* 538,5). This expression is difficult to interpret since no choice is made between the terms *ḥarf* and *ẓarf*.

"and if [ʾilāʾs] meaning is the reaching of the goal [i.e. 'until'], then add the word **dakin** to this marker" (*wa-ʾin kāna maʿnāhā intihāʾ al-ġāya fa-zid ʿalā hāḏihi l-ʿalāma lafẓa dakin,* 42,10).

This implies that **ġā** expresses the basic meaning, and that **dakin** is added as a complementary element.

2.3.3 Ibn al-Muhannā on 'until'

A third solution for this problem is given in *Ḥilya* (91,3), where **takī** (which reflects the same meaning as **daqin** and **dakin** in the other sources) is regarded as an equivalent of the Arabic *ḥarf al-ġāya* (particle of the goal), i.e. both *ḥattā* and *ʾilā* that express the goal *(al-ġāʾiyyatayni)*, e.g.,

13a	tāwuk	ʿIrāq-qā	takī	bār-miš
	so-and-so	ʿIraq-DAT	until	go-INFER/3sg
	'So-and-so went as far as ʿIraq.'			

Underneath some of the Turkic words their respective equivalents in Arabic are (in Rifʿat's edition) printed in a small font:

13b	tāwuk	ʿIrāq	qā takī	bār-miš
	fulān		*ḥattā*	*mašā*
	so-and-so		until	walk-PAST/3sg

However, Ibn al-Muhannā fails to describe the function and meaning of **qā**, and it seems that he regards **qā takī** as one single meaningful element. In addition, both *Qawānīn* and *ʾIdrāk* note that **ġā** may be added optionally after **daqin**, or **dakin**, resulting in the forms **daqin-ġā** and **dakinġā**, respectively (e.g. *ʾIdrāk* 145,14). In *Ḥilya* a similar form, **takinġā**, is given.[10]

In both *Qawānīn* and *ʾIdrāk* **dakin(ġā)** is analysed as a basic particle with an additional suffix (EDT 'equative suffix', see also *Fundamenta* I), although their opinions differ as to which of the two elements is used pleonastically. According to ʾAbū Ḥayyān, **ġā** is the complementary element, whereas in *Qawānīn* this is **dakin**.

In *Ḥilya* the analysis of **takinġā** is quite different from the one shared by *ʾIdrāk*, *Qawānīn* and, so to say, Clauson. In order to understand it fully, it is necessary to discuss briefly how, according to the Arabic grammarians, Arabic particles can be used with verbs, and, further the way the conjunction 'until' is conveyed in Turkic.

---

[10] EDT 477 and 484 distinguishes two postpositions.

I have already referred to the fact that *'ilā* is also used with the no-tion 'until a certain time'. However, the particle *'ilā* cannot be fol-lowed immediately by a verb. The reason for this is, according to the Arab grammarians, the fact that a verb cannot occupy the syntactic position of the genitive, because this is a privilege of nouns only. This problem can be solved by means of a special kind of noun, the *maṣdar*, translated with either 'verbal noun' or 'infinitive'. The *maṣdar*, apart from reflecting the meaning of a verb and possessing some verbal features, behaves like any other noun. As such the *maṣdar* may be sub-jected to governance by means of a particle: *'ilā maǧī'-i zayd-in* 'until Zayd's coming' or, 'until Zayd comes'. An alternative route, with a similar result, is the application of the particle *'an*, which, by governing the verb in the subjunctive, serves as an intermediate between the par-ticle and the verb. The combination, then, behaves as a verbal noun (cf. *'Uṣūl* II 297,1ff.):

14 *'ilā*　　　*'an*　　　*yaǧī'-a*　　　*zayd-un*
    until　　　*'an*　　　comes-SUBJ　　zayd-NOM
    'until Zayd comes'.

In Turkic the notion of the temporal 'until' is usually expressed by means of a suffix added to the stem of a verb. The resulting verbal form is not finite; it must be preceded by a pronoun:

15 [ṭur　　　　　mūn-dā]　　　man　　　kal-kinǧā
    [stand/IMP/2sg　here-LOC]　　I　　　come-until
    '[You stay here] until I come' (*'Idrāk* 151,3; also *Ḥilya* 91,9).

In the sources (i.e. *Qawānīn*, *'Idrāk*, *Ḥilya*, MG) the suffix **kinǧā** (after velar stems **ǧinǧā**) is, with this meaning, regarded as one of the equiva-lents of *'ilā* and *ḥattā*.

When this suffix is added to the stem **tak-** 'to reach', the result is **tak-kinǧā** (**tak-inǧā** in *Ḥilya*, 91,9f), a form quite similar to the one mentioned in *Qawānīn* and *'Idrāk*. In fact *'Idrāk* (150,20ff.) eventu-ally derives **daqin(ǧā)/dakin(ǧā)** from the verb **dak**, 'to reach', to which -**kinǧā** is added, allowing deletion of one **k** for alleviation. Returning to *Ḥilya*, in the sentence

16a bālīq　　yī-du-m　　　bāš-ī-n-ā　　　　takin-ǧā
     fish　　eat-PAST-1sg　head-POSS-n-DAT　until (POSTP)
     'I ate the fish until its head' (*'akaltu s-samakata ḥattā ra'sihā*),

Ibn al-Muhannā derives **takinǧā** directly from the verb **dak**, and, more importantly, he assigns it the status of a verb. This is obvious in his word-by-word analysis *(ta'rīb)* of (16a):

16b	bālīq	yī-du-m	bāš-ī-n-ā	takinǧā
	samak	ʾakaltu	ḥattā raʾsihi	balaġtu
	fish	I ate	until its head	I reached

2.3.4 Summary

Summarising, both *Qawānīn* and *ʾIdrāk* analyse the postposition **dak-inǧā/daqinǧā** in terms of a form **dakin** to which the optional suffix -**ǧā** may be added, in combination with -**ġā**. The basic meaning of *ʾilā* 'towards' is assigned to different elements, i.e. to **daqin** in *ʾIdrāk*, and to -**ǧā** in *Qawānīn*. In *Ḥilya* **takinǧā** has the status of a verb.

This difference of opinion is related to the fact that in Arabic grammar a sequence of two particles that govern the genitive is unacceptable. The reason for this is obvious: in such a case one of the particles would be subjected to governance by the other, or both of them would govern the same noun, which is impossible. We have seen above that both **ġā** and **daqin**(**ǧā**) are regarded as equivalents of Arabic particles; in this respect the fact that they follow one another is basically impossible to explain in terms of Arabic grammar, and the analysis the two sources give is, in fact, contrary to the principles of Arabic linguistic thinking.

Ibn al-Muhannā's approach to the form of **takinǧā** in terms of a verbal form conveying the meaning of 'reached', seems much more in agreement with these principles. Ibn al-Muhannā regards the sequence -**ā** + **takinǧā** as one of a verb and a particle, which is, of course, quite possible. His analysis of **qā takī**, however, which he—implicitly—regards as one element (cf. [13]), remains unexplained.

2.4. *The possessive construction*

A special case of the use of the particles of annexation is the possessive construction. It is special since in Arabic the particle is usually absent in the surface structure: it is an element the (later) grammarians posit between two nouns in the underlying structure of an *ʾiḍāfa* construction. The reason for this is the apparent tendency in late Arabic grammar to reduce the number of governors. They did not, like their predecessors, accept the possibility of one noun governing another, hence causing the genitive case ending on the governed noun. Instead, they posited a *ḥarf ǧarr* in the underlying structure that is to account for the genitive case.

2.4.1 *Li* and the possessive construction

The insertion a particle can also be explained in terms of semantic and syntactic case (cf. Chapter Four 4.1.1). For example, when the genitive occurs in the possessive construction *bayt-u zayd-in* 'Zayd's house', it could be interpreted as a semantic case (cf. Chapter Four 1.2.4).

However, it seems that the later grammarians preferred to analyse it as a syntactic case, i.e. in terms of government, thus enhancing the predictability of both governing capacities and declensional endings and smoothening their argumentation. The particle they posit is usu-ally *li* 'to', for example, the underlying structure of *bayt-u zayd-in* 'Zayd's house' is, in their opinion, something like *al-bayt-u li-zayd-in*, with *li* as the governing particle (see discussion above). The particle may also be *min* 'from', if the genitive conveys a partitive meaning, e.g. *ṯawb-u ḥazz-in* 'a cloth [made] of silk' is rephrased as *ṯawb-un min ḥazz-in.*

This principle in Arabic theory is explained clearly in the Margin Grammar, viz.,

> "Sunqur's slave' i.e. *'a slave to Sunqur'; 'that one's slave' i.e. *'a slave to that one' with the implied *li*; 'whose slave', with [the underlying structure] *'to whom his slave'. *Li* is implied in the underlying structure." (*mamlūku sun-qurin ʾay *mamlūkun li-sunqur wa-mamlūku ḏā ʾay *mamlūk li-ḏā bi-l-lām al-muqaddar ʾay mamlūku man bi-maʿnā li-man mamlūkuhu bi-taqdīr al-lām*, MG 44ʳmd.)

To put it in another way, the phrase *mamlūku sunqura* 'Sunqur's slave' is rephrased in the underlying structure as **mamlūkun li-sunqur* 'a slave to Sunqur', with the addition of the particle *li.*

2.4.2 The possessive construction

This section discusses how the grammarians explained the possessive construction in Turkic. It will be shown that the concept of a particle in the underlying structure is convenient for the analysis of Turkic possessive constructions too.

In the sentence *al-bayt-u li-zayd-in* 'the house [belongs] to Zayd', the phrase *li-zaydin* 'to Zayd' is the predicate *(ḫabar)* to the topic *(mubtadaʾ) al-baytu* 'the house'. Here the particle *li* cannot be omitted in the surface structure, for a phrase like, for example, **al-bayt-u zayd-in*, without the insertion of *li*, is regarded as non-grammatical.

In Turkic it is of course also possible to construct predicative sen-
tences of this type, but, unlike in Arabic, it suffices to put the predicate
in the genitive[11]:

17	bū	ʾat	zayd	niŋ	[durur]
	this/NOM	horse	zayd	GEN	[CORR]

'this is Zayd's horse' (hāḏā l-farasu li-zaydin, Ḥilya 93,4).

The sources usually interpret **niŋ** (with a variant -**iŋ** in Oġuz
[*turkmāniyya*]) as the equivalent of *li*; it "has the meaning of" *li*
(*ʾIdrāk*, 145,4). For *Qawānīn* the word *(lafẓa)* **nin** conveys the mean-
ing of *li* for possession *(milk)* and the specification *(al-iḫtiṣāṣ)*, which is
the way the function of *li* in sentences of the type in (1) is described in
Arabic grammar

The status of this element **niŋ** in (17) is described slightly differ-
ently in *Ḥilya*:

> "The last *n* [i.e. ŋ] belongs to the same category as the redundant [particle]
> *li*" *(fa-n-nūn al-ʾaḫīra bi-manzila al-lām az-zāʾid* [sic], *Ḥilya* 93,4).

This statement is interesting in two respects. In the first place it says
something about Ibn al-Muhannā's analytical approach to **niŋ**, in the
sense that he segments it into two parts, i.e. n + ŋ. This is possible in
the Turkic language he describes, since in Oġuz Turkic languages ŋ
alone represents the genitive case, i.e. **zayd-iŋ**. In the second place it is
of interest because of the application of the terms *manzila* and *zāʾida*
in this context. The term *manzila* refers to the status of **niŋ** compared
to *li* in Arabic, and the term *zāʾid* is used to specify the position of *li* in
the Arabic sentence.

In Arabic linguistic terminology the term *zāʾid* is used to indicate
elements that are inserted, and, as a result, exert governance, but
without adding any extra meaning to the sentence. For example, the
particle *li* in *ḍarabtu li-zaydin* 'I beat [to] Zayd', does not cause un-
grammaticality of the sentence, but it does not contribute significantly
to its meaning either (cf. *Irtišāf* II 435, 4ff.; cf. also Versteegh
1993:144-5 and 151 for this term in early grammatical terminology).
As a result, the particle *li* forms no part of the underlying structure,
since it is semantically redundant. Interestingly, according to Arabic
theory the *lām* in the translation of example (17) is not considered
zāʾid, it is an essential part of the predicate of the sentence. In regard
to the term *manzila*, Versteegh (1978: 264ff.) states

---

[11] The fact that in Turkic the word in the genitive can serve as a predicate may be in-
terpreted as yet another indication for the typology of its case.

"If a word A is said to be at the same *manzila* as a word B, it belongs to the same class as word B... it may even take over some of the functions... of word B".

The same definition applies in the case of Ibn al-Muhannā's use of *manzila*.

It seems that Ibn al-Muhannā compares **niŋ** to *li* in two respects. In the first place both denote the sense of 'property' *(milk)*, and secondly, both **niŋ** and *li* are, at least in Ibn al-Muhannā's view optional, i.e. they may be deleted. This analysis of *li* probably relates to Ibn al-Muhannā's variant interpretation of the same Turkic sentence else-where (85,12), which we shall deal with below.

2.4.3 The possessive ending

In Turkic a noun can be marked for a possessive relation to a pronoun by adding the appropriate pronominal marker:

18   'aṭ-im
        horse-POSS/1sg
        'my horse'.

Likewise, 'aṭ-iŋ 'your (sg.) horse'; 'aṭ-ī 'his/her horse', 'aṭ-imiz 'our horse', 'aṭ-iŋiz 'your (pl.) horse'. In the case of 'aṭ-larī, two analyses—or even three—are possible, viz., first 'aṭ-larī 'their horse' and second 'aṭlar-ī 'his horses'. In the first case **lar** is considered part of the reference to the possessor, which then must be plural (horse-their), whereas in the second it is considered part of to the possessed noun (horses-his). A third possible interpretation for this form is that both possessor and possessed are plural, e.g., 'their horses' (horses-their), because the suffix cannot appear twice in the same word for double plurality **'aṭlar-larī (cf. also *'Idrāk* 147,16f).

In order to stress (*taʾkīd, Qawānīn* 46,16) the possessive relation, the 'possessed' noun may be preceded by a personal pronoun which is marked for the genitive case, e.g. **ban-im** 'aṭ-im '*my* horse', **san-iŋ** 'aṭ-iŋ '*your* horse', 'a-niŋ aṭ-ī '*his/her* horse', **biz-im** 'aṭimiz '*our* horse', **siz-iŋ** 'aṭ-iŋiz '*your* (pl.) horses' and 'an-lar-niŋ 'aṭ-ī or 'aṭ-lārī '*their* horse(s)'.[12] The plural ending -**lar** is inserted between the noun and the possessive ending, e.g. 'aṭ-lār-im 'my horses'. Case endings follow after the possessive ending, e.g. **san-iŋ qaṭ-iŋ-dā** (you-GEN side-POSS/2sg-LOC) 'at your side'.

---

[12] The form **banim**, with b- and final -m is typically Oğuz; the sources also display variants from other Turkic languages, e.g. **maniŋ**, 'allarniŋ, **bizniŋ** etc..

In Turkic, two nouns can stand in a possessive relation for which both elements must be marked, the possessor with the genitive case, and the possessed with the possessive ending, respectively:

19   bī-niŋ   ʾaṭ-ī
        chief-GEN   horse-POSS/3sg
        'the chief's horse'

After vowels, an s is inserted between the word and the possessive marker, e.g.

20   bī-niŋ   ʾanā-s-ī
        chief-GEN   mother-s-POSS/3sg
        'the chief's mother'

In the following we shall discuss the status of the possessive ending, and the genitive case ending niŋ. We shall see that the sources had no problem with assigning the status of pronoun to the possessive ending, except for the ending of the third person. Further, we shall find that niŋ is generally regarded as equal to the particle li 'to'.

The Turkic possessive endings are described in terms similar to the Arabic construction of annexation (ʾiḍāfa) to pronouns:

> "the annexed noun precedes the [noun] it is annexed to, analogous to the Arabic." (wa-taqaddama [l-muḍāf] ʿalā l-muḍāf ʾilayhi qiyāsa l-ʿarabiyya, Qawānīn 46,2).

One would expect the suffixes to be analysed in a similar way. This as-sumption, however, holds only partially.

First of all, the Margin Grammar gives an analysis according to which the possessive endings are considered pronouns:

> "The pronoun for the first person is m, and for the first person plural it is miz; for the second person it is ŋ, and for him and for the second person plural it is ŋiz; for the third person it is a silent y, and for the third person plural it is lārī." (aḍ-ḍamīr li-l-mutakallim m wa-lahu wa-li-man maʿahu miz wa-l-li-muḥāṭab ŋ wa-lahu wa-li-man maʿahu ŋiz wa-li-l-ġāʾib y sākina wa-lahu wa-li-man maʿahu lārī, MG 59ᵛlt/bm; also 59ᵛrt.)

ʾIdrāk and Qawānīn give very similar analyses for the first and second persons (ʾIdrāk 118,15; 136,17).

For the ending ī (or i) of the third person, however, ʾIdrāk, Qawānīn and Ḥilya supply less explicit information. Even though they state in several instances that it is attached "you vocalise it with an ī" (taksiruhu), e.g. qul-ī 'his slave' (mamlūkuhu, ʾIdrāk 120,1f), nowhere is the status of this ending described. In fact, in some cases it is en-tirely disregarded. ʾAbū Ḥayyān, for example, writes with regard to qul-lari ṭur-du-lar

"i.e. 'their slaves stood up'. **lar** is attached to the singular if you annex to a singular [noun], and [its] **r** is vocalised with an *i*, this indicates the plural of the annexed; it may also indicate the plural of the ones to whom is annexed. This can be deduced from the context." (*'ay mamālīkuhum qāmū fa-lar talḥaqu l-mufrad fa-'idā 'udīfa 'ila mufrad wa-kusirat ar-rā' dalla dālika ʿalā ǧamʿ al-muḍāf wa-yadullu 'aydan ʿalā ǧamʿ al-muḍāf 'ilayhim wa-yatabayyanu dālika min siyāq al-kalām, 'Idrāk* 147,17f.)

From this statement it appears that he does not assign a distinctive meaning to **i** and hence, does not consider it a pronoun (*ḍamīr*).

In *Tuḥfa* a parallel is seen with the endings of the Turkic imperfect tense in **-dī**: "The marker of the third person is the silent **y** in nouns and verbs" (*ʿalāma al-ǧā'ib yā' sākina fī l-'asmā' wa-l-'afʿāl, Tuḥfa* 39ᵛ6). Indeed, the verbal endings are, apart from the first person plural, identical with the possessive endings in nouns: **kāldīm** 'I came', **kāldīṇ** 'you (sg.) came'; **kāldī** 'he/she came'; **kāldīk** 'we came'; **kāldīṇīz** 'you (pl.) came'; **kāldīlār** 'they came'. [13] However, because *Tuḥfa* applies the term ʿ*alāma* 'marker' indiscriminately to all words and morphemes, the status of these 'markers' remains unclear.

In *Ḥilya* (85,9) the possessive endings are considered markers too. For **qūl-um** 'my slave' *ǧulām-ī* Ibn al-Muhannā writes

"the silent **m** is the marker of the annexation to the speaker himself" (*wa-l-mīm as-sākina ʿalāma al-'iḍāfa 'ilā nafsi l-mutakallim, Ḥilya* 86,1).

According to *Qawānīn* (59,4) the **y** is a marker of annexation (ʿ*alāma li-l-'iḍāfa*), rather than a pronoun. *'Idrāk* assigns no function to the **y** exclusively.

The reason for the apparent hesitance in *'Idrāk* and *Qawānīn* with regard to assigning the status of pronoun of the third person to **ī** (or **i**) becomes more understandable when we recall what 'Abū Ḥayyān states in the first part of the grammatical section of his work:

"All three weak consonants [sc. *'alif, wāw* and *yā'*]... arise from the lengthening of the vowels." (*ǧamīʿ ḥurūf al-madd wa-l-līn aṯ-ṯalāṯa... 'innamā hiya nawāši' ʿan 'išbāʿ al-ḥarakāt, 'Idrāk* 101,10.)

No meaning can be assigned to the glides, 'Abū Ḥayyān intends to say, they are secondary since they originate from the vowels. Assigning the status of pronoun to a vowel is impossible within the framework of Arabic grammar, which concentrates on consonants rather than vowels. In Arabic grammar the *y* can have the function of a pronoun, for instance, in *ǧulāmī* /ǧulām-iy/ 'my slave', the *y* is the pronoun of

---

[13] In verbs like **'ur-sa-ṇ** (beat- COND-2sg) 'if you beat' the ṇ, too, is called pronoun of the second person (*ḍamīr al-muḥāṭab, 'Idrāk* 146, 20).

the first person. The preceding *i* merely follows the glide. In Chapter Four we showed that in Arabic grammar vowels can only serve as markers as the result of governance, and are never considered as governing elements themselves.

This indeed seems to be the status assigned to ī/i in *Qawānīn*, *'Idrāk* and *Ḥilya* i.e. 'a marker of annexation', i.e. a sign that the word is involved in an annexation construction. We shall discuss this in greater detail below.

2.4.4 The intermediate s: various opinions

We have already pointed to the fact that when the second word of the Turkic sequence ends in a vowel, it gets an intermediate **s** before the possessive ending ī, such as in **'anā-s-ī** 'his mother' (cf. above 20). In Arabic grammar the *n* can fulfil a similar function, e.g., *ḍaraba-n-ī* /ḍaraba-n-iy/ 'he beat me', in which *n* is inserted between the final vowel of the verb, i.e. *a*, and the *i* that precedes the pronoun *y*. In Arabic grammar this *n* is called *nūn al-wiqāya* 'protective *n*'. It is not considered a part of the pronoun that is at the position of the object, since it does not appear after particles, e.g. *lī* 'to me' and *maʿī* 'with me', and when the pronoun is governed by an active participle *ḍārib-ī* 'the one who beats me' (cf. *Irtišāf* I 470,8ff.). In the Margin Grammar, for instance, the function of the Turkic intermediate **s** is compared to that of this Arabic *n*.

> "You add to [the annexed element] an *s* vocalised with an *i* if its last consonant is weak [i.e. if it ends in /iy/, /aʷ/ or /uw/], it is [like] the protective *n* in Arabic in verbs or in particles that resemble [verbs], e.g. *ḍaraba-n-ī* 'he beat me'…. This is in nouns that end in an *a* with a lengthened vowel to which an *'alif* is added, [which serves to] protect [the *'alif*] against [contact with] the *i*." (*talḥaq [al-muḍāf] sīnan maksūra 'in kāna 'āḫiruhu muʿtallan wa-hiya nūn al-wiqāya ʿarabiyyatan fī l-fiʿl 'aw al-ḥarf allaḏī 'ašbahahu, naḥwa ḍarabanī… lākinna hāḏā fī l-ism al-maftūḥ al-'āḫir al-mušbaʿ allaḏī laḥiqathu al-'alif wiqāyatan lahu min al-kasr,* MG 59ʳrt.)

In the case of **'anā-s-ī** /'anāʷ-s-iy/ 'his mother', therefore, **s** is inserted to prevent the *'alif* /ʷ/ from being directly followed by *i*, since *'alif* cannot be vocalised.

A similar statement with regard to the **s** is found in *Qawānīn*, viz.,

> "The secret in this [i.e. the insertion of the **s**] is that the last consonant of the annexed noun is an *'alif* and the marker of the annexation is the **y**, both of them weak consonants. Therefore they insert the **s** between them in order to prevent a sequence of two weak consonants to occur." (*wa-s-sirr fī ḏālika 'anna l-ism al-muḍāf 'āḫiruhu 'alif wa-ʿalāma al-'iḍāfa yā' wa-humā*

ḥarfā ʿilla wa-ʾadḥalū as-sīn baynahumā li-ʾallā yatawālā ḥarfā ʿilla,
Qawānīn 44,20f.)

Here the intermediate **i** is not considered at all.

In *Dīwān* (536,13ff; copied in MG 39ʳrt) the insertion of **sī** is ac-
counted for with the same kind of phonological arguments. If the last
consonant of a word is vocalised in annexation, in the way the **l** in
ʾuġul 'son' is vocalised with an **i**, /ʾuġl-i/, only **y** is added instead of **sī**,
hence /ʾuġliy/, schematised this is reflected as follows:

    a. /ʾuġul/ → vocalisation /ʾuġul-i/ → addition of /y/ = /ʾuġul-iy/

In some cases the last consonant cannot be vocalised, for example if it
is *ʾalif* /ʾ/. Here Kāšġarī seems to presuppose the addition of two **y**'s,
$/y_1$-$y_2/$. If two **y**'s *(bi-yāʾayni)* were added after the *ʾalif*, he says, one
would get a sequence of three weak consonants, i.e. /ʾ-y_1-y_2/, viz.,
*/ʾataʾ-y_1-y_2/, which is impossible. The first /y_1/, then is replaced
(ʿuwwiḍat) with /s/, resulting in /ʾataʾ-s-iy_2/ *ʾatāsī* 'his father'. In a
schematised way his way argumentation can be represented as follows:

    b. Theoretical form */ʾataʾ-y_1-y_2/ → replacement of /y_1/ with /s/ →
/ʾataʾ-s-iy_2/ = *ʾatāsī*

The argument of two **y**'s seems also applicable to words that take a fi-
nal vowel, e.g.,

    c. Theoretical form */ʾuġul-y_1-y_2/ → deletion of /y_1/ → /ʾuġul-y_2/
(insertion /i/) → /ʾuġul-iy_2/ = **ʾuġlī**

There are still some problems, though. For example, it is not clear at
which point exactly the vowel /i/ is inserted. For (c) this does not mat-
ter much, but in (b) we might get the sequence /ʾi/, which we wanted
to avoid from the start. Therefore one is inclined to assume that in (b)
/s/ takes the vowel. In either case /y_1/ is either replaced or deleted.
Kāšġarī's argumentation with two /y/s rather than one probably relates
to the fact that in Arabic theory replacement of consonants is easier to
explain than insertion.[14]

---

[14] In *Dīwān* and MG the text runs as follows: "**sī** is a particle of annexation which is
attached to the last consonants of the nouns if the last consonant of the word is silent, but
if it is vocalised the **y** [alone] suffices to serve as its indicator. They say **ʾatā** (/ʾataʾ/), i.e.
'father. The *ʾalif* is silent, and if something is annexed to it they say **ʾaniŋ** **ʾatāsī** 'his fa-
ther' and they say' and they say **ʾanā** (/ʾanaʾ/) 'mother'. The *ʾalif* is silent too. It is not
possible to say **ʾaniŋ** *ʾatāyī** with two **y**s, because in that case three weak consonants
would stand in a sequence. The middle one is substituted with the **s** so that the speech
becomes euphonious and its softness increases. The vocalised is, e.g., **ʾuġul** ' son', they
say **ʾuġlī**, i.e. 'his son'. The **s** is not needed here, because the **l** is vocalised in the annexa-
tion...." (*sī ḥarf ʾiḍāfa yalḥaqu bi-ʾawāḥir al-ʾasmāʾ ʾiḍā kāna ʾāḥir al-kalima sākinan
wa-ʾiḍā kāna mutaḥarrikan yaktafī bi-l-yāʾ bayānuhu yuqāl ʾatā wa-huwa l-ʾab wa-l-
ʾalif minhu sākina fa-ʾiḍā ʾuḍīfa yuqāl ʾaniŋ ʾatāsī ʾay ʾabūhu wa-yuqāl ʾanā li-l-ʾumm
ṯumma yuqāl ʾaniŋ ʾanāsī al-ʾalif minhu sākina kaḍālika fa-lam yumkin ʾan yuqāl*

Another account of the insertion of **s** is found in *Ḥilya* (85,1ff):

"If the last consonant of the annexed noun is a weak consonant, then add an **s**
before this consonant, and vocalise the consonant *before* the **s** with an *ī*" *(fa-
ʾin kāna ʾāḥir al-muḍāf ḥarf ʿilla, fa-zid qabla l-ḥarf al-muʿtall sīnan
maksūran mā qablahā).*

He illustrates this with the word **qamǧī** /qamǧ-iy/[15] 'whip'. If /s/ is in-
serted between /ǧ/ and /y/, and subsequently /ǧ/ is vocalised with /i/,
the result is **qamǧ-is-ī** 'his whip'. However, the same procedure does
not hold for words like, e.g., **yā** /ya"/ 'bow'. In this case insertion of **s**
would result in */y-is-a"/, whereas the correct form is **yā-sī** 'his bow',
which —with some other examples that all end in -ā— is nevertheless
given by Ibn al-Muhannā.

ʾAbū Ḥayyān, not accepting ī (or i) as a pronoun, associates **s** with
the 'annexation' itself:

"**si** is the indicator of the annexation to the third person, if the last
consonant of the noun is vocalised" *(wa-si dalīl al-ʾiḍāfa li-l-ǧāʾib ʾiḏā
kāna ʾāḥir al-ism mutaḥarrikan, ʾIdrāk 147,5).*

Note that he uses the word *dalīl*, 'marker, indicator', for **s**, instead of
assigning a definite status. ʾAbū Ḥayyān does not regard **s** and **i** as
separate morphemes; in other instances the **s** is simply "for the annex-
ation" (*li-l-ʾiḍāfa, ʾIdrāk* 147,13) or "attached for the annexation"
(*lāḥiqa li-l-ʾiḍāfa, ʾIdrāk* 152,2), both of which amount to the same
thing.

In *Tuḥfa* the same evasive terminology is used, it seems, with regard
to this **s**. The principle of its attachment is described but its status re-
mains basically unexplained: "the **s** is added for the annexation" (*wa-
tuzād as-sīn li-l-ʾiḍāfa, Tuḥfa* 89ᵛ7, also 60ᵛ8).

A case similar to that of the **s** is the insertion of another consonant,
viz., **n**. This is put directly after the possessive ending of the third per-
son when it is followed by a locative or ablative case ending:

21	**ʾard**	**-i-**	**n-dā**
	space behind someone-	POSS/3sg-	n-LOC
	'behind him'.		

---

*ʾaniŋ *ʾatāyī bi-yāʾayni li-ʾannahu kāna yaǧtamiʿu fīhā ṯalāṯa ʾaḥruf min ḥurūf al-līn
fa-ʿuwwiḏat al-wāsiṭa minhā bi-sīn ḥattā ʿaḏuba al-lafẓ wa-rtafaʿa l-līn wa-ʾammā l-
mutaḥarrik naḥwa qawlihim ʾuǧul li-l-ibn ṯumma yuqāl ʾaniŋ ʾuǧli ʾay ibnuhu fa-lam
yuḥtaǧ hāhunā ʾilā s-sīn li-ʾanna l-lām qad taḥarrakat fī l-ʾiḍāfa... Dīwān 536,13ff =
MG 39ᵛrt.)*
[15] Cf. EDT 626.

To be sure, after the other possessive endings, no such **n** is inserted (22), and hence **'ard-imiz-dā** 'behind us' etc.:

22   **'ard-**   **iŋ-**   **dā**
         space behind someone-   POSS/2sg-   LOC
         'behind you (sg.)'.

This **n** (in 21) is described in *'Idrāk* as follows:

> "The pure **n** indicates annexation to the third person [and the nasal **n** (i.e. ŋ) indicates the second person]." (*fa-n-nūn al-ḫāliṣa tušʿir bi-l-'iḍāfa li-l-ġā'ib [wa-l-ḫayšūmiyya li-l-muḫāṭib], 'Idrāk* 136,18.)

In these contexts we have understood the references to the **i, s** and **n** with the terms ʿalāma al-'iḍāfa, 'marker for the annexation', or li-l-'iḍāfa, 'for the annexation', as deliberately vague, rather than as implicit indications for the annexation to the third person singular, and hence the acceptance of -ī (or i) as a pronoun.

2.4.5 The word sequence of the possessive construction

The other important element in the Turkic possessive construction is the genitive case ending and the fact that, compared with Arabic, the elements stand in the reverse order. I first deal with the sequence of the elements, and continue with the ending **niŋ**.

A striking difference between the Arabic and the Turkic possessive constructions is that the sequence of the elements involved is reversed, e.g. (19) **bī-niŋ 'aṭ-ī** (chief-GEN horse-POSS) versus *faras-u l-'amīr-i* (horse-NOM the chief-GEN). As has been shown above, this is not the case when possessive suffixes are added to a noun. In terms of Arabic theory this reverse sequence presents a problem, since in their view governance works in one direction only, i.e. from the beginning of the sentence towards the end, so that in principle the governance relations in the Turkic constructions would be the reverse of those in Arabic.

All sources pay repeated attention to this point, e.g. "The element to which is annexed precedes the annexed in this language." (*al-muḍāf 'ilayhi yataqaddamu fī hāḏihi l-luġa ʿalā l-muḍāf, 'Idrāk* 146,16; also *Ḥilya* 84,15f; MG 59ᵛlt/bm; *Qawānīn* 44,2). In the example above **bī** is the *muḍāf 'ilayhi*, and **'aṭ-ī** the *muḍāf*, reverse in sequence, but analogous to the analysis of the Arabic phrase *faras-u l-'amīr-i (muḍāf-muḍāf 'ilayhi)*. This remains without consequences for governance. The inverse sequence recurs in the case of the particles of the genitive, as has been shown above, and further, for example, with attributes *(naʿt)*, like adjectives in which the adjective precedes the qualified

noun. The grammarians seem to accept this awkward sequence in Turkic without much discussion.

2.4.6 Various interpretations of niŋ

We now continue the discussion of the interpretations the sources give of the genitive case ending niŋ. We have seen above that the grammarians equated the ending niŋ with the particle *li* when used in predicative sentences, which in that case seems quite obvious, since *li* is typically used in predicative sentences of this type in Arabic. In regular genitive constructions in Arabic things are different, since in the surface structure *li* does not appear. The grammarians are supported by their analysis of Arabic possessive constructions in which they posit the particle *li* in the underlying structure as a connective element.

’Abū Ḥayyān gives the following sentence:

23a	**kim-niŋ**	**qūl-ī**	**san**
	who-GEN	slave-POSS/3sg	you-NOM

'whose slave are you?' (*ġulāmu man ’anta*, *’Idrāk* 146,17).

Furthermore, he gives a word-by-word analysis of (23a), which sheds light on the way the respective functions of the Turkic words and morphemes are interpreted:

23b	**kim-niŋ**	**qūl-ī**	**san**
	li-man	*mamlūku-hu*	*’anta*
	to whom	his slave	you

In this analysis, it transpires that niŋ is regarded as the equivalent of the particle *li*.[16]

’Abū Ḥayyān is not the only one to equate niŋ with *li*, also in other sources this association is made, e.g.,

> "[The particle] *li* of annexation, i.e. nin is inserted to the noun to which is annexed" (*wa-tadḫulu lām al-’iḍāfa ʿalā l-muḍāf ’ilayhi wa-hiya nin*, *Qawānīn* 44,5; cf. similar statements in MG 59ᵛrt/ult; 43ᵛlt; also called *ḥarf al-’iḍāfa, Qawānīn* 73,2).

In *Tuḥfa* (89ʳ3), the Margin Grammar (59ᵛrt), and *Ḥilya* (85,9) the term ʿalāma al-’iḍāfa is applied to niŋ. In the case of *Tuḥfa* this is not very significant, since it uses the term ʿalāma for almost all Turkic nouns and morphemes. In *Ḥilya* it appears to be connected with Ibn al-Muhannā's view of the Arabic possessive construction.

Let us now return to Ibn al-Muhannā's twofold interpretation of the element niŋ in (17): **bū ’at zayd niŋ durur**. One of his interpreta-

---

[16] Also, the equivalent of qul-ī is *mamlūku-hu* 'his slave', but we concluded above that the *yā’* (ī) is not explicitly assigned the status of a pronoun.

tions is given above; Ibn al-Muhannā regards **niŋ** as the equivalent of the redundant particle *li*. This means that in his view both **niŋ** and *li* can be deleted from surface structure, which as a matter of fact does not agree with Arabic theory on this point. Ibn al-Muhannā's analysis (93,4) may be reconstructed as follows:

a) **bū 'at**   is equivalent to   *hāḏā l-farasu* (this horse)
b) **zayd niŋ**   *li-zaydin* (to zayd)
[c) the copula **durur**]

This amounts to the same as the analyses in *'Idrāk* and *Qawānīn* on the one hand, and the Western analysis on the other. In another instance (85,11), however, Ibn al-Muhannā translates this same Turkic sentence with *hāḏā faras-u zayd-in* 'this is Zayd's horse' (followed by a similar example, translated in the same way):

a) **bū**   is equivalent to   *hāḏā* (this)
b) **'at**   *faras-u* (horse-NOM)
c) **zayd niŋ**   *zayd-in* (zayd-GEN)
[d) the copula **turur**]

In the first sentence *hāḏā l-farasu* is the topic and *li-zaydin* the predicate, whereas in the second *hāḏā* is the topic, and *farasu zaydin* the predicate.

All this gives us reason to believe that Ibn al-Muhannā did not see any basic difference between the two Arabic sentences. In fact, he states explicitly that he regards *li* as *zā'id*, i.e. 'redundant', or 'optional', and in this respect, he probably considers the two Arabic sentences as variants that are equivalent to each other in meaning and construction. With this optionality of *li* he accounts for the fact that in Turkic **niŋ** is deleted in regular possessive constructions when constructed with proper names, e.g. **zayd 'atī** 'Zayd's horse' (*Ḥilya* 84,16). One could say that Ibn al-Muhannā regards the two sentences (i) *hāḏā farasu zaydin* and (ii) *hāḏā l-farasu li-zaydin* as equivalent, rather than viewing (ii) as the underlying structure of (i).

This approach not only has consequences for his analysis of both Arabic sentences, but also for that of the Turkic phrase, and the element **niŋ** in particular.

2.4.7 Summary

In the preceding sections I have tried to show that only one source directly applies the status of pronoun *(ḍamīr)* to the possessive ending ī, whereas the other sources seem to be more hesitant in assigning any

major function to it. The reason for this is probably that, in the view
of most authors, y is an extension *('išbā')* of the vowel, and in Arabic
grammar vowels cannot in principle assume the function and position
of a pronoun.

The term 'marker of annexation' *('alāma al-'idāfa)* also conjures
up associations with two related terms in Arabic grammar, namely
'marker of objectivity' *('alāma al-mafʿūliyya)* and 'marker of agency'
('alāma al-fāʿiliyya). In Chapter Four I showed that these terms are re-
lated to the abstract notions 'genitive' *(ǧarr)*, 'accusative' *(naṣb)*, and
'nominative' *(rafʿ)*, respectively, linking each of them to a specific syn-
tactic function. In this respect, *i*, originally only the marker of the
genitive case, is identified with the genitive case itself, and, by exten-
sion, it is also interpreted as a marker for the function of annexation
('idāfa).

In regard to the application of the term *'alāma al-'idāfa* to the
Turkic ending ī, it should not be interpreted as a reference to the
function of 'grammatically possessed' *(muḍāf)* and hence as an indica-
tor for governance, but rather as a reference to a more general notion,
namely the fact that the noun is involved in a possessive relation.

The labels assigned to s and n—the second referred to in *'Idrāk*—
which are not considered basic parts of the nouns, as they occur in
limited morpho(no)logical contexts only, give some support to this
idea. In *'Idrāk* this s is called "indicator of the annexation" *(dalīl al-*
'idāfa), which indicates the assumption that there are mor-
pho(no)logical reasons for its insertion, and the n "indicates [lit.
'makes feel'] the third person" *(tušʿir bi-l-'idāfa li-l-ǧāʾib)*. In neither
case is there direct assignment of function, although the relation with
the 'annexation' is indicated. As far as *'Idrāk* and *Qawānīn* are con-
cerned, no element is explicitly assigned the function of pronoun of
the third person.

The association of **niŋ** with the Arabic particle *li* is made by all
sources. They may have had two major reasons for this association. In
the first place the application of **niŋ** in predicative sentences as a direct
equivalent to *li*, and, secondly, the fact that *li* is posited in the underly-
ing structure of annexation constructions. In *Ḥilya*, Ibn al-Muhannā
does not differentiate between these different applications of *li*. He re-
gards the underlying structure as a variant of the sentence, considering
li a redundant particle *(zāʾid)* that can be deleted from the surface
structure without distortion of the meaning. In both *Ḥilya* and the
Margin Grammar, **niŋ** is called the 'marker of the annexation'
('alāma al-'idāfa).

2.5 *The locative*

The concept of *ẓarf* (locative) is related to three main themes in Arabic grammar. In the first place it is related to the notion of governance; more specifically the governance of optional objects by the verb. Secondly, it is connected to the principle of governance by the particles of the genitive, especially *fī*, further, the concept of annexation plays an important role. In our hypothesis the Arabic *ẓarf* is an instance of semantic use of the accusative case. It shall be seen that in Turkic locatives the accusative plays no role at all (An introduction on Arabic locatives is given above in Section 1.2).

2.5.1 Two types of locatives

In Arabic grammar the locative *(ẓarf)* is originally regarded as one of the optional objects to the verb, 'the object in which' *(al-mafʿūl fīhi)*, with the particle *fī* implied in the underlying structure and absent in the surface structure. Later grammarians, however, also admitted the appearance of *fī* in the surface structure, introducing semantic arguments for distinguishing a functional class of *ẓurūf* (see above 1.2). The locatives are typically divided into two classes: the locative of time *(ẓarf zamān)*, and the locative of place *(ẓarf makān)*. *Qawānīn* gives the following general definition of a locative:

> "In Arabic [the locative] is as a rule expressed with [a noun that] regularly
> includes [the meaning] *fī*. In this language it is as a rule expressed with the
> word *fī*, i.e. [dā], which they compulsorily put after [the locative], whether
> the latter is of time or place." *(wa-yuʿabbirūna ʿanhu fī l-ʿarabiyya mā
> taḍammana fī bi-ṭṭirād wa-ʾammā fī hāḏihi l-luġa fa-huwa mā ṣuriḥa fīhi
> bi-lafẓa fī bi-ṭṭirād fa-ʾinnahum yuṣarriḥūna wuġūban wa-yuʾaḫḫirūnahā
> ʾay [dā] ʿan aẓ-ẓarf zamānan kāna ʾaw makānan, Qawānīn,* 34,11.)

To start with the locative of time, however, *Qawānīn* does not give any examples with **dā** to illustrate this statement. ʾAbū Ḥayyān's description of the locatives of time too gives the impression that at least some of them are used without **dā**. For example, he gives the following locatives of time: **kunduz** 'during daytime' *(nahāran)*, **tunlā** 'at night' *(laylan)*, **kağā** 'in the evening' *(masāʾan)*, **qušluq** 'in the early morning' *(duḥan)*, **taŋlā** (with variant **taŋdā**) 'in the morning' *(ṣabāḥan)* *(ʾIdrāk,* 135; similar examples in mg 55ᵛrt/bm). The exemplifying sentences he gives are **tunlā ṭurdum** 'I stood at night' and **bukun sangar ṭurmiš daḳul** 'Today Sangar is not standing', in which **bukun** means 'today', and **ṭurdum ʾuŋ-dā** 'I stood up before' *(qumtu qablu)*, **ṣunrā ṭurdum** 'I stood up later' *(qumtu baʿdan [sic])* *(ʾIdrāk,* 136,6ff.).

Although the Turkic adverbs of the type **kunduz** and **kaǧa**, do not contain any suffix that conveys the meaning of *fī*[17], the sources do not hesitate translating them with Arabic equivalents which carry the accusative case.

The Arabic locatives of time, *baʿd* and *qabl* are, like other nouns —but unlike most other locatives of time—annexed to nouns and pronouns, i.e. *ǧiʾtu baʿda-ka* 'I came after you' and *ǧiʾtu qabla-ka* 'I came before you'. In this sense, the locatives *baʿda* and *qabla* are annexed (*muḍāf*) and the pronoun *-ka* 'you' is the noun that is annexed to (*muḍāf ʾilayhi*).

In the Margin Grammar the phrase ʿaṣr-dan burun 'before the evening prayer' in (24) is regarded as a regular annexation of a locative (**burun**) to a noun (ʿaṣr):

24   **ʿaṣr-dan**   **burun**   **kal-di-m**   **saǧā**
       evening prayer-abl   before   come-past-1sg   you/dat
       'I came to you before the evening prayer'
       (*ǧiʾtu ʾilayka min qabli l-ʿaṣri ...*, mg 56ʳtop).

25a   **kal-di-m**   **san-dan**   **ṣuŋ-rā**
       come-past-1sg   you-abl   after-loc ?
       'I came after you' (*ǧiʾtu baʿdaka*).

ʾAbū Ḥayyān must have realised that the construction of (25a) is quite different from the Arabic, and he gives an Arabic word-by-word analysis in which he paraphrases this sentence:

25b   **kal-di-m**   **san-dan**   **ṣuŋ-rā**
       ǧiʾtu   *minka*   *fī baʿdin*
       I came   from you   in later

In (25b) it is shown that **rā** in **ṣuŋrā** is regarded as synonymous with **dā**, which, in its turn translates *fī*, and **dan** as a particle that governs **san**. The same holds for locatives like ʾilkarū 'early' (*qabla*), in which ʾilk means 'first' (*ʾawwal*) and rū conveys the meaning of dā, e.g. **kal-di-m san-dan ilkarū**, meaning 'I came before you' (*ǧiʾtu qablaka*), and also for **lā** (*al-lām*) in **tunlā** 'at night' and **ṭaŋlā** 'in the morning' (cf. *ʾIdrāk* 136,10ff.).

When the words ṣuŋ and ʾuŋ are not annexed to a pronoun, ʾAbū Ḥayyān considers them locatives of time, meaning 'after' and 'before', respectively. The same holds when the particle dā—or its equivalent— is added: **ṭurdum ʾuŋ-dā** *qumtu qablu* 'I stood up early' and **ṣuŋ-rā / ṣuŋ-dā ṭurdum** *qumtu baʿdu* 'I stood up later'. When annexed to a

---

[17] Perhaps - duz could be interpreted as a suffix, although, if so, it is certainly not productive. In EDT kunduz is considered an early compound.

pronoun, they become locatives of place, ṣuŋ-i-n-dā šamālan 'at the left' and 'uŋ-i-n-dā yamīnan 'at the right', respectively (cf. *Idrāk* 136,7; 25 'uŋ *al-yamīn* 'right'; see discussion below).[18]

'Abū Ḥayyān's twofold interpretation of ṣuŋ, depending on whether or not -dā (or -rā) is attached, is only partially reflected by the etymological analyses provided in edt. The meanings edt (832) attributes to "soŋ"—his transcription of ṣuŋ—are: "soŋ originally perhaps physically 'the end' or 'back' of something, but normally used of time, as an Adverb or Postposition 'afterwards, after' (w. *Abl.*)..."

Clauson's remarks with regard to the respective etymologies of "oŋ" and "öŋ", leave little room for 'Abū Ḥayyān's interpretation of 'uŋ. According to Clauson (edt 166) "oŋ" means "left", whereas "öŋ [means] 'the front' of anything." Apparently 'Abū Ḥayyān confused the two words, because of the similar form they have in Arabic script, or, perhaps their pronunciation had—for some reason—merged in the language he described.

2.5.2 Later innovations in the concept

In late Arabic theory the locative of place (*ẓarf muḥtaṣṣ*) comprises all nouns that express a place (*ǧiha*) in themselves, such as *dār* 'house' and *masǧid* 'mosque' (see above Section 1.2). The verb is transitive to this type of locatives by means of the particle *fī*. If another particle is used, such as *'ilā* 'to', the construction is not considered a locative of place. 'Abū Ḥayyān applies the same criterion to Turkic:

> "The locative of place is only used with the particle that is synonymous with *fī*, as long as it is a locative" (*wa-ẓarf al-makān lā yustaʿmal 'illā bi-l-ḥarf al-murādif li-fī mā dāma ẓarfan*, *Idrāk* 136,4).

Let us see now how this principle is applied.

As explained above in 1.2, in *Irtišāf* 'Abū Ḥayyān divides the Arabic locatives of place into two categories. The first category is formed by the 'specified locatives' (*ẓurūf muḥtaṣṣa*). This category of locatives always involves the particle *fī* (or *bi*, in sofar as it is used to denote a locative) through which the verb governs the *ẓarf*. The second category involves the 'unspecified locatives' (*ẓurūf mubhama*). These locatives belong to the so-called six orientations (*al-ǧihāt as-sitt*, *Qawānīn* 35,6;

---

[18] The assumption that all these suffixes (i.e. -rā, -lā and -dā) convey the meaning of -dā, is close to Clauson's opinion (EDT 144 and xl); he regards 'ilkarū as an "abbreviated directive form of ilk" (>*"ilkgerü:"), in which -gerü: is considered a directive suffix, whereas ṣuŋrā can be interpreted as a combination of "soŋ 'end'... 'later'" and the suffix "-raː ... [that] forms Loc[ative] Adv[erbs]" (832 and xl). Similarly, the ending -lā in the words tunlā and ṭaŋlā is considered an adverbial suffix.

mg 55ᵛrt/bm; *Tuḥfa* 63ʳ6; *al-ġāyāt* in *Mufaṣṣal* 67,2ff.) and they typi-
cally occur annexed to a noun or pronoun (*lāzima* ʿalā l-ʾiḍāfa,
Mufaṣṣal 38,12).

The sources, especially mg and *Qawānīn*, list various words that can
be used as locatives of the first category, along with their respective
Turkic translations, i.e. ʾaw 'house' *(al-bayt)*, **bazar** 'market place' *(as-
sūq)* and **kant** 'town' *(al-balad)*. In this way the following locatives are
construed: **bazar-dā** 'on the market place' *(fī s-sūqi)*, **kantā** (for **kant-
tā**) 'in the town' *(fī l-baladi)*; **ʾaw-dā** 'in the house' *(fī l-bayti,
Qawānīn*, 35,1,10; similar examples in mg 55ᵛrt/bm *). The verb gov-
erns the noun, which in Arabic is inferred from the accusative case
ending in *-a*, e.g. *umkut hāḏā l-bayt-a* 'stay [in] this house!', the
Turkic of this sentence is **bu ʾaw-dā ʾaklān**[19] (mg 55ᵛrt/bm), with **dā**
as the equivalent of the implied particle *fī*.

For ʾAbū Ḥayyān—and probably also for Zamaḫšarī—the concept
of *ẓurūf al-makān* further includes all other instances in which the
particle *fī* is used or implied. This also holds for verbs that are typically
used with *fī*, e.g. *daḥaltu l-madīnata* 'I entered the city' for *daḥaltu fī
l-madīnati* 'I entered into the city',[20] even though semantically speak-
ing there is an important difference with other verbs. Verbs such as
daḥala 'he entered' express a movement or a direction, whereas this
notion is absent in the 'regular' locative of place. For Turkic this has
the consequence that with verbs that express a direction the dative case
must be used instead of the locative case.

The sources apply this extended concept of the *ẓarf* to Turkic too.
Therefore -**ġā**/-**kā** is not only considered the equivalent of *fī*, but—
analogous to -**dā**—also a particle which is used for expressing the loca-
tive of place, viz., mg (56ʳtop): *ṭalaʿtu l-qalʿata* [acc] 'I went out [to]
the castle' **qalʿa-ġā šiqtum**, and *daḥaltu l-madīnata* [acc] 'I entered
the city' **šar-ġā**[21] **kirdim, bāġ-qā kir** *udḫul fī l-karmi* 'go into the vin-
yard!' (mg 40ʳlt). The fact that these phrases are still regarded as *ẓurūf* is
inferred from the following statements in the Margin Grammar, viz.,

> "The locative of place is mostly only used with the particle that is synony-
> mous with *fī* as long as it is a *ẓarf...*" (*ẓarf al-makān fī l-ġālib lā yustaʿmal
> ʾillā bi-l-ḥarf al-murādif li-fī mā dāma ẓarfan*, MG 56ʳtop)

---

[19] Cf. TS (III 1399): "eğlenmek... 'beklemek, kalmak'", 'to wait'.

[20] ʾAbū Ḥayyān mentions a discussion among grammarians who ascribe to Sībawayh
the opinion that, e.g. *šām* 'Syria' in *ḏahabtu š-šāma* 'I went Syria' is a *ẓarf* because it gets
the accusative after the deletion of *ʾilā*, *ḏahabtu ʾilā š-Šāmi* 'I went to Syria', and other
similar instances.

[21] Short for *šahr* 'city' (Persian). The vocalisation of the Turkic phrases in this pas-
sage is by interpretation.

and

> "**qā** is a particle of the locative with the meaning of *fī*" (*qā ḥarf ẓarf bi-maʿnā fī*, mg 40ʾlt).

In the preceding sections (2.3), and in Chapter Four it is shown that according to Western grammar *ġā/qā* is, of course, a dative rather than a locative case. In the same sections we have also pointed out that the Arabic grammarians were aware of the meaning of -*ġā*, inasmuch as they equated it with the particle *ʾilā* 'to(wards)'. Nevertheless, they equate *fī* with this suffix as well. We shall discuss this in greater detail below.

The category of the 'unspecified' locatives of place includes, e.g., *ʾust-un-dā* i.e. 'above' (*fawqa*); *ʾalt-in-dā* i.e. 'underneath' (*taḥta*); *ʾun-in-dā* i.e. 'at the right' (*yamīnan*) and *ṣun-in-dā* i.e. 'at the left' (*šamālan*); *ʾaln-in-dā* i.e. 'in front of' (*ʾamāman*); *ʾard-i-n-dā*, i.e. 'behind' (*ḥalfan*) (*ʾIdrāk*, 135). The annexation is shown in *Qawānīn* (36,6) for **qāṭinda** 'with you' or 'next to you' (*ʿindaka*) and similar forms, e.g. **qāṭinizdā** 'with you [pl].' (*ʿindakum*), **qāṭlārindā** 'with them' (*ʿindahum*):[22]

> "There is only one word for 'with', i.e. **qāṭ**, there is no other, and **dā**—or **ḏā**—is the particle of the genitive; what stands between them are the pronouns of the referents." (*luġa ʿinda muǧarrada hiya qāṭ lā ġayra wa-dā ʾaw ḏā hiya ḥarf al-ǧarr wa-mā baynahumā ḍamāʾir yuʾtā bihā li-man hiya lahu, Qawānīn* 36,6; a similar statement on mg 55ʾrt/bm.)

In *ʾIdrāk*, *ʾAbū Ḥayyān* gives a similar analysis for *ʾard-i-n-dā* 'behind' (also partially quoted above):

> "The n indicates the second person if it is nasal [i.e. ŋ], and the annexation if it is pure [i.e. n]. These two can never be combined; they never say *ʾard-in-iŋ-dā with one of the two *nūns* being a nasalised *nūn* and the other a pure *nūn*. The pure *nūn* expresses the [annexation to] the third person and the nasalised *nūn* the [annexation to] the second person." (*wa-hāḏihi n-nūn... ʾin kānat ḥayšūmiyya fa-hiya li-l-ḥiṭāb wa-ʾin kānat ḥāliṣa fa-hiya li-l-ʾiḍāfa, lā yuǧmaʿ baynahumā, fa-lā yuqāl *ʾardiniŋdā wa-takūnu ʾaḥaduhumā ḥayšūmiyya wa-l-ʾuḥrā ḥāliṣa fa-n-nūn al-ḥāliṣa tušʿir bi-l-ʾiḍāfa li-l-ǧāʾib wa-l-ḥayšūmiyya li-l-muḥāṭab, ʾIdrāk* 136,17ff.)

In western grammar, forms like **qāṭlārindā** and, e.g. *ʾardimizdā*, and all other locatives of this type are analysed in a way that is very similar to the one prevalent in the sources:

---

[22] Cf. Lewis 1984 [1967] for a survey of similar adverbs (193-205) and postpositions (85-95) in Turkish.

26   **qāṭ-lāri-n-dā**
        with-poss/3pl-n-loc
        'with them'

In this analysis, **qāṭ** is a noun denoting a space next to or near some-
thing else.[23] However, in contrast to the analysis in the sources, the **n**
in **qāṭ-lāri-n-dā** is considered merely an intermediate between the
possessive suffix **-lari** and the locative case **dā**, and does not serve as a
pronoun.

The same holds for the locative **'ustuŋdā** 'above you', whose base
form, 'Abū Ḥayyān says, is **'ustun** (*'Idrāk* 135,19ff.; cf. also 13,6). He
goes to some pains in explaining that, since the form **'ustun-uŋ-dā** is
too heavy because of the two **n**'s, the basic **n** must be elided, thus re-
sulting in **'ustuŋ-dā**. In the Western analysis, however, **'ust** is the
base form, rather than **'ustun**.

27   **'ust-um-dā**
        top-1sg-loc
        'above me' or 'on top of me'

In this analysis "üst [means] upper surface, top" (edt 242). The word
'ustun 'Abū Ḥayyān refers to is interpreted by Clauson as an adverb
[and as an adjective] "connoting motion onto or a situation on
(something)... liable to be confused with oblique cases of üst" (edt
242). This confusion is caused by the insertion of **n** after the possessive
u in dat, acc and abl. Elsewhere (edt 130), Clauson suggests that **üstün**
is a form derived from **üst**. In other words, **üstün** cannot receive cases
or pronominal endings, since it is itself an adverbial form.[24]

In *Ḥilya*, the Turkic locative is paraphrased in a word by word trans-
lation:

> "the meaning of 'at, with' is **qātindā**, you say 'so-and-so has a horse' **tāwuk**
> **qāt-i-n-dā 'aṭ bār** ['So and so has a horse'] its analysis in Arabic is 'so-and-
> so with him in a horse existent'." (*wa-ma'nā 'inda qāt-i-n-dā taqūlu 'inda
> fulānin farasun tāwuk qātindā 'aṭ bār ta'rībuhu fulān 'indahu fī faras
> mawğūd, Ḥilya* 98,3.)

This is schematically represented in (28):

28   **tāwuk**   **qātin**   **dā**   **'aṭ**   **bār**
        fulān   *'indahu*   *fī*   *faras*   *mawğūd*
        so-and-so   with him   in   horse   existent

---

[23] For an etymology of **qaṭ** as a noun cf. EDT 593.
[24] In *'Idrāk* (22) 'Abū Ḥayyān gives a similar analysis of **'alṭindā** 'underneath', taking
'alṭin (*at-taḥt*) as its base form, whereas this is probably **alṭ**.

In Arabic, when a *ẓarf* is preceded by a particle that governs the geni-
tive, especially *min*, the *ẓarf* stops being a locative because in those
cases it acts as an ordinary noun, e.g. *min qabli* 'before'—lit. 'from be-
fore'.[25] The reason for this is the fact that *fī* is not longer implied, for
the new particle *min* takes its place. In terms of governance one could
say that *qabl* cannot simultaneously contain the meaning of one par-
ticle, i.e. *fī*, and be governed by a second one, i.e. *min.*

In *'Idrāk* it is suggested that some locative words may be construed
with a particle other than **dā**, although the resulting construction
"stops being a locative" *(fa-yaḫruǧu ʿan aẓ-ẓarfiyya)*. The example
mentioned is *'adraka min ḫalfika* 'he caught up from behind you [sg.]'
'ard-iŋ dan yat-tī (*'Idrāk* 136,6). In the same way the literal equiva-
lent of *min qablu* in Turkic, **burun-dan** 'later', is dismissed as a loca-
tive of time, because it is construed with the equivalent of *min*, i.e.
dan (MG 56ᵗᵒᵖ; also *Qawānīn* 36,18).

2.5.3 Summary

Summarising the findings with regard to the *ẓurūf,* it is possible to say
that in Arabic grammar the *ẓarf* is primarily a noun conveying a space
or time, which denotes that the action of the verb takes place in it. The
verb—which may be present in the underlying structure only—governs
the locative, which, as a result, is conceived of as an object to the verb
(mafʿūl fīhi). With regard to the syntactic elements they regarded as
locatives in Turkic, the Arabic grammarians made use of the fact that
most of them can easily be identified as compounds of a noun and a
locative marker, **dā**, which they equated with their particle *fī*. In this
sense they apparently accepted the fact that *fī* may have various realisa-
tions, viz., **rā**, **lā** and **rū**, although these were in the first place regarded
as synonymous with the Turkic particle **dā**, which in turn, is equivalent
to the Arabic particle *fī*. In some instances, however, *fī* is translated by
ġā/ḵā, the dative suffix. In other places **ġā/ḵā** is equated with the
particles *'ilā* 'to(wards)' and *li* 'to', but this apparently did not disturb
them.

The various realisations of *fī* also give some clues as to the way
Turkic phrases and sentences were constructed, in other words, the
Turkic sentences and phrases are translations from the Arabic. Arabic
issues were taken as points of departure and their respective transla-
tions into Turkic were analysed. If it had been the other way around,

---

[25] Not all locative nouns have a full declension.

the grammarians would have found that, for example, ġā/k̬ā, has different translations in Arabic, i.e. *ilā* and *fī*.

It seems that the notion of *ẓarf* in Arabic grammar, at least with the later grammarians, became related to the particle *fī* in a very mechanical way, regardless of its semantic interpretation. This was quite unlike the former conception of *ẓarf* in which it was regarded as denoting a more or less fixed place without the particle appearing in the surface structure. Especially interesting is that a semantic feature such as directionality apparently does not play a significant role.

The analytic approach the Arabic grammarians applied to the Turkic locatives is in accordance with their approach to Arabic; they segment every word into morphemes. In some instances their conclusions and analyses agree with western findings, in others the two approaches differ.

3. CONCLUSIONS

The main conclusion with regard to the translation and subsequent analysis of instances in which in Arabic the genitive case occurs is that the sources do not equate the Arabic case endings which are part of the system of *ʾiʿrāb* with Turkic, nor vice versa. Instead the Turkic morphemes that are regarded as case endings in Western analysis are considered particles, analogous to the Arabic particles of the genitive. In one instance, though, the possessive ending in ī /iy/, they seem to come close to a comparison with an Arabic case ending, in the sense that the y is considered a mere lengthening of i and cannot serve as a pronoun. In this way it only serves to indicate that the noun that ends in ī is involved in an annexation construction.

It is possible to draw another conclusion with regard to the relation of the Turkic and Arabic sentences. From the examples and the way they were analysed, it becomes obvious that the grammarians' primary material consisted of their usual Arabic sentences and phrases which they translated, subsequently discussing the translations. Turkic seems to be the original in very few cases. This is evident, for example, in the use of *fī*. In Arabic, *fī* can be applied with verbs that convey a rest, e.g., *qāma fī l-bayt-i* 'he stood up in the house' and verbs that express a movement, e.g., *daḫaltu fī l-madīnat-i* 'I entered into the city'. In Arabic grammar, the semantic difference between the two instances is not noted, and in both cases the construction with *fī* is considered a locative of place. Even when confronted with the different translations of *fī* in Turkic, -**dā** (LOC) -**ġā** (DAT) in this respect, they make

no attempts to generalise the meanings of *fī*. The only case in which they do so is in the translations of *ḥattā* and *ʾilā*, which already share a common meaning in Arabic, i.e. 'until'.

CHAPTER SIX

THE ACCUSATIVE CASE (NAṢB)

INTRODUCTION

The aim of this chapter is to give an insight into the way Arabic gram-
marians applied their concepts of objects and transitivity to Turkic lan-
guages, special attention will be paid to the views of 'Abū Ḥayyān al-
'Andalusī. It further contains a brief analysis of Turkic object con-
structions and transitive and intransitive verbs as they occur in the
sources. In the third section, it shall be seen how the Arabic grammari-
ans applied their theories to Turkic constructions and the role of their
notions of object and transitivity. I refrain from engaging in a detailed
discussion of transitivity in Arabic linguistics (for which cf. Owens
1988 and 1990, Bobzin 1983, Levin 1979 and Taha 1995[1]), but rather
confine the discussion to an introduction.

In this chapter I intend to point out that the Arab grammarians not
only assigned the same lexical meaning of Arabic verbs to their Turkic
equivalents, but also similar governing capacities. Although the objects
the verb governs take certain markers, these do not always have the
form the grammarians expected. This confrontation leads to inter-
esting observations where the Turkic verbs govern through the dative
case, two objects, or the socalled 'optional objects'.

The accusative case is of special importance, because it is the only
syntactic case in Turkic. It shall be seen that the analyses given by the
Arab authors is quite similar to the one given for Arabic.

1. ARAB GRAMMARIANS ON TRANSITIVITY VS. INTRANSITIVITY

The main goal Arab grammarians set themselves was to give an expla-
nation of the cases of declension in the Arabic language. In this study
it is argued that this approach is based on formal - syntactic rather than

---

[1] By the time I got the chance to study a copy of Zeinab Taha's most interesting dis-
sertation on the views of Sībawayh, al-Mubarrad and Ibn as-Sarrāǧ on transitivity, this
book was already in the final stage of preparation. Therefore the results of her research
could not be included in this study.

semantic criteria (cf. Chapter Four 1). The Arabic scholars' views on objects and transitivity originate from their concern to account for the ending *a* in nouns. In Arabic linguistic theory, this ending is typically caused by the governance of verbs, or particles which are said to have a resemblance to verbs and are therefore entitled to govern like verbs.[2]

The verb, thus, governs two cases, nominative and accusative. As a result of this government, the noun on the syntactic position of the nominative *(rafᶜ)* gets *u*, and the noun on the position of the accusative *(naṣb)* gets *a*. The function of agent *(fāᶜil)* position occurs on the syntactic position of the nominative *(rafᶜ)*, and the object *(mafᶜūl)* occurs on that of the accusative *(naṣb)*. In Arabic grammar there are various types of objects, all of which receive the accusative case ending.

One such object is the direct object *(mafᶜūl bihi)*. The direct object can be assigned to a special class of verbs only, the *ʾafᶜāl mutaᶜaddiya*, roughly equivalent to 'transitive verbs'. The term *mutaᶜaddin* is an elliptic expression for *fiᶜlun mutaᶜaddin fāᶜilahu ʾilā mafᶜūlin* 'a verb that passes by its agent to an object'[3] (cf. Bobzin 1983:95). An example of a transitive verb with its *mafᶜūl* is, e.g.,

1   *ḍaraba*   *ᶜamr-un*   *zayd-an*
     hit/PAST/3sg   ᶜAmr-NOM   Zayd-ACC
     "ᶜAmr hit Zayd.'

Not only transitive verbs may govern objects, for some objects may be governed by both transitive and intransitive verbs alike. These are the so-called optional objects, which include the absolute object *(mafᶜūl muṭlaq)*, the object of reason *(mafᶜūl lahu)*, the concomitant object *(mafᶜūl maᶜahu)*, the locative, i.e. object that expresses the time or place in which an action takes place *(mafᶜūl fihi* or *ẓarf)*.

2   *qāma*   *zayd-un*   *qiyām-a-n*   *al-yawm-a*   *ʾamām-a-ka*
     stand up   zayd-NOM   standing-   today-   front-
     /PAST/3sg   ACC-INDEF   ACC/DEF   ACC-you
                             mafᶜūl muṭlaq   *mafᶜūl fihi* *mafᶜūl fihi*

     'Zayd stood up in front of you today'
     (literally: 'stood up a standing').

In (2), both *al-yawma* 'today' and *ʾamāma* 'front' are locatives (a detailed discussion of the locative is found in *Irtišāf* II 225ff), and *qiyām-*

---

[2] The same holds for verb-like particles like *ka-ʾanna* 'as if', e.g. *ka-ʾanna zayd-an ᶜamr-un* 'as if Zayd (ACC) [were] ᶜAmr (NOM)', which will not be discussed here.

[3] For later Arabic grammarians this elliptic expression became the only name for transitive verbs.

an 'standing' is a verbal noun *(maṣdar)*, that occurs as an absolute object.

The object of reason expresses the reason for which an action is carried out, e.g.,

3a   *ǧiʾ-tu-ka*   *ḍarb-a*   *zayd-in*
       come/PAST-1sg-you beat/INF-ACC/DEF zayd-GEN
       'I came to you to beat Zayd' (*Irtišāf* II 221; 223,14).

This is paraphrased with

3b   *[ǧiʾ-tu-ka]*   *li-ḍarb-i*   *zayd-in*
       [I came to you]   for-beat/INF-GEN/DEF zayd-GEN

In the paraphrasis, introduced with *ʾay* 'that is', the implied particle *li* 'for, in order to' is shown in the surface structure.

Other objects are the excepted *(mustaṭnā)*, and the specification *(tamyīz)*, and the circumstantial expression *(ḥāl;* in Owens [1988] translated as 'condition'), which denotes the condition of the agent (or object). The circumstantial expression may have the form of the active participle marked with the accusative case, caused by governance of the verb:

4   *ǧāʾa*   *raǧul-u-n*   *ḍāḥik-a-n*
       come/PAST/3sg   man-NOM-INDEF   laughing-ACC-INDEF
       'A man came laughing' (*ʾIdrāk* 138,17; cf. Owens 1988:85).

1.1 Different types of government and direct objects

Intransitive verbs, such as *ǧāʾa* 'he came' and *qāma* 'he stood up', can-not govern a direct object. However, in Arabic linguistics there are three ways for an otherwise intransitive verb to become fit for govern-ing a direct object, all of which involve a so-called *ḥarf at-taʿdiya.* The first case is the use of the intransitive verb with a particle that governs the genitive case *(ḥarf ǧarr).* The other cases involve a change in the pattern of the verbal stem: the doubling of the middle consonant of the verb and, thirdly, an initial *hamza* added to the stem of the verb *(Mufaṣṣal* 115,8-20; also Owens 1988:175ff). In the first instance *ḥarf* must be translated as 'particle', and in the other two as 'consonant'. The following discussion deals with the way a particle is used to make an intransitive verb transitive.[4] After that follow the remaining two options.

---

[4] According to Ibn Ǧinnī, the particle may be considered a part of *(baʿd min)* the verb, or of the noun. When a part of the noun, the verb becomes transitive; this procedure is similar to the doubling of the middle radical *(takrīr)* or the prefixation of the *hamza.*

With regard to the first option, the *ḥarf ǧarr*, 'Abū Ḥayyān gives the following statement:

> "The intransitive verb may be implicitly connected to an object; it becomes especially transitive by means of a particle... e.g. 'I passed by Zayd' and 'I got angry with ʿAmr', if the verb does not especially require [the particle], e.g. 'I went out to Zayd'... the correct [opinion] is that it is called transitive." (*wa-qad yuʿallaqu l-lāzim bi-mafʿūlin bi-hi maʿnan fa-yuʿaddī bi-ḥarf al-ǧarr mahṣūṣan... nahwa marartu bi-zaydin wa-ǧadibtu ʿalā ʿamrin... fa-ʾin kāna l-fiʿl lā yaqtadīhi bi-ḥuṣūṣihi nahwa haragtu ʾilā zaydin wa-ṣ-ṣahīḥ ʾannahu yusammā mutaʿaddiyan..., Irtišāf* III 50,11f.)

In other words, 'Abū Ḥayyān's concept of transitivity *(taʿdiya)*, does not only include transitive verbs that have a direct object, but it also comprises all transitive verbs that are made transitive by means of a particle, regardless of whether or not the verb is typically used with a particle, such as *marartu bi-zaydin* 'I passed by Zayd', or incidentally, e.g. *haragtu ʾilā zaydin* 'I went out to Zayd'.[5] This opinion with regard to the direct object is also evident in Ibn Sarrāǧ's *ʾUṣūl* (II 65; see Owens 1988: 176 for further discussion).

'Abū Ḥayyān's argument for considering these nouns 'direct objects' is that the particle may be elided exceptionally *(šudūdan)* in poetry, regularly *(iṭṭirādan)*, or because of frequency of use *(katra al-istiʿmāl)*. Elision *(ḥadf)* means that the element disappears from the surface structure, but is still posited at the underlying level *(taqdīr)* (cf. on *ḥadf* Owens 1988:186 and Carter 1991). After elision of the particle, the verb governs the object directly which, as a result, adopts the accusative case. Examples are, e.g., *dahaltu d-dāra* 'I entered the house', vs. *dahaltu fī d-dāri* 'I entered into the house' and *dahabtu š-šāma* instead of *dahabtu ʾilā š-šāmi* 'I went to Damascus'. This process of elision of the particle and the subsequent direct governance of the intransitive verb is called *ittisāʿ*, 'flexibility', and it is comparable with other instances in which an intransitive verb governs a direct object. In his article Versteegh (1990), describes three instances of *ittisāʿ* in Sībawayh's *Kitāb*, two of which concern a change in the governance of

---

When it is considered part of the verb, the combination of particle and noun is in the syntactic position of the accusative *(kāna maʿa mā ǧarrahu fī mawdiʿ an-naṣb)*, which can be illustrated with the accusative of *ʿamr-an* in *marartu bi-zayd-in wa-ʿamr-an* 'I passed Zayd and ʿAmr' in which zayd-GEN is governed by the particle, and ʿAmr-ACC by the verb *(Ḥaṣāʾiṣ* I 341,6ff).

[5] Contrary to Owens (1988:298 n 219), I believe that the Arabic grammarians (including Ibn Ǧinnī *Ḥaṣāʾiṣ* I: 106,10; 342) did not consider *bi* and other particles 'marker[s] of objectivity', because this term is reserved for the accusative case and, by extension, the accusative case ending (see Chapter Four).

a verb, i.e., the use of optional objects such as the locative *(ẓarf)* and the verbal noun *(maṣdar)* as direct objects of a verb.

'Abū Ḥayyān assumes an elision of the particle in the case of the transitive verbs *'aḫaḏa* 'he took' and *ra'ā* 'he saw', which, according to his opinion, may be used either with or without a particle. He bases this analysis on occurrences of these verbs with the particle in many instances in the Qur'ān (e.g. 43): ✦ *'a-lam tara 'ilā lladīna ḫaraǧū min diyārihim* ✦ 'Did you not see the ones who went from their lands', in which *ra'ā* is transitive with the particle *'ilā* 'to'. A second example is (II 150) ✦ *wa-'aḫaḏa bi-ra'si 'aḫīhi* ✦ 'And he took his brother's head', in which *'aḫaḏa* is transitive by means of the *ḥarf bi* 'with'.[6] Although the use of *ra'ā* and *'aḫaḏa* with the particle is, in fact, exceptional, the grammarians may have considered it more economical to posit a particle that may be elided and reappear in certain instances, than one that must be inserted (see 3.2.4 for the consequences of the implied particle in the analyses of Turkic objects).[7]

1.1.1 Two types of objects

'Abū Ḥayyān makes a distinction between the two types of objects mentioned in the preceding section. The object that directly depends on the verb is called a 'proper object' *(mafʿūl ṣarīḥ)* and the object which is governed through a particle is called an 'improper object' *(mafʿūl ġayr ṣarīḥ)*. The object remains 'improper' even after elision of the particle, because the particle is still posited in the underlying structure (*'Idrāk* 121,22). Although the notions of *mafʿūl ṣarīḥ* vs. *mafʿūl ġayr ṣarīḥ* seem to have been developed within the context of Arabic linguistics, the terms do not occur in *Irtišāf* and *Manhaǧ*.

In practice there is no difference in status between the direct object of a directly transitive verb and one that is transitive by means of an (elided) particle. According to Owens (1988: 298 n. 219) the grammarian al-'Astarābāḏī (d. 686/1286), too, accepted objects that were governed by a verb by means of a particle as direct objects. To my knowledge, though, he did not use the terms 'proper object' *(mafʿūl bihi ṣarīḥ)* and 'improper object' *(mafʿū bihi ġayru ṣarīḥ)*.

---

[6] I thank Monique Bernards for her kind help in finding these places in her copy of *al-Baḥr al-Muḥīṭ*.

[7] With the particle, *ra'ā* conveys the meaning of admiration *(taʿaǧǧub)* (*al-Baḥr al-Muḥīṭ* II, 258, cf. *Lisān al-ʿArab* XIV 299^{b}11 and Bayḍāwī *Tafsīr* I, 93). In the case of *'aḫaḏa bi*, however, the particle *bi* does not seem to cause any fundamental difference in meaning.

The concept of *mafʿūl ṣarīḥ* versus *mafʿūl ġayru ṣarīḥ* is linked to yet other characteristics that ʾAbū Ḥayyān attributes to intransitive and directly transitive verbs. In *Manhağ* (126,9-10), he describes the directly transitive verbs *(ʾafʿāl mutaʿaddiya)* as strong *(qawiya)* and the intransitive verbs *(ʾafʿāl lāzima)* as weak *(ḍaʿīfa)*; only strong verbs are capable of governing a direct object, whereas weak verbs are not.[8] It is possible, though, to strengthen *(taqwiya)* a weak verb by means of a particle, as we have seen.

1.1.2 The strengthening particle *li*

Directly transitive verbs, such as *ḍaraba* 'he hit', may also be followed by a particle, i.e. *ḥarf ğarr li* 'to', although in this case it does not imply that they are weak:

5	*ḍarab-tu*	*li*	*zayd-in*
	beat/PAST-1sg	to	zayd-GEN
	'I beat Zayd.'		

Li is here regarded as a strengthening element *(muqwiya)* for the governance *(ʿamal)* of the governor *(ʿāmil)*. Unlike the particle used with weak verbs, in this case it is considered redundant *(ziyāda)*, since the meaning of the sentence is not influenced by it (*Irtišāf* II 435,4; cf. also Carter [on *ziyāda*] 1981:435, and 51; 111). In *Manhağ* (244, 20, 24-5) ʾAbū Ḥayyān mentions that this use of *li* is for expressing transitivity *(li-t-taʿdiya)*. Ibn Sīdah (d. 458/1066) states that a characteristic of a redundant particle is that it is not implied in the underlying structure *(taqdīr)*, because the "meaning does not create a need for it" (*wa-l-maʿnā lā yuḥwiğu ʾilayhi, al-Muḥaṣṣaṣ* 72,24f).

This use of the particle *li* is not comparable with *li* in the *mafʿūl lahu*, 'the object of reason', which can be elided from the surface structure, but remains posited in the underlying structure of the sentence (cf. 3; see discussion below in 3.2.2).

1.2 *Adding transitivity to the verbal stem*

We return now to the two other possibilities for adding transitivity to the verbal stem. In the first case, *taḍʿīf* 'doubling', the middle consonant of the verb is doubled; the verbal pattern *faʿa/i/ula* becomes *faʿʿala*.[9] In the second case the verb gets a supplementary *hamza* and

---

[8] In this particular context the doubling of the middle radical of the verbal stem is not explicitly mentioned.

[9] In *Mufaṣṣal* (115,9) *taṭqīl al-ḥašw*, 'reduplication of the middle radical'.

adopts the pattern *ʾafʿala*. For this option the term *naql* 'transfer' is used, an elliptic expression for *naql al-ḥaraka*, i.e. the vowel shifts from C_1 to the prothetic *hamza* (cf. Owens 1988:183).

After either procedure the new verb is transitive to one or, when applied to a transitive verb, to two or even three objects. Thus, *fariḥtu* 'I rejoiced' and *ḏahabtu* 'I went' become *farraḥtu-hu* 'I made him rejoice', and *ʾaḏhabtu-hu* 'I made him go', respectively. From a transitive verb, e.g. *kafala zayd-un ʿamr-an* 'Zayd supported ʿAmr', a new verb form is derived which is used with two objects, e.g.,

6 *ʾa-kfal-tu*   *zayd-an* *ʿamr-an*

  CAUS-support/PAST-1sg zayd-ACC ʿamr-ACC

  'I made Zayd support ʿAmr' (*Irtišāf* III 53,18).

These newly formed verbs that already were transitive before the process now have two object complements instead of one, both of which are equal in status. In other words, both *zayd-an* and *ʿamr-an* are direct objects to *ʾakfaltu*.

Transitivity is also conveyed with the pattern *istafʿala*, although this pattern is mainly used to express a demanding or asking for, e.g. *ḏarabtu zayd-an* 'I hit zayd'—*istaḏrabtu zayd-an ʿamr-an* 'I wanted Zayd to hit ʿAmr' (*Manhaǧ* 127,26; cf. also Wright 1986 [1896] I:44).

Although all bitransitive verbs, insofar as they have not been formed by a previous operation of transitivisation, may in theory become tritransitive, the procedure is limited to a few semantic entries only. It concerns basically those stems that express a transfer of information or knowledge, *ʾaʿlama* 'he made known', and *ʾaḫbara* 'he informed' and the like. A second constraint is the fact that tritransitivity only occurs with the verbal pattern: *ʾafʿala*:

7 *ʾa-ʿlamtu* *zayd-an* *al-faras-a* *musarraǧ-a-n*

  CAUS-know zayd-ACC ART-horse- saddled-ACC-INDEF

  /PAST/1sg   ACC/DEF

  'I informed Zayd [that] the horse is saddled'

It will be seen below (cf. 3.3.5) that the Turkic translation of this sentence contains a part that is direct speech. With regard to the Arabic sentence (7), therefore, it is important to note that the second part, i.e. *al-farasa musarraǧan*, is dependent on the verb *ʾaʿlamtu*, for which both elements are marked with the accusative case. This phrase cannot serve as direct speech, for in that case both elements should have been marked with the nominative case.

In this discussion we shall concentrate on verbs that are directly transitive to two objects and, further, discuss the exact status of the objects concerned in greater detail.

1.3 Three types of bitransitive verbs

ʾAbū Ḥayyān (*Irtišāf* III 55) distinguishes three categories of this type of bitransitive verbs. The first category he mentions concerns verbs that are transitive to two objects: the verb is directly transitive to one of them, whereas it is transitive to the other one by means of a *ḥarf ǧarr*. For example, *sammaytu zayd-an ʿamr-an* or *sammaytu zayd-an bi-ʿamr-in* 'I called Zayd ʿAmr'. The bitransitive verbs of the two other categories are directly transitive to both objects, the difference between them being relationship between the two objects. In one class of verbs the two objects do not refer to the same noun. This is the case with verbs of the type of *ʾa ʿṭā* 'he gave', and *kasā* 'he clothed', e.g.,

8	*ʾa ʿṭay-tu*	*zayd-an*	*dirham-a-n*
	give/PAST-1sg	zayd-ACC	dirham-ACC-INDEF
	'I gave Zayd a *dirham*.'		

9	*kasaw-tu*	*ʿamr-an*	*ǧubbat-a-n*
	dress/PAST-1sg	ʿamr-ACC	kaftan-ACC-INDEF
	'I dressed ʿAmr with a kaftan'.		

Zayd and *ʿAmr*, obviously, are not identical with *dirham* and *ǧubba*, respectively. Zamaḥšarī (*Mufaṣṣal* 116,21) calls the objects of such verbs *mafʿūlāni mutaǧāyirāni*, i.e. 'two differing objects'. In ʾAbū Ḥayyānʿs the verb is directly transitive *(bi-nafsihi)* to both of them, i.e. without mediation of a particle: "[the verb] is sometimes transitive by itself to two [objects]..." (*wa-[yata ʿaddā l-fi ʿlu] tāratan ʾilā iṯnayni... bi-nafsihi...*, *Irtišāf* III 55,15.).[10]
In *ʾIdrāk*: ʾAbū Ḥayyān gives a paraphrase of (8)

> "'I gave Sanǧar the *dirham*' i.e. 'to Sanǧar'. The giving actually occurs to the *dirham* and Sanǧar is the one to whom the *dirham* is handed. The equivalent of this in Arabic is 'I handed the *dirham* to Zayd'. It is not permitted [to

---

[10] There are two alternative ways to look upon the objects *dirham* and *ǧubba*. The first is considering them both instrumentals, much like in Russian *ya odaril ego rubl ēm* 'I bestowed him with a rubel" or *ya nakryl ego ǧ ubb oy* 'I covered him with a *ǧubba*', in which these objects take the instrumental case ending. From this perspective, the second accusative is here used semantically. The second is regarding *dirham* a direct object, and *zaydan* an indirect object (which is not possible for [9]), i.e. *zayd* is the semantic case. Based on these scarce seems that considering zayd an an indirect object is essentialle inspired by Western Grammar, not an 'absolute analysis'.

sangara[11] *fa-'inna l-'itā' 'innamā waqa'a ḥaqīqatan bi-d-dirham wa-sanğar huwa l-madfū' 'ilayhi ad-dirham wa-naẓīruhu fī l-lisān al-'arabī dafa'tu d-dirhama 'ilā zaydin wa-lā yağūzu *dafa'tu d-dirhama zaydan, 'Idrāk* 142,7f).[12]

A semantic paraphrase is permitted with an alternative verb that conveys the same lexical meaning as *'a'ṭaytu*, i.e. *dafa'tu* 'I handed'. This verb, however, is transitive to one of the objects with the particle *'ilā* 'to', viz.,[13]

10a	*dafaʿ-tu*	*d-dirham-a*	*'ilā*	*zayd-in*
	hand/PAST-1sg	ART-*dirham*-ACC/DEF	to	zayd-GEN
	'I handed the *dirham* to Zayd.'			

The role of the two objects, both marked with the accusative in Arabic, is explained with the aid of a paraphrase; the governance of one of the objects *(sangara)* by the verb is expressed by means of *li (li-sangara)*. We have seen that in Arabic linguistic theory the *ḥarf li* may be used as an extra element to strengthen the government of the transitive verb.

In the case of *dafa'tu* the particle cannot be deleted from the surface structure, i.e. the verb cannot directly govern both objects.

10b	**dafa'tu*	*d-dirham-a*	*zayd-an*
	hand/PAST-1sg	ART-*dirham*-ACC/DEF	zayd-ACC
	*'I handed the *dirham* Zayd.'		

Another possible reason for giving both sentences is that on a semantic level *zayd-an* in (8) has the same function as *li-zayd-in*, in which *li* expresses the meaning 'to', equivalent to *'ilā*, rather than serving as a strengthening element. However, the governance of the verb *'a'ṭā* 'he gave' is unlikely to be paraphrased with a particle, because the verb *'a'ṭā* only allows direct objects as complements.

In the second class of bitransitive verbs, the two objects refer to the same noun, e.g.,

11	*ẓanan-tu*	*zayd-an*	*ḥāriğ-a-n*
	think/PAST-1sg	zayd-ACC	leaving-ACC-INDEF
	'I thought Zayd [to be] leaving'.		

---

[11] The noun *sangar* probably belongs to the class of diptotic nouns that does not take the ending *i* in the genitive case, but *a* instead. Nouns of this class do not have final *n* either: *ḍarabtu sangara* 'I beat Sangar', instead of *ḍarabtu *sangar-an*.

[12] This statement is given in the context of the Turkic sentence *'aqğa-nī bir-du-m sangar-ğā*. It is possible that *li-sangar-a* is a mere translation of the Turkic **sangar-ğā** 'to Sangar' and hence not a paraphrase of the Arabic (see discussion below in Section 3.2).

[13] The term *naẓīr* in this context signifies that (8) has the same lexical meaning as (10a).

Verbs like *ẓanantu, ḥasibtu* 'I reckoned', *ʿalimtu* 'I knew' belong to the category of verbs that are *dāḥila ʿalā l-mubtadaʾ wa-l-ḥabar*, i.e. they are used with a nominal sentence consisting of a topic and a predicate, and govern both of them (see Owens 1988: 223ff. for a discussion on *daḥala* as a technical term; cf. also Saad 1975). According to Arabic theory, (11) is derived from a nominal sentence (12) in which one of the objects is the topic and the other a predicate:

12   *zayd-un*   *ḥāriǧ-u-n*
        zayd-NOM   leaving-NOM-INDEF
        'Zayd is leaving'.

In *Irtišāf* ʾAbū Ḥayyān states, "[the verb] may also [be transitive] to two [objects]; their origin is the topic and the predicate." (*wa-tāratan [yataʿaddā l-fiʿl] ʾilā iṯnayni wa-ʾaṣluhumā mubtadaʾ wa-ḥabar, Irtišāf* III 55,15 and 56ff; cf. *Manhaǧ* 90,3.)[14] The verb governs the connection *(nisba)* between the topic and the predicate, not the nouns themselves,

> "...the act originates from you and what is dependent on it is the relation between the two nouns, not the nouns [themselves]." (... *ʾinnamā l-fiʿl waqaʿa minka wa-mutaʿalliquhu an-nisba llatī bayn al-ismayni lā l-ismāni*, ʾAbū Ḥayyān in *Manhaǧ* 92,6).

According to ʾAbū Ḥayyān the semantic connection (cf. Owens 1988:304; Carter 1981:135) between the two objects, the former topic and predicate, constitutes the link between the verb and its objects, rather than a sort of direct government of each individual object by the one verb.

One of the objects to *ẓanna, ḥāriǧ-un*, can be replaced with a verb, e.g.,

13   *ẓanan-tu*   *zayd-an*   *ya-ḥruǧ-u*
        think/PAST-1sg   zayd-ACC   3sg-leave/PRES-IND
        'I thought Zayd [to be] leaving.'

The verbal form *yaḥruǧu* stands on the syntactic position of the accusative because it is governed by the verb, but it does not take an accusative ending.

This link between *zayd* and its predicate *ḥāriǧ / yaḥruǧu* is related to another difference between the classes of bitransitive verbs. In nominal sentences, e.g., *zayd-un ḥāriǧ-un* 'Zayd is leaving', and *ʾinna*

---

[14] This category of verbs is also called *ʾafʿāl al-qulūb* 'verbs of the heart', for most of these verbs "signify an act that takes place in the mind" (Wright 1986 [1898] II:48); they may also express doubt, an opinion or an act of learning. In many grammatical works the *ʾafʿāl al-qulūb* are treated in a separate chapter (*Irtišāf, ʾIdrāk*).

zayd-an ḫāriǧ-un '[topicalisation] Zayd is leaving', it is impossible to delete either the topic or the predicate. The same holds when both elements are governed by a verb, such as *ẓanna*:

14a   *ẓanantu   zaydan
      *'I thought Zayd.'

14b   *ẓanantu   ḫāriǧan/yaḫruǧu
      *'I thought leaving/leaves.'

For *ʾaʿṭā* 'he gave' and other verbs of the same class, on the other hand, it is possible to elide either one of the objects

15a   *ʾaʿṭaytu   zaydan
      'I gave [to] Zayd.'

15b   *ʾaʿṭaytu   dirhaman
      'I gave a *dirham*.'

(*Mufaṣṣal* 118,1-7; also *Irtišāf* III 56-7; *ʾUṣūl* I 78ff; Owens 1988:174 and 240).

This indeed shows a basic difference between the two classes of bitransitive verbs. Arabic grammarians further developed a test to determine whether a given object is a direct object *(mafʿūl bihi)* or an optional one *(ḥāl, ẓarf)*. The test consists of passivising the verb and seeing which one of the former objects takes the nominative. The noun that can take the nominative is a direct object. The passive voice of the verb is called 'the verb whose agent is not mentioned' *(al-fiʿl allaḏī lam yusamma fāʿiluhu)*. In the view of the Arabic grammarians 'the object substitues for the agent' *(al-mafʿūl bihi yanūbu ʿan al-fāʿil*, Ibn Mālik apud *Manhaǧ* 111,4; see also Owens 1988: 180-5).

16   *ḍaraba   zayd-un   ʿamr-an   al-yawm-a*
     beat/PAST/3sg Zayd-NOM   ʿAmr-ACC   ART-day-ACC/DEF
     'Zayd beat ʿAmr today.'

When the agent is not mentioned, i.e. in the passive, the direct object takes its place, taking the nominative, while the optional object remains in the accusative:

17   *ḍuriba   ʿamr-un   al-yawm-a*
     beat/PAST/PASS/3sg ʿamr-ACC   ART-day-ACC/DEF
     'ʿAmr was beaten today.'

The fact that the term 'object' *(mafʿūl)* is maintained for the function of *ʿamr* in (17) indicates that the underlying case roles are what determines the construction and, hence, the case ending that is typically assigned to an agent (cf. also Owens 1988:57ff).

Optional objects cannot take the place of the agent of the passivised
verb:

18   *ḍuriba   al-yawm-u
        beat/PAST/PASS/3sg   ART-day-NOM/DEF
        *'Today was beaten.'

Zamaḫšarī (*Mufaṣṣal* 117,1-3) applies this test to sentences with two
objects. Only a direct object may take the nominative. Thus, in the
passive sentence (8) becomes either

19a   ʾuʿṭiya   zayd-un   dirham-a-n
        give/PAST/PASS/3sg   zayd-NOM   dirham-ACC-INDEF
        'Zayd was given a dirham.'

or

19b   ʾuʿṭiya   dirham-u-n   zayd-an
        give/PAST/PASS/3sg   dirham-NOM-INDEF   zayd-ACC
        'A dirham was given [to] Zayd.'

This time, both tests result in acceptable sentences and thus the two
objects are direct objects. In Zamaḫšarī's view—and that of most other
Arabic grammarians—both nouns in the passivised construction retain
their function of object, since the relation with the act expressed by the
verb does not change essentially. In regard to (19a) and (19b)
Zamaḫšarī mentions a preference for (19a):

> "But the construction with the [element that] is the semantic agent is
> preferable, i.e. Zayd, for he is [the] recipient." (*ʾillā ʾanna l-ʾisnāda ʾilā mā
> huwa fī l-maʿnā fāʿil ʾaḥsanu wa-huwa zaydun li-ʾannahu ʿātin, Mufaṣṣal*
> 117,1-3; cf. Owens 1988:181 for a discussion of the term *ʾisnād*.)

Zamaḫšarī has as similar preference for *kusiya ʿamr-un ǧubbat-an*
"ʿAmr was dressed [in] a kaftan', in which *ʿamr* serves as the agent of
the passive verb, rather than *kusiyat ǧubbat-un ʿamr-an* 'a kaftan was
dressed on ʿAmr', because ʿAmr is "the one that is dressed" *(muktasin)*.
The reason for this preference is explained in terms of case roles, i.e.
the nouns Zayd and ʿAmr are the ones that carry out the action of
'receiving' and 'being dressed'—i.e. they are the agents—of the active
verb that expressed the basic notion. Zamaḫšarī uses these terms to ex-
press a preference for a construction, and not to judge its grammatical-
ity.

In this section we have seen that the main point in the analysis of the
Arab grammarians is that they base their analysis of all objects on for-
mal-syntactic principles. In fact, in Arabic grammar all occurrences of

the accusative case are regarded as a result of governance by a verb or verb-like particles. The differences between the objects (direct object, locative, circumstantial expression), whose only common point is an ending in *a* are not explained semantically, but in terms of distribution. Regarding the two objects in (8), *zaydan* and *dirhaman* as equal elements, too, a morpho-syntactic approach is followed, rather than a semantic one, for it is solely based on the fact that both carry the accusative case ending.

In Section 3 below it shall be shown how the Arab grammarians dealt with this and other issues related to the marking of objects in Turkic.

2. OBJECTS AND TRANSITIVITY IN TURKIC

In this second section I give a brief survey of Turkic object constructions. All Turkic languages have a set of endings, traditionally called cases, of which only the nominative case is unmarked. We shall concentrate here on the accusative and the dative cases as far as they occur in object constructions.

2.1 *The accusative case*

The accusative case is in the sources generally marked with **ni** or **nī**:

20 **ʾaġaš-ni**   **ʾal-ḍu-m**
       stick-ACC   take-PAST-1sg
       'I took the stick' (*Qawānīn* 31,6).

The accusative suffix is also put after pronominal possessive suffixes, e.g.,

21 **qul-um-nī**   **ʾur-du-m**
       slave-POSS/1sg-ACC   beat-PAST-1sg
       'I beat my slave' (*ʾIdrāk* 147,14f).

Likewise, **qul-umuz-nī ʾur-du-q** 'we beat our slave', **qul-uŋ-nī ʾur-du-ŋ** 'you [sg.] beat your slave', etc.

The accusative is generally omitted when the object is undetermined, which is expressed by the use of the numeral **bir** 'one', e.g.,

22 **bir**   **ʾaqġā**   **bir-di-m**
       a/one   coin   give-PAST-1sg
       'I gave a coin' (*ʾIdrāk* 147,10).

It may also be omitted when referring to a collective, e.g. ʾalṭun bir- ʿto give gold' (ʾIdrāk 132,11).

In one instance, namely when it occurs after the possessive ending of the third person singular, nī is reduced to n: [15]

23a   ʾaš-i-n   ya-dī
         food-POSS/3sg-ACC   eat-PAST/3sg
         'he ate his food' (MG 60ᵛrt).

But this elision of the i apparently does not always take place, since we also find nī:

23b   ʾaš-i-nī   ya-dī
         food-POSS/3sg-ACC   eat-PAST/3sg
         'he ate his food' (MG 60ᵛrt).

The elision of i seems to be arbitrary, that is to say, it follows no obvious rules. The occurrence of nī after a possessive may be a dialectal feature.

2.2 The dative case

Some verbs require the dative case, which is marked by the suffix GA ([ġa/qa] in 'back' words and [ke/ge], cf. Chapter Three—in 'front' words, depending on the rules of vowel harmony and consonant assimilation. The dative case is typically used with verbs that express a movement or a direction, e.g.,

24   bay-kā   baq-ṭu-m
         chief-DAT   look-PAST-1sg
         'I looked at the chief' (Qawānīn 31,10).

25   ʾav-ġā   čiq-ti-m
         hunt-DAT   go out-PAST-1sg
         'I went out hunting' (ʾIdrāk 145,14).

Turkic verbs may be connected to two noun complements too. In this case the indirect object is always marked with the dative case:

---

[15] This assumption is confirmed by many examples in the Qipčaq text Kitāb fī ʿIlm an-Nuššāb (cf. Öztopçu [1990]), in which both accusatives occur, e.g., bāšin 'his head' (31a), kirišin 'his bow' (32b) vs älini 'his hands' (38a) (transcr. according to Öztopçu). There are also examples of this form of the accusative in Dīwān (366,5) ʾar at-i-n turġur-dī 'the man stopped his horse'.

26   'aqǧa-nī   bir-du-m   sanǧar-ǧā
        coin-ACC   give-PAST-1sg Sanǧar-DAT
        'I gave Sanǧar the coin' ('Idrāk 142,7).

2.3 Verbal forms: Konverbs

One of the complements may be a verbal form:

27   sanǧar-nī   čiq-miš   'uranla-di-m
        Sanǧar-ACC   go out-PAST   think-PAST-1sg
        'I thought Sanǧar had left' ('Idrāk 139,7).

The first part of (27) is a subordinate sentence, for which in English
usually the conjunction 'that' is commonly used, e.g., 'I thought that
Sanǧar had left'.

In (27), **sanǧar-nī**, however, is an object to **'uranladim**, and
čiqmiš is a predicate to **sanǧar**. In *Qawānīn* the accusative ending on
the direct object is omitted in this construction, e.g., **bay min-miš
ṣāġin-di-m** 'I reckoned [that] the chief was riding' (*ḥasibtu l-'amīra
rākiban, Qawānīn* 31,18). Instead of a past tense ending in **miš**, it is
possible to use another tense, e.g., the 'aorist' in (A)r, which in the
sources is used as the present tense,[16] e.g., **bay-nī bin-ar 'uranla-di-m**
'I thought the chief [was/is] riding'.[17]

In the Turkic language(s) described in the sources, the suffix **miš** af-
ter the stem denotes the so-called inferential, which indicates that the
speaker has not witnessed the action expressed by the verb.[18] ('Idrāk
106,21f; MG 45ᵗtop; Dīwān 297,15), viz.,

28   'aybak   ṭur-miš
        'aybak   stand-INFER
        '·'Aybak—apparently—has stood up', 'is apparently standing.'
        ('is surely standing', qā'imun yaqīnan, 'Idrāk 107,1).

---

[16] (A)r stands for ir, ır, ar, er, ur and ür. The vowel is omitted if the stem of the verb
ends in a vowel. Further, (I)p— following—stands for ip, ıp, üp, up, the various forms
of both suffixes may have, according to the vowel harmony. In the sources it is generally
reflected as (i)b, or (u)b, because of the limitations of the Arabic script.

[17] A similar construction is possible in Turkish: "**ben seni öldü biliyordum** 'I was
thinking you were dead' ('I was considering you <as> "he has died"')" (Lewis 1984, 274).
It is not clear whether or not other tenses of the verb may be used in the language(s)
described in the sources.

[18] The feature described here is also typical of modern Oġuz languages such as
Turkish (Lewis 1984: 122) and Turkmen (Hanser 1977: 82; 115). In modern Qipčaq
(and other Turkic) languages this function is usually expressed by means of the verbal
suffix ġan/gen.

By extension, **miš** is also regarded as a general marker of the (active) participle in the past tense (*'Idrāk* 106,12), viz.,

29	**kun**	**duġ-miš**
	sun	be born-PAST

'The sun has risen.'

In 'Abū Ḥayyān's terms (29) means 'the sun was rising in the past' *(aš-šams ṭāliᶜa fīmā maḍā)*. In *Ḥilya* (101,15), **miš** is basically regarded as the 'active participle' *(ism al-fāᶜil)*, with no reference to temporal value (also in *Dīwān* 298,3f).

In the language described in *'Idrāk* and *Tuḥfa* the stem of the verb in (27) may also bear the ending of a so-called 'Konverb' or 'gerund' in (I)p, e.g., **bī-nī ġīq-īb ṣā'n-dī-m** 'I reckoned the chief had gone out' *(Tuḥfa* 77ᵛ6). 'Konverb' and 'gerund' are terms used in Western grammars of Turkic languages for infinite verbal forms whose agent and tense must be deduced from the syntactical or semantic context. The conjugated verb typically follows at the end of the sentence. The converb in (I)p usually signifies that the action of the verb in (I)p pre-cedes the action expressed by the conjugated verb.[19] In the Margin Grammar (53ᵛbm/lt; also 61ᵛbm) this function of (I)p is described in terms of resemblance in function to the connective particle in Arabic *wa-* 'and', e.g.,

30	**kal-ib**	**bar-dī**
	come-KONV	go-PAST/3sg

'He came and went away' *(ġā'a wa-rāḥa, MG 53ᵛbm/lt).*

Alternatively, the verb in (I)p is occasionally interpreted as expressing a simultaneity, and is analysed as the equivalent of an Arabic *ḥāl*, which takes the accusative, viz.,

31	**yikir-ib**	**suwla-dī**
	to be angry-KONV	speak-PAST/3sg

'He spoke angrily' *(taḥaddaṯa ġaḍbāna, Qawānīn* 38,1).

In the language(s) of our sources the gerund in -(I)b is also used as a fi-nite form. With this function, the verb in (I)p can also be used pred-icatively, especially when followed by the corroborative element -DIr, e.g.,

---

[19] In modern Qipčaq languages such as Tatar (Poppe 1963: 76; 102) and Kazakh (SKJa 318) this is still a main function of the suffix.

32   sanğar   ṭur-ub-ṭur

      sanğar/NOM   stand-KONV-CORR

      'Sanğar is standing' (*sanğaru qāʾimun, ʾIdrāk* 122,4).

The predicative use in (I)p(-DIr) however, is perhaps limited to the
third person singular. The fact that the form in (I)p(-DIr) may be used
predicatively is for the Arabic grammarians an important argument for
equating it with the function of the active participle in Arabic: *qāʾim*
'standing' (see discussion below). In this sense, its use is comparable to
that of **miš**, which they also relate to the Arabic active participle.
According to the Margin Grammar -b-tur is used by some speakers to
express the past tense (*ʿalāma al-māḍī*, MG 38ʳlt; 38ʳtop; *Qawānīn*
11,7-8).[20] (See for occurrences of these suffixes in other Turkic lan-
guages Fundamenta I; esp. Doerfer 1959:388 and Pritsak 1959:84 for
Armeno-Qipčaq).

     The verb in (I)p, with an optional suffix -An also serves to denote
what is called in Araabic grammar the 'circumstantial expression' (*ḥāl*,
ʾIdrāk 137,9ff; also MG 54ʳrt/ult; 54ᵛbm/lt; 54ᵛlt/bm; *Qawānīn* 37,
14ff).[21] The circumstantial expression denotes the condition of the
agent (or the object) during the action expressed by the verb, e.g.,

33   'ar   kal-dī   kul-uban[22]

      man/NOM   come-PAST/3sg   laugh-KONV

      'The man came laughing'.

2.4 Denominal and causative verbs

After this summary of object constructions and dative cases, we shall
now give an account of the ways to derive verbs and construct
causatives in Turkic.

     In Turkic verbs can be derived from some nouns by adding the suf-
fix **lā**:

kīt	'lock'	**kīt-lā**	'to lock'
ṭāš	'stone'	**ṭāš-lā**	'to stone'
baš	'head'	**baš-lā**	'to start'

---

[20] The numerous instances in *ʾIdrāk* in which Turkic verbs in **miš** (e.g., 112,13,
117,19, 126,5ff; 131,18, 132,5, 137,20) or (I)b occur are translated into Arabic with an
active participle (or vice versa), in some cases with indication of the past tense (cf.
122,4f; 129,5; 137,1ff).

[21] "The *nūn* expresses the intensification of the circumstantial expression, as if you
repeat it." (*ʾIdrāk* 137,20).

[22] The 'Konverb' in (I)p-An is found in Crimean Tatar (Doerfer 1959: 386).

(cf. MG 38ᵛtop; also *Qawānīn* 38,3ff; *'Idrāk* 121,10; *Ḥilya* 125-6; *Dīwān* 15,16ff.).

In Turkic the notion of causative is conveyed by adding a suffix to the verbal stem. The form of this suffix depends on whether the stem ends in a vowel or a consonant, and vowel harmony and consonant assimi-lation (although these are not always expressed in our sources). Causativity is expressed by means of four types of suffixes, i.e. -DIr, -Ir/Ar, -t, and -GIz/ zIr, of which DIr is probably the most frequent and productive, and therefore usually regarded as the basic form, viz.,

'ul- 'to die'—'ul-dur- 'to kill'

sawin- 'to rejoice'—sawin-dur- 'to make happy'.[23]

'išit- 'to hear'—'išit-tur- 'to make listen' (*'Idrāk* 14; 116,19)

In some verbs causativity is expressed by means of the suffix -Ir, viz.,

čiq- 'to leave'—čiq-ar-dī 'to make leave' (*'Idrāk* 44)

'iğ- 'to drink'—'iğ-ur-dī 'to make drink' (*'Idrāk* 110,3).

The suffix may have various other forms, of which GIz and zIr occur in our sources (for these and other causative suffixes in Turkic languages see further *Fundamenta* I):

ṭur- 'to stand up'—ṭur-ġuz- 'to raise'

kur- 'to see'—kur-kuz- 'to show'

'am- 'suckle'—'am-zur- 'to make suckle'.

When the stem ends in a vowel or **r**, -t must be used, e.g.,

yuru- 'to walk'—yuri-t- 'to make walk'[24]

2.5 *Causative verbs in syntax*

After a transitive verb has been equipped with a causative particle, it takes two objects, one of which is an indirect object. The direct object is marked with the accusative, the indirect object takes the dative case:

34	'ul	'anī	bak-kā	qirġa-t-tī
	he	he-ACC	chief-DAT	be angry-CAUS-PAST/3sg

'He₁ incited the chief to be angry with him₂,' (*Dīwān* 433,7).

35	'ul	maŋā[25]	sūz	kizla-t-tī
	he	I/DAT	word	hide-CAUS-PAST/3sg

'He urged me to keep the words secret' (*Dīwān* 437,12).

---

[23] Vowel harmony, e.g., (öldür), (sewindir) is not always, and mostly only partially reflected in Arabic script (cf. Chapter Three).

[24] For a comparative inventory of causative suffixes in *Dīwān*, *Ḥilya*, *'Idrāk*, *Tuḥfa*, *Tarğumān*, *Bulġa*, and *Qawānīn*, see Čaykovskaya (1981:52ff).

[25] maŋā is spelled /manka"/ in *Dīwān*.

When the verbal stem **bil-** 'to know' is equipped with the suffix **dir**, it means 'to make know' or 'to inform'. This verb, too, takes two objects, one of which takes the dative case, e.g.,

36   ʾul   maŋā   ʾīš   **bil-tur-dī**

        he   I/DAT   matter   know-CAUS-PAST/3sg

        'He informed me of the matter' (*Dīwān* 354,17).

The second object may be replaced with a verbal from, e.g.,

37   bī-ḵā   bil-dir-du-m   ʾāṭ   ʾayarla-n-ub-tur

        chief-DAT   know-CAUS-   horse   saddle-PASS-

                PAST-1sg   KONV-CORR

        'I informed the chief the horse has been saddled' (*ʾIdrāk* 129,5).

The word order of (37) is rather puzzling, for in Turkic languages the verb is usually found at the end of the sentence, such as in (36). The sentence in (37) very much resembles the word order of the Arabic equivalent (see discussion below) and it cannot be ruled out that the wording of this sentence has undergone an influence of Arabic. This is discussed in greater detail below in Section 3.3.3.

Another reason for the inverse word order may be the fact that the second part is direct speech.[26] In modern Turkic languages some form of de- 'to say' is usually placed between the direct speech part and the conjugated verb. This form of de- is di-y-e (say-KONV) in, for example, Turkish (cf. Lewis 1984: 175), and de-p (say-KONV) in Qazaq, a Qipčaq language, e.g., Bastyq-qa at jertte-l-gen de-p ait-ty-m 'I told the chief [DAT] "the horse [NOM] has been saddled" [saddle- PASS-PAST]'. Another way to separate the conjugated verb, **bildirdum**, from the direct speech part is to change the word order. This inverse word order is also evident in yet another example in which direct speech is used without a separating element in a derivative form of **de-**, e.g., **sanǧar ʾayit-tī sunqur ṭur-miš** 'Sanǧar said Sunqur is standing' (*qāla sanǧaru sunquru qāʾimun*, *ʾIdrāk* 154,14f). The second part of the Arabic equivalent, i.e. *sunquru qāʾimun* 'Sunqur is standing' is also direct speech.

After this exposé of constructions in Turkic languages we shall see in the following section how they were interpreted by the Arabic grammarians.

---

[26] It is not clear to what extent sentences with indirect speech were possible in fourteenth-century Qipčaq. No example is found in our sources. It is possible to have constructions with two object complements: the indirect object marked with a dative case and the direct object generally consisting of a nominalised verbal form (or a subject participle) marked with an accusative case ending. For Turkish it would result in a sentence like *At-ın eyerle-n-diği-ni bey-e söyledim 'I told the chief that the horse has been saddled' (cf. similar forms in Old Anatolian; Guzjev 1990:124).

3. ARAB GRAMMARIANS ON TURKIC OBJECT CONSTRUCTIONS

The first part of this section deals with the analyses of Turkic sentences whose Arabic translations consist of a verb and an object, either with or without a *ḥarf ǧarr*. After that follows a discussion of constructions in which two or more objects are involved.

3.1 *The status of the Turkic marker of the accusative*

The five main sources (*Ḥilya, Tuḥfa, Qawānīn*, MG and *ʾIdrāk*) agree that the marker of the direct object is **nī** placed after the object:

38   **bitik-ni**   **yaz-ḏu-m**

        book-ACC   write-PAST-1sg

        'I wrote the letter' (*katabtu l-kitāba, Qawānīn* 31,7).

"The marker [of the direct object] is an **n** vocalised with an **i** [following] immediately after the direct object' (*ʿalāmatuhu nūn maksūra ʿaqīb al-mafʿūl bihi, Qawānīn* 31,1). In *Ḥilya* **ni** is explicitly followed by **y**, resulting in /niy/: "You add to its [sc. the object] basic consonants an **n** vocalised with an **i**, and [you add] a **y**" (*tuzīd ʿalā ḥurūfihi l-ʾaṣlīya nūnan maksūratan wa-yāʾan, Ḥilya* 88,12-6), e.g.,

39   **zayd-nī**   **qīlīğ**   **bīlā**   **ʾūr-du-m**

        zayd-ACC   sword   with   beat-PAST-1sg

        'I beat Zayd with the sword' (*Ḥilya* 88,14).

The status of the ending **nī** (or a part of it, as we shall see below in 3.1.1) is described in various terms:

1. *ʿalāma al-mafʿūliyya* 'marker of the objectivity' (*ʾIdrāk* 139,2).
2. *ʿalāma al-mafʿūl bihi* 'marker of the direct object' (*Ḥilya* 88,16;[27] MG 55ʳlt/md).
3. *ʿalāma an-naṣb fī l-mafʿūl bihi* "marker of the accusative as far as the direct object is concerned" (MG 55ʳtop/rt).
4. *ʿalāma an-naṣb hāḏā fī l-mafʿūl bihi aṣ-ṣarīḥ* "marker of the accusative as far as the proper object is concerned" (*ʾIdrāk* 139,3).[28]
5. *ʿalāma al-mafʿūliyya wa-ʿalāma an-naṣb ʾayḍan* 'both the marker of objectivity and the marker of the accusative' (*Qawānīn* 31,2); "this [holds] if [the verb] is transitive by itself" (*hāḏā ʾiḏā kāna l-fiʿl*[29] *yataʿaddā bi-nafsihi, Qawānīn* 31,8).

---

[27] In *Ḥilya* (88,6) the term 'marker of agency' (*ʿalāma al-fāʿiliyya*) occurs too.
[28] In *Tuḥfa* (77ʳ10) the term *mafʿūl bihi aṣ-ṣarīḥ* 'proper object' is used.
[29] Em. for *al-mafʿūl*

Before going into further detail in regard to these definitions and de-scriptions, it is useful to recall the Arabic theory concerning the end-ings of declension *('i'rāb)*. In Chapter Four we saw that in Arabic grammar the ending *a* is basically regarded as a marker for the ac-cusative *(naṣb)*, which, in its turn, is a marker for a function, i.e. the object *(mafʿūl)*. The term *naṣb* is used for the ending *a* in nouns and verbs; it is the formal general term to describe all instances in which the ending *a* occurs, i.e. governance by verbs and particles.

The term *mafʿūl*, on the other hand, refers to the function of the noun in the sentence, i.e. the object of a verb, including the optional objects. The term *al-mafʿūl bihi* is reserved for those instances in which the verb directly governs the object, e.g., *ḍarabtu zayd-an* 'I beat Zayd'. It is also used in instances of flexibility of the language *(ittisāʿ)* for verbs that govern a direct object in some instances but are normally used with a particle, e.g., *naẓartu zayd-an* 'I looked [at] Zayd' instead of *naẓartu ʾilā zayd-in* 'I looked at Zayd'. The expression *al-mafʿūl bihi aṣ-ṣarīḥ* 'the proper object' excludes all instances of governance of ob-jects where a particle is involved. The functions (agent, object and ele-ment to which is annexed) have no visible marker in the surface struc-ture; their respective markers are the abstract notions of the syntactic position. In other words, the ending *a* itself is not the marker of objec-tivity *(mafʿūliyya)*. We have seen in Chapter Four that in Arabic grammar the ending *a* was (at least by Ibn as-Sarrāǧ) identified with the syntactic function of the noun it occurs on, i.e. 'accusative' *(naṣb)*. Only in this sense is it possible to understand *a* as a 'marker of objec-tivity' *(ʿalāma al-mafʿūliyya)*.

In regard to the terms used for the Turkic suffix nī, 'marker of ob-jectivity' *(ʿalāma al-mafʿūliyya)* in descriptions (1) and (4), or the like, is a general reference to the function of the noun in the sentence, the object in its most general sense. In most descriptions (2, 3, 4) it is fur-ther specified that nī is used for 'the direct object' *(al-mafʿūl bihi)* or 'the proper object' *(al-mafʿūl bihi aṣ-ṣarīḥ, ʾIdrāk)*, thus excluding other types of objects. *Ḥilya* and MG apply the term 'marker of the di-rect object' *(ʿalāma al-mafʿūl bihi)*.

In Arabic grammar the term *ʿalāma al-mafʿūl bihi* does not exist, simply because the direct object has no exclusive marker. In this sense, it can be considered an innovation, albeit based on a development in Arabic grammar itself. The specification with regard to the direct object as opposed to another type of object is necessary since, as we shall see, in Turkic the marker of the object may vary according to its function.

The terms ʿalāma an-naṣb and ʿalāma al-mafʿūliyya (1, 3, 4, 5) are far less innovative; they combine the regular descriptions of accusative (naṣb) and direct object (mafʿūl bihi) in the same way as they are found in works on Arabic grammar. In these sources the specification for the direct object, in the other sources expressed with the term mafʿūl bihi which is lacking in Qawānīn's first definition (5), is found in the additional statement. The use of the term 'accusative' (naṣb) for Turkic further refers to governance relations, i.e. governance of a verb on its object.

The use of the term mafʿūl can be interpreted as a functional reference. The use of naṣb, however, is of far greater significance: it is a formal term and therefore a clear indiation that nī is here perceived as a the marker of governance.

A term which is also used in this context without being a technical term is dalīl al-mafʿūliyya, 'indicator of objectivity', (ʾIdrāk 147,6;11; 142,6 'indicates' [tadullu]). Here is no reference to the Arabic declensional system. (See 3.2.3 for an interpretation of the dative as marker of the object.)

3.1.1 Morphological analyses of nī

Before proceeding with the discussion of objects that are governed by a verb with implication of a particle, it may be interesting to pay attention to the various morphological analyses of nī in the Margin Grammar, Ḥilya and ʾIdrāk.

3.1.1.1 The Margin Grammar

According to the analysis in the Margin Grammar, for instance, the object marker is the n alone, whereas the y (i.e. i) is added to express an additional meaning. The evidence lies in object markers that occur after a possessive suffix of the third person singular. We have seen that in those cases the final vowel can be deleted, e.g.

40/23a   ʾaš-i-n   ya-dī
         food-3sg-ACC   eat-PAST/3sg
         'he ate his food.'

In the Margin Grammar the Arabic equivalent of this sentence is

23a"   ʾakala   ṭaʿām-a-hu
        eat/PAST/3sg   food-ACC/DEF-POSS/3sg
        'he ate his food' (MG 55ʳbm).

But it is possible to have both the possessive suffix and nī:

41/23b   'aš-i-nī   ya-dī
         food-3sg-ACC   eat-PAST/3sg
         'he ate his food.'

In the Margin Grammar this difference is compared to the insertion of
the *ḥarf li* with transitive verbs in Arabic. Thus the Arabic equivalent of
(41) is:

23b"   *'akala*   *li*   *ṭaʿām-i-hi*
         eat/PAST/3sg   to   food-GEN/DEF-POSS/3sg
         'he ate his food.'

In Arabic grammar, as is pointed out above in 1.1.2, a transitive verb
may be followed by an additional particle *li (lām zāʾida)*, which serves
as a strengthening element. This particle is not posited in the underly-
ing structure. In the Margin Grammar, the function of **y** (**ī**) in the ac-
cusative ending **nī** is compared to that of *li* in Arabic:

> "You add a **y** at the end of the word which takes the place of *li* of transitivity
> in Arabic". (*zidta yāʾan fī ʾāḥir al-kalima taqūmu maqām lām at-taʿdiya
> fī l-ʿarabīya*, MG 60ʳrt; also 55ʳbm.)

The fact that the **y** may be omitted in this one occasion is taken as evi-
dence for its being a mere additional element *(ḥarf zāʾid)*, whose func-
tion is to strengthen the government of the verb. This point of view is
also applied in other instances in which the suffix **nī** occurs (although
in these cases **ī** /y/ cannot be omitted):

> "If [**miz**] occurs on the syntactic position of an object, you ad **nī** to it... and
> it only occurs with *li* of transitivity, you say **'aṭ-imiz-nī min-dī** which
> means 'he rode our horse'." (*fa-ʾin waqaʿat mawqiʿa mafʿūlin zidta fī
> ʾāḥirihā lafẓa nī... wa-lā yakūnu ʾillā fīmā lām at-taʿdiya, fa-taqūlu
> ʾāṭimiznī mindī ʾay rakiba li-farasinā*, MG 61ʳtop; also 60ʳrt.)

We summarise this statement as follows. The Turkic sentence in (42)

42   'āṭ-imiz-nī   min-dī
         horse-POSS/1pl-ACC   ride-PAST/3sg
         'He rode our horse,'

has the following equivalent in Arabic:

43   *rakiba*   *li*   *faras-i-nā*
         ride/PAST/3sg   to   horse-GEN/DEF-POSS/1pl.

This implies that in the Margin Grammar all instances in which **nī** oc-
curs are analysed as the object marker **n** on the one hand and a *ḥarf
zāʾid*, i.e. /y/, on the other, which serves to strengthen the governance
of the verb.

3.1.1.2 Ḥilya

Another interesting morphological analysis of nī is found in Ḥilya, based on the suffixing of nī after the possessive ending of the first person singular, e.g.,

44a   'āṭ-im-nī   'ar-dur-ġay   man
       horse-POSS/1sg-ACC   be tired-CAUS-FUT   I
       'I will exhaust my horse' (ʾutʿibu farasī, Ḥilya 132,14).

According to Ibn al-Muhannā's analysis, in 'āṭ-im-nī, 'āṭ is the noun used for a 'horse' (ism al-faras), man means 'I' (bi-maʿnā 'anā), and the y is an addition for the object (ziyāda li-l-mafʿūl). This analysis is schematised in (44b):

44b   'āṭ-man-ī
       faras-'anā-ziyāda li-l-mafʿūl
       horse-I-addition for the object

In other words, the combination of the possessive suffix (i)m and the accusative case ending nī, both considered morphemes in Western analysis and in most of our sources, is segmented in the pronoun man and the accusative ending ī /iy/. Moreover, the y /iy/ is regarded as the distinctive marker of the object, instead of nī (although elsewhere [Ḥilya 88,15-6] it is stated that the n and the y together form the marker of the direct object). A possible explanation for the analysis in (44b) may be the fact that Ibn al-Muhannā regards the n and the y as separate morphemes, and in this way accounts for the occurrence of both markers. On the other hand, this is the only instance in which he gives this description of the accusative.

    Another point is that Ibn al-Muhannā's analysis involves considerable shifts and changes of vowels that are quite unlikely to have occurred in real usage. The regular form is (a), whereas (b) is inferred from Ibn al-Muhannā's description: (a) 'atimnī → (b) *'atmanī. His analysis involves the deletion of i before m, and the insertion of a vowel a between m and n. This focus on consonants is typical of Arabic linguistic thinking. Because of the imperfect vocalisation of the text, these differences in vocalisation are not shown in Rifʿat's edition of the text.

3.1.1.3 ʾIdrāk

The respective analyses in the Margin Grammar and Ḥilya are quite different from the one found in ʾIdrāk. 'Abū Ḥayyān regards nī as a derivative form of the regular marker in i, with n as the basic marker of

the object (also *'Idrāk* [147,6]), with a possibility to delete **i** or the **y** (**ī**). In this sense, **i** (or **ī**) is not regarded as a morpheme:

> "I beat [like] Sangar's beating'. **'ur-du-m sangar uruš-i-n**... But as to the ending in a **n**, which is attached here after the **ī** of the annexation, its base form is **'ur-uš-i-nī**. **nī** is the marker of the accusative and **ī** [indicates] the annexation for the pronoun of the third person... Then you elide the *y* and leave the **n** unvocalised. It is permitted to pronounce the base form, so you say **'uruš-i-nī**." (*ḍarabtu ḍarba sangara 'urdum sangar urušin... wa-hāḍā t-tanwīn al-lāḥiqa hunā ba'd kasra al-'iḍāfa 'aṣluhu 'urušinī fa-nī 'alāma an-naṣb wa-l-kasra li-l-'iḍāfa li-ḍamīr al-ġayba ṯumma ḥaḍafta al-yā' wa-'abqayta n-nūn sākina wa-yaġūzu n-nuṭq bi-l-'aṣl fa-taqūlu 'urušinī, 'Idrāk* 135,3-5.)

This coincides with 'Abū Ḥayyān's statement that, as a rule, the weak consonants *('alif, wāw, yā')* arise from the lengthening of the vowels *(nawāši' 'an 'išbā' al-ḥarakāt)* and form no part of the root of the word *('Idrāk* 101,11-3). We have seen in Chapter Five 2.4.3 that this point of view also plays a role in his analysis of **ī**, the possessive ending of the third person.

3.1.2 Attachment of the accusative to the possessive ending

When the possessive ending is attached to the noun, the Turkic accusative suffix is added to the end of the compound, thus referring to the case of the combination, e.g. **'āṭ-imiz-nī** 'our horse (ACC)'. This sequence of the morphemes differs from the one in Arabic, in which the case ending precedes the possessive suffix:

45		**'āt-imiz-nī**	[**min-dī**]
		horse-POSS/1pl-ACC	ride-PAST/3sg
	'He rode our horse.'		

46	[*rakiba*]	*faras-a-nā*	
	ride/PAST/3sg	horse-ACC/DEF-POSS/1pl	
	'He rode our horse.'		

In the Margin Grammar this leads to the following analysis:

> "If [**miz**] occurs in the place of an object, you ad **nī** to it... **'āṭ-imiz-nī min-dī**[30] which means 'he rode our horse'." (*fa-'in waqa'at mawqi' maf'ūlin zidta fī 'āḥirihā lafẓa nī... fa-taqūlu 'āṭimiznī mindī 'ay rakiba li-faras-inā*, MG 61ʳtop.)

According to this statement, **miz**, the possessive suffix of the first person plural, stands on the place of the object *(maf'ūl)*. This analysis is

---

[30] This Turkic fragment in MG is imperfectly vocalised.

obviously based on Arabic grammatical theory in which each mor-
pheme is followed by the appropriate case ending. [31] From this perspec-
tive, *faras* 'horse' is marked for the accusative *(naṣb)* with *a*, while it is
followed by the pronoun *nā* 'our' (which stands in the position of the
genitive *[ǧarr]* because of the annexation). In Turkic the order of the
morphemes is reversed and the marker of the object **nī** is attached to
the pronominal ending **miz**, 'our'. This has lead the author of the
Margin Grammar to the assumption that in Turkic the pronoun **miz**
stands on the place of the object and is accordingly marked, rather
than 'horse'.

3.2 *Transitivity by means of a* ḥarf ǧarr

We have seen in the first section of this chapter that Arabic intransitive
verbs *(ʾafʿāl lāzima)* may become transitive to an object by means of a
particle of the genitive *(ḥarf ǧarr)*. This holds, for example, for the verb
waqafa 'he stood still' which may be used with the particle *li* 'for':[32]

47	*waqaf-tu*		*li*	*s-sulṭān-i*
	stand still/PAST-1sg		for	ART-sultan-GEN/DEF

'I stood still for the sultan' (*Qawānīn* 31,13).

Note that this particle cannot be deleted from the surface structure. If
the verb is transitive with *li*, in Turkic **ġā** is used, viz.,

48	sulṭān-		ġa	ṭur-du-m
	sultan		DAT	stand-PAST-1sg

'I stood still for the sultan' (*Qawānīn* 31,13)

In Arabic some intransitive verbs are typically used with a particle (see
also above 10a). The verb is regarded as transitive by means of this par-
ticle. An example of such a verb is:

49a	*naẓar-tu*		*ʾilā*	*l-ʾamīr-i*
	look/PAST-1sg		to	ART-chief-GEN/DEF

'I looked at the chief.'

In some instances the particle *ʾilā* may be deleted from the surface
structure, after which the object is assigned the accusative case ending,
because the verb governs it directly. This process is called *ittisāʿ* (or

---

[31] In those instances in which the case endings, i.e. governance relations, are not evi-
dent in the surface structure, they are posited in the underlying structure, e.g. the ac-
cusative ending *a* is not visible in *ʾatʿabtu faras-ī* /faras-a-ī/ (horse-ACC-my) 'I ex-
hausted my horse'.

[32] See for a full account of the meanings of *li* Irtišāf II 433-4. In *Qawānīn* the expres-
sion *waqafa li* 'he stood still for' is considered an example of transitivity by means of a
particle. It is discussed under the heading *al-mafʿūl bihi* 'the direct object' (31).

tawassuᶜ), 'flexibility', which is needed here, because an intransitive verb cannot normally govern a direct object, viz.,

49b *naẓar-tu* *l-ʾamīr-a*
     look/PAST-1sg ART-chief-ACC/DEF

In our sources, too, this possibility of governance of a direct object by an intransitive verb is described:

> "If it is [a verb] which is transitive by means of 'to', either overt or deleted, according to linguistic flexibility, then the marker [of the object] is the element *ġā* or *kā*..." (*fa-ʾin kāna mimmā yataᶜaddā bi-ʾilā malfūẓan bihā*[33] *ʾaw maḥḏūfatan ᶜalā t-tawassuᶜ fa-ᶜalamatuhu lafẓa ġā ʾaw kā...*, *Qawānīn* 31,8-10.)

This implies that also in those cases where in Arabic there is no *ḥarf* visible in surface structure, the Turkic uses *ġā/kā*, the equivalent of the particle *ʾilā* 'to', and that there is only one Turkic equivalent for both (49a) and (49b), i.e.

50 **bay** **kā** **baq-ṭu-m**
    chief DAT look-PAST-1sg
    'I looked at the chief.'

3.2.1 The status of *ġā/kā*

From the quotation above it is difficult to draw any conclusions regarding the status of *ġā/kā*; the use of the term *ᶜalāma* points at an interpretation as a marker of the object, albeit not very explicitly.

In *Tuḥfa*, however, *ġā/kā* is regarded unequivocally as a particle, rather than as a marker of the object:

> "If it is [a verb] which is transitive by means of 'to' (*ʾilā*) in the Arabic language, whether it is stated or deleted, according to linguistic flexibility, they use the particle when [*ʾilā*] is stated." (*wa-ʾiḏā kāna fī l-ᶜarabiyya mimmā yataᶜaddā bi-ʾilā manṭūqan bi-hā ʾaw maḥḏūfa ᶜalā t-tawassuᶜ ʾataw fī l-manṭūq bihā bi-ḥarf al-ġarr*, *Tuḥfa* 77ᵛ8f.)

The term 'marker of the improper object' for *ġā* is not used in all instances in which the particle is elided in Arabic. Let us consider, for example, the case of a verb that is typically used with a particle, i.e. *fī* or *ʾilā*, *daḥala fī/ʾilā* 'he entered [into]'. The verbal noun governs the object in the same way, viz.,

---

[33] In *ʾIdrāk* (139,10) *manṭūqan bihā* 'uttered' is used instead of *malfūẓan bihā* 'stated'.

51   ʾāman-tu   li   duḥūl-i-n   ʾilā   l-ǧannat-i
        believe/   for   entering-   to   ART-heaven-
        PAST-1sg   GEN-INDEF   GEN/DEF
        'I believed in order to go to heaven.'

The verbal noun can be placed in an annexation construction with the
object when the original particle is deleted, viz.,

52   ʾāman-tu   li   duḥūl-i   l-ǧannat-i
        believe/   for   entering-   ART-heaven-
        PAST-1sg   GEN/DEF   GEN/DEF
        'I believed in order to go to heaven' (ʾIdrāk 142,6).

Even though the noun al-ǧanna stands in an annexation construction
with the verbal noun duḥūl, al-ǧanna is regarded as the object to duḥūl
(cf. Irtišāf III 174; also Wright 1986 [1898] II 57; 61).

The verbal noun can exert governance on the object when it is
made independent from it. This is possible by means of the article or
the annexation of the verbal noun to the pronoun of the agent, viz.,

53   ʾāman-tu   li   duḥūl-ī   l-ǧannat-a
        believe/PAST-1sg   for   entering-POSS/1sg   ART-heaven-ACC

In (53) the verbal noun duḥūl-ī governs al-ǧanna in the accusative,
whereas in (52) it does not.[34]

In ʾIdrāk (54a) is given as the Turkic equivalent of (52), viz.,

54a   kirtun-du-m   kir-mak-um   ʾuǧun   ʾuǧmaq-qā
        believe-PAST-1sg   enter-INF-POSS/1sg   for   heaven-DAT
        'I believed in order to go to heaven.'

3.2.2 Two analyses of the verb and its object

In terms of Arabic linguistics there are two possible ways to analyse the
relationship between **kirmakum** and **ʾuǧmaqqā**. In the first place it is
possible to consider **ʾuǧmaqqā** the object of **kirmakum**, since it is de-
fined by the annexation to the pronoun of the first person (**m**).
Further, there are other instances where, according to ʾAbū Ḥayyān,
ǧā/qā serves to indicate the object (cf., for example, 50). This first op-
tion is reflected in the word-by-word-analysis of (54a) which ʾAbū
Ḥayyān gives, viz.,

---

[34] The particle can also be reinserted: ʾāmantu li-duḥūlī ʾilā al-ǧannat-i, in which case
the verbal noun governs through the particle.

54b	kirtundum	kirmak-um	ʾuġun	ʾuġmaq-	qā
	ʾāmantu	duḥūl-ī	lām	al-ġanna	tadullu
			al-ʿilla		ʿalā l-mafʿūl
	I believed	entering-1sg	li of	ART-heaven-	indicates
			reason		the object

Here, **qā** is regarded as an element that indicates the object *(tadullu ʿalā l-mafʿūl)*. Note that he does not use the term ʿalāma 'marker'.

The second analysis is to regard **qā** as the Turkic equivalent of the particle ʾilā, which is an association made by ʾAbū Ḥayyān elsewhere (145,15; see also Chapter Four 3.3). In fact, the verb *daḥala*—and therefore also the verbal noun *duḥūl*—is normally transitive by means of the particle ʾilā. The particle is deleted from the surface structure of the Arabic sentence (52), but it remains in the underlying structure. In this sense, the Turkic phrase would reflect the underlying structure of the Arabic sentence.

Both analyses are reflected in ʾAbū Ḥayyān's interpretation of the status of **qā**, viz.,

> "The q that is vocalised with a... indicates the object; qā—or ġā—are used to indicate the meaning of 'to' whose meaning is the goal." *(wa-l-qāf al-maftūḥa... tadullu ʿalā l-mafʿūl wa-qā ʾaw ġā ʾinnamā takūnu fīmā yušʿiru bi-maʿnā ʾilā maʿnāhā al-ġāya, ʾIdrāk 142,6.)*

It seems that there is indeed some confusion as to which status should be assigned to **qā**. ʾAbū Ḥayyān uses rather vague descriptions as 'indicates the object' *(tadullu ʿalā al-mafʿūl)* and 'indicates the meaning of' *(yušʿiru bi-maʿnā,* lit. 'makes feel'), rather than the term 'marker' *(ʿalāma)*. From these rather vague descriptions one could conclude that ʾAbū Ḥayyān comes close to perceiving it as a marker for an improper object (see 3.2.3 for a similar analysis of a dative case as a marker).

A different analysis of GA is given for the object of **ʾuqšadī**. In ʾIdrāk it is translated with the Arabic verb *ʾašbaha* 'he resembled', a verb that governs its object directly in the accusative. In Turkic, though, the object of **ʾuqšadī** is marked with **ġā**, the dative, viz.,[35]

55	sangar	ʾarslān-ġā	ʾuqša-r
	sangar/NOM	lion-DAT	resemble-PRES/3sg

'Sangar resembles a lion' *(ʾIdrāk 128,3)*.[36]

---

[35] In Western terms, ʾuqša- governs its object in the dative case (EDT 97).

[36] In ʾIdrāk (55) is given as the translation of *ka-ʾanna sangara ʾasadun* 'as if Sangar [were] a lion'.

In (55), "ǧā has the meaning of *li* that conveys [the meaning] of objec-
tivity" (*bi-maʿnā l-lām allatī tuʿṭī l-mafʿūliyya, ʾIdrāk* 128,3). Here *li*
must be interpreted as the so-called 'strengthening particle' whose
most important characteristic is that it is redundant; the verb does not
really need it in order to govern (see discussion above in 1.1.2).

3.2.3 The object of reason

In the preceding section we have seen that after deletion of the particle
li when it denotes a cause or a reason, the verb directly governs the ob-
ject, 'the object of reason' *(mafʿūl lahu)* (cf. [3a]). For Turkic the
sources give a similar construction:

56a   kal-du-m   ʾāš   ya-ma-kā[37]
      come-PAST-1sg   food   eat-INF-DAT
      'I came in order to eat the food'
      (*ǧiʾtu li-ʾakl-i ṭ-ṭaʿām-i, Qawānīn* 33,16).

In the analysis which is given of this sentence, **yamak** is the verbal
noun *(maṣdar)*, and "the a in [the verbal noun] is the marker of objec-
tivity" *(al-fatḥa allatī fīhi ʿalāma al-mafʿūliyya)*, viz,

56b   kaldu-m   ʾāš   yamak-ā
      fiʾl māḍī-fāʿil   *ism li-ṭ-ṭaʿām*   *maṣdar-ʿalāma al-mafʿūliyya*
      verb of-agent   noun 'food'   verbal noun-marker of
      past tense   objectivity

Of great significance here is that ā is regarded as a marker, rather than
the equivalent of a particle, which also would have been possible. After
all, in Arabic the meaning of the *mafʿūl lahu* is by definition para-
phrased by means of a particle. Possible reasons for this analysis are in
the first place that weak consonants, such as *ʾalif*, are only reluctantly
interpreted as governing particles. Second, for Arab grammarians the
link between ā and kā/ǧā may not have been obvious (even though af-
ter possessive endings 3sg the dative is reduced to ā too). The same may
have been true for the possibility that **yama** in itself may serve as a ver-
bal noun and as such be governed by particles. This analysis of ā is very
similar to that of the dative in Turkic translations of Arabic sentences
from which the particle is elided. In that instance too, there seems to
be a preference for analysing the Turkic object as being directly gov-
erned rather than conveniently positing the elided particle.

---

[37] In Western analysis this infinitive in mA is a so-called short infinitive, which can
take the necessary case and possessive endings.

Furthermore, note that here not the term 'marker of the accusative' is used, but rather a 'marker of objectivity' (cf. 3.2.1-2; for further discussions of the object of reason with different examples *Idrāk* 141-2; MG 56ʳrt; on the object marker see 3.1).

3.2.4 Structurally deleted particles

The opposite case—i.e. the deletion of a particle from the Arabic sentence, while the particle in question remains absent from the Turkic construction—is possible too. It concerns those Arabic verbs that have a *ḥarf* in the underlying structure that only occasionally appears in the surface structure, e.g., *ʾaḫaḏa* 'he took' (*Tuḥfa* 77ᵛ10), *raʾā* 'he saw' (*Tuḥfa* 77ᵛ9; *Idrāk*, 139,11), and *rakiba* 'he mounted' (*Tuḥfa* 77ᵛ10).[38] We have seen above in 1.1 that the Arabic verb *ʾaḫaḏa* can be used with the *ḥarf bi* 'with' and, albeit with a more specific meaning, *raʾā* with *ʾilā*.[39] In this sense, we can reconstruct (57a) as the basic sentence of (57b):

57a	*raʾaytu*	*ʾilā*	*sanğar-a*
	see/PAST-1sg	to	sanğar-GEN
	'I saw Sanğar (with amazement).'		

After deletion of the particle, the verb directly governs the direct object:

57b	*raʾaytu*	*sanğar-a*
	see/PAST-1sg	sanğar-ACC
	'I saw Sanğar.'	

The Turkic equivalent of both sentences is:

58	**sanğar-nī**	**kur-du-m**
	sanğar-ACC	see-PAST-1sg
	'I saw Sanğar' (*Idrāk* 139,11).	

In both *Idrāk* and *Tuḥfa* (58) is regarded as an example of deletion of the particle, and is therefore considered an exceptional instance of nī occurring in an 'improper object.' In the Turkic sentence, however, no particle is found. It appears that the attribution of Arabic governing capacities is given priority over a reinterpretation of Turkic material.

---

[38] Unfortunately, *Tuḥfa* only gives the Turkic translations of the Arabic verbs meant here. But since these translations are fairly consistent in comparison with those in the other sources, it seems safe to interprete **kūrdūm** as a translation of *raʾaytu* 'I saw', **mīndīm** for *rakibtu* 'I mounted' and **ʾāldīm** for *ʾaḫaḏtu* 'I took'.

[39] The verb *rakiba* 'he rode' may be used with the particle *ʿalā* 'on', although some Arab grammarians consider this as a *ẓarf* rather than a *ḥarf* (*Irtišāf* II 451,10).

3.3 *Transitivity by means of internal change, and bitransitive verbs*

The other way to make an intransitive verb transitive is a morphological change, i.e. the addition of a morpheme to the verbal stem. We
have pointed out above that this applies to both Arabic and Turkic.

Because the morphological process of adding an element to the
stem in order to gain an additional meaning is familiar to the Arabic
grammarians, the sources deal quite briefly with this issue. In the first
place, it is possible to derive a verb from a noun by means of the suffix
lā: kīt 'lock' → kīt-lā 'to lock' (see above 2.4).

3.3.1 Opinions on the denominal suffix lā

In *Ḥilya*, Ibn al-Muhannā explicitly describes the views of the Arabic
grammarians with regard to the derivation of verbs from nouns:

> "Know that, while according to the Kufans the Arabic verbal noun origi
> nates from the verb, and that the verb originates from the verbal noun, ac
> cording to the most famous of the Basran grammarians, in the Turkic lan
> guage some verbal nouns do not originate from the verb but from a noun
> or what resembles it." (*wa- ʿlam ʾannahu lammā kāna l-maṣdar al-ʿarabī
> yanšaʾu ʿan al-fiʿl ʿalā raʾy al-kūfiyyīn wa-yanšaʾu al-fiʿl ʿanhu ʿalā raʾy aš-
> šahīr min nuḥāt al-baṣriyyīn fa-fī l-luġa t-turkiyya yanšaʾu baʿḍ al-
> maṣādir lā ʿan al-fiʿl bal ʿan ism wa-mā kāna sabīlahu, Ḥilya* 125,14-17;
> the same discussion is found in *Irtišāf* II 202,7ff; cf. also Bohas 1982:189ff)

In some other sources, the suffix is said to serve *ʾiʿmāl* or *ʿamal*, i.e.
'governance'. The governance meant here is obviously not syntactic
governance but some other kind of influence:

> "lā reflects governance... the governance in it is that it changes[40] its *ʾalif*
> [i.e. ā] in a *y*... which you vocalise with an i." (*tušʿir bi-l-ʿamal... fa-l-
> ʿamal fīhi ʾan taqliba ʾalifuhu yāʾan... wa-taksiruhā, Qawānīn* 38,3ff;
> similar statement in *ʾIdrāk* 121,10.)

Indeed, unlike other stems ending in a vowel, the stem of denominal
verbs changes in the 'present tense' and the 'converb' in (I)b: suz
'word' suz-lā-y-ur 'he talks' (*ʾIdrāk* 121,11)—suz-la-y-ib 'talking'
(*Qawānīn* 38,5; *ʾIdrāk* 113,12). In *ʾIdrāk*, however, these forms are also
explicitly said to occur without y, viz., suz-lā-r and baš-lā-r.

In the Margin Grammar lā is related to the patterns *tafʿīl* and *ʾifʿāl*
that stand for references to the doubling of the second consonant of
the root, and the addition of the *hamza* before the stem, respectively:

---

[40] Cf. Chapter Three 1.3.2 on *qalb*.

"Know that the l in all [these] positions is an expression for the particle of transitivity from the characteristic of [the pattern] *ʾifʿāl*...." (*wa-ʿlam ʾanna l-lām fī ǧamīʿ al-mawāḍiʿ ʿibāra ʿan ḥarf at-taʿdiya min mīza al-ʾifʿāl...*, MG 38ᵛtop; cf. also *Dīwān* 597,4.)[41]

This supposed relation of **lā** with the Arabic patterns *tafʿīl* and *ʾifʿāl* is interesting. The Margin Grammar gives the following examples: **qūltūq-la-dī** 'he took (something) under his arm' *(ʾabbaṭa)*, **qiliğ ladī** 'he hit with a sword' *(sayyafa)*, **ṭāš-la-dī** 'he stoned' *(ḥaǧǧara)*. In Arabic most roots may serve as a basis for a verb. In this sense the verbs *ʾabbaṭa, sayyafa* and *ḥaǧǧara* can theoretically be derived from the roots /ʾ-b-ṭ/, /s-y-f/ and /ḥ-ǧ-r/, respectively. In practice, however, the meanings associated with these verbs in the Margin Grammar seem at best very rare, and perhaps even non-existent. The verb *ḥaǧǧara*, for instance, means 'he limited' rather than 'he stoned', and the meaning of 'he hit with a sword' is better expressed with *sāfa* (although it must be noted that the verbal form with a doubled middle consonant is also used to express an intensified action).[42] The verb *ʾabbaṭa*, too, does not exist in the given sense, the meaning 'he took under his arm', is expressed with *taʾabbaṭa*. It seems that these and some other verbal forms given in the Margin Grammar are denominative verbs rather than derivations from the root itself (cf. also Bohas 1982:171).

3.3.1.1 Excursus: Ibn al-Muhannā on denominal verbs
In *Ḥilya*, Ibn al-Muhannā continues his discussion of denominal verbs with the description of a suffix to derive verbs from nouns. This suffix, though, is not **lā** but, instead, an unvocalised **l**, viz.,

"The marker of this is a silent l after the basic consonants of the verb and before the marker of the verbal noun." (*fa-ʿalāmatuhu lām sākina baʿd ḥurūf ʾaṣl al-fiʿl wa-qabla ʿalāma al-maṣdar, Ḥilya* 125,16.)

The 'marker of the verbal noun' is, of course, the suffix **maq/mak**, which is attached to the stem.

However, most examples Ibn al-Muhannā gives can be traced back to a compound of a noun and the verb **ʾal-maq** 'to take', e.g., **ʾurīnğ-al-maq**[43] *(ar-rišwa)* 'to take bribes', **tusū-l-maq** 'to benefit', for **tusū**

---

[41] It is also associated with the passive and reflexive forms of the verb: **ṭāš-la-dī** 'he stoned' - **ṭāš-la-n-dī** 'he was stoned' *(taḥaǧǧara)*; **kītlādī** 'he locked' - **kit-la-n-dī** 'it was locked' *(ʾuǧliqa)*.

[42] In *Lisān* and Lane. In *Dīwān* (573,3) a paraphrasis is used instead for **ṭāš-lā-dī** *ḍaraba... bi-l-ḥiǧārati* 'he hit... with the stone' (cf. also *Dīwān* 586ff for many more Turkic verbs in -lā.).

[43] Contrary to EDT 234 "urːnčlamak".

al-maq *(an-naf^c)* lit. 'to take benefit', (cf. EDT:554), ṣāṭ[i]n **al-maq** *(aš-širā', Ḥilya* 126,1ff) 'to purchase', lit. 'to take as a buy'.[44] These are, once more, examples in which Ibn al-Muhannā comes to surprising conclusions due to his interpretations of the vocalisation of the text.

3.3.1.2 Particles of transitivity attached to verbal stems

The respective shapes of causative suffixes of Turkic are outlined in 2.4 and 2.5. The suffixes are divided into four groups, DIr, Ir/Ar, and -t after vowels, and one fourth group of less common suffixes, such as GIz, zIr and others. Most sources discuss the causative sufffixes quite briefly (*'Idrāk* 110,3ff; *Qawānīn* 68,7ff). They are recognised as morphemes that in function resemble certain Arabic particles and therefore are called 'particles of transitivity' (*ḥurūf at-ta^cdiya, 'Idrāk; Qawānīn* 68,6), '*'af^caltu* [1sg.] for transitivity' (*'af^caltu li-t-ta^caddī, Ḥilya* 130,13), 'particles that indicate transitivity' (*ḥurūf tadullu ^calā t-ta^cdiya,* MG 37^1t). The first denomination is exactly the one applied to the Arabic equivalents.

The sources typically select one basic particle. In a number of instances this is -dur, e.g., **min-dur** 'make mount!'. The choice for any of the other suffixes is made on different grounds; -t is usually associated with stems that end in a vowel, but the other forms cause analytical problems.

Most sources state that after vowels the suffix is **t**: "If... the last consonant of the verb is vocalised, then add a silent **t**" (*wa-'in kāna... 'āḫir al-fi^cl mutaḥarrikan fa-zid ^calayhi tā'an sākina, Qawānīn* 69,9; similar statement in *'Idrāk* 113,17), e.g., **yuru-t** 'make walk!'.

The suffix DIr is indeed regarded as the base form. From this it is possible to derive -Ar/-Ir, by means of deletion of -D, viz., **'iğ-** 'to drink'—**'iğ-ur** 'make [him] drink!', **čiq-** 'to leave'—**čiq-ar** 'make leave!',

> "They make verbs transitive with the r alone, and delete the preceding consonant. This is heard in a limited number of words that have to be learned by heart [in their language], and [these words] should not be taken as a basic rule." (*wa-qad ^caddaw 'af^cālan bi-r-rā' faqaṭ wa-ḥadafū l-ḥarf alladī qablahā sumi^ca dālika fī 'alfāẓ ma^cdūda tuḥfaẓ wa-lā yuqās ^calayhā, Qawānīn* 69,1-3.)

In other words, **dur** is regarded as the basic form, from which in these verbs the **d** is deleted. In Arabic grammar, each consonant that is inserted into the basic root is considered a separate morpheme that adds

---

[44] Most of these words are imperfectly vocalised.

a meaning. In this concept it is not possible to have suffixes consisting of more than one consonant (less even to have suffixes of different shapes that convey the same meaning). A good example is the dis-agreement among Arab grammarians as to the status of /w/, /y/ and /"/ in the regular plural and dual forms of the noun, the basis of which is that in Arabic theory each one of these cannot signify gender, number and case at the same time (cf. Chapter Four 1.2.3).

For Turkic, the possibility of deleting -**d** must mean that it is basi-cally superfluous, and that the meaning of the suffix is carried by -**r** alone (but see following discussion), viz.,

> "The **t**, the **ġ** and the **k** [of yet other suffixes] are not what causes the verbs to be transitive, but rather it is the **r** alone that does this." (*fa-t-tā' wa-l-ġayn wa-l-kāf hunna lasna bi-muʿaddiyāt li-l-'afʿāl wa-'innamā l-muʿaddiya ar-rā' faqaṭ, Dīwān* 366,10.)

The fact that the two suffixes mentioned above are considered the usual ways to add 'transitivity' to the verb, as opposed to others is evi-dent in the Margin Grammar (38ʾlt). There is a fourth type of causative suffixes, i.e. GIz/zIr which cannot be that easily derived from the basic suffix. For this it is necessary to distinguish between 'basic rule' *(qiyāsī)* and 'practical usage' *(samāʿī).*

The forms in -**dur** and -**t** are said to be *qiyāsī* 'analogous with the basic rule' whereas other suffixes are said to be *samāʿī* 'depending on practical usage' (on *qiyās* in Arabic grammar see Versteegh 1980). Examples of these verbs are, e.g., **ṭur-ġuz** 'raise!'; **kur-kuz** 'show!' and -**zur**, e.g., **'am-zur** 'suckle!'. The same distinction between basic rules and practical usage is also made in *Qawānīn* and *'Idrāk*, viz.,

> "A few words do not obey [the rules] we mentioned, their [forms] can be determined from what is heard." (*qad ḥaraġa [sic!] ʿammā dakarnā 'alfāẓ qalīla ġiddan fa-sabīluhā s-samāʿ, Qawānīn* 69,12.)

According to 'Abū Ḥayyān, who gives the same verbs,

> "it is allowed to use **dur** with these verbs. The use of **dur** is the rule." (*wa-yaġūzu l-maġī' fī hādihi l-'afʿāl bi-dur wa-huwa l-qiyās, 'Idrāk* 110,10.)

He gives the following alternatives: **ṭur-dur**, **kur-dur**, **ṭam-dur** and **'am-dur**. In this way he escapes from giving rules for either deriving the exceptional forms from the basic ones, or devising certain rules that would explain why certain verbs take an irregular suffix. Although *Qawānīn's* opinion differs, he does not give further explanations either, viz.,

"These forms are not heard otherwise, and other [forms] are not permitted [with these words]." (*wa-hāḏihi l-ʾalfāẓ lam tusmaʿ minhum ʾillā hākaḏā fa-lā yaǧūzu fīhā ġayruhā, Qawānīn* 69,16; cf. also *Ḥilya* 131,11; on causatives with -z cf. *Dīwān* 312.)

We now return to Kāšġarī, who in addition to DIr, discerns an additional form, i.e. GAr (394,2). As we learned from the quotation above, he considers -r sufficient for conveying causativity. Instead of arguing that in certain circumstances D- is deleted, as the other sources do, Kāšġarī says the opposite, namely that -ṛ is basic, and that any other, secondary accompanying consonant is inserted, viz.,

"... the reason for the insertion of the t, ġ or k is for ease of pronunciation." (*...wa-ʿilla duḫūl at-tāʾ wa-l-ġayn wa-l-kāf li-suhūla taqaʿu fī l-lafẓ, Dīwān* 366,8ff.)

A reason Kāšġarī mentions for the insertion of t, ġ or k (before for r) is, for example, an unacceptable sequence of three rs. Such a sequence may result when a transitivised verb is conjugated for the 'future tense' ('aorist'), e.g. *ṣūwra-r-ur* 'he will give water':

"One of these rs would belong to the root, the second would be the r of transitivisation, and the third would be the r of the future tense." (*yakūnu ʾiḥdā ar-rāʾāt ʾaṣlīya wa-ṭ-ṭāniya yakūnu rāʾ at-taʿdiya wa-ṭ-ṭālita takūnu rāʾ al-istiqbāl, Dīwān* 366,15f.)

Therefore, it is necessary to insert an additional consonant to this stem, viz., **sūw-ġar-ur.** [45]

Statements in *Ḥilya* too point to a perception of d and r as distinct morphemes: "The d and the r are two markers of transitivity" (*wa-d-dāl wa-r-rāʾ ʿalāmatā t-taʿdiya, Ḥilya* 132,14) or "two additions for transitivity" (*zāʾidā t-taʿdiya,* 132,11-2), although Ibn al-Muhannā does not elaborate further on this.

3.3.1.3 An unusual analysis of causative sentences in *Dīwān*
In his syntactic analysis of sentences with transitivised verbs, Kāšġarī exposes a view that differs from the usual one found in the works on Arabic grammar. Instead of assuming two objects to the newly formed verb, he posits two agents.

The regular transitive verb expresses an

---

[45] For unclarified reasons the first r disappears. As an alternative one could regard this procedure as an instance of change of r into ġ. Such an argumentation, though, is far from being likely because of the difference in status between the two consonants: r is basic whereas ġ is secondary, as it is associated with the suffix.

"action which originates from one agent to an object" (*fī 'l yanša'u min fā 'il wāḥid fa-yaqa 'u 'alā mafʿūl...*),

viz.,

59	'ar	bitik	bitī-dī
	man/NOM	book	write-PAST/3sg

'The man wrote a book' *(kataba r-raǧulu l-kitāba).*

When a transitivising particle is added, the result is the following,

60	'ul	'aŋar	bitik	biti-t-tī
	he/NOM	he/DAT	book	write-CAUS-PAST/3sg

'he₁ made him₂ write a book' *('aktabahu l-kitāba).*[46]

For these causative sentences Kāšǧarī gives the following analysis:

"...the verb has become transitive to one object from two agents." (*yata 'addā l-fī 'l 'ilā mafʿūl min fā 'ilayni*, Dīwān 353,8; 13)

One of the said agents is the one that orders *('āmir)*, in (60) **'ul** 'he', and the other agent is the one that performs the action *(mubāšir)*, **'aŋar** 'him' (*Dīwān* 366,2; also 15,14 en 445,3ff.). In this view, the only object in the sentence is **bitik** 'book'.[47]

---

[46] *Dīwān* is the only source in which sentences containing a causative verb and two objects are given.

[47] There are a number of differences regarding the terminology between Kāšǧarī's *Dīwān* on the one hand, and the later sources on the other. The terms used in *Dīwān* are quite different from the ones used in regular treatises on Arabic grammar and in the other descriptions of Turkic. Some of these terms clearly stand outside the canonical (Basran) tradition (for a recent discussion of the Basran and the Kufan traditions, see Bernards 1993). The ones used here (sc. *'āmir* and *mubāšir*) are not the only examples. The same holds for the terms used for the vowels, i.e. *rafʿ* (280,2,7) for *u*, *ḥafḍ* (280,2) for *i*, and *naṣb* (280,1) for *a*, which alternate with *ḍamm*, *kasr* and *fatḥ*, respectively, even within a word. In the Basran tradition, starting with Sībawayh, *rafʿ*, *ǧarr* and *naṣb* are used for the declensional endings, whereas *ḍamm*, *kasr* and *fatḥ* are non-declensional; the Kufan grammarian al-Farrā' (d. 207/822) does not make this distinction in the same way (cf. Owens 1990:159; also Versteegh 1993:18). Al-Farrā'"s system is as follows:

	declensional	non-declensional
u	*rafʿ*	*ḍamma/rafʿ*
a	*naṣb*	*fatḥa/naṣb*
i	*ḥafḍ*	*kasra/ḥafḍ*

The use of these terms in *Dīwān* appears to stand very close to the practice in the Kufan tradition.

Further, the unusual term *ǧābir* occurs quite often in *Dīwān* in the sense of 'future tense' (18,14; 280,6; 283,2,4,5; 284,6; 280,10ff) alternating with the regular term *istiqbāl*, or *muḍāriʿ*. In the instances mentioned here (which is far from being a full inventory of all occurrences) it occurs in the following combinations: *'alāma ai-ǧābir* 'marker of the future tense', *al-fī 'l al-ǧābir* 'the verb of the future tense' and *al-'afʿāl al-ǧābira* 'the verbs of the future tense' (In one instance the diacritical dot on the *ǧayn* is not legible because of the *fatḥa*, and one could read *'ābir* 'past tense' [18,14] instead, although this is un -

The addition of -t is referred to as "a proces of transitivisation of intransitive verbs to two agents" (*ta'diya al-'af'āl al-lāzima 'ilā fā'ilayni*, 121,1; cf. also 365,12 and 444,15; also 426,11 and 17). In practice this is correct; after transitivisation instead of one, two agents are involved, whereas the verb may still be intransitive, e.g. bat- 'to sink'—bat-ur- 'to cause to sink'.

In his analysis, Kāšġarī does not refer to the Arabic translation of the sentence, *'aktaba-hu l-kitāb-a*, in which both -hu 'him' and *al-kitāb-a* 'book' are marked as objects, and according to the canonical theory regarded as such. Kāšġarī, instead gives an analysis which from a semantic perspective is possible too, but is not in accordance with the way the elements are marked.

In summary, the grammarians regarded both - **dur** and - **t** after vowels as the basic 'particles of transitivity'. Furthermore, the other 'particles' are either derivatives from **dur**, e.g. **r**, or exceptional forms, e.g. **ġuz** and **ġur**. In this respect their consensus is that **r** alone suffices to express the meaning of transitivity.

In what follows other types of verbs are on issue, i.e. verbs transitive to two and three objects, respectively.

3.3.2 The first category of bitransitive verbs: *'a'tā*

As we have seen above in 1.3, in Arabic grammar there are two categories of verbs with two objects. The first category includes verbs whose objects do not refer to the same noun, e.g., *'a'tā* 'he gave' (cf. [8]), in which both objects take the accusative case. The second category involves verbs whose objects do refer to the same noun, e.g. *ẓanna*. With regard to the first category, according to Western analysis the indirect object in Turkic gets the dative case, and the direct object the accusative (cf. 2.2). As for Turkic sentences of the second category, one of the 'objects' may be a verbal form and as such never get a case ending.

As a starting point in the discussion we take the following remark in *Qawānīn*, viz.,

---

likely in the context.). Biesterfeldt (1990) notes that the term *ġābir* is noted in Ibn Fāriġūn's *Ġawāmi' al-'Ulūm*, together with some other unusual linguistic terms. The term *ġābir* also often occurs in K. *al-'Ayn* (cf. Talmon 1997). It is too early yet to draw any conclusions, but this scanty evidence seems to indicate that *Dīwān* stands in the Kufan tradition, rather than in the Basran.

"Know that they never make the verb transitive by itself to two objects."
(*wa-ʿlam ʾannahum lā yuʿaddūna l-fiʿla bi-nafsihi li-mafʿūlayni ʾabadan,*
Qawānīn 31,15; cf. also MG 55ʿlt/md.)

In other words, according to the author of *Qawānīn*, Turkic verbs
never govern two objects directly. We have seen above that the notions
'directly transitive' *(mutaʿaddin bi-nafsihi)* or 'proper object' *(mafʿūl*
bihi ṣarīḥ) imply for Turkic that the object is marked with **nī**, whereas
in Arabic, there is no exclusive marker for the direct object, since all
objects are indistinctively marked with *a*. The author of *Qawānīn*
probably comes to this statement because a Turkic verb can govern
only one object in the accusative with **nī** at the time.

In *ʾIdrāk* the same phenomenon is recognised, albeit in less explicit
terms:

"[The verbs] that are transitive to two objects in the language of the Arabs
are, in this language, transitive to one of them with **nī** and to the other with
ǧā or with **kā**." (*wa-ʾammā fimā yataʿaddā fī l-lisān al-ʿarabī ʾilā iṯnayni*
fa-ʾinnamā taʿaddā [sic] fī hāḏā l-lisāni bi-ʾaḥadihimā bi-nī wa-li-l-
ʾāḫar bi-ǧā wa-kā..., *ʾIdrāk* 139,3ff).

This is exemplified by

61	ʾaʿṭaytu	sanǧar-a	ṯawb-a-n
	give/PAST-1sg	sanǧar-ACC	cloak-ACC-INDEF
	'I gave Sanǧar a cloak.'		

which is translated into Turkic as

62	**sanǧar-ǧā**	**ṯūn-nī**	**bir-du-m**
	sanǧar-DAT	cloak-ACC	give-PAST-1sg
	'I gave Sanǧar the[48] cloak.'		

The criterion ʾAbū Ḥayyān applies in the assignment of the respective
markers is the sequence of the two objects:

"In [the case of] the first [word] which is the first object in the Arabic lan-
guage, **ǧā** is attached to the base form, and **nī** [is attached] in [the case of]
the second [word], which is a second object in the Arabic language, and the
opposite is not permitted." (*fa-talḥaqu fī l-ʾawwal allaḏī huwa mafʿūl*
ʾawwal fī l-ʿarabī ǧā wa-fī t-ṯānī ʾallaḏī huwa ṯānin fī l-lisān al-ʿarabī nī
ʿalā l-ʾaṣl[49] wa-lā yaǧūzu l-ʿaks, *ʾIdrāk* 139,6-7.)

---

[48] In our discussion of this sentence we disregard the fact that the object is indefinite
in Arabic, whereas it is probably definite in the Turkic phrase.

[49] In this context *ʾaṣl* cannot be understood as reference to the theoretical notion of
'root', which does not exist as such in Turkic, but something more like 'stem', or 'base
form'.

Abū Ḥayyān's description is limited, for it only refers to the syntactic sequence of two objects, rather than their respective functions, and no explicit terminology is applied to describe the function of ġā.

Elsewhere in ʾIdrāk, the status of ġā is more explicitly described in the context of the passivised bitransitive verb ʾaʿṭā 'he gave', viz., ʾuʿṭiya sangar-u dirham-an 'Sangar was given a *dirham*' (19a). Note that in Arabic *sangar* is marked with the nominative case. For (19a) the following translation into Turkic is given:

63	bir-il-dī	sangar-ġā	bir-aqġā
	give-PASS-PAST/3sg	sangar-DAT	one-coin

'A coin was given to Sangar' (*ʾIdrāk* 134,12f).

For (63) the relation of the respective objects and the verb are analysed as follows:

> "[The verb] is transitive to the first object with the particle of annexation, which is ġā, and the verb dominates the second object." (*fa-taʿaddā li-l-ʾawwal bi-ḥarf al-ʾiḍāfa alladī huwa ġā, wa-tasallaṭa l-fiʾl ʿalā l-mafʿūl aṭ-ṭānī, ʾIdrāk* 134,13f.)

It follows that here ġā is explicitly assigned the status of a 'particle of annexation' *(ḥarf al-ʾiḍāfa)*. In view of the fact that ʾaʿṭā 'he gave' is directly transitive to both objects, the term *ḥarf al-ʾiḍāfa* must be interpreted as a reference to the strengthening particle *li* which can only be posited with transitive verbs. Any other interpretation, such as positing it as a translation of the particle ʾilā 'to', would violate the Arabic principle of direct governance of the verb ʾaʿṭā.

This assumption is confirmed in *Qawānīn*, where ġā is described in a very similar way:

> "If the second [object] of the two is not identical with the first one, you attach to the first the marker of transitivity by means of *li*, and [you attach] to the second one the marker of objectivity." (*wa-ʾin kāna t-tānī minhumā ġayra l-ʾawwal ʾalḥaqta bi-l-ʾawwal ʿalāma t-taʿdiya bi-l-lām wa-[bi]-t-tānī ʿalāma al-mafʿūliyya bihi, Qawānīn* 31,19; also very similar in MG 55ʾlt/md.)[50]

From this quotation it follows that, according to *Qawānīn*, too, ġā indicates that the object is governed by means of the strengthening particle *li* that is posited in the underlying structure of all transitive verbs.[51]

---

[50] Although in MG (55ʾtop/rt) both ġā/kā and nī are called ʿalāma an-naṣb fī l-mafʿūl *bihi* in a general sense, this term is not applied in a specific context.

[51] It may be of importance to note that this statement follows shortly after a brief discussion of sentences of the type *katabtu li-sunqura* sunqur-ġā yaz-ḍu-m 'I wrote to Sunqur', in which the particle is used in the surface structure.

In practice, the two terms, *ʿalāma al-mafʿūliyya bihi* and *ʿalāma at-taʿdiya* in *Qawānīn* and the Margin Grammar, respectively, amount to the same. A verb that is transitive with the particle of transitivity is by definition transitive by itself, i.e., transitive to a direct object. The difference between the terms is that 'marker of transitivity' (*ʿalāma at-taʿdiya*) refers to governance of the verb, rather than to the type of object, and indicates the way the verb is transitive to the object.

3.3.3 The second category of bitransitive verbs, *ẓanna* e.a.

We have now come to the second category of bitransitive verbs, i.e. the verb whose both objects refer to the same noun. We have already seen above in 1.3 that this category mainly contains verbs that belong to the subclass of the *ʾafʿāl al-qulūb*, i.e. *ẓanna* 'he thought', *ḥasiba* 'he reckoned' and *ʿalima* 'he knew'. ʾAbū Ḥayyān dedicates a separate chapter in his *ʾIdrāk* to the Turkic equivalents of these verbs, i.e. **ʾuranla-dī**, **ṣaġin-dī** and **bil-dī**, respectively (*ʾIdrāk* 128-9), which he begins as follows:

> "As for ʾuranladī, its meaning is 'he thought' and in this language it is se-
> mantically connected to two objects as in Arabic, even though in reality the
> connected element is the relation [between the two objects]." (*ʾamma*
> *ʾuranladī fa-maʿnāhu ẓanna wa-tataʿallaqu bi-mafʿūlayni fī hāḏā l-lisān*
> *ka-l-lisān al-ʿarabī wa-ʾin kāna fī l-ḥaqīqa ʾinnamā mutaʿalliquhā [sic]*
> *an-nisba, ʾIdrāk 128,12f.)*

With the last part of this statement ʾAbū Ḥayyān means that the connection *(taʿalluq)* between the two objects and the verb is the fact that the two nouns together form a nominal sentence consisting of a topic and a predicate.[52] In Arabic both objects are governed by the verb, which is made explicit with the accusative case, viz.,

64	ḥasib-tu	sanġar-a	ḥāriǧ-a-n
	reckon/PAST-1sg	sanġar-ACC	leaving-ACC-INDEF
	'I reckoned Sanġar leaving'.		

In Turkic the objects to these verbs do not get the same marker:

> "You suffix nī to the first object and you leave the second [object] without
> [any marker], suffixing neither nī nor ġā to it." (*fa-tudḥilu nī ʿalā l-mafʿūl*

---

al-ʾawwal wa-tuhmilu t̲-t̲ānī fa-lā tudḫilu ʿalayhi nī wa-lā ǧā, ʾIdrāk
139,7.)[53]

This can be exemplified with (65a) and (65b):

65a **sanǧar-nī**   **čiq-miš**   **ṣaǧan-du-m**
     sanǧar-ACC   leave-PAST   reckon-PAST-1sg
     'I reckoned Sanǧar leaving' (*ʾIdrāk* 139,7).

65b **bī-nī**   **ǧiq-īb**   **ṣāʾan-dī-m**
     chief-ACC   leave-KONV   reckon-PAST-1sg
     'I reckoned the chief leaving' (*ḥasibtu l-ʾamīra ṭāliʿan, Tuḥfa* 77ʳ5).

Qawānīn has a similar analysis:

> "If the verb is transitive to two objects in Arabic, and the second one of the
> two is identical with the first one, it is permitted, not obligatory, to attach
> the marker of the direct object to the first, without attaching any marker to
> the second [object]." (*fa-ʾid̲ā kāna l-fiʿl yataʿaddā li-mafʿūlayni fī l-*
> *ʿarabiyya wa-t̲-t̲ānī minhumā huwa l-ʾawwalu fa-ʾinnaka tulḥiqu ʿalāma*
> *l-mafʿūl bihi li-l-ʾawwal minhumā ǧawāzan lā wuǧūban wa-tuhmilu t̲-t̲ānī*
> *min al-ʿalāma, Qawānīn* 31,15-8.)

The point both sources make is that although the Turkic forms ending
in **miš** and **ib** both serve as the equivalent of the Arabic participle (see
discussion on suffixes in mIš and (I)p above), they do not get the
marker of the direct object. Nevertheless, they are considered objects.

3.3.4 Objects without apparent marker

The apparent lack of any marker of declension in **ib** and **miš** (in
65a/b) is accounted for in various ways. One is to assign to them the
status of marker of the object, or a similar status. This is possible be-
cause of its association with the *ḥāl*, the circumstantial expression. The
ḥāl is one of the optional objects to the verb and, hence, receives the
accusative case, viz.,

> "Know that the **b** in **min-ib-an** ['riding'] is the marker of the
> circumstantial expression..., that is, the marker of the object." (*fa-ʿlam*
> *ʾanna l-bāʾ fī miniban ʿalāma al-ḥāl... ʾay ʿalāma al-mafʿūl,* MG 54ʳlt/ult;
> similar association in *ʾIdrāk* 129,21 and 137,1ff.)

In the early stages of the Arabic linguistic tradition, the Kufan gram-
marian al-Farrāʾ (d. 207/822) interpreted the accusative case of the se-
cond object as a circumstantial expression *(ḥāl)* (cf. *Irtišāf* 56,4f; cf.
Owens 1990 for Farrāʾ's theory on *ḥāl*). Although this does not reflect

---

[53] From this it is possible to conclude that both nī and ǧā are markers of the object;
see discussion in the previous section.

the opinion of most later grammarians, it is a theory which is possible within the framework of the Arabic linguistic tradition. Moreover, it is not in contradiction with their interpretation of the Turkic data.

The second object of the sentence can also have the form of the 'present tense',

> "[nī] is attached to the first object and you use the second [object] in the form of the imperfect tense." (*tulḥiqu fī l-mafʿūl al-ʾawwal wa-taʾtī bi-t̠-t̠ānī bi-ṣīġa al-muḍāriʿ, ʾIdrāk* 128,13.)

An example of this is, e.g.,

66   **bay-nī**   **bin-ar**[54]   **ʾuranla-du-m**
      chief-ACC   ride-PRES   think-PAST-1sg
      'I thought [that] the chief [was] riding.'

67a   *ẓanan-tu*   *l-ʾamīr-a*   *rākib-a-n*
      think/PAST-1sg   ART-chief-ACC/DEF   riding-ACC-INDEF
      'I thought the chief [was] riding.'

With the first object (*al-mafʿūl al-ʾawwal*) ʾAbū Ḥayyān obviously refers to **bay-nī** 'chief' and with the second (*at̠-t̠ānī*) he means the verbal form, in this case **bin-ar** 'he rides'. The fact that the second object has the form of the *muḍāriʿ* 'imperfect tense' is no problem, since in Arabic, too, *rākiban* 'riding', the second object, can be replaced with the *muḍāriʿ* form of the same verb, without consequences for the meaning of the sentence, e.g.,

67b   *ẓanan-tu*   *l-ʾamīr-a*   *ya-rkab-u*
      think/PAST-1sg   ART-chief-ACC/DEF   3sg-ride/IMPF-IND
      'I thought the chief [was] riding.'

The fact that the second object does not bear the accusative marker is of minor importance: it takes the form of the imperfect tense[55] which discharges it from taking case endings. On the underlying level, however, the verbal form **bin-ar** takes the syntactic place of the object, just like *yarkabu* in (67b).

It is also interesting that both sources describe the two objects in term of 'first' and 'second'. Indeed, in Arabic theory the only way to distinguish syntactically between the two objects, sc. *zaydan* and *ḥārigan*, is their sequence. The reason for this is that, syntactically speaking, the objects are equal in status, both being governed by the same verb. The remarks ʾAbū Ḥayyān makes with regard to the link

---

[54] In some Turkic languages verbs with **m-** have **b-** instead.
[55] Both *ʾIdrāk* and *Qawānīn* give more examples of this class of verbs but with different verbal forms, e.g. **min-miš** 'had ridden', **min-ib** 'rode'.

(muta ʿalliq) between them refer to the connection between the for-
mer topic and predicate and not to the way the verb governs the ob-
jects.

3.3.5 Tritransitive verbs

We have seen in 1.2 that, according to the Arabic grammarians, some
verbs can take three direct objects. In Arabic, all objects are marked
with the accusative case (cf. above [7]), e.g.,

68	ʾa-ʿlam-tu	al-ʾamīr-a	al-faras-a	mulǧam-a-n
	CAUS-know/	ART-chief-	ART-horse-	bridled-
	PAST-1sg	ACC/DEF	ACC/DEF	ACC-INDEF

'I informed the chief [that] the horse [has been] bridled'
(*ʾIdrāk* 129,21).

The grammarians assumed a similar relation between the verb and its
three complements in the Turkic equivalents of the Arabic originals.
Their discussions concern Turkic sentences of the type already men-
tioned above in (37), which is repeated below for the sake of conve-
nience:

37	**bī-ḵā**	**bil-dir-du-m**	**ʾāṭ**	**ʾayarla-n-ubtur**
	chief-DAT	know-CAUS-PAST-1sg	horse	saddle-PASS-PAST

'I informed the chief "the horse has been saddled"' (*ʾIdrāk* 129,21)

In Turkic, however, only **bī** 'chief' gets a case ending: in our terms a da-
tive case. Although they are not marked as objects, ʾAbū Ḥayyān con-
siders all Turkic equivalents of the Arabic objects as objects:

> "To the first object, [sc. **bī**] which was the agent before transitivity [i.e. be-
> fore the verb was made transitive by addition of the suffix], is added the
> marker for the improper object... As for the second object [sc. **ʾāṭ**], it re-
> mains without a marker of either the proper or the improper object. For the
> third object [sc. **ʾayarlanubtur**] you use the form of the circumstantial ex-
> pression..." (*tudḥilu ʿalā l-mafʿūl al-ʾawwal alladī kāna fāʿilan qabla n-
> naql ʿalāma al-mafʿūli ġayri ṣ-ṣarīḥ... wa-tuhmilu ṯ-ṯānī min ʿalāma al-
> mafʿūl aṣ-ṣarīḥ wa-min ʿalāma ġayr aṣ-ṣarīḥ wa-taʾtī bi-ṯ-ṯāliṯ*[56] *bi-ṣīġa al-
> ḥāl...*, *ʾIdrāk* 129,5-8).

In short, according to ʾAbū Ḥayyān, the sentence has three objects, **bī**
'chief', **ʾāṭ** 'horse' and **ʾayarlanubtur** 'saddled'. The only object to get
a marker is **bī**, which takes **ḵā**, which, in its turn, is defined as the
marker of the improper object. The ending in (I)btur is regarded as ex-
pressing the *ḥāl*, the circumstantial expression, in Arabic. The *ḥāl* is

---

[56] Em. for *aṭ-ṭānī* 'second'.

governed by the verb too and is therefore assigned the accusative case, albeit not as a direct object complement but as an optional object (see discussion above in Section 1). The term 'form of the circumstantial expression' *(ṣīġa al-ḥāl)* probably suffices to associate it with an accusative case.

In western terms, however, the second part of the Turkic sentence does not show a dependency relation; it is direct speech. More precisely, **'ayarlanubtur** stands in a predicative relation to **'aṭ**, and therefore neither of them is marked with the accusative case. In Arabic too, *mulġaman* is a predicate to *al-farasa*, but both of them are objects to the verbal form *'aʿlamtu*.

In *Tuḥfa* another version of this same sentence (cf. [37]) is given, e.g.,

69	**bī-nī**	**bīl-dīr-dī-m**	**'āṭ-nī**	**'iyārlā-n-ībtīr**
	chief-ACC	know-CAUS	horse-ACC	saddle-PASS-PAST
		PAST-1sg		

'I informed the chief the horse has been saddled' (*Tuḥfa* 70ʳ10).

In this case, both **bī** and **'aṭ** are assigned the accusative case, which in Turkic is typically used for direct objects. In our terms this can be analysed in terms of the respective functions. In sentences of this type one expects one object to be marked with the dative case for the function of indirect object (cf. [37]). In (69), however, these functions are not distinguished. For example, **bī**'s function of indirect object is not evident here, and it is difficult to picture how it should be understood as a direct object of the verb **bīldīr-**. The other accusative case, **'āṭ-nī**, is suspect, too. In this part of the sentence that is direct speech there is only one instance in which the noun **'āṭ** could get an accusative case, namely when it would depend on an active verb, such as **'iyārlā-** 'to saddle', which is not the case here.

In brief, it appears that the surface structure of the Arabic sentence (68) *'aʿlam-tu al-'amīr-a al-faras-a mulġam-an*, with three objects marked with an accusative case, shines through the Turkic sentence in (69). The differences between the examples given in 'Abū Ḥayyān's *'Idrāk* and the one in *Tuḥfa* can be shown in a more schematic way. Since the Arabic verb *'aʿlamtu*, 'I informed', may govern three direct objects in Arabic, these same capacities are assigned to its Turkic equivalent, **bildirdum**, e.g.,

'aʿlamtu	*al-'amīr-a*	*al-faras-a*	*musarraġ-an*	surface structure
verb	obj. 1 *naṣb*	obj. 2 *naṣb*	obj. 3 *naṣb*	underlying structure
verb	obj. 1 -** kā**	obj. 2 ø	obj. 3 *ḥāl*	Turkic sentence *'Idrāk*
verb	obj. 1 -**nī**	obj. 2 -**nī**	obj. 3 (?)	Turkic sentence *Tuḥfa*

ʾAbū Ḥayyān applies the underlying structure of Arabic to the Turkic
sentence, without trying to impose markers from the Arabic surface
structure on it: he accepts that on the surface level Turkic has a differ-
ent realisation than on the underlying level. *Tuḥfa*'s author, on the
other hand, sticks to the surface structure of the Arabic sentence, thus
creating a Turkic translation that is probably ungrammatical.

In *Dīwān* the only instances in which the verb **biltur-** occurs with
two objects, one of them is marked with the dative case too, e.g., ʾ**ul
maŋā** (DAT) ʾ**īš bilturdī** 'He informed me- DAT of the matter' (ʿ*arrafa-
nī al-ʾamr-a, Dīwān*, 354,17; also 368,5: **bilduz-**).[57]

4. CONCLUSIONS

In this concluding section I present some general observations on the
analysis of the Arab grammarians on governance in terms of semantic
and syntactic case. Furthermore, I give a summary of the discussion on
some terminological innovations made in order to describe the unfa-
miliar Turkic constructions.

4.1 *Analyses of different types of objects*

Theoretically speaking, the grammarians could have limited the as-
signment of the function of object to those nouns that actually bear
the same object marker, as in Arabic. Instead, to some extent they
seem to accept different types of object markers for Turkic.

A possible interpretation may be that they were used to assigning
the function of object *(mafʿūl)* to nouns, even if they do not show the
marker of the accusative *(naṣb)*. In Chapter Four it is shown that this
occurs in Arabic, for instance, when the word ends in a glide, or when
the case endings are omitted in pausal forms. In those cases the gover-
nance relations between the verb and its object(s) can be made explicit
in the underlying structure of the phrase. The underlying structure of
the Arabic sentence thus could serve for both Turkic and Arabic.

A similar analysis would, indeed, do for instances where there is no
marker at all (such as in ʾ**aṭ** in 37), but it does not cover the manifold
realisations of the object marker. The fact that the terminology used to
describe **nī** to a large extent resembles the terms applied to the Arabic

---

[57] According to EDT (334) the verb ʿbiltür- is typically used with two objects: "'to
make (something *Acc.*) known (to someone *Dat.*)'." In EDT the sentence in *Tuḥfa* (69) is
emended without further reference as "biye [sic] **bildirdim atnı eyerleniptir**".

marker of the object is, according to our hypothesis, no coincidence, because it is, in fact a syntactic case (cf. 3.1).

In Chapter Four it was shown that the Arabic genitive case does not have a corresponding Turkic case that behaves in a similar way; in terms of our hypothesis the explanation is that the Turkic oblique cases cannot be explained as syntactic case. As we have seen in the present chapter, things are different for the accusative. In some instances, no-tably the endings nī, and those of the ḥāl (here -ib), these are particular markers, whose use, although unusual, is restricted to one type of object complement. In this way, they can nonetheless still be related to governance of a verb.

The main problem constitutes ġā, whose function as a governing element (as demonstrated in Chapter Five 2.3) is incompatible with that of mere marker it is supposed to have here. In Arabic linguistic ar-gumentation semantic considerations as a rule are not allowed; there-fore ġā cannot simply be considered 'marker of the indirect object' (although ʾAbū Ḥayyān appears to come close to such an interpreta-tion).

Instead, the problem is partially succesfully approached with the help of an (existing) subdivision of the object into 'pure' and 'non-pure' objects in which the criterion is the use of a particle. In this way they attempt to solve it—as usual—by means of positing certain ele-ments in the underlying structure.

4.2 Innovations in terminology

In this chapter we have seen that the Arabic grammarians applied their morphological principles to Turkic objects and verbs. With regard to verbs, they posited a number of morphemes that—in their view—add transitivity. These analyses lead to some interesting innovations in terminology.

As to objects, they found that the two types of 'direct objects' they recognised in Arabic grammar are marked differently in Turkic. The morphemes that serve as object markers are called 'marker of the ac-cusative' (ʿalāma an-naṣb) in some sources (MG, ʾIdrāk Qawānīn). This is a term used in Arabic linguistics for the 'case ending' a. Another term, 'marker of objectivity' (ʿalāma al-mafʿūliyya), was originally ap-plied to express the function of words that occupy the syntactic posi-tion of the 'accusative' (naṣb), but was then associated with all nouns with the 'case' ending a. The Turkic ending nī is especially associated with the 'accusative' (naṣb), although the sources indicate that nī typi-

cally occurs in those instances in which the verb directly governs the
object, i.e., without the implication of a particle. The use of the terms
'marker of the direct object' (ʿalāma al-mafʿūl bihi) or 'marker of the
proper object' (ʿalāma al-mafʿūl (bihi) aṣ-ṣarīḥ) in this respect are in-
novations, since in Arabic the direct object has no exclusive marker.

For the same reason the term 'marker of the improper object'
(ʿalāma al-mafʿūl ġayri ṣ-ṣarīḥ) can be considered innovative. It is ap-
plied to the morpheme ġā/k̲ā, and associated with objects that are gov-
erned by means of a particle, whether this is implied or deleted.

Inasfar Arabic theory does not permit an analysis as an underlying
particle this analysis therefore seems to come close to the Western no-
tion of 'marker of indirect object'. This is even the case for -ā in the
object of reason (cf. 3.2.3).

It has become clear that the Arabic grammarians assigned the func-
tions and governing capacities of Arabic verbs to their Turkic equiva-
lents. Verbs like ḍaraba / ʾurdī 'he beat' are transitive to one object,
ʾaʿṭā / birdī 'he gave', ẓanna / ṣaġandī 'he thought' to two objects, and
ʾaʿlama / bildirdī 'he informed' to three. In other words, the governing
capacities of Arabic verbs are transferred to their Turkic equivalents.

The morphemes to add transitivity—in our terms 'causativity'—to a
verb, are called 'particles of transitivity' (ḥurūf at-taʿdiya), analogous to
the Arabic particles and consonants that are used to indicate transitiv-
ity. In both languages this process can be described in terms of the ad-
dition of a morpheme to the verbal stem. Even though construction of
denominal verbs is basically impossible in Arabic linguistic theory, for
this proces the same terminology is used. In this sense, then, the suffix
-lā 'adds transitivity' or 'governance' to a noun (cf. 2.3 and 3.3.1).

The tendency to a morphological analysis leads the author of the
Margin Grammar to separate nī into two morphemes, n as a marker of
the object and y (ī) as an optional strengthening element, equivalent
to the Arabic particle li. This analysis is based on the fact that nī
changes into n after possessive suffixes of the third person.

CHAPTER SEVEN

CONCLUSIONS

INTRODUCTION

In this study I have dealt with various of issues relating to Arabic gram-matical analyis of Turkic. After a general outline in Chapter One, the sources were classified according to external and internal evidence in Chapter Two. It was established that there is not only internal evidence that links *'Idrāk* to *Tuḥfa*, the connection between them is also evident from the way both works are organised, which is quite similar to that of 'Abū Ḥayyān's *Irtišāf*. In this structure, the first division is into the domains of lexicology, morphology, and syntax. All other grammars show a different division which is primarily organised according to the three elements of speech, nouns, verbs and particles. This division is the canonical organisation of Arabic grammar.

In Chapter Three it is shown that the Arab grammarians systemati-cally followed the principles of Arabic phonology and phonetics in as-signing labels for velarisation and palatalisation to Turkic words. It was pointed out that these labels were assigned for morphological reasons, rather than phonetic ones. The main criterion for the assignment of the labels is to determine which form of the suffix has to be attached to the word in question. An indication for this is the only partial coverage of the distinction between front and back suffixes, since suffixes that do not contain any of the distinctive consonants are not indicated as either front or back. Turkic consonants and vowels that do not occur in Arabic are described with terms used for Arabic allophones.

In Chapters Four, Five and Six, the most important conclusions are that Arabic case is predominantly syntactic whereas most Turkic cases are semantic. Furthermore, the Arab grammarians regarded the seman-tic Turkic case endings basically as particles, although in some cases the function of some endings (especially ḵā/ḡā) merges with that of a marker. The hypothesis I propose is that the Arabic concept of case is based on this typological feature. In this chapter more observations are advanced on the use and the development of a descriptive model based

on syntactic case, followed by some comments on the attribution and development of linguistic concepts.

1. TURKIC AND ARABIC

The description of the application of the Arabic model on Turkic, forces one to think about the differences and agreements that may exist between the two languages. Here, various aspects come to one's mind, such as for example, definiteness, verbal forms (i.e. the manifold non-finite verbal forms of Turkic), word order, etc..

1.1 *Case and segmentability*

In Chapter Four 4.1.2, I briefly discussed the fact that segmentability is an important element in Arabic theory, although it is of course not exclusively linked with syntactic case. Nevertheless, in order to develop a theory based on declension, the ability to distinguish a word from its ending is a prerequisite. At an early point in the development of Arabic linguistic thinking this important step must have been made.

The grammarians based their theory on nouns both singular and with the socalled broken plurals and imperfect verbs, which take the three different endings, which they subsequently correlated with various types of governance. In later morphological analyses they attempted to account for instances of verbal and nominal endings that do not match the tripartite pattern. Most of these exceptions were associated with roots whose final consonant is a glide, for which the grammarians conceived a set of morphological rules with which they could expain the apparent lack of a declensional ending in the surface structure. Thus, with the help of the concept of 'underlying structure', the grammarians are nearly always able to reduce the numerous apparent deviations from the rule to instances of a mere morphologically conditioned feature.

In no instance the fourfold concept of declensional endings was given up, not even for the so-called regular forms of the plural and dual, which in fact, are highly irregular. The regular plurals have only two forms, i.e. the masculine nominal endings *-ūna* NOM/*-īna* GEN/ACC, and the feminine *-ātun* NOM/*-ātin* GEN/ACC, and the dual, viz. *-āni/-ayni* GEN/ACC. The opinions of the grammarians differed on how these and similar endings should be analysed in further detail (although they did agree on how to segment them). In particular, they were occupied with the status of /w/ and /y/ in /uwna/ and /iyna/, and

/"/ and /y/ in /a"ni/ and /ayni/, respectively (see Chapter Four 1.1.1).
The starting point in these discussions is the principle that each
element represents only one meaning, according to which, for in-
stance, /w/ and /y/ cannot not express simultaneously number, case
and gender.[1]

These analyses are based upon and analoguous to those of singular
nouns and broken plurals which are more easily to segment into word
and declensional ending.

1.2 *Syntactic case and* ʾiʿrāb

An appropriate point of departure for our study is the very base of
Arabic grammatical theory, i.e. ʾiʿrāb. It seems reasonable to presup-
pose that the Arab grammarians applied it to Turkic in exactly the
same way as they used it for Arabic. This immediately conjures up the
question 'what exactly is ʾiʿrāb?'. The answer to this question not only
has to be in accordance with definitions that do right to Arabic case
and verbal endings itself, it is also to cover Arabic interpretations of
Turkic case (and, perhaps verbal endings). Thus, there are three main
points of attention. In the first place a precise definition of ʾiʿrāb in
Arabic theory. Second, a study of the typological features of Arabic de-
clension which are covered by ʾiʿrāb —i.e. Arabic case and verbal end-
ings— and, thirdly, a typological analysis of Turkic case.

In the previous chapters of this study I pointed out that in Arabic
grammar ʾiʿrāb is conceived as a set of markers that indicate gover-
nance by other linguistic elements. This definition itself not only
matches Babby's description of syntactic case on all points, there is also
evidence that Arabic case is, indeed, basically syntactic. The second
type of case proposed by Babby is semantic case, which does not occur
in Arabic (if we exclude the semantic use of the accusative). As for
Turkic, out of its six cases, the genitive, dative, locative and ablative are
semantic, whereas the accusative is syntactic. In both languages the
nominative is difficult to classify as either semantic or syntactic, which
may be a general typological feature of the nominative.

---

[1] Equally impossible, albeit for other reasons, is the assumption that the endings are
mere lengthenings of their short pendants, i.e. *-un/ -in* and *-atun/ -atin* in the singular
and broken plural forms of nouns.

1.3 *Syntactic case as the base of the Arabic concept of ʾiʿrāb*

It is not unreasonable to postulate that the Arabic concept of ʾiʿrāb is based on a language which has exclusively syntactic case—and, incidentally, verbal endings that can be accounted for in similar terms. (There are, however, some instances in which Arabic cases can be used semantically, see discussion below.)[2] For Arabic theory this means on the one hand that it has inherent difficulties in analysing semantic case in terms of ʾiʿrāb, not only to inasfar as case is used semantically in Arabic (see the various types of optional objects, Chapter Six), but also semantic case in other languages. On the other hand, since the theory was especially deviced for analyses of syntactic case, Turkic syntactic case should be recognised without many difficulties.

Although at first sight the picture appears somewhat blurred because of the differences between pure *(mafʿūl ṣarīḥ)* and non-pure objects *(mafʿūl ġayr ṣarīḥ)*—which, in fact, is a useful specification of the object in the later Arabic linguistic tradition—only the Turkic accusative is, indeed, described in similar terms as the Arabic cases.

The semantic cases of Turkic are quite straightforwardly considered equivalent to Arabic particles. Or, in different terms, it is impossible to incorporate Turkic semantic case in the concept of ʾiʿrāb. Such an interpretation is impossible, since there would then be no governor to account for their occurrence; the the endings of ʾiʿrāb are mere markers of governance, never governors themselves.

The absence of a governor from the surface level is apparently more serious than that of a declensional marker. Markers can easily be omitted from the surface level, or merge for (mor)phonological reasons, whereas without governors the whole structure becomes meaningless. It appears that governors are primary to markers (cf. az-Zaǧǧāǧī in *ʿIlal*, Versteegh 1995:95ff.).

1.4 *Syntactic case and underlying structure*

As we have seen, traditional grammar is basically semantically oriented, whereas modern linguistic theories, such as the Government and Binding theory and Case Theory, tend to analyse case in syntactic terms. In the latter theories, syntactic relations between linguistic elements are described in terms of government on a theoretical level,

---

[2] It goes, of course, too far to claim that a theory based on a language with almost exclusively syntactic case should *inevitably* develop a concept of case that only admits syntactic case.

which determine the shape of the surface structure. Whether or not case endings in fact appear in the surface structure is, in this type of argumentation, of secondary importance. On this abstract level case is also posited for languages that do not have case in surface structure, such as, for example, English. One could say that the concept of an abstract level is a very useful, if not indispensible aspect of any theory that posits governance relations between syntactic elements.

Similar statements can be made for Arabic theory. The concept of an underlying level, *taqdīr*, permits the grammarians to posit governance relations between elements even though their effect might be not visible on the surface level. Whether or not the markers of governance appear in the surface structure can is discussed in less basic morphonological terms. The grammarians did differ in their opinions of, for example, which meanings are to be assigned to segments of certain verbal and nominal endings (see Capter Four 4.1.2, and 1.3 below), but as a rule they agreed on the basic concepts of four types of governance and how governance is indicated.

When giving an analysis of those aspects in a foreign language the concept of an underlying structure is very useful too (although the Arab grammarians actually never claimed universal application of their model). With the underlying level one is able to account for unusual word sequences, awkward morphological forms and the absence of markers of governance without adapting the model to the new language. It allows one to analyse the foreign syntax and morphology according to a familiar model without having to seek recourse in weak ad hoc interpretations. There is no need to discuss or redefine matters such as governance and related subjects, because they are already extensively dealt with in regular grammatical treatises. For instance, it does not really matter which marker the object takes on the surface level, as long as it can be attributed to a type of governance that yields a similar distribution as the one attibuted to Arabic verbs.

Indeed, only in instances in which the original model does not have any solutions, simply because they do not occur in the language on it is based, such as when a governing element or a marker of govenance appear on unexpected positions (or with a verbs they never occur with in Arabic), when two governors appear in a sequence, or when verbs apparently have governing capacities Arabic verbs do not have, problems do arise. It is in the analyses of these aspects of Turkic that the sources give different solutions, and each of them is plausible within the Arabic model.

This procedure is not unlike what happens in modern linguistic the-
ory. Originally exclusively based on English, its general rules are con-
stantly adapted based on new data from languages that had not been
analysed in this way before. The main difference between the two the-
ories are, of course, their respective goals. The aim of modern linguis-
tics is eventually to build a theory with universal validity, whereas the
Arabic grammarians did not have such aspirations.

1.5 *Semantic case and governance*

Thus if we accept that governance is the only way to account for mor-
phological syntactic case, a theory which is based on this is likely to
have a certain concept of governance. Even if the cases have no mor-
phological realisation, such a theory may have its merits (although it is
unable to satisfactorily account for semantic case; cf. Babby 1991:2
n.3). Conversely, semantic concepts are of little use for analyis of syn-
tactic case. In this respect I refer to the attempts of Hjelmslev (1972),
Lyons (1971) and Jakobson (in Waugh and Halle 1984) to describe
case in general abstract semantic terms, e.g.

> Un cas, comme une forme linguistique en général, ne signifie pas
> plusieurs choses différentes; il signifie une chose, il porte une seule no-
> tion abstraite dont on peut déduire les emplois concrets (Hjelmslev
> 1972:85).

and

> Although each case of the noun was given a label suggestive of at least
> one of its principal semantic functions (...), it was impossible to give a
> satisfactory general definition of case itself in semantic terms... (Lyons
> 1971:289).

Western grammar developed models in order to show theoretical links
between these abstract notions, among which the semantic concept of
'direction' was taken as a base. From this point of view the dative case
expresses a movement toward the subject, an ablative a movement di-
rected from the subject, and a locative would convey a rest. The ideal
was to link each theoretical case to a single meaning. Such an ideal,
however, is difficult to achieve, since, as we have seen, not all cases can
be satisfactorily explained in semantic terms.[3]

---

[3] Babby (1991:2) already refers to the different ways semantic case is used in lan-
guage; in addition I would suggest that in some languages semantic cases can be used
syntactically, and vice versa, whereas this use may be restricted in others.

2. CONCEPTS

It seems safe enough to assume that the concepts applied in the descriptions of Turkic are exactly the ones developed in the Arabic lin - guistic tradition. Indeed, there are no significant differences in structure and style from regular works on Arabic, and therefore, one should be careful in attributing special insights to their authors.

2.1 *Adoption of other concepts*

In previous studies Mahmūd al-Kāšğarī and Abū Ḥayyān al-'Andalusī, the compilers of *Dīwān* and *Kitāb al-'Idrāk*, respectively, have been atributed other or better knowledge than that of their contemporaries. However, the sole scientific framework they and all other scholars knew was the Arabic tradition; it formed the very basis of their linguis - tic insights and their only starting point to describe language.[4] Unlike the modern western scholar who can choose among a variety of models to apply in certain studies, and who is able to understand the pros and cons of a given approach, it is difficult to imagine Arabic grammarians such as Kāšğarī and Abū Ḥayyān pondering on the most suitable theo - retical framework for the description of Turkic. True, they did had a profound knowledge of foreign languages and scripts, but this fact alone may not necessarily have brought them to abandon their schol - arly concepts. In the case of Arabic linguistics this is even less likely be - cause of the religious associations with linguistic studies.[5]

2.2 *Development of new concepts*

Concepts of syntactic functions are only developed if there is a need to do so. For example, the concept 'indirect object' is not needed if both indirect and direct object are marked with the same marker, such as in Arabic. The marker of both objects can be accounted for by positing governance of the verb, affecting both of them in the same way. Different markers, though, are linked to a different function, such as *u* for the agent *(fāʿil)*, and *a* for the object *(mafʿūl)*, respectively.

In Chapter Six we have shown that the Arabic grammarians found that in Turkic objects are marked differently, a fact that they relate to

---

[4] It goes without saying, however, that within the framework of Arabic linguistic each author could exert his creativity in the application of the instruments supplied.

[5] The question of whether or not the Arabic tradition as a whole was influenced by Greek or Syriac grammatical thinking (for Greek influence see, e.g., Versteegh 1977) is yet another matter.

the way the objects are governed. Thus they related nī to the direct ob-
ject, and ġā/ḵā to the object that in Arabic is transitive by means of a
particle.[6] In later Arabic theory these objects are distinguished too, even
in cases in which the particle is absent from the surface structure. In
some instances, as with the verb ʾaʿṭā 'he gave', though, ġā/ḵā marks
the object without the possibility to refer to a deleted particle, and the
combination of both notions brings them—at least ʾAbū Ḥayyān—
very close to positing a new kind of object, perhaps what in our system
is called an 'indirect object'. One could say that that there is a
beginning of a trend to extend the concept of object. The same can be
said of the possessive ending, which is regarded as a mere marker of the
possessive rather than a possessive pronoun.

Notwithstanding all insights, it quite unlikely that ʾAbū Ḥayyān
applied these new concept to Arabic as well. In his view and that of all
other scholars of his time, the descriptive model they used for Arabic
was not in want of any modifications or new concepts: it already per-
fectly suited its goals.

2.3 A transferred concept of language varieties

A common point in all sources is the reference to the Turkic language
as turkiyya, which is often contrasted with other variants of Turkic. The
sources that were compiled during the Mamlūk period (Qawānīn,
ʾIdrāk, Tuḥfa, Tarǧumān, Bulġa) especially mention Oġuz, which they
call turkmāniyya, while the features of turkiyya itself seem to point at a
Western Qipčaq language. In Šuḏūr it is called al-luġa at-turkiyya al-
ʿuṭmāniyya, and the same source refers to the language described by
ʾAbū Ḥayyān as tatariyya. Kāšġarī uses the term turkiyya too, but in re-
spect to ḫāqāniyya. All this is quite confusing: each author has his own
preference.

The attitude towards Turkic points to a preference for one di-
alect/language in particular which each author considers the 'real'
Turkic. One could say that turkiyya has the status of a prestigious vari-
ant, and the attitude towards it is comparable to the approach to
Classical Arabic in treatises on Arabic, which is, needless to say, held in
very high esteem against less prestigious dialectal variants. In terms of
concepts one could say that the concept of a 'pure', prestigious lan-
guage versus non-prestigious variants is applied to the linguistic situa-

---

[6] The fact that verbs which govern by means of a particle in the later linguistic tradi-
tion are regarded as transitive is already an important extension of the concepts of
transitivity and object.

tion of Turkic peoples (whether or not it is a genuinely reflection of the same is quite a different question).

The sources indeed display a normative attitude in which words and forms are given as 'correct' *turkiyya*. In *Qawānīn* the preference for *turkiyya* is accompanied by severe warnings against the use of *turk-māniyya* (cf. Chapter Two 1.5). In many cases the subjective prefer- ences of the compilers or those of their informants may have played a role in the decision of whether or not to assign to a given word or form the label *turkiyya* (and hence 'correct'). At least in one fourteenth cen- tury source (*ʾIdrāk*), *turkiyya* in some instances refers to typically Qipčaq forms and to Oġuz in others (although in general the Qipčaq forms seem to be preferred). In most cases words and forms explicitly labeled as either Qipčaq or Oġuz can indeed be identified as such. Nevertheless, in spite of this seeming precision, it is not exactly clear what the terms *qibǧāqiyya* and *turkmāniyya* in reality stand for, and with which modern languages they are related. For this is necessary to develop a set of linguistic and lexical criteria to distinguish between Qipčaq and Oġuz.

3. THE COMPILATION OF TURKIC GRAMMARS

There is one issue which has not been dealt with so far, i.e. the motiva- tion of each compiler for writing a grammar of Turkic. This subject is related to two matters. First, the readership, and the way the grammars are conceived as didactical instruments, and, secondly, the interest in compiling manuals of Turkic.

3.1 *The readership*

To start with the first point, *Šuḏūr* is set up as a didactic instrument (cf. Chapter Two 1.7), and as such it stands in a tradition in which word lists of Ottoman, Arabic and Persian were especially developed for learners of those languages. With *Dīwān* Kāšġarī probably intended to show his fellow scholars that in many respects Turkic can compete with Arabic. In regard to the other sources, however, the intended goal can be deduced less easily. The grammatical treatises compiled in Mamlūk times provide a complete means to learn a language (a word list, morphology [in *ʾIdrāk*] and an elaborate discussion of syntax). Although the purposes for their compilation may to some extent have been didactic, this is not explicit in the structure of the works, which points in the first place to a profound scholarly interest.

In Chapter Two I pointed out that the manuals are set up like regular studies of Arabic grammar, dealing with issues that are familiar to the Arabic scholar. In this way, the grammars take features of Arabic as starting points, rather than giving a gradual introduction to the new language. The fact that the Arabic linguistic tradition itself was initiated in reaction to mistakes committed by new (and old) members of the Islamic community does not automatically imply that special learners' grammars were also developed. On the contrary, one might say, the contents of Arabic linguistic manuals are usually quite complicated and deal with highly abstract matters. The necessity of having or developing a special didactic approach for second language learners probably did not exist as a concept in the minds of the compilers. Instead, they structured their grammars of Turkic according to the familiar patterns of Arabic linguistic treatises, leaving basic complicated concepts, such as underlying structure and governance without any reference. Apart from the fact that the readership must have been sufficiently acquainted with these issues, works of this type are indeed hardly the place to discuss or even elaborate new theories. Furthermore, it is hardly conceivable to envisage any need or intention on the part of the Arab grammarians to learn from descriptions of other languages in order to improve their descriptive model.

3.2 The interest in Turkic

With regard to the second point, for *Šuḏūr* and *Dīwān*, again, reasons for compilation can be found in the historical context. *Dīwān* was compiled at a time in which Turkic peoples played important political roles, while *Šuḏūr* dates from Ottoman times, in which Ottoman Turkish, which already had a flourishing literature, served as the main language of official communication. Furthermore, the Turkic grammars that date from Mamlūk times can easily be related to the ruling class in Cairo, which was of Turkic origin and which was constantly strengthened with new young recruits from Central Asia. However, there were almost no direct relations between the arabophone population which was largely illiterate, on the one hand, and the military aristocracy on the other. Between these two layers in Mamlūk society stood — more or less isolated as well — the class of the ʿulamāʾ (see, for example, Haarmann 1988), which to a large extent consisted of descendants of the Mamlūks. Children of Mamlūk soldiers and Arab women were usually kept outside the military establishment. Nevertheless, they were given a good education, and, as a result, some of them

must have known both Turkic and Arabic at a high level. This situation itself did not immediately call for manuals of Turkic, but it does not seem too far-fetched to suppose that some members of this class of ʿulamāʾ combined both the practical knowledge and theoretical linguistic background that were needed for the compilation of grammars of Turkic.[7] Scholarly interest shown by non-Turkic ʿulamāʾ may have been the main inducement to put down this knowledge in writing. It is interesting that this resulted in a 'boom' of grammars in the 13th and 14th centuries.

One could release this study from the rather tight bonds of linguistics, put it in a more general context, and regard it as a report on the way descriptive tools or concepts are applied to other issues than they were originally developed for. Then its issue would become the degree of flexibility in people's concepts in a general sense, or rather how flexible people are when it comes to describing the habits of people with differ- ent beliefs or cultural backgrounds.[8] In this respect one could concen- trate on religious aspects, but also approach the matter from a sociolog- ical perspective.[9] This leads to yet another question, namely to what extent people are aware that they perceive reality from within a certain framework. The description of foreign languages with a model that has been developed for another language is only one of the many possible approaches.

---

[7] In this context, the fact that ʾAbū Ḥayyān, who was not of Turkic origin, succeeded in mastering so many languages is even more surprising.

[8] For discussions of the religious practices of Jews as described by Muslim theolo- gians, cf. Adang (1993).

[9] In this sense the accounts by Ibn Ḥaldūn, Ibn Baṭṭūṭa, and al-Ǧāḥiẓ have already served as sources for a large number of Western studies (to mention only some as a ref- erence, art. Djughrāfiyā by Taeschner in EI²; Miquel 1967-88; Lewis 1990).

PART TWO

TRANSLATION OF

ʾABŪ ḤAYYĀN AL-ʾANDALUSĪ'S

KITĀB AL-ʾIDRĀK LI-LISĀN AL-ʾATRĀK

INTRODUCTION

More than a decade ago, I reckoned it would be 'convenient' to have a Dutch translation of *Kitāb al-Idrāk*. I intended to append the translation to the MA thesis which, at the time, I was preparing at the University of Nijmegen. After graduating, in the summer of 1988, it seemed an equally 'logical step' to start translating the text into English in anticipation of engaging in a Ph.D. research project on Arabic grammars of Turkic. Naturally, I had very much underestimated the effort this enterprise would demand. My Dutch translation turned out to be completely mistaken at many points, not to mention the numerous new problems of interpretation and translation of Arabic terminology which I was confronted with.

Without any hesitation, I sent the first drafts of my work to my supervisor, Kees Versteegh, then director of the Netherlands' Institute in Cairo, who must have been quite surprised to find them in his mail. He managed to find the time for writing detailed comments on virtually all aspects of the translation and sent them to me. Later on, when I was preparing my Ph.D. thesis, on more than one occasion Prof. Versteegh and I had the opportunity to discuss passages from *Idrāk* again.

Even though he always maintains that his students should not explicitly thank him for his advice, I nevertheless would like to express my gratitude for his help and encouragement without which it would have been impossible indeed to translate this complicated text. Any mistakes are, of course, my own.

The present text has no further pretensions than to give a reliable and readable translation of ʾAbū Ḥayyān al-ʾAndalusī's *Kitāb al-Idrāk li-Lisān al-ʾAtrāk*. Although *Idrāk* also contains a large Turkic-Arabic vocabulary, the present translation is limited to the second and third sections on morphology and syntax, respectively. Information about the author, ʾAbū Ḥayyān al-ʾAndalusī, and the context of *al-ʾIdrāk* in the Arabic linguistic tradition is given in Part One of this book.

In the translation references are given to two manuscripts and their edition by Caferoğlu (1931). The first and most important manuscript

is the Velieddin ms 2896 (Beyazıt National Library, abbreviated VD), and the second is the İstanbul Üniversitesi Halis Efendi ms 6597 (Süleymaniyye Library, abbreviated İÜ). References to page numbers in Caferoğlu's edition are not preceded by any abbreviation. I checked Caferoğlu's edition against both manuscripts, and my conclusion is that apart from some minor points (such as occasionally omitted *hamzas*) it is a reliable reflection of the text as it appears in VD.

In some respects, *Kitāb al-ʾIdrāk* could be characterised as a rather dry but solid grammatical description of Turkic. There are almost no references to Arabic grammar, nor is there hardly any elaboration on the principles on which the author based his desscription. In fact, for a study of ʾAbū Ḥayyān's theories on grammar his *Irtišāf*, *Manhaǧ* or even his *tafsīr*, *al-Baḥr al-muḥīṭ* might be more appropriate choices. Nevertheless, I believe that the text in all its conciseness is as accurate as a fourteenth-century professional linguist in the Arabic tradition could write it; a grammar of a foreign language is not the place to discuss matters that relate to the describing model. The basis of the model is still Arabic, and any discussion of elaboration of the model should be related to that language.

The present translation may of interest for arabists in order to have an insight into how exactly the Arabic linguistic model is applied to foreign words and morphemes. Turcologists may be pleasantly surprised with the linguistic insights ʾAbū Ḥayyān displays as a non-native speaker in his descriptions, even though those are not necessarily identical with ours, and through them get a better understanding of fourteenth-century Turkic. The translation may enable general linguists to more easily compare the Arabic descriptive model with other linguistic traditions. For more details on ʾAbū Ḥayyān's analyses of Turkic the reader is referred to the first part of this book.

I have attempted to give translations for Arabic words and terminology that are as unified as possible. For example, 'equivalent' translates *murādif* but also the verbs *rādafa*, and other derived forms. Likewise, 'indicates', 'indicator' and the like are translations of *dalāla*, *dāll* or *yadullu*. In some instances, especially inasfar as more technical terminology is concerned, apt translations were difficult to find. For example, in *Idrāk* the regular term used for 'present tense' is the *muḍāriʿ*. In many instances, though, ʾAbū Ḥayyān applies the term *fī l al-ḥāl* or, elliptically, *al-ḥāl*. Here confusion with the *ḥāl*, here translated as 'circumstantial expression', is likely to occur. Therefore, wher-

ever 'present tense' translates *ḥāl*, the original Arabic is given in a label in the text.[i]

As translations of the Arabic terms used for the 'declensional' endings I have chosen the terms 'accusative' for *naṣb* and 'genitive' for *ǧarr*. The term *rafʿ*, nor the verbal pendants of these terms occur in the text. Furthermore, I choose to translate *ẓarf* as 'locative' (following Talmon 1997), although I am aware that this might add to the terminological confusion.[ii]

For a better comprehension of the text two further remarks may be appropriate. In the first place, the ʾAbū Ḥayyān quite often states that a certain feature in verbs (such as, e.g., the addition of the Turkic passive suffix) is used in the present, future and past tenses alike. The reason that he keeps repeating this point is that in Arabic this type of adaptations typically takes place inside the root of the verb rather than suffixed to a verbal stem. Examples are, e.g., k-t-b 'the root write': *kataba - kutiba - kātaba* 'he wrote' - 'it was written' - 'he corresponded', etc.. Accordingly, all tenses have conjugations in the present and imperative that differ according to the voice of the verb, e.g., *yaktubu* 'he writes' - *yuktabu* 'it is being written' - *yukātibu* 'he is corresponding'. Furthermore, the Arabic tradition does not have a concept of 'stem', and hence the sometimes akward use of different terms (*nafs al-kalima, ʾaṣl al-kalima*) by the author. Second, in some of his morphological analyses ʾAbū Ḥayyān applies other Arabic morphological principles. For example, whenever in so-called 'lengthened vowels' are involved, they must be understood as combinations of (ā) *fatḥa+ʾalif* /aʺ/, (ū) *ḍamma+wāw* /uw/, or (ī) *kasra+yāʾ* /iy/, respectively. Only then expressions such as "as substitution of *ʾalif* with *yāʾ*" make sense.

A different problem was to find apt translations of ethnonyms that refer to Turkic peoples, such as *ʾAtrāk, Qibǧāq* and *Turkmān*. I chose to translate them with 'Turks', 'Qipčaq' and 'Oǧuz', respectively. Arguments for the proposed translations here are given in Part One, Chapter One.

In the text additional information that might facilitate reading is given between brackets. For example, I have included the Arabic originals of all exemplifying phrases and sentences, and original Arabic technical terms are inserted where they appear first and in other instances where it may be useful to know which term is used.

---

[i] At Zaǧǧāǧī's time (appr. 670 - appr. 350) the term *fiʾl al-ḥāl* was still controversial (cf. Versteegh, 1995: 145 and 207).

[ii] Carter (1981) uses "space/time qualifier", Owens (1988, 1990) "circumstantial words" and Versteegh (1995) "adjunct".

The translation of the exemplifying sentences is based on the Arabic examples; where no Arabic version is given by the author, of if this is apparently incorrect, a direct translation from the Turkic is provided between rectangular brackets. For sake of convenience, Turkic mor-phemes are separated from each other by means of a dot, except if they were already written separate in the Arabic text, in which case they are separated by a space. The translation starts with ʾAbū Ḥayyān's intro-duction to the work and then continues with the part on morpholoyg and syntax. The first part, the Turkic-Arabic word list is skipped. All chapters and paragraphs have been numbered in order to facilitate searching and reference.

As pointed out above, there are three versions of the text, i.e. two manuscripts, i.e. the Veli ed-Dīn MS (VD) and the University of Istanbul MS (İÜ). The third is Caferoğlu's edition based on both. The three versions relate to each other in the following manner:

TRANSLATION OF
KITĀB AL-IDRĀK LI-LISĀN AL-ʾATRĀK

INTRODUCTION

<5> In the name of God the merciful, the Compassionate,

5   O Lord, help us, Thou Respectable,

The wise, dilligent and precise worker, the learned grammarian that
sets example, ʾAṯīr ad-Dīn ʿAbū Ḥayyān Muḥammad ibn Yūsuf ibn
ʿAlī ibn Yūsuf ibn Ḥayyān al-ʾAndalusī the Šāfiʿite, immigrant in the
Land of Egypt —may God protect it and cherish it— said:

10   "Praise be to God Who is praised in all languages, Who is free from
signs of imperfection, Who comprises the good and beneficience of all
creatures, Who has endowed the world of man with reason. Prayer be
upon him whom were given all elements of speech and revelation, who
was sent to the nations with the best of religions, who was preferred

15   above the rest of common people, kings and ghosts: Muḥammad, the
Guide to the path of delight. [Prayer be upon] his good kind, the first
ones to be guided and to believe. May He be merciful with his compan-
ions, the supporters of the just and the suppressors of the unjust.
        The precise knowledge of every language is obtained by knowledge of

20   three things: the first are the meanings of all simple words, which is
called 'lexicology' (*ʿilm al-luġa*). The second are the rules of those
simple words before their construction, which is called 'morphology'
(*ʿilm at-taṣrīf*). The third are the rules in case of a construction, which
is called by those that speak about the Arabic language: 'syntax' (*ʿilm*

25   *an-naḥw*).
        I have already written and compiled a large number of books on the
science of the language of the Arabs: *Kitāb at-Taqrīb, Kitāb al-*
Mubdiʿ, al-Mawfūr, Ġāya al-Iḥsān wa-Nukat al-Ḥisān, Kitāb at-
Tadrīb and other scientific works that please the ear and honour the

30   paper.
        The purpose of this book is to record a large part of the language of
the Turks in lexicography, morphology and syntax. I registered this
language consonant by consonant and arranged the lexicographical

part about according to the consonants of the alphabet in the Turkic
language. In this book I first mention the Turkic word <6> and its
equivalent *(murādifihā)* in the Arabic language. After this follow mor-
phology and syntax.

5　　What is in it of lexicography is taken from those I confide in with
regard to the transfer [of knowledge]; I may have used a strange ar-
rangement and peculiar summarisations. As to what is in it of mor-
phology and syntax, I have not followed a [known] method but I ex-
tracted what was known latent by investigating and questioning. I

10　reached with a nice attainment the best of goals and I gained with
many questions the widest search and goal.

　　I have harvested the sweet fruits of heavenly inspiration from the or-
chards of this language. I extracted precious pearls from the width of
knowing it, the objects of utmost desire. In this respect I reached my

15　goal of compiling a book on this language; thus I attained a great succes
in the competition [of compiling books on Turkic].

　　Therefore I entitled my book 'Book for the Comprehension of the
Language of the *'Atrāk*.[1] I put a sign for what is palatalised (q), one for
what is velarised (ḫ), one for what is mixed (š). I also put a sign for

20　what is transferred from the language of the Persians (f) and from that
of the Oğuz[2] *(turkmān)* (t), i.e. what I found suitable for my book.

　　I found disagreement among people talking in the language of the
Turks in that they add or elide a consonant, or change a vowel in an-
other, vocalise a silent consonant or silence a vocalised consonant

25　etcetera. Know that these are variants among them in this language, for
much of this language has changed in these realms, due to the mingling
with arabicised peoples and other foreigners.

　　I divided my book into three parts: the first part deals with the lexi-
cology, the second part with morphology and the third part deals with

30　with syntax."

---

[1] *'Idrāk* also means 'competition; the reaching or attainment of a goal'.

[2] In this translation, *turkmān* and related adjective *turkmānī* are translated as 'Oğuz'.
The use of 'Turcoman' is too strongly associated with present-day Turkmen; this asso-
ciation was probably not intended by the author.

<101> <VD 32ᵛ,1>

II MORPHOLOGY

1. [Morphology *(taṣrīf)* means] knowledge of the rules of a word *(kalima)* before its conjunction *(tarkīb)* with another word.

5   2. The word consists of the consonants *(ḥurūf)* of the alphabet. In this language *(lisān)* the alphabet has twenty-three consonants, i.e. the hamza, the pure *(ḥāliṣa)* b, the mixed *(mašūba)* b, the t, the pure ǧ, the mixed ǧ, the d, the r, the z, the s, the š, the ṣ, the ṭ, the ġ, the q, the pure k, the Bedouin k, *(badawiyya)* the l, the m, the pure n, the post-

10   palatal n *(ḥayšūmiyya)*, the w and the y. If you find in a word another consonant than these, know that the word is not Turkic but transferred from [another] language into this, e.g. ʾaḫšam, farman, firišti·lār and kulaf and others.

    3. The word is an utterance *(qawl)*, or rather what is intended by the

15   utterance. The utterance is created *(mawḍūʿ)* to indicate a meaning, namely a noun *(ism)*, a verb *(fiʿl)* and a particle *(ḥarf)*.

    4. The noun is uniradical *(ʾaḥādī)*, biradical *(ṯunāʾī)*, triradical *(ṯulāṯī)*, tetraradical *(rubāʿī)* or pentaradical *(ḫumāsī)*.

    5. The uniradical noun is vocalised *(mutaḥarrik)* with u *(ḍamma)*, a

20   *(fatḥa)* or with i *(kasra)*, e.g., ṣū and yā and ǧī. The consonants *(ḥurūf)* following [the vocal signs] are a lengthening *(ʾišbāʿ)* and are no part of the root. All three weak consonants *(ḥurūf al-madd wa-l-līn)* [sc. ʾalif, w and y] are like this; none of them is part of the root in this language *(luġa)*; they all arise from the lengthening of the vowels *(ʾišbāʿ al-*

25   *ḥarakāt)*.

    6. The biradical noun: logically *(al-qisma al-ʿaqliyya)* there are twelve possibilities *(qism)*, all of which are used: *faʿ* like san³ and *fiʿ* like kim, *fuʿ* like yuz, *faʿu* like qapu, *faʿa* like ʾara, *faʿi* like ʾari, *fuʿu* like ʾurū, *fiʿi* like bitī, *fuʿa* like buġa, *fiʿa* like qiṣa, *fiʿu* like bisū, *fuʿi* like

30   ṭulī.⁴

    7. The triradical noun: logically there are many possibilities, of which twenty-six are used: *faʿl* like ʾard, *fiʿl* like ṣirt, *fuʿl* like kurt, *faʿlā* like barǧā, *faʿlā* like ṣaġrī, *faʿlū* like qarġū, *faʿalā* like yaraṣā, *faʿalī* like

---

³ The method of using consonantal patterns is, of course, a basic point in Arabic linguistic theory. The patterns can be reflected as follows (random examples): *faʿ* R₁aR₂ / CAC, *fiʿl* R₁iR₂R₃ / CICC, *fuʿalū* R₁uR₂aR₃u / CUCACU *fuʿlul* R₁uR₂R₃uR₄ / CUCCUC, etc.

⁴ Note that the author, in agreement with his preceding remark, does not distinguish between the reflection of long and short vowels in the script. Instead, he used the same basic patterns.

ṭarazī, *faʿalū* like ʾayakū <*102*> *fiʿlā* like sirkā, <VD 33ʳ> *fiʿlī* like dirkī, *fiʿlū* like ʾinğū, *fuʿlā* like ʾuzkā, *fuʿlū* like kuprū, *fuʿlī* like bukrī, *fuʿlū* like ʾuyurū, *fuʿalū* like buqawu, *fiʿalū* like bilaḵū, *faʿal* like ʾatak, *faʿil* like ʾadik, *faʿul* like kazuğ, *fuʿal* like kuzak, *fuʿul* like
5 ʾuzum, *fiʿil* like kiğit, *fiʿal* like kišan and *fiʿul* like diluk.

8. The tetraradical noun: logically there are many possibilities, of which twenty-three are used: *faʿlal* like ʾatmak, *fuʿlul* like bursuq, *fiʿlil* like čibğiq, *fuʿlal* like ʾuğmaq, *fiʿlal* like bizkak, *fiʿlul* like kirbuk, *faʿlil* like balṭir, *faʿlul* like ṭašqun, *fuʿalul* like ʾuraġat, *faʿilil* like ʾanilik,
10 *fuʿalul* like ʾuyanuq, *fuʿulal* like ʾuyukan, *faʿalul* like babağuk, *faʿalal* like ṣarağan, *fuʿalil* like bušariq, *fiʿalil* like bilazik, *fiʿalal* like čiğanaq, *fiʿilil* like ʾisillik, *fiʿillā* like ʾisirğā, *fiʿillī* like ʾikindī, *faʿillī* like ʾalingī, *fuʿullā* like supurkā, *faʿullā* like ṣaqurğa, *faʿallū* like qaraldū, *fuʿullū* like quburğū, *fuʿullī* like quyunğī, *faʿilā* like qaṭirğā,
15 *fuʿalā* like qurbağā, *faʿalī* like šapšapī, *fuʿalī* like qulnağī, *fuʿalū* like qunrawū, *faʿalū* like damrakū, *faʿill* like šakird.

9. The pentaradical noun: logically there are many possibilities, of which twenty-three are used: *faʿlalal* like ʾarqaṭaq, *fuʿlalal* like ʾurmağak, *fiʿlalal* like ṣirṭalan, *faʿlilil* like yaltirik, *faʿlalul* like čatlawuk,
20 *faʿlalil* like qaṣṭaliq, *fuʿlulul* like muštuluq <*103*>, *faʿlulul* like yaldurum, *fiʿlalal* like bildaraq, <VD 33ᵛ> *faʿallal* like ʾaraslan, *fuʿullal* like ʾuṣurmaq, *fuʿullul* like ʾuğurğuq, *faʿullal* like ʾağuršaq, *faʿillal* like ʾağinğağ, *fuʿallil* like yuzarlik,[5] *faʿillil* like yağirliq, *fuʿallul* like čuwalduz, *fiʿillal* like sibildaq, *fuʿullil* like ṭuqurğin, *faʿallul* like
25 ṭağarğuq, *faʿlullā* like qabṭurğā, *faʿullulā* like qamurṣuğā and *faʿallalī* like kasaltakī.

10. They use nouns consisting of six consonants (*sudāsī*), although those are few, and some of them are in accordance with the following patterns: *fiʿlilil* like bildirğin, *faʿillilal* like čağilṭilaq, *faʿlillal* like sal-
30 kinğak, *fiʿallalal* like ṣaringaqan, *fiʿlullul* like qilquyruq, *faʿullulul* like kamulduruk, *fuʿlullal* like ʾulṭurqağ. These are the nouns of six radicals that need to be studied[6] and analysed (*ištiqāq*) in order to discover whether they are compound (*murakkaba*) or simple (*basīṭa*). Some of the words which are clearly[7] compound, are qilquyruq, which
35 is composed of qil which means 'hair' (*aš-šaʿr*) and quyruq which means 'tail' (*danab*), and ṣaringiqan, which is a composition of ṣarī which means 'yellow' (*ʾaṣfar*) and čiqan 'yellow colour' (*lawn ʾaṣfar*).

---

[5] Em. from İÜ for buzarlik.
[6] In İÜ *ḥafḍ*?
[7] *ittaḍağa* prob. err. for *ittaḍaḥa*

11. The verb is 1. uniradical, 2. biradical, 3. triradical and 4. tetra-radical.

12. The uniradical verb is vocalised with *u* or *a* or *i*. Examples of this are: **yū** and **yā** and **yī**[8] its letters are lengthened *(ʾišbāʿ)* just like we saw in the case of the uniradical nouns.

13. The biradical verb: logically there are twelve possibilities, of which ten are used: *fuʿ* like **ʾup**, *faʿ* like **ʾaṭ**, *fiʿ* like **ʾiǧ**, *fuʿu* like **ʾulu**, *faʿa* like **baza**, *fiʿi* like **biši**, *fiʿa* like **ʾira**, *fuʿi* like **čuzi**, *fuʿa* like **buda** and *faʿi* like **dabi**.

14. The triradical verb: logically there are many possibilities, of which nineteen are used: *faʿl* like **ṭart**, <VD 34ʳ> *fiʿl* <104> like **qirb**, *fuʿl* like **ʾurt**, *faʿal* like **ʾaman**, *faʿul* like **ʾanuq**, *faʿil* like **ṭagil**, *fuʿul* like **supur**, *fiʿil* like **kiriš**, *fuʿal* like **čumal**, *fiʿal* like **qizar**, *faʿla* like **ʾaqša**, *fiʿla* like **ʾiqša**, *fuʿla* like **bulga**, *faʿli* like **barki**, *fuʿli* like **bulqi**, *fiʿala* like **ʾiḳala**, *fiʿila* like **ʾikila**, *fuʿila* like **bunila** and *faʿala* like **dapala**.

15. The tetraradical verb: logically there are many possibilities, of which fifteen examples are used: *faʿlul* like **ṣabṭur**, *faʿlil* like **ṣarṣil**, *faʿlal* like **tabran**, *fuʿlul* like **qurṭul**, *fuʿlal* like **ʾuḳrat**, *faʿlula* like **čaṣṭula**, *fuʿlila* like **muštila**, *fiʿlil* like **bildir**, *faʿlala* like **yaǧmala**, *faʿalla* like **baǧašla**, *fuʿulla* like **sumulda**, *fiʿilla* like **qimilda**, *fuʿalla* like **ʾuyanla**, *faʿilla* like **yarilǧa**, *faʿallal* like **čakalan**.

16. Sometimes they use pentaradical verbs, although these are very few in number. Some of them are [constructed] according to the fol-lowing patterns: *faʿlalla* like **ʾaykanla**, *fiʿlilla* like **tizkinla** and *faʿlulal* like **ṣanqulan**.

17. We have set up *(wazannā)* these patterns *(ʾabniya)* and applied them to these examples in order to classify them. It is necessary to study the structures of each of them, so that the primary radicals *(al-ḥarf al-ʾaṣlī)* may be distinguished from the augmented radicals *(zāʾid)*. Only then can the primary radical be compared with the primary radi-cal and the augmented radical with the augmented radical. Perhaps we will learn this in what is to come, with God's permission.

III THE RULES TO WHICH THE WORD IS SUBJECT

1. The status of the simple words consists of two categories. One cate-gory concerns the nouns and the other concerns the verbs.

2. The first category contains the diminutive *(taṣġīr)*, the reference *(nasab)*, the plural *(al-ǧamʿ)*, the active participle *(ism al-fāʿil)* and its

---

[8] Addition from ĪU.

intensive form *(al-mubālaġa fihi)*, the addition for the superlative *(ziyāda li-t-tafḍīl)*, the passive participle *(ism al-mafʿūl)*, the verbal noun *(maṣdar)*, the [noun of] place *(al-makān)*, the instrument *(ʾāla)*, the manner *(hayʾa)* and the destination *(ʾiʿdād)* <VD 34ᵛ>.

5   3. The second category contains the particle of transitivity[9] *(ḥarf al-fiʿl)*, the particle of the passive-reflexive form *(muṭāwaʿa)*, the particle of the medio-passive form *(ittiḥāḏ)*, the particle of reciprocity *(mušāraka)*, the particle of the imperfect tense *(muḍāraʿa)* and the particle of the past tense *(maḍī)* <105>.

10   1. Chapter on the diminutive

1. The marker *(ʿalāma)* for the diminutive *(taṣġīr)* is **kinā** in Qipčāq, and in Oġuz it is **ġuq** and **ġuk**. [The marker] is attached *(tulḥaqu)* to the last consonant of the noun *(ʾāhir al-ism)*, in the singular *(mufrada)* as well as the plural *(maġmūʿa)*, e.g. **qul·kinā** and **qul·lar·kinā** and like

15   **čaġan·ġuq** and **kuġuġuk**. **Kinā** is attached to the noun without addition *(muṭlaqan)*.

2. But when **ġuq** and **ġuk** are attached to a biradical noun whose second consonant is silent *(sākin)*, they may be followed by a *s*, whereas in other cases they may not. You say **qul ġuqas**, **ʾat ġuqas** and

20   **ʾit ġukas**. **Yā** means 'bow' *(al-qaws)*, as its diminutive form you say **yā ġuġas**. They change the *q* into a *ġ*. This is the usage *(masmūʿ)*, but it is not the rule *(muqīs)*. **ʾAz** means 'little' *(al-qalīl)*. They may also use the diminutive forms **ʾaz ġuq** and **ʾaza ġuq** 'very little' *(qalīl qalīl)*. The last form is irregular in two respects. In the first place the *s* is not at-

25   tached. Secondly, the consonant before the marker *(ʿalāma)* of the diminutive bears an *a*. The rule *(qiyās)* would be **ʾaz ġuqas**. There is no alternative to choosing either **ġuq** or **ġuk**. **Ġuq** is attached when the word is pronounced velarised *(mufaḥḥam)* and **ġuk** when it is pro-nounced palatalised *(muraqqaq)*.

30   3. The diminutive occurs in determined nouns *(ʾasmāʾ al-ʾaʿlām)*, in undetermined nouns *(an-nakirāt)* and in demonstrative pronouns *(ʾasmāʾ al-ʾišāra)*.

4. In the case of **bū** ['this'] you say **bū·kinā** and in the case of **ʾan·lar** you say **ʾan·lar·kinā** with the meaning of 'these little' *(dayyā)*

35   and 'those little' *(ʾuwalīyā)*, respectively. In Oġuz [the diminutive of the demonstrative pronoun is not used *(lam yusmaʿ)*].[10] You say

---

[9] In these cases is difficult to determine whether *ḥarf* must bee understood as 'particle' or as 'consonant'.
[10] Addition from ÍÜ (50ᵛ9).

bu·n·ǧuq, bu·n·ǧuqas, bu·n·lar·ǧuq and bu·n·lar·ǧuqas; the *q* [in these words] may be changed into a *ǧ*[11]

5. The pronoun is not subject to the diminutive.

2. Chapter on the reference

5   1. The reference indicates either an occupation *(ṣanʿa)* or something else.

     2. When the reference indicates an occupation, you add *ǧī* to the last consonant of the noun. Thus you say ʾaš·ǧī or 'cook' *(ṭabbāḥ)* and yā·ǧī 'archer' *(qawwās)*.

10   3. When the reference indicates something other than an occupation, you add **luǧ** to the last consonant of the noun. Thus you say rūm·luǧ 'Byzantine' *(rūmī)* and šām·luǧ 'Syrian' *(šāmī)*. The Oǧuz change the *ǧ* into a *w*[12] <106>. In that case you say rūm·lū and šām·lū, respectively. The *l* in these words, namely **luǧ** and **lū**, often receives an *i*, [resulting in **liǧ** and **lī**, respectively].

15   4. <VD 35ʳ> [The word] **luǧ** may also occur with the meaning of 'owner' *(ḏī)* or 'possessor' *(ṣāḥib)*. Thus you say for 'money-owner' *(ḏī māl)* **mal·luǧ·dur** i.e. 'money-owner' *(ḏū māl)*. **Luǧ** is both the marker for the reference and for the possessor *(ḏī)* with the meaning of 'owner' *(ṣāḥib)*. The [element] to which it is linked *(ittaṣala)* makes clear which one of the two meanings is meant.

20

3. Chapter on the plural

     1. The marker for the plural is **lar**. The dual form in this language is a plural. Thus you say **qul·lar** ['slaves']. This is [also] applicable to two persons. If you want to emphasise *(tanṣīṣ)* that there are two men, you use the number that indicates *(dālla)* [their number]. This is yet to come in the [Chapter on] the Number *(ʿadad)* [below Section 47.3a-b]. You attach this marker to the verb as well—in accordance with what will be said in the [Chapter on] the Verb and the Agent [below Section 31]—to indicate that the verb is connected with a plural *(ʾisnād al-fiʿl ʾilā l-ǧamʿ)*. The plural of the singular noun is formed like ʾaṭ·lar ['horses'] and the plural of the collective noun *(ism al-ǧamʿ)* like **baliq·lar** ['fishes'].

25

30

---

[11] Not in İÜ.

[12] The argumentation is as follows: /luǧ/ → /luw/.

4. Chapter on the active participle

1. The marker for [the active participle] in the future tense *(al-mustaqbal)* is **daġī**. Thus you say **barġa·miz 'ul·daġi·biz** 'we will all die' *(kullunā mā'itūn)*. Its marker for the past tense is **miš**. Thus you

5 say **kun duġ·miš** 'the sun was rising in the past' *(aš-šams ṭāli'a fīmā maḍā)*. Both markers are attached to the last consonant of the verb —which is the original root *(al-'aṣl)*— as an indication *(dalāla)* of the agent in the two tenses *(zamānayni)*. If you want to negate the active participle, you add the particle of negation *(ḥarf nafy)* [sc. **ma**] before

10 **miš** in the past tense and **daġī** in the future tense. Thus you say **sanġar ṭur·ma·daġī** i.e. '[Sanġar] is not standing' *(ġayru qāyim)* or 'the money' *(al-māl)* **qal·ma·miš·dī**. **'I·dī** is attached to the last consonant of **miš** i.e. 'the money was not remaining' *(al-māl kāna ġayra bāqin)*, i.e. 'it remained not' *(mā baqiya)*.

15   2. Sometimes **ġī** is attached to the word as an indication of the permanence of the adjective *(dalāla 'alā luzūm aṣ-ṣifa)* since it behaves like an occupation. Thus you say **kil·aġī** [and] **bar·aġī** 'who keeps coming' *(dāyim al-maġī')* and 'who keeps going' *(dāyim ar-riwāḥ)*, respectively.

20   3. Their utterance **bil·kā** means 'a wise man' *('ālim)*. This is a permanent adjective *(ṣifa lāzima)*; the bedouin's *k* [ḵ] *(al-kāf al-badawiyya)* is a substitute *(badal)* for **ġī**.

4. The last consonant of a word is either vocalised or silent. If it is vocalised, you add a *y* and you vocalise it with an *i*. You do this with the

25 permanent adjective, e.g. **čuru·yi·ġī** ['rotting']. If [the last consonant of the root] is silent, you vocalise it with an *i* <VD 35ᵛ> if the preceding consonant is vocalised with either an *a* or an *i*, e.g. **bar·i·ġī** ['going'] and **bir·i·ġī** ['giving']. You vocalise the last consonant of the root with an *u* if the consonant before it is vocalised with an *u*, e.g. **ṭur·u·ġī**

30 ['standing'] and **'ulṭur·u·ġī** ['sitting'].

5. **Miš** is also attached as an indication of an adjective that is meditated about *(ṣifa muḥbar 'anhā)* <107> when the speaker did not witness it personally but only heard of it. You say **'aybak ṭur·miš** i.e. 'is surely standing' *(qāyim yaqīnan)*, even if you did not see him standing.

35 This meaning is observed *(malḥūẓ)* in every adjective to which **miš** is attached and in every instance in which it is attached. If they say e.g. **kun duġ·miš** 'the sun is rising' *(aš-šams ṭāli'a)*, this meaning is still preserved in its original sense *('aṣl al-waḍ')*, even though it is used differently here as a figure of speech *(ṭarīqa al-maġāz)*. Its original mean-

40 ing is that it is used for what one knows but has not seen oneself.

5. Chapter on the intensive form of the active participle

1. You add **kan** to the last consonant of the verb when it is pro-
nounced palatalised, and **ġan** when it is pronounced velarised. The
consonant before the marker ends in *a*. Thus you say **kal·a·kan** 'often
coming' (*ğayūʾ*) and **bar·a·ġan** 'often going' (*ḏahūb*). When the pre-
ceding consonant is silent, the form does not belong to this chapter; its
meaning becomes 'the one that did' (*allaḏī faʿala*). This is yet to come,
God willing, in the Chapter on the Relative [see below Section 25.11].

2. There is also a form **kul·aǧ**, with a *ǧ* as an alternative (*ʿiwaḍ*) for
the [combination of the] *k* with the *n*, i.e. 'much laughing' (*ḏahūk*), the
origin (*maʾḫaḏ*) is the spoken language (*samāʿ*), but the rule is
kul·a·kan, which is used too.

6. Chapter on the addition for the comparative[13]

1. The marker for the comparative (*tafḍīl*) is **rak** when the word is
pronounced palatalised and it is **raq** when the word is pronounced ve-
larised. The markers are attached to the permanent adjective, that is in
the same category (*manzila*) as the occupation, being a pure noun,
whether the last consonant of the noun carries **ǧī** or not. Examples of
words ending in **ǧī** are **bar·ǧī·raq** 'more going' (*ʾaktar ḏahāban*), and
bir·ǧī·rak 'more giving' (*ʾaktar ʾiʿṭāʾan*).

2. Examples of a noun that does not carry **ǧī** on the last consonant
are their utterances **yik·rak** 'better' (*ʾaḫyar*);[14] **yik**, meaning 'good'
(*ḫayr*); **yaman·raq**—'meaner' (*ʾašarr*); **yaman**—'mean' (*šarr*);
bilka·rak—'more knowing' (*ʾaʿlam*); **bilḫā**—'who knows' (*ʿālim*) and
yawuz·raq—'worse' (*ʾašarr*). **Yawuz** means 'bad' (*šarr*); <VD 36ʳ> this
word is Oġuz, for they say **yawuz** instead of **yaman**.

7. Chapter on the passive participle

1. The passive participle is formed from a passive verb only. You say
ʾur·ul·dī for 'he was beaten' (*ḍuriba*) and **bir·il·dī** for 'he was given'
(*ʾuʿṭiya*). When you mean the passive participle, you elide the marker
of the verb in the past tense and you put **miš** instead of it. <108> You
add either a silent *l* or a *n* before it, according to the detailed explana-

---

[13] In Arabic the form of the comparative is identical to the that of the superlative. The
only difference is that the superlative is aways defined, e.g *ṣaġīr* 'small'; *al-ʾaṣġar* 'the
smaller one' or 'the smallest one.'

[14] The comparative form of *ḫayr* is *ḫayr*, not *ʾaḫyar*.

tion that is to come in the Chapter on the Construction of a Verb to
the Passive Participle [see below Section 35]. Thus you say **saṅar 'ur·**
ul·miš 'i.e. is beaten' *(maḍrūb)*, and 'the coat is' *(aṭ-ṭawb)* **bir·il·miš**
'given' *(mu'tan)*. [You also say] 'killed' *(mumāt)*.[15] and **dapala·n·miš**
'killed' *(maqtūl)*. This is the case in the past tense. With regard to the
future tense, you say **'ur·ul·daġī, bir·il·daġī, 'ul·dur·ul·daġī** and
dapala·n·daġī.

2a Either a silent *k* with a palatalised word or a silent *q* with a ve-
larised word indicate the passive and the active participle. This is the
case when it is a permanent adjective, although this is not the rule for
each verb. What actually occurs is **'uz·uk** 'cut' *(al-mafṣūl)* from **'uz** i.e.
'split!' *('ifṣil)*, and **'aġ·uq** 'opened' *(al-maftūḥ)* from **'aġ** i.e. 'open!'
('iftaḥ). You say **čuru·k** 'rotten' *(al-bālī)* which is derived from **čuri·dī**
i.e. 'he rotted' *(balā)*.

2b. You may not apply this as a general rule; they do not say
****'ur·uk**[16] for 'beaten' *(maḍrūb)* in the case of **'ur·dī**, nor ****ṭur·uq** for
'standing' *(qā'im)* in the case of **ṭur·dī**.

2c. But [on the other hand] it is permitted to apply the general rule
[s.c. the attachment of **miš**] to verbs whose passive participles may also
be formed by means of a *k* or a *q*. So you say **'uz·ul·miš** 'cut' *(al-*
mafṣūl) and **'aġ·il·miš** 'opened' *(al-maftūḥ)*. In the case of the active
participle you say **čuri·miš** 'rotten' *(al-bālī)*; this is in the past tense. In
the future tense you say **čuri·daġī**, which has already been mentioned
in the Chapter on the active participle [see above Section 4].

8. Chapter on the verbal noun

1. The marker [for the verbal noun] is the attachment of **maq** to the
imperative form of the verb if the preceding consonant is velarised. The
marker is **mak** if the preceding consonant is palatalised. It is permitted
to add the form **liq** to **maq** and the form **lik** to **mak**. It is also
permitted to reduce [both markers] to a silent *m*. These three possibili-
ties are the basis of the analogy *(muqīs)*. <VD 36ᵛ> You may say either
kal·mak 'coming' *(maġī')* or **kal·mak·lik** or **kal·im**. And, likewise, you
may say **bar·maq** 'going' *(ḍahāb)*, **bar·maq·liq** and **bar·im**. The con-
sonant before the silent *m* is vocalised with an *i*.

---

[15] This apparently strange form, where *maqtūl* would have seemed more plausible,
contains all the features shown by **'ul dur ulmiš**: a causative in the construction of a pas -
sive participle.
[16] One would expect **'uruq** instead, since **'ur-** is pronounced velarised.

2. Another marker for the verbal noun is a silent š attached to the final consonant of the verb [i.e. verbal root]. This indicates the mean-ing of the intensive form in the verbal noun. Thus you say **ʾur·uš** and **ṭur·uš**. The final consonant of the verb is either vocalised or silent.

5   3. If the final consonant is vocalised, you add a *y* with an *i*. So in the case of **yi·dī** you say **yi·y·iš**. If the final consonant is silent, and the consonant before it is vocalised with an *u*, you give it an *u*, e.g. **ʾur·uš** and **ṭur·uš**. [If the penultimate consonant is] vocalised with an *a*, you give it an *i*, e.g. **baq·iš** and **bir·iš**, 'look' *(naẓar)* and 'giving' *(ʾiʿṭāʾ)*, re-

10   spectively.

        4. If you wish to give a noun the meaning of a verbal noun, you add **liq** or **lik** to the last consonant of the noun, in the same way you add it to the verbal noun. Thus you say **bāliq·liq** in the case of **bāliq** ['fish'] ; this means 'being a fish' *(samakiyya)*. And in the case of **bay** <109>

15   ['chief'] you say **bay·lik** which means 'principality' *(ʾamīriyya)*. In the case of **yak** you say **yak·lik** which means 'goodness' *(ḫayriyya)*. In the case of **yuq** you say [**yuq·luq**][17] which means 'lacking' *(maʿdūmiyya)* and they use it to express 'poverty' *(faqr)*.

9. Chapter on the noun of place

20   1. You add to the imperative form of the verb **ǧak** or **ǧaq**, adding an *a* before the added element. [The addition is] according to the rules of velarisation and palatalisation. Examples of this are **kal·aǧak** and **ṭur·ǧaǧaq** i.e. 'place of arrival' *(makān al-maǧīʾ)* and 'place of stand-ing' *(makān al-qiyām)* respectively. This is the rule for each verb. In this

25   language there is no noun of place derived from the verb.[18]

10. Chapter on the instrument

1. The instrument is indicated either by a *k* vocalised with an *u* or a *q* vocalised with an *u*. For 'he cut' *(qaṭaʿa)* you say **kas·tī**. For the in-strument [constructed] from this verb you say **kas·kū**. [Another form

30   for] 'he cut' *(qaṭaʿa)* is **biǧ·tī** For the instrument [constructed] from it you say **biǧ·qū**. For 'he whetted' *(sanna)* you say **bila·dī** and for the instrument [constructed from it], which is the whetstone, you say **bila·kū**. The *k* in **bila·kū** is bedouin. And **ʾalaḵ**,[19] which is used for the

---

[17] Addition from IÜ.

[18] ʾAbū Ḥayyān here means the derivation of a noun of place in the same way nouns are derived from roots in Arabic, e.g. *madrasa* 'place of learning', derived from the root *d-r-s* 'to learn'.

[19] It is difficult to decide whether k here stands for ḵ or k̰

instrument 'sieve' *(munḫal),* [is derived] from **'ala·dī** <VD 37ʳ> i.e. 'he
sieved' *(naḫala).* Its base form is **'ala·ḳū,** analogous *(miṯl)* to **bila·ḳū,**
but it was alleviated by silencing the *ḳ.* Thus they say **'alaḳ.** The first
form [sc. **·ḳū**] is the rule *(qiyās)* for those verbs from which an
5   instrument is constructed.

11. Chapter on the manner

1. The manner *(hayʾa)* is indicated by a silent *š.* Thus you say **min·iš**
i.e. 'manner of riding' *(hayʾa ar-rukūb)* and **'uṭur·uš** i.e. 'manner of
sitting' *(hayʾa al-quʿūd).* The consonant before the *š* receives an *u* if
10   the consonant before it is vocalised with an *u,* e.g. **'uṭur·uš.** But it re-
ceives an *i* if [the preceding consonant] is vocalised with either an *i* or
an *a,* e.g. **min·iš** and **čal·iš,** in agreement with what we said in the
Chapter on the verbal noun [see above Section 8]. You say
'uṭur·uš·im **'aybak** **'uṭur·uš·ī** 'my manner of sitting is 'aybak's'
15   *(qiʿdatī qiʿdat 'aybak).*[20]

12. Chapter on the destination

1. The marker for the destination *(ʾiʿdād))* is either **liq** or **lik.** Thus you
say **ṭun·luq**[21] 'what is meant for clothes' *(al-muʿadd li-l-libās)*;
yama·liq 'a patch meant for mending' *(ar-ruqʿa al-muʿadda li-t-*
20   *tarqīʿ)* <110> and **biti·lik** i.e. 'what is meant for written things' *(al-*
muʿadd li-l-kitāb). In the Arabic language there exists no particular ex-
pression for this meaning.

13. Chapter on the particle of transfer and transitivity

1. [The particle of transfer and transitivity *(ḥarf an-naql wa-t-taʿdiya))*
25   is as follows.] The last consonant of a verb is either silent or vocalised.
If it is silent, you add [the particle of transitivity] **dur** to its final
consonant. Thus you say **min·dur** for 'make [him] mount' *(ʾarkib).* If
[the last consonant] is vocalised, you add a silent *t.* Thus you say **yuri·t**
for 'make [him] walk' *(maššī).* You say **dapala·t** for 'make [him] kill'
30   *(ʾaqtil)* and **'uqi·t** for 'make [him] read' *(ʾaqriʾ).*

2. The particle of transitivity *(ḥarf at-taʿdiya)* is obligatory *(yalzamu)*
in all categories *(ʾaqsām)* of the verb, i.e. the imperative —as in the
examples above— the past tense, e.g. **min·dur·dī,** the imperfect tense,

---

[20] Em. for VD 37ʳ7 *'abīka* 'your father'.
[21] The suffix **liq** here changes into **luq,** due to vowel assimilation.

e.g. **min·dur·ur** and the future tense, e.g. **min·dur·kay.** Sometimes
the *d* is elided and only the *r* remains, but this is not the rule *(lā
yanqās).* You say for instance 'make [him] drink' *('asqi)* **'iǧ·ur** and
'make [him] leave' *('aḫriǧ)* **čiq·ar.** But **čiq·tur** is the regular form.

5   3. Sometimes words are made transitive with something other than
dur. The speakers say **ṭur·ġuz** for 'raise' *('aqim)* and **kur·kuz** for
'show' *('arī).* For 'drip' *(naqqiṭ)* they say **ṭam·zur** and for 'suckle'
('arḍiʿ) they say **'am·zur.** You may also use **dur** with these verbs. <VD
37ᵛ> The use of **dur** is the rule *(qiyās).* Thus you say **ṭur·dur, kur·dur,**
10 **ṭam·dur** and **'am·dur.**

14. Chapter on the consonant of the passive-reflexive form

1. The consonant of the passive-reflexive form *(al-muṭāwaʿa)* is a
silent *(sākina) l.* You say **kas·dī** for 'he cut' *(qaṭaʿa)* and **kas·il·dī** for
'he was cut' *(inqaṭaʿa).* You say **'uz·dū** for 'he split' *(faṣala)* and for 'it
15 was split' *(infaṣala)* you say **'uz·ul·dū.**
    2. If the verb consists of one consonant, e.g. 'he broke' *(kasara)*
si·dī, in the passive-reflexive form a silent *n* is used instead of a *l.* Thus
for 'he was broken' *(inkasara)* you say **si·n·dī** .

15. Chapter on the consonant of the medio-passive form

20 1. The consonant of the medio-passive form *(ittiḥāḏ)* is a silent *n.* You
say **yasṭa·dī** 'he propped him up on a pillow' *(wassadahu)* and
yasṭa·n·dī 'he used it as a pillow' *(tawassadahu)* i.e. 'took it as a pil-
low' *(ittaḫaḏahu wisādatan).* This is derived from their expression
yasṭuq which means 'pillow' *(wisāda).* An equivalent *(naẓīruhu)* for
25 this is **'urtū** 'covering' *(al-ġiṭā').* When the speaker intends to say 'he
covered' *(ġaṭṭā)* he says **'urt·tī.** And when he means 'he was covered'
(taġaṭṭā), he says **'urt·un·dī.** This is the rule in those words that accept
medio-passivity. You say **ṭun** for 'coat' *(ṭawb)* and then you say **ṭu-**
nan·dī 'he got dressed' *(taṭawwaba)* i.e. 'he took a coat' *(ittaḫaḏa*
30 *ṭawban),* i.e. 'he dressed himself' *(iktasā)* <111>.

16. Chapter on the consonant of reciprocity

1. The consonant of reciprocity *(mušāraka)* is a silent *š,* like the *š* of
the manner. You say **'ur·dī** for 'he beat' *(ḍaraba)* and **'ur·uš·dī** for
'he fought' *(taḍāraba).* In case of 'he stabbed' *(ṭaʿana)* you say **ṣanǧ·dī**
35 and for 'he battled [with someone]' *(taṭāʿana)* [you say] **ṣanǧiš·tī.**

Hence the expression ṣanǧiš kun·ī 'day of the attack' *(yawm aṭ-ṭaʿān)*. In the imperfect tense you say 'uruš·ur and in the future tense 'uruš·ǧā. If the verb is constructed for the passive form, the rule would demand that they say 'uruš·ul·dī for 'he was attacked' *(tuḍūriba)* or

5   ṣanǧ·uš·ul·dī for 'it was fought' *(tutūʿina)*, but neither one of these forms is actually used.

17. Chapter on the consonant of the Imperfect Tense

1. <VD 38ʳ> The consonant of the imperfect tense *(muḍāraʿa)* is a silent *r*, like they say ṭur·ur i.e. 'he stands up' *(yaqūm)*, kal·ur i.e. 'he

10   comes' *(yaǧīʾ)*, ya·r i.e. 'he eats' *(yaʾkul)* and si·r 'he breaks' *(yaksir)*.

2. One may also use yi·y·ur and si·y·ur as if the [speakers] strengthen the verb with an extra radical, when it consists of only one. They would say yu·y·ur 'he washes' *(yaǧsil)*, but more often they say yu·r.

15   18. Chapter on the consonant of the past tense

1. The consonant of the past tense *(maḍī)* is a *d* that is vocalised with an *i*, if the preceding consonant is vocalised with an *i* or an *a*. [Furthermore, the consonant of the past tense] is vocalised with an *u* if the preceding consonant is vocalised with an *u*, e.g. ṭur·dū, 'ultur·dū,

20   bar·dī and bir·dī. In Oǧuz the *d* always receives an *i*. The *d* may be changed in a *ṭ* or a *t* because their place of articulation is near that of the *ḍ*.

IV

This is a chapter in which the consonants of addition and their places

25   are dealt with, and [further] the consonant of substitution, the elision, the transfer of the vowel and the assimilation.

19. Chapter on the addition

1. The elements indicating the addition *(ziyāda)* are divided into two categories *(šayʾān)*. The first one is etymology *(ištiqāq)* and the second

30   one morphology *(taṣrīf)*. Etymology means the deduction of the derived form *(farʿ)* on the basis of the base form *('aṣl)*. Morphology means the deduction of the base form from the derived form. This will

become clear <*112*> when the places of addition are explained. The consonants of addition are *r, n, s, b, t, ğ, d, z, q, k, ġ, y, m, š, l* and *w*.

2a. A *r* is added to the verb in the case of the present tense *(fi ʾl ḥāl al-muḍāriʿ)*, like **kal·ur** ['he comes'] and **ʾal·ur** ['he takes'].

5   2b. The *r* is also added when a verb is derived from a noun e.g. [the noun] **qayġū** ['grief']. When you derive something from this noun, you say **qayġur·dī; balkū** ['shine'] **balkur·dī; ʾaq** ['white'] **ʾaġ·ar·dī; qizil** ['red'] **qizar·dī; yašil** ['green'] **yašar·dī; kuk** ['blue'] **kuk·ar·dī** <VD 38ᵛ> and **buz** ['grey'] **buz·ar·dī**. This is the general rule for the
10   colours.

2c. A *r* is added preceded by a *l* in the plural. In the case of **ʾaṭ** ['horse'] you say **ʾaṭ·lar** and in the case of **ʾit** ['dog'] you say **ʾit·lar**. **Lar** is also attached in the adaptation of the verb to the plural *(nisba al-fiʾl ʾilā l-ğamʿ)*, e.g. **kal·di·lar** ['they came'].

15   2d. A *r* is added with a *k* in **rak** and with a *q* in **raq**, as an indication of the superlative *(dalāla ʿalā t-tafḍīl)*, e.g. **yak** [vs. **yik·rak**] and **yaman** [vs. **yaman·raq**].

2e. A *r* is added preceded by a *d*. [Together they] form the particle **dur** which is used to express transitivity. [For example,] in the case of
20   **min·dī** ['he drove'] you say **min·dur·dī** ['he let drive']. The *d* may be elided while the *r* remains; in the case of **ʾiğ** ['drink'] you say **ʾiğ·ur** ['make drink']. Its base form, though, is **ʾiğ·dur**.

3a. A silent *n* is added to express the medio-passive form *(ittiḥād)*. In the case of **yasṭa·dī** ['he took as a cushion'] you say **yasṭa·n·dī** and
25   in the case of **ʾurt·tī** ['he covered'] [you say] **urt·un·dī**.

3b. A *n* is added to the verb when it is constructed for the passive form *(mā lam yusamma fāʿiluhu)*, according to the detailed information that will be given in the appropriate chapter [see below Section 35]. In the case of **yi·dī** ['he ate'] you say **yi·n·dī** ['he was eaten'] and
30   in the case of **ṣina·dī** ['he tested'] you say **ṣina·n·dī** ['he was tested']. The *n* is added likewise in the case of the passive participle *(mafʿūl)* whose agent is not mentioned, e.g. **ṣina·n·miš**.

3c. A *n* is added in the case of the plural *(ğamʿ)* of **ʾar** ['man']; in the case of 'I saw sensible men' *(raʾaytu riğālan ʿuqalāʾ)* you say **ʾuṣlū**
35   **ʾaran·lar kur·du·m** and similarly you say **ʾa·n·lar ʾar·an ʾuṣlū·lar· dur**.

3d. A silent *n* is added preceded by a *k* that is vocalised with an *a* [i.e. **kan**]. This serves as an indication of the intensive form of the active participle *(dalāla ʿalā l-mubālaġa fī sm al-fāʿil)*. The *n* is also added
40   preceded by a *ġ* [i.e. **ġan**]. In that case you say **kal·akan** and **bar·aġan**.

An exception *(šadda)* [to this rule] is **kul·aǧ**. The rule is **kul·akan**, which is also used *(sumiʿa)*.

    3e. A *n* is added together with these two [sc. **kan** and **ǧan**] while the preceding consonant remains silent. This serves as an indication of

5   both the active participle *(ism al-fāʿil)* and the passive participle *(ism al-mafʿūl)* in the meaning of the past tense used as a relative clause *(bi-maʿnā al-maḍī mawṣūlan)*. Thus you say **kal·kan** and **bar·ǧan** 'the coming' *(al-ǧāʾī)* and 'the going' *(aḏ-ḏāhib)*, i.e. 'the one that came' *(allaḏī ǧāʾa)* and 'the one that went' *(alaḏī ḏahaba)*, respectively. You

10   say **dapala·n·kan**, i.e. 'the one that was killed' *(alaḏī qutila)*; **ʾur·ul·ǧan** i.e. 'the one that was beaten' *(alaḏī ḍuriba)* <VD 39ʳ> and **bir·il·kan** i.e. 'that was given' *(alaḏī ʾuʿṭā)* respectively.

    3f. A *n* is added with an *a*, when it is preceded by a *k* vocalised with an *i*. This serves as an indication of the diminutive. Thus in the case of

15   **qul** you say **qul·kinā** and in the case of **ʾaṭ** you say **ʾaṭ·kinā**.

    3g. In the circumstantial expression *(ḥāl)* a silent *n* is added preceded by a *b* vocalised with an *a* [sc. -**b·an**]. This will be dealt with in the Chapter on the circumstantial expression [see below Section 39].

    3h. A silent *n* is added in the case of **ṭuqṣan** ['ninety'] and **saksan**

20   ['eighty']. The base forms *(ʾaṣl)* [of these words] are **ṭuquz** ['nine'] and **sakiz** ['eight'], respectively. <113> The *n* indicates multiplication *(taḍʿīf)* of the number by ten since **ṭuquz** means 'nine' *(tisʿa)* and together with the *n*, it means 'ninety' *(tisʿīn)*. **Sakiz** means 'eight' *(ṯamāniya)*; together with the *n* it means 'eighty' *(ṯamānīn)*.

25   3i. A *n* is added in the case of **ṣamaṣ** ['obstinate'] when they want to derive a verb from it. They say **ṣamaṣ·na·dī** for 'he was obstinate' *(laǧǧa)*, in order to differentiate between the noun and the verbal stem *(fiʿl)*.

    4a. A silent *s* is added after **ǧuq** and **ǧuk**, the markers for the

30   diminutive in Oǧuz, in every biradical noun whose second consonant is silent. You say **qul·ǧuqas** ['little slave'] and **ʾit·ǧukas** ['doggie'].

    4b. A *s* is added to the last consonant of an annexed noun [in the case of] a pronoun of the third person singular, when [the last consonant] of the noun carries a vowel, like **quǧa·s·ī**, **ʾaṭa·s·ī**, **ʾulu·s·ī** and

35   **qari·s·ī**.

    4c. A *s* is added for the third person singular *(ǧāʾib)* or plural *(ǧāʾibīn)* in the negative form of the verb of the present tense *(al-ḥāl al-manfī)*, e.g. **kal·ma·s** ['he will not come'] and **kal·ma·s·lar**. And also for the second person both in the singular *(al-muḥāṭab)* and plural

40   *(al-muḥāṭabīn)*, e.g. **kal·ma·s·san** ['you [sg.] do not come'] and **kal·ma·s·siz** ['you [pl.] do not come']; and for the first person both in

the singular *(al-mutakallim)* and plural *(al-mutakallimīn)*, e.g.
kal·ma·s man ['I will not come'] and **kal·ma·s biz** ['we will not
come'].

4d. A doubled *s* *(mušaddada)* is added and it is vocalised with an *a*
preceded by a *m* vocalised with an *i* [i.e. **missā**]. An example of this is
the use of **kaškā** when its predicate *(ḥabar)* is not a verb, e.g. **kaškā
sangar kul·ar·mis·sa i·dī** ['If Sangar had only laughed'] and also
kaškā sangar 'aw·dā·mis·sa 'i·dī ['If Sangar had only been at home']
and the equivalent *(murādif)* for 'if not' *(law lā)*, like **sangar
dakul·mis·sā sunqur ṭur·miš 'i·dī** ['If it had not been for Sangar,
Sunqur would have stood up'].

5a. A silent *b* is added as a general rule *(qiyāsan muṭṭaridan)* in the
circumstantial expression *(al-ḥāl)*, e.g. **ṣur·ub** ['asking']; **kal·ib**
['coming'] and **kir·ib** ['entering'].

The *'alif* in **lā**, which is added to the last consonant of the verb to
indicate the governance *('amal)*, changes into a *y*. Thus you say
suz·la·y·ub ['saying'] and **'ag·la·y·ub** ['weeping']. It is permitted to
elide *(ḥaḏf)* the *'alif*, because of its clustering *(iltiqā')* with < VD 39ᵛ>
the silent *b*. Thus you say **suz·la·b** and **'ag·la·b**.[22] We already referred
to the addition of the *n* [to this form] in the Section on the *n* [see
above Section 19.3g and below Section 39].

5b. A silent *b* is added, with repetition of the first consonant of the
word, in the case of the colours. In the case of **ṣarī** ['yellow'] you say
ṣab·ṣarī and in the case of **qizil** ['red'] you say **qib·qizil**. In the case of
kuk ['blue'] you say **kub·kuk**. This indicates the intensive form of the
adjective: 'violent yellow' *(aš-šadīd aṣ-ṣufra)*, 'violent red' *(aš-šadīd al-
ḥumra)* and 'violent blue' *(aš-šadīd az-zurqa)* respectively.

5c. The *b* is sometimes substituted by a *m*, but this will be dealt with,
God willing, in the Chapter on the consonants of Substitution [see be-
low Section 20].

6. A *t* is added as a general rule to indicate transitivity in those verbs,
of which the last consonant carries a vowel. In the case of **dapala**
['kill!'] you say **dapala·t** ['make kill!'] and in the case of **yuri** ['walk!']
[you say] **yuri·t** ['make walk!'].

7a. An isolated *ǧ* is added in the case of reference to an occupation
(nasab 'ilā ṣ-ṣanʿa). You say **'aš·ǧī** ['cook'] and **yā·ǧī** ['archer'] as an
indication of the permanence of the adjective *(luzūm aṣ-ṣifa)*, so that it
becomes an occupation, e.g. **bar·iǧī** ['much going'] and **kal·iǧī** ['much
coming']. The expression *(qawluhum)* **bil·kā** with a Bedouin *k* is an

---

[22] The argumentation can be represented as follows (" represents *'alif*): /laʺ/ → /lay/;
elision of /ʺ/ before *b*: /laʺ/ → /laʺb/ → /lab/.

exception *(šadd)*; the rule is **bil·ǧī** and this has been dealt with in the Chapter on the Active Participle [see above Section 4].

7b. Exceptionally, a *ǧ* is added in the case of **kul·aǧ** ['much laughing'] as an indication of the intensive form in the active participle.

7c. A *ǧ* is added together with a *q* and a *k* in **ǧuq** and **ǧuk** in the case of the diminutive which has been discussed earlier [see above 1].

7d. With both consonants [sc. *q* and *k*] the *ǧ* is used vocalised with an *a*, as a marker for the noun of place *(ism al-makān)*, like <114> **kal·aǧak** ['place of coming'] and **ṭur·aǧ·ǧaq** ['place of standing'].

7e. A *ǧ* vocalised with an *i*, is added preceded by a *d* vocalised with an *a* [i.e. -**daǧī**]. [We discussed this] in the Chapter on the Active Participle in the future tense [see above Section 4.1]. Thus you say **ʾul·daǧī** ['he will die'].

7f. A *ǧ* is added together with an *a*, lengthened *(mušbaʿa)* by an *ʾalif* in **ʾaq·ǧā**, **qarā·ǧā** and **ṣarū·ǧā** as an indication of the diminutive which also conveys tenderness *(taḥbīb)*.

8a. An isolated *d* is added, vocalised with an *i* or an *u*, as an indication of the past tense of the verb e.g. **kal·dī** and **ṭur·dī**.

8b. A *d* is added together with another consonant as has been explained above in the paragraph about the addition [of the *d*]. It is added <VD 40ʳ> preceding a *ǧ* which is vocalised with an *i* as an indication of the active participle in the future tense [, i.e. **daǧī**].

8c. A *d* is added together with a silent *r* in **dur** which is used to express transitivity. Sometimes they change it into **tur**.

8d. When the *d* indicates the past tense, [it is substituted for either a] *ṭ* or *t* in some cases, although this is not the rule. We claim that the base form is a *d*, because most of what is used is with a *d*. It only occasionally occurs with *ṭ* or *t*. Frequency *(al-kaṭra)* is an indication that something is the base form *(al-ʾiṣāla)* and rare occurrence *(qilla)* is an indication that something is the secondary form *(al-farʿiyya)*.

9. A silent *z* is added together with either a *ǧ*, a *k* or a *r* to express transitivity in exceptional cases, e.g. **ṭur·ǧuz** ['make stand up'], **kur·kuz** ['make see'], **ṭam·zur** ['make drip'], and **ʾam·zur** ['make suck'].

10a. The addition of the *q* has already been discussed in the case of **ǧuq** and **ǧaq**.

10b. A *q* is added together with a *l* that is vocalised with an *i* to indicate the destination *(al-ʾiʿdād)*, e.g. **yamā·liq** ['patch for mending']. It serves also as an indication of the 'verbal noun' of the nouns *(maʿnā al-maṣdar min al-ʾasmāʾ)*, like **bāliq·liq** and further as a corroborative

(taʾkīdan) after the marker for the verbal noun, like **bar·maq·liq** ['going'].

10c. A silent *q* is added preceded by a *m* that is vocalised with an *a*, to serve as an indication of the verbal noun, e.g. **bar·maq** ['going'].

10d. A silent *q* is added occasionally as an indication of the active participle, e.g....[23] and also as an indication of the passive participle, e.g. **'aǧ·uq** ['open'].

10e. A *q* vocalised with an *u* is added to indicate the instrument, e.g. **biǧ·qū** in the case of every verb that is pronounced velarised *(mufaḫ-ḫam)*.

11a. A single *k* vocalised with an *u* is added as an indication for the instrument, e.g. **kasa·kū** and **bila·ku̬** in each word that is pronounced palatalised *(muraqqaq)*. In exceptional cases it occurs silent, e.g. **'ala·k̭**.

11b. A silent *k* is added preceded by a *m* vocalised with an *a* to indicate to the verbal noun, e.g. **kal·mak**.

11c. A *k* is added preceded by a *l* as a corroborative of the verbal noun, like **kal·mak·lik**. It is also added to serve as an indication of the 'verbal root' of the nouns, e.g. **bay·lik**, or as an indication of the destination, e.g. **biti·lik**.

11d. A silent *k* is added preceded by a *ǧ* that is vocalised with an *u* in **ǧuk** to indicate the diminutive.

11e. A silent *k* is added preceded by a *ǧ* vocalised with an *a* in **ǧak** in the case of the noun of place.

11f. <VD 40ᵛ> A *k* vocalised with an *a* is added to indicate the intensive form of the active participle [, e.g. in **bil·k̭ā** 'wise'].

11g. A *k* vocalised with an *a* is also added in the cases of **kinā** and **yak·rak**, which already have been discussed.

11h. A *k*, vocalised with an *u* is added with the preceding consonant vocalised with an *a* to serve as a marker for a collective of numbers *(al-iǧtimāʿ fī l-ʿadad)* e.g. **'īka·kū**,[24] **'iǧa·kū, durd·a·kū, bīš·a·kū, yida·kū**, i.e. 'the collection of two' *(al-'iṯnāni al-muǧtamiʿāni)* <115>, 'the collection of three' *(aṯ-ṯalāṯa al-muǧtamiʿa)*; 'the collection of four' *(al-'arbaʿa al-muǧtamiʿa)*; 'the collection of five' *(al-ḫamsa al-muǧtamiʿa)*; 'the collection of seven' *(as-sabaʿa al-muǧtamiʿa)* respectively. This form is not used with **sakiz**. The rule would demand **saksa·kū**. The *k* is also added in the case of **bira·kū**. Its meaning is the singular *par excellence (al-mufrad bi-ḏātihi)*, as if it were a collective of singularity.

---

[23] The Turkic example expected here is missing. An example that would fit the description is **'uyan·uq** 'awake'.
[24] Em. for 'ilakuw.

11i. A *k* is added after the verb of the past tense, followed by ğā [, hence **k·ğā**]. Between them [i.e. the verb and **k·ğā**] *(mutawassiṭan)* stands any pronoun that is implicitly *(yastakinnu)* attached to the verb. This construction indicates that the *k* is the equivalent *(turādif)* to the meaning 'whenever' *(mahmā)*. The verb is preceded by **tayma** and this together gives the meaning of 'whenever' *(kullamā)*, e.g. **tayma kaldu·k·ğā**[25] **ʾağirlağa·man** ['Whenever I come, I show respect'] and **sanğar tayma kaldu·k·ğā ʾağirlağa·man** ['Whenever Sanğar comes, I show respect'].

11j. A *k* is added together with a silent *l* as a corroborative element in the imperative form of verb in the second person singular *(fiʾl al-ʾamr li-l-muḥāṭab)*, e.g. **kal·ḳil** and **kul·ḳil**.

11k. A *k* is added in **supur·kā**, which is derived from **supur**.

12a. A *ğ* is added exceptionally followed by a *z* in the case of [the particle of] transitivity, e.g. **ṭur·ğuz**.

12b. A silent *ğ* is added together with a *l*, vocalised with an *u* as a reference to something else than an occupation. You say e.g. **rūm·luğ**.

12c. A *ğ*, vocalised with an *a*, is added followed by a silent *n* [, hence **ğan**] as an indication *(dalāla)* of the intensive form of the active participle, e.g. **bar·ağan**. It is also added as an indication of the active participle and the passive participle as relative clauses *(mawṣūlayni)*, e.g. **yarat·ğan** and **yaraldi·ğan**.

12d. A *ğ*, vocalised with an *u*, is added to indicate the meaning of the collective of the number. They say **ʾalta·ğū** and **ṭuqs·ağū** 'the collection of six' *(as-sitta al-muğtamaʿa)* and 'the collection of nine' *(at-tisʿa al-muğtamaʿa)*.

12e. A *ğ* is added together with a silent *l* as a corroborative element in the form of the imperative verb, <VD 41ʳ> in the second person singular, e.g. **ṭur·ğul** and **bar·ğil**.

13a. A *m* is added with a silent *š* as an indication of the active participle in the past tense [, i.e. **miš**].

13b. A *m* is added in the case of **yattī** ['seven'] and **ʾaltī** ['six']; they would say **yatmiš** ['seventy'] and **ʾalṭmiš** ['sixty'] as an indication of the multiplication [of the number][26] by ten.

13c. The addition of the *m* has been treated above in the paragraph about **maq** and **mak** in the case of the verbal noun; it is permitted to reduce these to a silent *m*.

---

[25] ʾAbū Ḥayyān here in detail describes the form given in ĪŪ, i.e. without n (as in VD **kinğā**).
[26] Addition from ĪŪ.

14a. The addition of the *š* has already been dealt with. A silent *š* is added without any other consonant as an indication of the mutual action *(tafāʿul)*, as in **ʾur·uš·dī** and **sanğ·iš·tī**.

14b. A *š* is added to indicate the manner, e.g **min·iš** and **ʾulṭur·uš**
i.e. 'manner of riding' *(hayʾa ar-rukūb)* and 'manner of sitting' *(hayʾa al-quʿūd)*, respectively.

14c. A *š* is added to indicate the intensive form of the verbal noun, e.g. **ʾur·uš** and [**ṭur·uš**].[27]

15a. The addition of the *l* with the *q* in **liq**, and with the *k* in **lik**, has already been discussed.

15b. A *l* vocalised with an *a* is added to the last consonant of the noun if you want to form a verb from that noun, e.g. **ʾutruk·la·dī** ['he lied']; **ʾut·la·dī** ['he advised'] and **suz·la·dī** ['he said']. This is a general rule.

15c. A silent *l* is added as an indication that the verb is formed for the passive form *(bināʾ al-fiʿl li-l-mafʿūl)*, e.g. **ʾur·ul·dī** ['he was beaten'] and **bar·il·dī** [pass. form of 'he went']. And also that it is formed for the passive participle, e.g. **ʾur·ul·miš** ['he was beaten']. And likewise are <116> **bar·il·miš**, **ʾur·ul·dağī** and **bar·il·dağī**. It [also] occurs in a relative participle when the perfect tense is meant, e.g. **bir·il·kān** and **ʾur·ul·ğān** i.e. 'who was given' *(allaḏī ʾuʿṭā)* and 'who was beaten' *(allaḏī ḍuriba)*, respectively.

15d. A *l* is added to the number which is preceded by the indicator of collectivity. The *l* is added vocalised with an *a* [i.e. **la**]when it is preceded by a silent consonant. This only occurs in the case of the annexation *(ʾiḍāfa)*. If you annex to a third person, you attach the *s* to the *l* [i.e. **la·s·ī**]. Thus you say **ʾīka·w·la·s·ī** i.e. 'both of them' *(kilāhumā)* and **ʾuğ·aw·la·s·ī** i.e. 'all three of them' *(ṯalāṯuhum)*. If you annex to a second person, you say **ʾika·w·la·ŋuz** i.e. 'both of you' *(kilākumā)* and **ʾuğ·aw·la·ŋuz** i.e. 'the three of you' *(ṯalāṯatukumā)*. If you annex to a first person, you say **ʾika·w·la·muz** 'both of us' *(kilānā)* and **ʾuğ·aw·la·muz** <VD 41ᵛ> 'the three of us' *(ṯalāṯatunā)* respectively. You use it in this way up to **sakiz**. We shall see whether it is used in the cases of **sakiz** ['eight'] and **ʾun** ['ten'], and whether it is attached to **sakiz** and to **ʾun** with the meaning 'a ten' *(ʿušra)*. [We shall also see] whether this is used *(yuqāl)* in the case of the compound number *(murakkab)*, the related number *(al-maʿṭūf)*, the hundred and the thousand.

---

[27] Taken from İÜ (60ᵛ1) as em. of **ʾulṭuruš** in VD.

16a. A *w* vocalised with an *u* is added if the preceding consonant is vocalised with an *a*. It serves as an indication to the collective of the number. They say 'īk·aw·u 'collective of two', etc.] 'uğ·aw·u, durd·aw·a [sic], baš·aw·u, 'alt·aw·u, yad·aw·u and tuquṣ·aw·u. It

5   is not used in the case of **sakiz** ['eight']; the rule would demand [saks·awū].[28] The *w* is added also in the case of **bir·awu·wu**; its meaning is the perfect singular *(al-munfarid bi-ḏātihi)*. The indication of the collective of the numbers is not used with 'un or with any number higher [than it]. [In those cases] no *w*, *k*, or *ġ* [is added].[29]

10   16b. A *w* is added in the case of **čatla·wu·k** ['hazelnut'], since it is taken from **čatla·dī** i.e. 'he made a cracking sound' *(ṣawwata bi-farqaʿa)*.

20. Chapter on the [Phonological] Substitution

1. The [phonological] substitution *(badal)* is rare in this language and

15   the letters it occurs with are few, i.e. *hamza, q, s, d, ṣ, t, b* and *z*.
2. The *hamza* is substituted by a *b*; they say 'abbāq ['violent white']; its base form is 'ab·'aq. The *hamza* is substituted for a *b* and assimilates with the *b*.
3. The *q* is substituted by a *ġ*; they say 'aġar·dī, its base form is

20   'aqar·dī because 'aq means 'white' *(al-'abyaḍ)*. They say **buz·ġuš**, its base form is **buz·quš** i.e. 'white bird' [lit. 'grey bird'].
4. The *s* is substituted by a *ṣ*, they say 'utuṣiz.[30]
5. The *d* is substituted by a *t*, they say 'išit·tur·dī ['he made listen']; its base form is 'išit·dur·dī.

25   6. [The *ṣ* is substituted by a *š*, they say **bašmaq** ['shoe']; its base form is **baṣmaq**].[31]
7a. The *t* is substituted for a *d*, they say **damur** ['iron']; its base form being **tamur**. They say **dutun** ['tobacco'], the base form of which is **tutun**. They say **diz** ['knee']; its base form is **tiz**. <117> They say **dakā**

30   ['he-goat']; its base form is **takā**. They say **dalū** ['mad']; its base form is **talū**. They say **dilkū** ['fox']; its base form is **tilkū** and as for **diltaq** ['reason'], its base form is <VD 42ʳ> **tiltaq**.
7b. The *t* is substituted by a *ṭ*; they say **ṭuṭ·ṭī** ['he held']; its base form is **ṭuṭ·ṭī**.

---

[28] Instead of VD **saksakuw**, 'Abū Ḥayyān in fact describes the form given here. The final - *u*/-*a* in all forms can be explained as a possessive suffix.
[29] Em. of *ʿayn*.
[30] İÜ (61ᵛ4) is not fully vocalised either: 'ṭṣuz.
[31] Addition from İÜ 61ᵛ4.

9. The *b* is substituted by a *m*. They say **kum·kuk** ['violent blue'], whereas its base form is **kub·kuk**, and they say **yam·yašil** ['violent green']; its base form is **yab·yašil**.

10. The *z* is substituted by a *ṣ*, they say **ṭuqṣan** ['ninety']; its base form is **ṭuqzan**.

21. Chapter on the elision

1. Elision *(ḥaḏf)* is rare in this language and it is not a general rule *(yaṭṭarid)*. The consonants that may be elided are: *d, y, r, hamza* and *q*.

2. The *d* is elided from 'iǧ·**dur** ['make drink']; **kaǧ·dur** ['make pass'] and **qayt·dur** ['to turn back']. They say 'iǧ·**ur**, **kaǧ·ur** and **qayt·ur**, respectively.

3. The *y* is elided from '**aykir** ['to twist']; they say '**akir**.

4. The *r* is elided from **birlā** ['with']; they say **bilā**.

5. The *hamza* is elided from '**uš** which is used to call attention *(at-tanbīh)*. They say **šimdī** whereas its base form is '**imdī**; the *hamza*'s vowel *(ḥaraka)* [in the word] '**imdī** [sc. *i*] is transferred *(nuqilat)* to the silent consonant that precedes it, i.e. the *š* [sc. in '**uš**] and then the *hamza* is elided.

6. The *q* is elided from **ṣiǧqan** ['mouse']; they say **ṣiǧan**.

22. Chapter on the transfer of the vowel

[There is no text here]

23. Chapter on the assimilation

1. Assimilation *(ʾidǧām)* [occurs] both within the word *(kalima)* and between two words. Within one word it is very rare; it is only preserved in '**allī** which means 'fifty' *(al-ḥamsīn)*, and in **yaṣṣī**, which means 'a plate of iron or stone' *(aṣ-ṣafīḥ min ḥadīd wa-ḥaǧar)* and other materials.

V [Syntax]

24. Chapter on the rules of construction

1. Speech is an utterance which indicates a syntactical relation *(qawl dāll ʿalā nisba ʾisnādiyya)*. Its categories are the request *(ṭalab)*, the
5 predicate *(ḫabar)* and the originative sentence *(ʾinšāʾ)*.

1a. The request is a request to do something or to refrain from something. An example of the first is **bar·ġil** 'go!' *(iḏhab)* and an example of the second is **bar·mā** 'don't go!' *(lā taḏhab)*.

1b. An example of the predicate is **ṣanġar ṭur·muš** 'Ṣanġar is
10 standing' *(ṣanġar qāʾim)*.

1c. An example of the originative sentence is **ṣaṭ·ṭim** 'I have sold' *(biʿtu)*. <118>

2. <VD 42ᵛ> Speech *(al-kalām)* is divided into two sentences; a nominal sentence *(ǧumla ismiyya)* and a verbal sentence *(ǧumla
15 fiʿliyya)*.

2a. The nominal [sentence] is a compound of a topic and a predicate. The copulative verbs *(an-nawāsiḥ)* originate from [the nominal sentence]; these are **ʾi·dī** and its sisters; **kaškā** and its sisters; and **ṣaġan·dī** and its two sisters.

20 2b. The verbal sentence consists of a verb and an agent. The sentence that consists of a verb and an object whose agent is not mentioned originates from this.

3. The introduction[32] consists of the division of the noun into the undetermined word and the determined word on the one hand, and
25 the division of the verb into the imperative form, the past tense, the imperfect tense and other forms [on the other].

25. Chapter on the undetermined and the determined word

1. The undetermined word *(nakira)* is a word that is used for a general reference *(mā wuḍiʿa šāyiʿan)*. It is divided into two divisions: a simple
30 word *(mufrad al-lafẓ)* and a compound one *(murakkab)*.

2. The simple word is like **ʾar** ['man'] and **ʾat** ['horse'].

---

[32] Instead of *at-taqaddum*, Caferoğlu proposes *li-nuqaddim*. We prefer reading as *wa-t-taqdīm*.

3. The compound word may either consist of two nouns, like **dawā qušʰ**[33] for 'ostrich' *(naʿām)*, or of more than two, e.g. **qabarǧuq lū baǧā**[34] for 'turtle' *(sulaḥfā)*; as opposed to undetermined words in the language of the Arabs, in which compound words basically do not exist.[35]

4. The determined word is a word that is used for a special reference *(mā wuḍiʿa ḥāṣṣan)*: the personal pronoun *(muḍmar)*, the proper name *(ʿalam)*, the demonstrative pronoun *(ism ʾišāra)*, the relative *(mawṣūl)* and what is annexed to any of these *(muḍāf)*.

5. The personal pronoun occurs independently *(munfaṣil)* and suffixed *(muttaṣil)*.

6. The independent pronoun for the singular of the first person is **man** and the plural of the first person **bīz**. Its base form is **mīz** with a *m*, to which **lar** may be attached to emphasise the plurality. The independent form for the second person singular is **san** and the plural of the second person **sīz**. All of **man, mīz, san** and **sīz** occur both obligatorily *(ʿumdatan)* and optionally *(faḍlatan)*. The independent forms for the third person singular are **ʾul** and **ʾun**, and for the plural of the third person it is **ʾan·lar**.[36] Its base form is **ʾal·lar**, but the [first] *l* is substituted by a *n*. The *hamza* is vocalised with a *a*, because of the *a* of the *l* in **lar**. They do not take the silent *n* into account, since it is separative *(ḥāǧiza)* rather than firm *(ḥaṣīn)* because it carries no vowel. It is pronounced in its base form, when the first *l* is elided; in that case they say **ʾu·lar**. In Oġuz they [usually] say **bu·lar**, substituting the *hamza* [in **ʾu·lar**] by a *b*.

7. The suffixed noun of the first person singular <VD 43ʳ> to which something is annexed, is a silent *m*, e.g. **qul·um** 'my slave' *(ġulāmī)*. For the second person singular to which something is annexed it is a silent post-palatal *n (nūn sākina ḥayšūmiyya)*, e.g. **qul·uŋ** 'your [sg.] slave' *(ġulāmuka)*. [In the case of] the plural of the second person to which something is annexed, it is the same *n*, [i.e. the post-palatal *n*] together with a *z*, like in **qul·uŋuz** 'your [pl.] slave' *(ġulāmukum)*. The

---

[33] Namely **dawā** 'camel' and **quš** 'bird'.

[34] Namely **qabarǧuq** 'blister', **lū** 'owner', and **baǧā** 'toad'. In the author's view **lū** is the equivalent of the Arabic *ḏu*, which is a noun (cf. 3.3. and 43.3.).

[35] Al-Ḥadīṯī (1966:182), quoting from the Cairo Ms (?), in addition has the following line : *wa-laysa fī t-turkiyya ʾadā miṯlu al li-t-tarʿrīf* "Turkic does not have a particle like the the definite particle *al*."

[36] ʾAbū Ḥayyān here clearly describes the form **ʾanlar**. In the printed text, however, it is spelled as **ʾunlar**. This difference may be due to a misreading by the copyist or by Caferoğlu.

way in which these personal pronouns are used will be dealt with—God willing—in numerous chapters to come.

8. The name is what is completely identical with its nominatum *(musammāhu)*. It may be either simple, like **Sunqur** ['falcon'] or a compound of two nouns, like **'Aq·buġā** ['white bull'], or of a noun and a verb, like **'Ay·duġ·dī** [lit.: 'the moon has arisen']. [It may also be] improvised *(murtaġal)*, like **Qalāwūn**, or taken *(manqūl)* either from a noun, like **Lāġīn** ['falcon'], or from a verb, like **Buk·tī** ['he bent'] and **Sanǧ·ar** ['he stabs'].[37]

9. The demonstrative pronoun for a simple noun that is nearby is **bū**, with the meaning of 'this' *(ḏā)*, and for the simple noun that is far away [it is] **'ul**, for 'that' *(ḏāka)*. This expression is identical with that of the personal pronoun with the meaning 'he' *(huwa)*, which has been mentioned previously in the Section on the suffixed pronouns *(al-muḍmarāt)* [see above Section 25.6]. For the nearby plural it is **bu·lār**; and for the far away **'an·lar**. We have already mentioned that **'an·lar** is used for the plural of the third person of the pronouns. It has come to mean both <119> a pronoun with the meaning 'they' *(hum)* and a demonstrative pronoun with the meaning of 'those' *('ulā'ika)*.

10. To the locative demonstrative pronouns belong **bun·dā** meaning 'here' *(hunā)* and **'an·dā** meaning 'there' *(hunāka)*.

11. The relative has three expressions *('alfāz)*, depending on the relative clause, as it [sc. the relative clause] can be either a locative *(ẓarf)* or a sentence *(ǧumla)*.

12. If the relative clause is a locative, the relative is a *k* vocalised with an *i* [sc. ki]. An example of this is **bū 'aw·dā·ki·niŋ qul·ī·dur** 'this is the slave of the one in the house' *(hāḏā ġulāmu l-laḏī fī d-dāri)* and **bū 'aw·dā·ki·lār·niŋ qul·ī·dur** 'this is the slave of the ones in the house' *(hāḏā ġulāmu l-laḏīna fī d-dāri)*.

13. If [the relative clause] is a sentence, it can be either nominal or verbal.

14. <VD 43ᵛ> If it is nominal, the singular relative is **'ul kīm quġa·s·ī kurk·lū·dur 'aw·dā·dur**, 'the one whose master is good is in the house' *(allaḏī sayyiduhu ḥasan fī d-dāri)*; and the plural [form] is

---

[37] These names are probably not randomly chosen. They may refer to contemporaries of Abū Ḥayyān, e.g., Sulṭān Qalāwūn, al-Malik al-Manṣūr, Mamlūk ruler of Turkic origin (r. 678/1279-689/1290); Sulṭān Lāġīn, al-Malik al-Manṣūr Ḥusām ad-Dīn, slave of Qalawūn's (d. 698/1299), Probably Sanǧar ad-Dawādārī (d. 699/1299-1300) cf. Haarmann 1988:97ff.). It is more difficult to determine who the name Sunqur may refer to; Sunqur al-Bursuqī, Qāsim ad-Dawla (d. 507/1113) may be a candidate, although he can hardly be considered a contemporary of 'Abū Ḥayyān's (cf. also Part One, II note 32).

'an·lar kim qul·lar·i[38] kurk·lū·dur 'aw·dā·dur ['those whose slaves
are good are in the house'].

15. If it is verbal, it can be either in the past tense or in the future
tense.

16. If it is in the past tense, the relative is ġan, which is attached to
the last consonant of the verb when it is pronounced velarised. It is
kan when [the last consonant of the verb] is pronounced palatalised.
You say kur·du·m san·ī 'ur·ġan·ī 'I saw the one that hit you [sg.]',
yarat·ġan 'the one that created' *(alladī ḥalaqa)*; and kal·kān 'the one
that came' *(alladī ǧā'a)*. It is permitted to use a verb in the past tense
after 'ul kim ['the one that'] and 'an·lār kim ['the ones that'], e.g.
'an·lār kim ṭur·di·lār 'aw·dā 'the ones that stood up are in the house'
(alladīna qāmū fī d-dāri).

17. If [the relative clause] is in the future tense, the relative is 'ul kim
and 'an·lār kim, which depends on the meaning you intend. The
condition of the relative is that it is either a predicative sentence, a
locative or a word in the genitive case. Examples of this have been
shown previously.

18. When the relative is a *k* vocalised with an *i*, the relative clause can
only be a locative or a word in the genitive case *(maǧrūran)*. It is
necessary to place the relative clause before [ki]. In the case of other
relatives [than ki], the relative clause must follow, whether it is a sen-
tence or something else. One example of a sentence has already been
given. Examples of the genitive particle *(ǧarr)* and a word in the geni-
tive case *(maǧrūr)* are 'an·lar kim 'aw·da·dur kurk·lū lar·dur 'the
ones in the house are good' *(alladīna fī d-dāri ḥisān)* and 'an lar kim
'al·niŋ dā·dur bak·lar dur 'the ones in front of you [sg.] are chiefs'
(alladīna 'amāmaka 'umarā').[39] It is possible to separate *(tafṣīl)* 'ul·kim
and 'an·lar·kim from their relative clauses with the object. For
instance, you say 'ūl kim sangar·nī 'ur·dī kal·dī 'the one that hit
Sangar came' *(alladī ḍaraba sangar ǧā'a)*.

19. The conditional sentence may occur *(taqi')* as a clause to the
relative. Thus you say kal·dī 'ul kim kur·sa·ŋ saw·ka·san 'the one you
[sg.] would love if you saw [him] came' *(ǧā'a lladī 'in <VD 44ʳ> ra'ayta
'aḥbabta)*. Sā is the conditional particle *('adā aš-šarṭ)* and the *ŋ* is the
pronoun of the second person.

20. The annexed *(muḍāf)*. The annexation sometimes occurs with a
personal pronoun, e.g. qul·um, qul·umuz, qul·uŋ, qul·uŋuz qul·ī

---

[38] Em. for qullar.
[39] In these phrases dā is considered the equivalent of *fī*, a particle that governs the
genitive case.

and **qul·lar·ī**. <*120*> This will be dealt with in the Chapter on the an-
nexation [see below Section 51]. The last consonant of a word you
annex to a third person is necessarily either silent, in which case you
vocalise it with an *i*, as in **qul·ī** 'his slave' *(mamlūkuhu)*, or vocalised, in
5   which case a *s*, vocalised with an *i*, is attached to it, as in **quğā·s·ī**,
qarī·s·ī and **'ulwa·s·ī** 'his master' *(sayyiduhu)*, 'his grandfather'
(ʿağūzuhu) and 'his lord' *(ʿaẓīmuhu)*, respectively. Sometimes you an-
nex something to a proper name, like **sunqur qul·ī** ['Sunqur's slave'],
or to something which is demonstrated, like **bu·nuŋ qul·ī** ['this one's
10   slave'], or to a relative: **kim·niŋ qul·ī** ['whose slave']. The manner of
the annexation will be treated later—God willing.
    21. In this language there is no particle to determine the undeter-
mined word as in Arabic; however, when the noun is considered to be
understood by the interlocutors, they confine themselves to mention-
15   ing it as an undetermined noun, trusting the understanding of the in-
terlocutors. They do not use **'ul** before it. The prevalent custom in
their language is that they use the expression for both the personal
pronoun of the third person and the demonstrative pronoun, which is
'ul. [For example,] when there is an agreement about a fish between
20   you and a second person, you say with the meaning 'the fish is moist'
(as-samak ṭarī) **'ul baluq yaš·dur** and with the meaning 'the fishes are
moist' *(al-'asmāk ṭarīya)* you say **'ul baliq·lār yaš·dur**. It is permitted
to say **yaš·lār·dur**.

26. Chapter on the verb

25   1. The verb is divided into three categories: the imperative form, the
past tense and the imperfect tense.
    2. The imperative form is the base form. The past tense, the imper-
fect tense, the active participle, the verbal noun, the action, the noun
of place, and the instrument are all separate branches and they all de-
30   rive from the imperative form. The imperative form is intended for ei-
ther a third person, a second person or a first person.
    3. If the imperative is intended for a third person it must <VD 44ᵛ>
contain the consonants of the imperative, like **sangar kal·sun** 'Sangar
must come' *(sangar li-yağiʾ)* and in the case of the plural of the third
35   person, like **sangar kal·sun·lār** 'they must come' *(li-yağī'ū)*. **Sun**[40] is
the particle of the imperative, the equivalent *(naẓīr)* to the *l* of the im-
perative in Arabic [cf. the preceding examples].

---

[40] In Caferoğlu's edition the final consonant is a *kāf* with three dots. İÜ reads **sun**.

4. If it is used for a second person, it is either singular or not. If it is singular, the most correct use is the imperative of the verb without any addition. But it is permitted to add ġil or ḳil, with a Bedouin ḳ, as a pausal form. If the word is pronounced velarised, you use ġil; if it is

5 pronounced palatalised, you use ḳil. The use of velarisation and palatalisation is determined by what is actually heard. We already analysed this in 'The Book of the Verbs' *(Kitāb al-ʾafʿāl)* that we compiled about this language. If the first consonant of the imperative form is vocalised with an *u*, then the penultimate consonant [sc. ġ

10 resp. ḳ] is vocalised with an *u* too, except if there is a *a* in the verb. Examples of this are ṭur·ġul, kul ḳul, kustar ḳil and ʾur·ġul ['stand up!', 'laugh!', 'show it!' and 'hit!']. If [the final consonant] is vocalised with an *a* or an *i*, [then] the penultimate consonant <121> is vocalised with an *i*, except if there is a *u* in the verb. Examples of this are **bar·ġil**

15 ['go!'] [41] ʾišit·ḳil ['listen'], and **tak·tur·ġil** ['make plant!']. [42]

5. If the imperative is not a singular, you add a single *n*; thus you say ṭur·uŋ. If you wish, you may add a *z*. The *z* reflects a feeling of respect *(taʿẓīm)*, e.g. you say ṭur·uŋ·uz; the *z* is the remainder *(baqiya)* of siz. It is permitted to use siz, you say ṭur·uŋ·siz. Siz is used as a corrobora-

20 tive form.

6. If it is for a first person, it is either singular or not. If it is singular, you say **bar·ġā·y·im** and **kal·kā·y·im** 'I must go' *(li-ʾaḏhaba)* and 'I must come' *(li-ʾaǧiʾa)*. If it is not singular, you say **bar·ġā·lim** and **kal·kā·lim** 'we must go' *(li-naḏhaba)* and 'we must come' *(li-naǧiʾa)*.

25 The imperative of the first person singular is very rare in Arabic but in this language it is very frequent.

7. As for the past tense and the imperfect tense, in the part about the morphology <VD 45ʳ> we have already discussed their special features and their distinctive markers. The consonant that precedes the *r*,

30 which is a marker for the imperfect tense, is either vocalised or silent. If it is vocalised, you add the *r* and the vowel undergoes no change. If it is silent, you vocalise [the consonant before the *r*] with either an *u* or an *a* [, hence **ar/ur**]. The criterion is the actual use, as we have explained in *The Book of the Verbs*.

35 8. This is the case as far as the verbs that do not end in **lā**[43] are con-
cerned, which serves to indicate governance *(ʾiʿmāl)*, since in that case

---

[41] Em. for ġal.

[42] Surprisingly, this front verb has a back suffix. This must be a mistake.

[43] As explained by Abū Ḥayyān, Turkic verbs ending in lā are usually derived from nouns. In Standard Arabic the derivation of verbs from nouns rarely occurs, since most verbs can be related to a verbal root.

the *ʾalif* changes into a *y* vocalised with an *u*. You say **suz·lā·yur** ['he speaks'] and **baš·lā·yur** ['he begins']. It is permitted to elided the *y*. In that case you say **suz·lā·r** and **baš·lā·r**, [although] the base form [includes] the *y*.[44]

5   9. As to the future tense that is used as a predicate,[45] if it is pro-nounced velarised its particle is a *ǧ* vocalised with an *a*, e.g. **tur·ǧā** 'he will stand up' *(sa-yaqūmu)*. If it is pronounced palatalised, its particle is a *k* vocalised with an *a*, like **kal·kā** 'he will come' *(sayaǧīʾ)*. We shall —God willing— discuss the verbs in the Chapter on the Verb and its

10   Agent [see below Section 32] as far as affirmation *(ʾiṭbāt)*, negation *(nafy)*, interrogative *(istifhām)* and prohibition *(nahy)* are concerned.

10. The verbs are also divided into inflected verbs *(mutaṣarrifa)* and non-inflected verbs *(ǧāmida)*.

11. The inflected verb is [the verb] of which the form differs accord-

15   ing to the difference in tense *(zamān)*, like **tur** ['stand up!'], **tur·ur** ['he stands up'], **tur·ǧā** ['he will stand up'] and **tur·dī** ['he stood up'].

12. The non-inflected verb is the verb that has only one form, like **ʾi·dī** with the meaning 'he was' *(kāna)*. Neither the imperative form nor the imperfect tense of this verb is used. If it means 'he sent'

20   *(ʾarsala)*, it is inflected and an imperative form of it may be used. In that case they say **ʾi·bu·nī**[46] 'send this' *(ʾarsal hāḏā)*. So **ʾi·dī**][47] has both the meaning of 'he was' *(kāna)* and the meaning of 'he sent' *(ʾarsala)*. And in the same way **dak̲ul**, with the meaning 'he is not' *(laysa)*, is not inflected in their speech *(kalām)*. This will —God will-

25   ing— be explained in the Chapter on **ʾi·dī** and its Sisters [see below Section 28].

13. [The verbs are] also divided into intransitive and transitive verbs. The intransitive verb, for example, is **tur**. The transitive [verb] is di-vided into two divisions: transitive to <122> a proper object *(mafʿūl*

30   *ṣarīḥ)* and transitive to an improper object *(mafʿūl ǧayr ṣarīḥ)*. The transitive [verb] is [further] divided into transitive to one object, <VD 45ᵛ> transitive to two objects and transitive to three objects; we will deal with this later.

---

[44] See above n. 21.

[45] To be distinguished from the active participle for the future **daǧī**.

[46] Initial **d** has probably been elided from this word, the original form being /id·bu·niy/

[47] Em. for VD **ʾi·d̲ī**.

27. Chapter on the topic and the predicate

1. First you mention the topic *(mubtada²)* and you put the predicate *(ḫabar)* after it. Thus you say **saŋar ṭur·ub·ṭur** 'Sangar is standing' *(sanǧar qāyim)* and **saŋar ṭur·taǧī·dur** 'Sangar will be standing to-
5   morrow' *(sanǧar qāyim ġaddan)*. It is also permitted to say **ṭur·ub·an·ṭur** or **ṭur·miš·dur**. If the predicate is compared with the topic, you say **saŋar ᵓarslān·dur** i.e. 'Sangar is a lion' *(sanǧar ᵓasadun)*.

2. If [the predicate] is a locative of place *(ẓarf makān)* you say **saŋar ᵓaw·da·dur** 'Sangar is in the house' *(sanǧar fī d-dāri)*, and
10   **saŋar ᵓard·iŋ·da·dur** 'Sangar is behind you [sg.]' *(sanǧar ḫalfaka)*. The locative of place can only occur with the particle that is equivalent *(murādif)* to 'in' *(fī)*, in the sense of containing something, namely **da**.

3. If [the predicate] is a locative of time *(ẓarf zamān)*, in some cases the particle [that expresses the sense] of containing something *(ḥarf al-*
15   *wiᶜāᵓ)* does not occur, such as **saŋǧiš bu·kun·dur** 'the fight is today' *(aṭ-ṭaᶜān al-yawma)* and **qunuq·luq bu·dun·dur** 'the visit is tonight' *(aḍ-ḍiyāfa al-layla)*. In other cases, the particle [that expresses the sense] of containing something does occur,[48] like **saŋǧiš ṭaŋ·da·dur** i.e. 'the fight is tomorrow' *(aṭ-ṭaᶜān ġaddan)*. In all these [sentences]
20   **dur** is used as a corroborative element. It is customary to use this cor-roborative element and not to elide it.

4. If the predicate is a sentence, you say **saŋar ᵓata·s·ī ṭur·dī** i.e. 'Sangar's father is standing' *(ᵓabū sanǧar qāyim)* and **saŋar ᵓata·s·ī ṭur·ub·ṭur** 'Sangar's father is standing' *(ᵓabū sanǧar qāyim)*. It is per-
25   mitted to elide the predicate because it is implied by the meaning [of the sentence]. They say **ᵓarmaǧan**, meaning **qanī** i.e. 'the present, where [is it]' *(al-hadīya ᵓayna)*.

5. If the predicate is an interrogative noun, it is permitted in this language to put it either at the beginning or at the end of the sentence,
30   and this is the principle. Thus you say **saŋar qanī** ['where is Sangar?'] and **qanī saŋar**.

6. Similarly, if the topic is a noun, it is permitted to put it before or after the predicate. You say **kim·dur ᵓaw·dā** 'who is in the house' *(man fī d-dāri)* and **ᵓaw·da kim·dur** 'in the house is who' *(fī d-dāri man)*.[49]
35   7. If the topic is a conditional noun *(ism šarṭ)*, it is necessary to put it at the head. Thus you say **kim ṭur·sā ṭur·ǧā·man ᵓa·niŋ bilā** 'whoever will stand up, I'll stand up with him' <VD 46ʳ> *(man yaqum*

---

[48] *Mā yaᵓtī bi-l-ḥarf* em. of *mā lā yaᵓtī bi-l-ḥarf*.
[49] The Arabic sentence is not as ungrammatical as its English translation may seem.

ʾaqum maʿahu). Kim is used both as an interrogative [noun] *(istifhām)* and as a [noun of] condition *(šarṭ)*.

8. If the topic is a relative, you say ʾul kim ṭur·dī ṭur·du·m ʾa·niŋ bilā 'I stood up with the one that stood up' *(allaḏī qāma qumtu ma ʿahu)*. It is not permitted to put this topic at the end. ʾUl kim came to mean 'the one that' *(allaḏī)* and ʾā·lār kim 'the ones that' *(alladīna)*. This has already been treated in [the Chapter on] the rela - tive [see above Section Section 25.11]. The topic of a nominal clause may have more than one predicate. You mention the predicates one after the other and you put dur at the end. <123>

28. Chapter on the copulative verbs

1. [The copulative verbs *(an-nawāsiḫ)*] are ʾi·dī and its sisters. ʾI·dī conveys the meaning 'he was' *(kāna)*, and ḅul·dī means 'he became' *(ṣāra)*; tuna·dī means 'he spent the night' *(bāta)*. Daḳul means 'he is not' *(laysa)*—this is with the Bedouin *k*. For the meaning of nearness in time you use yaz·dī in the sense of 'he was about to' *(kāda)* and baš·la·dī with the meanings 'he started' *(ʾanšǎʾa)* and 'he began' *(ibtadaʾa)*.

2. As for ʾi·dī, you say sangar ʾurū·ṭur·ur ʾi·dī with the meaning 'Sangar was standing' *(sangar kāna qāyiman)*. It is permitted to say ʾurū·ṭur·miš ʾi·dī. The meaning of ʾurū is 'upright' *(muntaṣib)* and ṭur·ur is a verb with the form of the present tense *(fī'l ḥālī)*. The meaning of ṭur·miš is 'standing' *(qāyim)*. It is also permitted to say sangar ṭur·miš ʾi·dī, but in the sentences mentioned here ʾurū is used, since ṭur·miš means both 'he stood' *(waqafa)* and 'he stood up' *(qāma)*. If they want to make clear [that its meaning is] 'he stood up' *(qāma)*, ʾurū is used with ṭur·miš, which is an indication of being upright *(intiṣāb)*. This is how they use ṭuraq·iŋ qanī 'where is your [sg.] place of standing up' *(ʾayna maqāmuka)* and 'your [sg.] standing-place' *(mawqifuka)*. If you mean 'I was standing' *(kuntu qāyiman)*, you say ʾurū·ṭur·ur ʾi·di·m and ʾurū·ṭur·miš ʾi·du·m; 'you [sg.] were standing' *(kunta qāyiman)* ʾi·di·ŋ with the post-palatal *n*; 'they were standing' *(kānū qāyimīn)* ʾurū ṭur·ur·lār ʾi·dī and ʾurū ṭur·miš ʾi·dī 'you [pl.] were standing' *(kuntum qāyimīn)* ʾurū ṭur·miš ʾi·di·ŋiz. The personal pronoun of the second person is attached to the verb [i.e. ʾid·iŋ]; the marker of the plural is not attached to the predicate, except [in the case of] the third person *(bi-ḫilāf al-ġayba)*. The marker of the plural <VD 46ᵛ> is attached to the predicate and not to the verb. The verb remains as it is in the third person singular.

3a. If the predicate is an interrogative noun, you say **saŋar qaydā ʾidī** with the meaning ['where was Saŋar'].[50]

3b. ...the difference was dealt with in what we said about the singular noun. For 'when did you [sg.] become <124a> wise' *(matā ṣirta* ⁵ *ᶜāliman)* you say **qaǧan bul·du·ŋ bilkā** and [for] '[when] did you [sg.] become wise' *([matā] ṣirtum ᶜulamāʾ)* you say **qaǧan bul·du·ŋuz bil·kā·lar**. And it is permitted [to say] **bilkạ·lar qaǧan bul·du·ŋuz** ['When did you [pl.] become wise'] and also what is similar to it. With the meaning 'where did you [sg.] become' *(ʾayna ṣirta)* you say **qaydā** ¹⁰ **bul·du·ŋ**.[51] It is not permitted [to say] **bul·du·ŋ qay·dā**. <125a>

4. With the meaning 'who became the one that was standing up' [you say] *(man ṣāra l-qāʾim)* **kim bul·dī ṭur·ǧān** and it is permitted to say **ṭur·ǧān kim·bul·dī**. For 'who became the ones that were standing up' *(man ṣārū al-qāyimīn)* you say **ṭur·ǧān·lar kim bul·di·lār**. The ¹⁵ marker of the plural is attached to the verb, unlike **ʾi·dī**, to which it is not attached.

5. For 'whoever become the standing one, I will love him' *(man ṣāra l-qāyim ʾuḥibbuhu)* you say **kim bul·sā**[52] **ṭur·ǧān saw·ka·man ʾānī**. **Kim** means 'who' *(man)*, **bul** is the verb and **sā** is the consonant of the ²⁰ particle of condition *(ʾadā aš-šarṭ)*. They do not have a noun that includes the meaning of the particle of condition. When they mean the condition they use the conditional particle, attaching it to the verb, putting it after it. It is preceded by the noun. **Ṭur·ǧān** means 'the standing one' *(al-qāyim)* and **saw·kā** is the verb of the pure future ²⁵ tense *(fiᶜl al-mustaqbal al-ḫāliṣ)*. In this construction the past tense by which the future tense is intended is not used, as it is in Arabic.[53] **Man** is the pronoun of the verb *(ḍamīr al-fiᶜl)* in 'I will love him' and **ʾā·nī** is the pronoun of the direct object *(ḍamīr al-mafᶜūl)* in 'I will love him' *(ʾuḥibbuhu)*.

³⁰ 6. In the case of the plural you say **kim·lar bul·sa·lar ṭur·ǧan·lar saw·ka·man ʾan·lar·nī** ['whoever be the standing ones, I will love them'] and with the meaning 'the people got leaving' *(an-nās ṣārū ḫārigīn)* you say **kiši·lar čiq·miš·lar bul·dī lar**. It is permitted [to say]

---

[50] After comparison with İÜ these paragraphs (which are here numbered 3-14) in VD should be re-arranged as follows: 9, 10, 11, 12, 13, 14, 3a-b, 4, 5, 6, 7, 8. The text in İÜ (Caferoğlu 124,11-125,5), indeed, has this sequence and therefore appears more coherent. The difference between the two texts is probably due to a copyist's misplacement of one folio of the authograph.

[51] In his attempt to adhere to a consistent translation of **bulduŋ**, Abū Ḥayyān may not be right in this case. The meaning of the Turkic sentence is rather 'where were you [sg.]'.

[52] Em. for **bil·sa**.

[53] Em. *al-murā bihi* for *al-murād bihi*.

bul·dī lar čiq·miš·lar kišilar.⁵⁴ 'He became' (*ṣāra*) is mostly used in
connection with generic nouns (*ʾasmāʾ al-ʾaǧnās*) and with attributes
that are used in the same way (*ṣifāt*) but not in combination with un-
stable attributes (*ṣifāt muʿāraḍa*). Analogy determines what is permit-
5   ted.

7. With regard to the use of **tuna·dī**; for 'Sangar spent the night
laughing' (*bāta sangar ḍāḥikan*) you say <VD 47ʳ> **sangar kul·ā
tuna·dī** and for 'where did you [sg.] spend the night' (*ʾayna bitta*) you
say **qan·dā tuna·di·ŋ**. The status (*ḥukm*) of **tuna·dī** is exactly the
10 same as the status of **bul·dī**.

8. **Dakul** is a word that conveys both the meaning 'is not' (*laysa*)
and the meaning of the negation *mā*. If you mean 'Sangar is not leav-
ing' (*laysa sangar ḥāriǧan*), you say **sangar čiq·miš dakul** with this
meaning and for 'people are not leaving' (*an-nās laysū ḥāriǧīn*) you say
15 **kiši·lar čiq·miš dakul·lar**. For 'I am not leaving' (*lastu ḥāriǧan*) you
say **čiq·miš dakul man**. For 'we are not leaving' (*lasnā ḥāriǧīn*) you say
čiq·miš·lar dakul·biz. If **dakul** means 'he is not' (*laysa*), you attach
the marker of the plural which is attached to the verbs. If **dakul** means
'not' (*mā*), it is not attached. For 'you [sg.] are not leaving' (*lasta
20 ḥāriǧan*) you say [...].

9. For 'where was Sangar' (*ʾayna kāna sangar*) you say **sangar qaydā
ʾi·dī** and **qaydā ʾi·di·ŋ** for 'where were you [sg.]' (*ʾayna kunta*). It is
not permitted to put **qaydah** [sic] after the verb. For 'when was the
fight' (*matā kāna t-ṭaʾānu*) you say **sangiš qaǧan ʾi·dī**. The verb may
25 not precede **qaǧan** <126>. Neither **sangar** nor **sangiš** in the two pre-
ceding examples may be put after the verb. Thus you say **qaydā ʾi·dī
sangar** ['where was Sangar'] and **qaǧan ʾi·dī sangiš** ['when was the
fight']. The same applies to those constructions that resemble these
two.

30 10. For 'who was the standing one' (*man kāna l-qāʾim*) you say
ṭur·ǧan kim ʾi·dī. It is permitted to say **kim ʾi·dī ṭur·ǧan**, which
means 'who was the one that stood up' (*man kāna llaḏī qāma*). For
'who were the standing ones' (*man kānū l-qāʾimīn*) you say **ṭur·ǧan·lar
kim ʾi·dī·lar**. It is permitted to say **kim·lar ʾi·dī·lar ṭur·ǧan·lar**. In the
35 case of 'the people were leaving' (*an-nās kānū qāyimīn*) you say
čiq·miš·lar ʾi·dī kiši·lar; it is permitted to say **kiši·lar čiq·miš·lar
ʾi·dī·lar**.

11. If you negate **ʾi·dī** and then affirm it, for 'Sangar was nothing
but standing' (*mā kāna sangar ʾillā qāyiman*) you would say **sangar**

---

⁵⁴ Em. for kiški.

daḵul·dī makar ṭur·miš. The base form is with ʾi·dī, i.e. 'he didn't be' *(laysa kāna)* with the meaning of 'he was not' *(mā kāna)*. It will be shown later that daḵul can be used in two ways. You say for 'Sangar will not be but standing' *(mā yakūnu sangar ʾillā qāyiman)*[55] sangar

5 bul·ma·ġā makar ṭur·miš. <VD 47ᵛ> In this construction 'he does not become' *(mā yaṣīru)* [sc. bul·ma·ġā] is used for 'he is not' *(mā yakūnu)*, because ʾi·dī has no imperfect tense.

12. bul·dī is pronounced with velarisation of the *b*. For 'Sangar became wise' *(ṣāra sangar ʿāliman)* you say sangar bilḵā bul·dī. It is also

10 permitted to say bul·dī bilḵā sangar. bul·dī may be put at the beginning [of the sentence], at the end or in the middle, unlike ʾi·dī, since its predicate does not come at the end [of the sentence], e.g. *ʾi·dī ʾurū ṭurur. ʾI·dī must always be suffixed to the noun, for example kim ʾi·dī ṭur·ġan ['who was the one that stood up'] or to the predicate,

15 e.g. ʾurū ṭurur ʾi·dī ['he was standing upright']. For 'you [sg.] became wise' *(ṣirta ʿāliman)* you say bilḵā bul·du·ŋ and for 'I became wise' *(ṣirtu ʿāliman)* you say bilḵā bul·du·m. It is also permitted to say bul·du·m bilḵā.

13. A distinction is made with respect to a generic noun used as a

20 predicate, like 'judge' *(al-qāḍī)* or 'chief' *(al-ʾamīr)* or 'minister' *(al-wazīr)*. In that case they allow it to be put either at the beginning or at the end. However, if the predicate is an unstable adjective *(ṣifa ʿāriḍa)*, like 'standing' *(al-qāyim)* or 'sitting' *(al-qāʿid)*, in their speech it is almost solely preserved with the adjective preceding the verb. Thus you

25 say ṭur·ur bul·du·ŋ ['you [sg.] were standing up']. The same applies to adjectives like 'black' *(al-ʾaswad)*, 'blue' *(al-ʾazraq)* and 'long' *(aṭ-ṭawīl)* and the like, which may stand at both the beginning and at the end, unlike unstable adjectives.

14. With the meaning 'you [pl.] became wise' *(ṣirtum ʿulamāʾ)* you

30 say bul·du·ŋuz bilḵa·lar. With the meaning 'you [pl.] became standing up' *(ṣirtum qāyimīn)* you say ṭur·ur·lar bul·du·ŋuz.[56] According to...

15. ...čiq·miš daḵul·san ['you [sg.] are not leaving'], and for 'you [pl.] are not leaving' *(lastum ḫārigīn)* čiq·miš·lar daḵul·lar·siz and

35 čiq·miš·lar daḵul·siz and čiq·miš daḵul·siz. [These two sentences] both mean the same, i.e. that the meaning of daḵul can be both 'he is not' *(laysa)* and 'not' *(mā)*. As a rule, neither the predicate nor the preceding noun is put after daḵul. Likewise, the predicate does not precede the noun.

---

[55] This sentence conveys the meaning of the future tense as intended by Abū Ḥayyān.
[56] Here some text is missing.

16. If you wish to affirm the predicate of **daḳul**, you say for ʿSanǧar is only a chief' *(laysa sanǧar ʾillā ʾamīran)*[57] **sanǧar daḳul·dur makar baḳ**. Makar, which has the meaning of ʿexcept' *(ʾillā)*, is taken from the Persian and is used by the Oǧuz. With the meaning <VD 48ʳ> ʿI [will] not come to you [sg.] but tomorrow' *(ʾanā mā ʾātīka ʾillā ǧadd-an)* you say **man kal·ma·s·man sa·ǧa makar taŋ·dā**. <127> Kal is the verb, and **mā** is the particle of negation; the *s* is put there for the negation of the present tense *(al-ḥāl al-manfī)* and **man** means ʿI'. The literal meaning of **kal·ma·s** is ʿcomes not I' *(mā yaǧīʾu ʾanā)*, which is a manner of expressing ʿI do not come' *(mā ʾaǧīʾu)* in Arabic. As far as **ṣaǧā** [ʿto you'] is concerned, its base form is **ṣan·ǧā** from which the *n* was elided because of the frequency of its use; the base form is used less often. Ǧā has the meaning of ʿfor' *(al-lām)* and ʿto' *(ʾilā)*. **Makar** means ʿexcept' *(ʾillā)*, as has been said already, **taŋ** means ʿmorrow' *(al-ǧad)* and **dā** ʿin' *(fī)*.

17. **Yaz·dī** has the meaning of ʿhe was about to' *(kāda)* and **baṣla·dī** has the meaning of ʿhe started [to]' *(ʾanšaʾa)*. Thus for ʿSanǧar was about to stand up' *(sanǧar kāda yaqūmu)* you say **sanǧar ṭur·a yaz·dī** and for ʿSanǧar is about to stand up' *(sanǧar yakādu yaqūmu)* you say **sanǧar ṭur·a yaz·ar**. For ʿSanǧar started to stand up' *(sanǧar ʾanšaʾa yaqūmu)* you say **sanǧar ṭur·ā baṣla·dī**. The predicate in these sentences [sc. **ṭur·ā**], only occurs in the form of the circumstantial expression *(ḥāl)*. It does not occur with the form of the imperfect tense, nor the past tense. Moreover, **ṭur·ā** cannot precede **sanǧar**. The base meaning of **yaz·dī** is ʿhe made an error' *(ʾaḫṭaʾa)*, but it is used for ʿhe was about to' *(kāda)* and it may also have the meaning of ʿhe wrote' *(kataba)*.

29. Chapter on **kaška** and its sisters[58]

1. The meaning of **kaška** is ʿwould that' *(layta)*. It [sc. **kaška**]is synonymous *(murādif)* with it [sc. *layta*], being taken from the Persian. You say ʿwould that Sanǧar [were] laughing' *(layta sanǧar ḍāhikan)* **kaška sanǧar kul·ar·mis·sā yi·dī**. Kaška has the meaning ʿwould that'; **kul·ar** means ʿhe laughs' *(yaḍḥaku)*, [expressing] the meaning ʿlaughing' *(ḍāhik)*. **Mis·sa** is a word that emphasises *(tuʾakkid)* the meaning of **kaška** and it is not permitted to elide it. **Yi·dī** has the

---

[57] Lit.: ʿSanǧar is not but a chief'.
[58] The sisters of **kaška** musst be understood as a references to its Arabic equivalent, i.e. *layta* ʿperhaps' which, along with some other particles govern the noun in the accusative case.

meaning 'he was' *(kāna)*, its *hamza* [sc. in ʾi·dī] was substituted for a *y.* Its meaning is 'would that Sangar laughed' *(layta sangar ḍāhikan).* You say **kaškā sangar ʾaw·da·miš·sa ya·dī** with the meaning 'would that Sangar were at home' *(layta kāna sangar fī d-dāri).* If the predicate of

5 'would that' *(layta)* is a verb, then **miš·sa** is not attached to it. You say with the meaning 'would that Zayd left' *(layta zaydan yaḥrugu)* **kaškā zayd čiq·sā·y·dī** [Lit.: 'would that Zayd had left']. **Sā** conveys the meaning of 'if *(law)*, its literal meaning *(maʿnā)* is <VD 48ᵛ> 'would that Zayd if he was [he] left' *(layta zaydan law kāna ḥaraǧa).*[59]

10 2. **Kim** has the meaning of an interrogative, the conditional noun, and a relative that follows ʾul. [Kim] occurs literally with the meaning 'so that' *(kay)*, the rules of which will be dealt with later, and with the meaning of 'that' vocalised with an *a* *(ʾanna al-maftūḥa).* You say for 'I learned that Sangar was laughing' *(ʿalimtu ʾanna sangar ḍāhik)*[60]

15 **bil·du·m kim sangar kul·ar·dur.** You say with the meaning 'know that your [sg.] master is Sangar' *(ʾiʿlam ʾanna sayyidaka sangar)* **bil·ḳil kim quǧa·ŋ sangar·dur.** In these examples **kim** conveys the meaning of 'that' *(ʾanna)* vocalised with an *a.*

3. As far as the meanings of 'perhaps' *(laʿalla)* and 'as if' *(kaʾanna)*

20 are concerned, there exists no particle in this language that expresses it. This [meaning] is expressed by a verb; [in that case] you say 'may be Sangar comes' *(laʿalla sangar ǧāʾa)* <128> **bul·ǧā kim sangar kal kay.** **bul·ǧā** means 'he becomes' *(yaṣīru)*; **kim** is an affirmative particle to the meaning 'he becomes' *(yaṣīru)* and **kal·kay** means 'he comes'

25 *(yaǧīʾu).*

4. The meaning 'as if' *(kaʾanna)* is expressed with the form *(ṣīga)* **ʾuqša·r** which means 'he looks like' *(yušbih).* Thus you say 'as if Sangar were a lion' *(kaʾanna sangara ʾasad)* **sangar ʾarslān ǧā ʾuqša·r** [lit. 'Sangar resembles a lion']. **Ǧā** has the meaning of the *lām* [i.e. the par-

30 ticle *li*] that conveys [the meaning of] objectivity *(mafʿūliyya).* It is not permitted to put **bul·ǧā** at the end in the preceding example but it is permitted to say **ʾuqša·r sangar ʾarslān·ǧā** and **ʾuqša·r ʾarslān·ǧā sangar.** **ʾUqša·r** remains a verb, even when it means 'as if' *(kaʾanna)* and the same goes for its conjugations *(at-taṣarruf fīhā)*, unlike **bul·ǧā**

35 **kim**; for [this word] includes *(ḍamanat)* the meaning 'may be' *(laʿalla)* and it is not conjugated *(lam yutaṣarraf fīhā).*

5. There are two words *(kalimatayn)* that convey the meaning 'as if' *(kaʾanna)*, [namely] **kībī** and **layin**, which literally mean 'image' *(miṯl).*

---

[59] All these verbs are reflected in the Turkic sentence too, but only **ydī** indicates the past tense.
[60] The Arabic sentence means 'is laughing', rather than 'was laughing'

Thus you say **sangar ʾarslān biki·dur dakī**[61], **sangar ʾarslan daqin dur**[62] [and] **sangar ʾarslan layin·dur** i.e. 'Sangar is like the lion' *(sangar miṯlu l-ʾasadi).* It is permitted to elide **dur** from the two examples. **Kibī** and **daqī** must be placed at the end of the sentence be
5   cause of the [principle] of annexation in the Turkic language *(al-lisān at-turkī).*[63]

30. Chapter on ʾuranla·dī, ṣaġin·dī and bil·dī

1. As for **ʾuranla·dī**, its meaning is 'he thought' *(ẓanna)* and in this language it is semantically connected *(tataʿallaqu)* to two objects like
10   in Arabic, even though in reality the connected element *(mutaʿalliq)* is the relation *(nisba)* [between the two objects]. You say **bay·nī ʾuranla·du·m ṭur·ur** with the meaning 'I thought the chief [to be] standing' *(ẓanantu l-ʾamīra qāʾiman)* and you say **bay·nī bin·ar ʾūranla·du·m** with the meaning 'I thought the chief [to be] riding'
15   *(ẓanantu l-ʾamīra rākiban).* [Nī][64] is attached to the first object and you use the second object in the form *(ṣīġa)* of the imperfect tense. It is permitted <VD 49ʳ> to put **ʾūranla·du·m** after the two [words], and this is the most correct. Only in a few cases is it permitted to put it in the middle. It is rarely put at the beginning. Who speaks thus is known
20   as a novice *(daḫīl)* in this language.

2. As to **ṣaġin·dī**—the Oguz say **ṣan·dī** instead—its meaning is 'he reckoned' *(ḥasiba).* Its principles are the rules *(ḥukm)* of **ʾuranla·dī**.

3. As to **bil·dī**, its meaning is 'he knew' *(ʿalima).* Its rules are the rules of **ṣaġin·dī** and **ʾuranla·dī**. It is also used transitive to one ob
25   ject. <*129*> In that case you say **sangar·nī**[65] **bil·du·m**, i.e. 'I knew Sangar' *(ʿariftu sangara).* **Kim** with the meaning 'that' *(ʾanna)* and the [words that] depend on it take the position of two objects...[66] These verbs necessarily either precede or follow.[67] If they precede, they are connected to the sentence that begins with **kim** with the meaning
30   'that' *(ʾanna).* Thus you say **ʾuranla·r man kim bay ʾaw·da·dur** 'I think that the chief is in the house' *(ʾaẓunnu ʾanna l-ʾamīra fī d-dāri).*

---

[61] Em. of **dibidur**.
[62] ʾAbū Ḥayyān here extend the meaning of **daaqin**; it is not likely to be a copyist's mistake for **layin**.
[63] That is, in the case of **daqī**, but not in the case of **layin**, which is just a suffix.
[64] Addition by the translator; not present in either one of the Mss.
[65] Em. of **sangar bildum**.
[66] The connection between these sentences is grammatically incorrect.
[67] The meaning of this sentence remains unclear because some words are missing.

If they follow, they are transitive to two objects, as we have explained previously.

4. It is permitted to attach the particle of transitivity (ʾadāt an-naql) to bil·dī. In that case it becomes connected (mutaʿalliqa) to three ob-
5   jects. [Thus for 'I informed the chief that the horse [had been] brindled' (ʾaʿlamtu al-amīra al-farasa mulǧaman) you say bī·ḳā bil·dir·du·m ʾāṭ ʾayar la·n·ub·tur].[68] To the first object, [sc. bī] which was the agent before transitivisation, is added the marker for the improper object and that is the one to which [the verb] is
10   [semantically] connected (yutawaṣṣalu) with ʾilā ['to'] in the language of the Arabs (lisān al-ʿarab). As for the second object [sc. ʾāṭ], you omit (tuhmil) any marker, either of the proper or the improper object. For the third[69] object [sc. ʾayarlan·ub·tur] you use the form of the circumstantial expression and because of this the b, which [serves to
15   express] the circumstantial expression, is added to it. The base form of tur is dur, which serves to indicate the corroborative. It is permitted to put either the verb before the three [other words] or the three other words before the verb.

31. Chapter on the verb and the agent

20   1a. The most correct is the overt noun (al-ism aẓ-ẓāhir) preceding the verb. This does not belong to this chapter, but forms [part] of the Chapter on the Topic and the Predicate [see above Section 27]. Thus you say sangar kal·dī ['Sangar came'].

     1b. It is permitted to put [the overt noun] at the end, saying kal·dī
25   sangar. This does not belong to this [that?] chapter. If you construct (ʾasnadta) a verb in the past tense [according] to an overt singular noun that precedes it, you say sangar kal·dī and sangar·lar kal·di·lar for the plural. In the case of the imperfect tense you say sangar kal·ur, for the plural you say sangar·lar kal·ur·lar. In the future tense you say
30   sangar kal·kay. Kal·ā is Oġuz and [in the case of the plural they say][70] kal·kay·lar and in Oġuz kal·a·lar. It is permitted both to elide the y in kal·kay, and to leave it. The pronoun of the third person in all these examples is implicit; the attached lar is the marker of the plural, not a pronoun.

---

[68] Addition from İÜ.
[69] Both manuscripts and Caferoğlu's edition read ṭānī 'second', which must be an error.
[70] Addition from İÜ.

2. If you adapt the verb to a second person in the past tense, you say
kal·di·ŋ i.e. 'you [sg.] came' *(ǧi'ta)* and **kal·di·ŋiz** i.e. 'you [pl.] came'
(ǧi'tum). It is permitted to say **kal·di·ŋiz·lar**, with the marker of the
plural for reason of esteem *(tafḫīman)* and respect *(ta'ẓīman)*. In the
5   case of the present tense *(ḥāl)* <VD 49ᵛ> [you say] **kal·ur·san** i.e. 'you
[sg.] come' *(taǧī'u)* and in the case of the plural **kal·ur·siz**. It is per-
mitted to say **kal·ur·siz·lar**. In the future tense [you say] **kal·kay·san**
and **kal·ka·san**. In Oġuz [you say] **kal·ā·sin**. In the plural [you say]
kal·kay·siz. It is permitted to say **kal·kay·siz·lar**. <*130*> The silent *n*
10   [sc. ŋ] and **ŋiz** are two suffixed pronouns, whereas **san** and **siz** are two
independent pronouns, considering the fact that they both occur as a
topic.

3. If you construct *('asnadta)* the verb in the past tense with a first
person you say **kal·du·m** and **kal·du·man**. In the case of the plural
15   [you say] **kal·du·k**, **kal·du·k·lar**, **kal·du·biz** and **kal·du·miz** —with
the *b* substituted *('ibdāl)* for a *m*. In the case of the present tense *(ḥāl)*
you say **kal·ur·man**, and **kal·ur·um** in Oġuz. In the case of the plural
[you say] either **kal·ur·biz** or **miz**. It is permitted to attach **lar** to both
of them. In the case of the future tense you say **kal·ka·man**, and
20   **kal·am** in Oġuz. Its base form is **kal·ā·man**, but the *n* is elided and the
m becomes silent. [Thus] [71] two silent consonants accumulate: namely
the *'alif* and the *m*. The *'alif* is elided because of the accumulation of
two silent consonants. [72] In the case of the plural [you say] **kāl·ka·biz**
and **miz** according to [the rules of that language] and in Oġuz [you say]
25   **kalā·wuz**, by substituting the *b* for a *w* vocalised with an *u*. **man** and
miz are two independent pronouns, it is obvious that they both occur
as topics.

4. As far as the construction of an imperative verb [to a person] is
concerned, it was already discussed in the Chapter on the Verb [see
30   above Section 26]. To the meanings that are attached to the verb be-
long the negation the interrogative, the prohibition. All of this is sub-
ject to rules that we—God willing—shall discuss.

32. Chapter on the negation in the verb

1. [The particle of the negation in the verb] is a **m** vocalised with an *a*,
35   [placed] after the verbal root *(nafs al-fiʿl)*. In the past tense you say

---

[71] Addition from İÜ.
[72] The argumentation can be schematically summarised as follows: **kal·ā·man**: base
form /kal·a"man/ elision of an → /kal·a"m/ elision of *'alif* → /kal·am/.

sangar kal·ma·dī ['Sangar did not come'] [and in the plural][73]
sangar·lar kal·ma·dī·lar ['the Sangars did not come']. In the present
tense *(ḥāl)* you say sangar kal·ma·s ['Sangar doesn't come'] and
kal·ma·s·lar in the plural. In the future tense [you say] kal·ma·kay
5   and in the plural you say sangar kal·ma·kay·lar ['the Sangars will not
come']. This is the case as far as the third person is concerned.
     2. As to the second person, in the past tense of the singular you say
kal·ma·di·ŋ and kal·ma·di·ŋiz of the plural. In the present tense
(ḥāl) [you say] kal·mas·san and in the plural kal·mas·siz·lar
10   kal·mas·lar·siz. In the future tense [you say] kal·ma·kay·san. In
Oğuz [you say] kal·mi·yā·sin. Its base form in their language is in the
affirmative form: kal·a·san. Before the attachment of san, the base
form in [their language] was kal·ā. When they attach san to the verb
they elide the *'alif* [sc. *'ā*]. When they negate [the verb], they insert the
15   particle [sc. m ā] between the last consonant of the verb and
[between][74] the particle of the future tense, which is the *'alif*, as if the
'alif were the base form instead of a *y*[75], because its base form is
kal·kay. The k was elided in the language of the Oğuz. The y was
changed into an *'alif*, that they elided in the affirmative form. They put
20   the *'alif* to its [sc. the verb's] base form, instead of the y, so that the
negative form and the affirmative form are different. They vocalise it
with an *a* because of the sequence of the vowels *(li-tawālī al-*
ḥarakāt).[76] In the plural [you say] kal·ma·kay·siz.
     3. As for <131> the first person, you say kal·ma·du·m and in the
25   case of the plural biz kal·ma·du·k <V D 50[r]> and biz·lar
kal·ma·du·k·lar. In the present tense *(ḥāl)* [you say] kal·ma·s·man
and in the plural kal·ma·s biz. In the future tense [you say]
kal·ma·kay man and in the plural kal·ma·kay·biz.

33. Chapter on the interrogative

30   1. The marker *('adā)* of the interrogative may either be a particle or a
noun.
     2. If it is a particle, the interrogative may either refer to a verbal sen-
tence or a nominal sentence.

---

[73] Addition from İÜ.
[74] Addition from İÜ.
[75] *wa-ka'anna al-'aṣl ['alif] badalan 'an al-yā'*: em. of *wa-kānat al-'aṣl ['alif]...* in
Caferoğlu's edition makes little sense. *'Alif* is an addition from İÜ.
[76] In Arabic linguistic argumentation *'alif* is always understood as a consonant. The
argumentation deals with the affirmative form, rather than the negation. It can be sum-
marised as follows: basic form /kalkay/ → /kalay/ → /kala"/.

3. If it refers to a verbal sentence, and this is what is meant in this chapter, the verb can be in the third person, the second person or the first person.

4. If it is a verb in the third person, the particle is always put after the
5 verb, regardless of whether it is in the past tense, the imperfect tense or the future tense and whether the verb is adapted to a singular or a plural. For 'Sangar, did he come' *(sangar hal ǧāᵓa)* you say **sangar kal·di·mū**, with the *m* vocalised with an *u* or an *i*. The base form is *u*, and *i* is used for vowel harmony *(li-l-ᵓitbāᶜ)*. You say **sangar·lar**
10 **kal·di·lar·mū** ['did the Sangars come'] [and for 'Sangar, does he come' *(sangar hal yaǧīᵓu)* you say **sangar kal·ir·mū**].[77] In the case of the future tense, you say **kal·ka·mū** and **sangar·lar kal·ka lar·mū**.

5. If it is [a verb in the] second person, you say **san kal·di·ŋ·mū** ['did you come'] in the past tense, with a post-palatal *n*. The *n* does
15 not assimilate with the particle of interrogation. In the plural [you say] **kal·di·ŋiz·mū**; in the present tense *(ḥāl)* **kal·ur·mu·san** and in the plural **kal·ur·mu·suz**. In the future tense [you say] **kal·ka·mu·san** and in the plural **kal·ka·mu·suz**.

6. If it is [a verb in the] first person, in the past tense you say
20 **kal·di·m·mū**; the *m* in **kal·di·m** assimilates with the particle of interrogation. In the plural [you say] **kal·du·k·mū**; and in the present tense *(ḥāl)* [you say] **kal·ur·mu·man** and **kal·ur·mu·niz** in the plural.[78] In the future tense [you say] **kal·ka·mu·man** and **kal·ka·mu·niz** in the plural.[79]

25 7. If you ask a question about a negative verb and all [forms] of the imperative, the negation or the prohibition [...].[80] If the interrogative[81] refers to an object, in this language the principle is to put the object at the beginning, but it is permitted to put it at the end, unlike the Arabic language as far as the interrogation is concerned, since it is not
30 permitted to put the object at the end. <VD 50ᵛ>

8. But as for the imperative and the prohibition, in their case the base rule is to put [the object] at the end, but is permitted to put it at the beginning.

9. If [the interrogative] refers to a nominal sentence, the particle is
35 attached to its last [word]. Thus you say 'is the sun rising' *(ᵓa š-šams ṭāliᶜa)* **kun duġ·muš mi·dur**.

---

[77] Additon from İÜ.
[78] Prob. err. for **kal·ur·mu·biz** or the like.
[79] Prob. err. for **kal·ka·mu·biz** or the like.
[80] This passage is difficult to interpret; some words are missing.
[81] *fa-l-istifhām* em. for *wa-l-istifhām*.

10. If the marker is a noun, it can either refer to a singular or a plural [form]. If it refers to a singular, you say **qay·dā·dur sangar** for 'where is Sangar' *('ayna sangar)*, **qagan·dur ṣangiš** for 'when is the fight' *(matā al-qitāl)* and **kim·dur sangar** [for] 'who is Sangar' *(man sangar)* and **nā·dur bū** [for] 'what is this' *('ayyu šay' hāḏā)*. <132> It is permitted to put the interrogative noun after the singular [word].

11. If [the interrogative] refers to a sentence, [it is] either a verbal or a nominal sentence.

12. If [the interrogative] refers to a verbal sentence, you say **qagan ya·di·ŋ** for 'when did you [sg.] eat' *(matā 'akalta)*; and **qay·dā ya·di·ŋ** for 'where did you [sg.] eat' *('ayna 'akalta)*; and for **nā ya·di·ŋ** 'what did you [sg.] eat' *(mā 'akalta)*; and **kim·nī 'ur·du·ŋ** for 'whom did you [sg.] beat' *(man ḏarabta)*. It is obligatory to put the interrogative noun at the beginning. **Nā** is more commonly used than **kim**, because **nā** is being used for rational and nonrational entities alike, and **kim** is especially [used] *(muḫtaṣṣ)* for rational entities.

13. If [the interrogative] refers to a nominal sentence *(ǧumla ismiyya)*, you say **qagan kun·duǧ·miš dur** for 'when is the sun rising' *(matā š-šams ṭāliʿa)*, and **qay·dā kun·duǧ·miš dur** [for] 'where is the sun rising' *('ayna š-šams ṭāliʿa)*. After **nā**, nor after **kim**, a nominal sentence can follow.

14. One the nouns of interrogation is **naǧā**, which is a question about a number with the meaning 'how many' *(kam)* in Arabic. When [the speaker] means 'how many did you [sg.] beat' *(kam ḏarabta)*, he says **naǧā 'ur·du·ŋ** [in which] **naǧā** must precede. For 'how many are your [sg.] slaves' *(kam mamālīkuka)*, he says **qul·lar·iŋ naǧā·dur**.[82] In this example **naǧā·dur qul·lar·iŋ** is permitted. If you use a specification *(tamyīz)*, it is put in the middle. Thus you say **naǧā kiši 'ur·du·ŋ** for 'how many men did you [sg.] beat' *(kam raǧulan ḏarabta)* and **naǧā kiši·lar 'ur·du·ŋ** [for] 'how many men did you [sg.] beat' *(kam riǧālan ḏarabta)*. It is specified both in the singular and in the plural.

15. There is also a predicative *(ḫabariyya)* **naǧā** that conveys the meaning of the valuation *(istikṯār)*. This can only be understood in the context *(qarīna)* of their speech: **sulṭān naǧā san·iŋ·kibī kiši·lar·ḳā 'alṭun bir·ma·dī**. <VD 51ʳ> Its literal meaning *(murādif)* is 'the sulṭān how many like you [sg.] people gold gave not' *(as-sulṭān kam miṯluka 'unāsan aḏ-ḏahab mā 'aʿṭā)*. **Naǧā** has the meaning of 'many' *(kaṯīr)*; **san·iŋ·kibī** [has the meaning of] 'like you [sg.]', **kiši·lar** [means] 'people', and **ḳā** is the marker of objectivity *(ʿalāma al-mafʿūliyya)*.

---

[82] Em. for **qulllarin**.

Bir·ma·dī has the meaning 'did not give it' *(mā ʾaʿṭāhu)*, and although it looks *(ṣūra)* like a negation, its meaning is affirmative *(ʾiṯbāt)*. The meaning is 'The sultan has given gold to many [people] like you [sg.]' *(kaṯīrun miṯluka as-sulṭānu ʾaʿṭāhu aḏ-ḏahaba).*

5   16. **Qač** has the same meaning as the interrogative **naǧā**, it is used in questions, and rules are identical to those of **naǧā**, except for the fact that it requires mention of the specification. You say **qač kaz ʾur·du·ŋ** 'how many times did you [sg.] beat' *(kam marra ḍarabta)*, when you pose a question about a verb, e.g., **qač qul·uŋ bar** 'how many slaves are

10   to you [sg.] existent' *(kam mamlūk laka mawǧūd)*. And [as for] **bar**, its meaning is 'existing' *(mawǧūd)*. When **qač** precedes the noun, you put the particle **bar** after the noun. If you put [**qač**] at the end [of the sentence], you say **qul·uŋ qaǧ·ṭur**, without using the particle *(lafẓa)* **bar**; **ṭur** is used as a corroborative [form].

15   17. One of the interrogative nouns is **naǧuk**, which is used to pose a question about the manner *(kayfiyya)* with the meaning of 'how' *(kayfa)*. Thus you say **naǧuk ṭur·du·ŋ** for 'how did you [sg.] stand up' *(kayfa qumta)*, and **sanǧar naǧuk ṭur** for 'how is Sanǧar' *(kayfa sanǧar)*. It is obligatory to have it precede the verb, not the noun; e.g.

20   **sanǧar naǧuk ṭur·miš dur** 'how is Sanǧar standing' *(kayfa sanǧar qāʾim)*. In the language of the Oġuz it **natā** is used with the meaning of 'how'. Its rules are the same as those of **naǧuk**.

       18. You say **naluk** for 'why' *(lima)*; in Oġuz [they say] **nišā** and **nā ʾuġun** for 'what for' *(li-ʾayyi šayʾin)* and **nā·dan** for 'from what'

25   *(mimmā)*. You say <133> **naluk ṭur·du·ŋ** ['What for did you [sg.] stand up'], **nišā ṭur·du·ŋ**, **nā ʾuġun ṭur·du·ŋ** and **na·dan ṭur·du·ŋ**.

       19. You say **nā kurku·lū kiši dur** for 'Sanǧar, what a good man he is' *(sanǧar mā ʾaḥsanahu raǧulan)*. **Nā** has the meaning of 'what' *(mā)* and 'which' *(ʾayyu)*. **Kurku·lū** [means] 'good' *(ḥasan)*, **kiši** [means]

30   'man' *(ʾinsān)*, whereas **dur** is a corroborative element. And [in the case of] 'what an ugly man he is' *(mā ʾaqbaḥahu raǧulan)* [you say] **nā yaman ʾar·dur**. This is not the exact equivalent of what we said, but it conveys its meaning. It is equivalent to the Arabic expressions: 'what a beautiful man Zayd is' *(ʾayyu ḥasanin zaydun)* and 'what an ugly man

35   ʿAmr is' *(ʾayyu qabīḥ ʿamr)*. In reality they do not intend a question but they use [this] to express surprise *(maḥraǧ at-taʿaǧǧub)*, as in their expression 'what a man Zayd is' *(ʾayyu raǧulin zaydun)*, or 'he is re-spectable' *(huwa ʿaẓīm)*, e.g. **sanǧar nā ʾuzun ʾar·dur** [means] 'Sanǧar, what a tall man he is' *(sanǧar ʾayyu ṭawīlu raǧulin)*. <VD 52ʳ>

40   When you mean 'which [thing] is your [sg.] letter' *(ʾayyu šayʾin kitābuka)*, you say **nā nasā na·dur biti·ŋ**. **Nā** means 'which' *(ʾayy)*,

nasā 'thing' (šayʾ); the second nā is a corroborative element and so is
dur, whereas biti·ŋ means 'your [sg.] letter' (kitābuka). It is permitted
to put biti·ŋ at the beginning and it is also permitted to say nā nasā
na·dur bū [lit.: 'what thing, what is this']. It is also permitted to put
5   bū, which has the meaning of 'that' (ḏā), at the beginning.

34. Chapter on the prohibition

1. [The particle of the prohibition] is a m vocalised with an a, like the
particle of negation. But the depending verb has the future tense, un-
like the negation, which is inserted into the past tense, the future tense
10   and the present tense (ḥāl). If you construct a prohibition for a verb
that is constructed for a third person, you say sanğar kal·ma·sun for,
'Sanğar must not come' (lā yağiʾ sanğar) and sanğar·lar kal·ma·
sun·lar in the plural.
    2. And for a verb that is constructed for a second person, you say
15   san kal·mā for 'do not come' (lā tağiʾ). Baylik has [already] dealt with
the meaning of this. It is permitted [to say] bar·mā·mā 'do not go' (lā
taruḥ) [and] ʾulṭur·mā·mā 'do not sit down' (lā taqʿud), [and]
yaṭ·mā·mā 'do not sleep' (lā tanam). This has a contemptuous con-
notation (iḥtiqār), with the meaning of 'woe unto you' (waylaka). In
20   the plural [you say] siz kal·ma·ŋiz, and it is permitted [to say] siz
kal·ma·ŋiz·siz.
    3. For a verb that is constructed for a first person, you say
kal·ma·kā·y·im for 'I will not come' (lā ʾağiʾ) and kal·ma·kā·lim for
'let us not come' (lā nağiʾ).

25   ### 35. Chapter on the replacing agent

1. For the construction of the verb in the passive participle (an-nāʾib
ʿan al-fāʿil) one uses either a silent l or a silent n, according to the
following explanation. [As a premisse] we say that the verb is either
uniradical or biradical or something else.
30   2a. If [the verb] is uniradical or biradical, and its second consonant
is either vocalised or a silent l, or if [the verb ends in] lā, which serves
governance (ʿamal), [in all these cases] the addition (al-mazīd) is a
silent n.
    2b. [Past tense] Thus if you construct ya·dī i.e. 'he ate' (ʾakala) for
35   the passive form, you say ya·n·dī i.e. 'he was eaten' (ʾukila); yu·dī 'he
washed' (ğasala)—yu·n·dī; si·dī 'he broke' (kasara) — si·n·dī 'he was
broken' (kusira); ṣina·dī 'he tested' (ğarraba)—ṣina·n·dī 'he was test-

ed' (ǧurriba); **duša·dī** <134> 'he spread out' (faraša)—**duša·n·dī** i.e.
'he was spread out' (furiša); **dapala·dī** ['he killed']—**dapala·n·dī** 'he
was killed' (qutila).[83]

3. [Present tense] For 'he is being eaten' (yuʾkal) you say **ya·n·ur**;
5   'he is being washed' (yuġsalu)—**yu·n·ur**; 'he is being broken'
(yukassaru)—**si·n·ur**; 'he is being tested' (yuġarrabu)—**sina·n·ur**; 'he
is being killed' (yuqtalu)—**dapala·n·ur**; <VD 52ʳ> 'he must be eaten'
(li-yuʾkal)—**ya·n·sun**; 'he must be washed' (li-yuġsal)—**yu·n·sun**; 'he
must be broken' (li-yukassar)—**si·n·sun**; 'he must be tested' (la-
10 yuġarrab)—**sina·n·sun**. In the case of '**al·dī** meaning 'he took'
('aḫaḏa) you say '**al·in·dī** i.e. 'he was taken' ('uḫiḏa); **ṣal·dī** 'he sent'
('arsala)—**ṣal·in·dī** i.e. 'he was sent' ('ursala).

4. But if the verb is biradical with a silent second consonant —not
an *l*— or something else [?], you add an *l*. Thus in the case of '**ur·dī**,
15 with the meaning of 'he beat' (ḍaraba); if you construct it for the pas-
sive form, you say '**ur·ul·dī** 'he was beaten' (ḍuriba); 'he is being beat-
en' (yuḍrabu)—'**ur·ul·ur**; for 'he must be beaten' (la-yuḍrab) you say
'**ur·ul·sun**; and in the case of 'he will be beaten' (sa-yuḍrabu) you say
'**ur·ul·ǧā**. The addition of the *l* and the *n* is a general rule in these
20 cases, no matter whether it concerns the past tense, the imperfect
tense or the imperative.

5a. If the agent is not mentioned, one of five things may take its
place (qāma maqāmahu) [sc. in the sequence of the sentence, not
functionally]:

25 5b. The object. For 'the meal was eaten' ('ukila aṭ-ṭaʿām) you say
'**āš ya·n·dī** or **ya·n·dī '**āš.

5c. The verbal noun: you say e.g. 'the beat was beaten' (ḍuriba aḍ-
ḍarabu) '**ur·maq '**ur·ul·dī.

5d. The locative of time, **bu kun '**ur·ul·dī ['he was beaten today'].

30 5e. The locative of place, '**ard·iŋ dā '**ur·ul·dī i.e. 'he was beaten be-
hind you [sg.]' (ḫalafaka ḍuriba) and '**aln·iŋ dā '**ur·ul·dī i.e. 'he was
beaten in front of you [sg.]' (qaddāmaka ḍuriba).

5f. The genitive particle and the word in the genitive case: you say
e.g. 'it was taken from Sangar' ('uḫiḏa min sangar) **sangar dan**
35 '**al·in·dī**.

6. In case the verb has two objects, the verb may belong to those
whose first object is not identical with the second. In that case you say
bir·il·dī sangar·ǧā bir·aqǧā for 'Sangar was given a dirham' ('uʿṭiya
sangaru dirhaman). [The verb] is transitive (taʿaddā) to the first object

---

[83] Here 'Abū Ḥayyān does not mention a biradical stem ending in l; for the present
tense he gives '**al** 'to take'.

with the particle of annexation, which is **ġā**, and the verb dominates
(tasallaṭa ʿalā)[84] the second object. Wherever you put it, at the be-
ginning or at the end, it is always correct. Neither **ṣaġan·dī** nor
bil·dur·dī are actually construed in the passive.

5   VI CHAPTER ON THE VERB'S REQUIREMENTS
            WITH REGARD TO OPTIONAL PARTS OF THE SENTENCE

1. These are the verbal noun, the locative of time, the locative of place,
the circumstantial expression, the direct object, the object of reason,
the concomitant object and the excepted [elements]. And as to the
10   specification, its rules will be treated after the discussion of the excep-
tion.

36. Chapter on the verbal noun

1. You say **yiyiš**[85] **ya·du·m** for 'I ate food' *(ʾakaltu aklan)*. They have
the verbal noun *(maṣdar)* preceding the verb and they say < VD 52ᵛ>
15   **ʾuruš ʾur·du·m, ʾur·maq ʾur·du·m, ṭur·uš ṭur·du·m**. They also use
ṭur·maq [instead of **ṭur·uš**]. The **š** indicates the meaning of intensity
(mubālaġa) <135> to the verbal noun. This has been treated in the
part on morphology *(taṣrīf)*.
    2. The verbal noun is divided into two categories: unspecified
20   *(mubham)*, which has been discussed, and specified *(muḫtaṣṣ)*.
    3. It is either specified by an adjective, you say e.g. 'I beat a harsh
beat' *(ḍarabtu ḍarban šadīdan)* **qaṭī ʾur·maq ʾur·du·m**, or by annexa-
tion, e.g. 'I beat Sanġar's beat' *(ḍarabtu ḍarba sanġar)* **ʾur·du·m
sanġar ʾur·uš·i·n**. It is permitted to use **ʾur·maq**. But as to the ending
25   in a *n (tanwīn)*, which is attached here after the *i* of the annexation, its
base form is **ʾur·uš·i·nī**. **Nī** is the marker of the accusative *(naṣb)* and
the *i* [sc. after **š**] [indicates] the annexation to the personal pronoun of
the third person. Then you elide the *y* and leave the silent *n*. It is
permitted to pronounce *(nuṭq)* the base form, so you say **ʾur·uš·i·nī**

30   37. Chapter on the locative of time

1. The locative of time *(ẓarf az-zamān)* consists of two categories
(qismān): unspecified *(mubham)* and specified *(muḫtaṣṣ)*.

---

[84] From IÜ em. of *yusalliṭu.*
[85] Em. for biyiš.

2. Examples of the unspecified locative are e.g. **kunduz** i.e. 'in day-time' *(nahāran)*; **tun·lā** i.e. 'at night' *(laylan)*; **taŋ·lā** or **taŋ·dā** i.e. 'in the morning' *(ṣabāḥan)*; **kaǧä** i.e. 'in the evening' *(masāʾan)*; **qušluq** i.e. 'in the forenoon' *(ḍuḥan)*; **kindin** i.e. 'in the late afternoon' *(ʿaṣran)*; **duštä** i.e. 'at noon' *(ẓuhran)*; in Oġuz they say **ʾuylā** 'at noon' *(ẓuhran)*.

3. The specified locative of time is the word to which any one of these locatives is annexed, or which is described [with an adjective]. The verb dominates *(yatasallaṭu ʿalā)* these locatives, whether it is transitive or intransitive. You say **tun·lā ṭur·du·m** ['I stood at night'] in the case of the intransitive verb and **tun·lā ʾur·du·m sanǧar·nī** ['I beat Sanǧar at night'] in the transitive verb. It is permitted that the verb precedes the locative or that it comes after it, whether the verb is conjugated *(mutaṣarrif)*, as in our example, or not conjugated. Thus you say **bū kun sanǧar ṭur·miš daḵul** ['today Sanǧar is not standing'] and **sanǧar ṭur·miš daḵul bū·kun**.

38. Chapter on the locative of place

1. To the [locative of place] *(ẓarf al-makān)* belong **ʾust·un·dā** i.e. 'above' *(fawqa)*; **ʾalt·in·dā** i.e. 'underneath' *(taḥta)*; **ʾuŋ·in·dā** i.e. 'at the right' *(yamīnan)* and **ṣuŋ·in·dā** i.e. 'at the left' *(šamālan)*; **ʾaln·in·dā** i.e. 'in front of' *(ʾamāman)*; **ʾard·in·dā** <VD 53ʳ> i.e. 'behind' *(ḫalafan)*. And also **qarši·ŋ·dā** i.e. 'in front of you [sg.]' *(tilqāʾaka)* and **yan·iŋ·dā** i.e. 'next to you [sg.]' *(ḥiḍāʾaka)*. Some of these locatives are used as an annexed element only [sc. annexed to a pronoun], therefore this is exemplified with the *n* that reflects the pronominal form of the second person.

2. The base form of **ʾust·uŋ·dā** is **ʾustun·iŋ·dā**, because the first *n* is a basic consonant of the word, the second *n* is the *n* of the second person to which it is annexed. It is permitted to <136> elide one of the two *n*s in order to alleviate *(taḫfīfan)*. It is clear that the elided one is the first from the root of the word *(nafs al-kalima)*, because the second *n* is inserted to indicate a meaning. Their utterance **ʾust·um·dā**, 'on me' *(ʿalayya)*, shows that it is definitely the first *n* which has been elided. They elide the first *n* because the *m* indicates the first person. They never say *****ʾustun·um·dā** with both the *n* [and the *m*]. After he had spoken about the permanence of the *n*, the compiler of this book said '*al-mawlā Tāǧ ad-Dīn* informed me of this'.

3. The locative of place is only used with the particle that is equivalent to 'in' *(fī)* [sc. **dā**], as long as it is a locative. It is permitted to use

another particle *(ḥarf ǧarr)* than **dā**, although in that case it stops being a locative. [For example,] you would say **'ard·iŋ dan yat·tī** i.e. 'he caught up from behind you [sg.]' *('adraka min ḥalfika)* —'**ard·iŋ** may be replaced by **'uŋ·uŋ** meaning 'before' *(qabla)* or by **ṣuŋ·uŋ** with the

5 meaning of 'after' *(ba'da)*. In these two cases the annexation is not obligatory because they express a locative of time rather than a locative of place. Thus you say **ṭur·du·m 'uŋ·dā** 'I stood up before' *(qumtu qablu)*. **Burun** too means 'before' *(qablu)*. It is permitted to say **ṣuŋ·rā ṭur·du·m**, with a *r*, or **ṣuŋ·dā** with a *d*, e.g. 'I stood up later'

10 *(qumtu ba'dan)*. The *r* and the *l* in the examples **tun·lā** ['at night'], **taŋ·lā** ['in the morning'] and **ṣuŋ·rā** ['later'], respectively, both ex- press the meaning of the *d* [sc. the particle **dā**], which, [in its turn] corresponds to 'in' *(fī)* in Arabic.

4. If you wish to annex [to a pronoun], as in the expression 'I came

15 after you [sg.]' *(ǧi'tu ba'daka)*, you say **kal·di·m san·dan ṣuŋ·rā**. This means literally 'I came from you [sg.] in later' *(ǧi'tu minka fī ba'din)*. For the example 'I came before you [sg.]' *(ǧi'tu qablaka)* you may say either **kal·di·m san·dan burun** or **kal·di·m san·dan 'ilk·arū**[86] or **kal·di·m san·dan 'uydin**. **'Ilk** means 'first' *('awwal)*. The *r*, in **'ilk·**

20 **arū**, which is vocalised with an *u*, has the meaning of **dā**, which is an equivalent of 'in' *(fī)*. <VD 53ᵛ>

5. On these locatives may depend both intransitive and transitive verbs and conjugated and non-conjugated verbs, like on the locative of time *(ḥāl ẓarfu z-zamān)*. Thus you say **'ard·iŋ·dā sanǧar ṭur·miš**

25 **daḵul** ['Sanǧar is not standing behind you [sg.]'] and **sanǧar ṭur·miš daḵul 'ard·iŋ·dā**.

6. The locatives of place only occur especially with the annexation. The *n*, which has already been mentioned, indicates the second person if it is nasalised [sc. *ŋ*], whereas [it indicates] the annexation if it is

30 pure *(ḥāliṣa)* [sc. *n*]. These two can never be combined; they never say *'**ard·iniŋ·dā** with one of the two *n*s being a nasalised *n* and the other a pure *n*. [The reason is that] the pure *n* expresses [annexation to] the third person and the nasalised *n* [annexation to] the second person.

7. If you annex to a first person, you use either the personal pro-

35 noun of the first person singular or plural, as you do in other combina- tions *('iḍāfāt)*. Thus you say **'ard·im·dā** i.e. 'behind me' *(ḥalafī)*; **'ard·imiz·dā** i.e. 'behind us' *(ḥalfanā)* and **'ard·in·dā** i.e. 'behind him' *(ḥalfahu)*. You may use it for both the second and the third per- son since it is explained by the preceding pronoun. In the case of the

---

[86] Em. for **sanu**.

second person you say **san·iŋ 'ard·iŋ·dā** ['behind you [sg.]'] and in the case of the third person you say **'a·niŋ 'ard·in·dā** ['behind him']; and you say **'an·lar·niŋ 'ard·in·dā** for 'behind them' *(ḫalfahum)* <137>.

<p style="text-align:center">39. Chapter on the circumstantial expression</p>

1a. The circumstantial expression *(ḥāl)* is indicated by either a *w* or an *'alif*, which are added to the end of the word. Thus you say **yūrī kal·ū** i.e. 'walk coming' *('imši ġā'iyan)* and **yūrī ṣur·ā** i.e. 'go asking' *(sir sā'ilan)*.

1b. If you intend to construct *(binā')* a circumstantial expression from a verb, you first look at the imperative form *(fiᶜl al-'amr)*.

1c. If the final consonant [of the stem] is silent, then in the imperfect tense it must be vocalised with either an *a*, viz., **kul·ar**—in which case you attach an *'alif* [to the stem]; thus you say, for example **kul·ā kal·dī** i.e. 'he came laughing' *(ġā'a ḍāḥikan)*, or with an *u*, viz., **kal·ur**. In the latter case you attach a *w* to the imperative form; you say for example **kal·ū**, i.e. 'coming' *(ġā'iyan)*.[87]

1d. If [in the imperative form the final consonant] is vocalised with any vowel, you put a *y* vocalised with an *u* after it. Thus you say **yu·y·ū** i.e. 'washing' *(ġāsilan)*; **ya·y·ū** i.e. 'eating' *('ākilan)*; **'aġla·y·ū** i.e. 'crying' *(bākiyan)*; **yuri·y·ū** i.e. 'walking' *(māšiyan)*.

2. The general rule is to allow attachment of a silent *b* in all these instances. The preceding consonant [i.e. the last consonant of the verb] is vocalised with an *u* <VD 54ʳ> if the penultimate consonant is vocalised with an *u*. Thus you say **ṣur·ub**. [The last consonant of the verb] is vocalised with an *i* if [the penultimate consonant] is vocalised with an *a* or an *i*, in that case you say **kal·ib** and **kir·ib**, and if [the penultimate consonant] is silent —except for the *'alif* in **lā**, which remains silent— you say **yakir·ib**[88] 'mad' *(ġaḍbān)*. [Finally,] as to verbs that end in **lā** which serves governance, the *'alif* [of **lā** = /la"/] is replaced by a *y* vocalised with an *u*. Thus you say **suz·la·y·ub** i.e. 'speaking' *(mutakalliman)* and **'aġ·la·y·ub** 'crying' *(bākiyan)*. You may elide the *y*,[89] in which case you say **suz·la·b**.

3. Sometimes a silent *n* is attached to the *b*. In those cases the *b* itself is vocalised with an *a*. Thus you say **kul·ub·an** ['laughing'],

---

[87] I.e. /kul/ → /kul+"/ addition of a: /kal-a"/; /kal/ → /kal+w/ addition of u: /kal-uw/

[88] 'Abū Ḥayyān's description here would point to the form **yakrib** instead, which is not likely.

[89] Em. of *'alif*.

ṣur·ub·an ['asking'] and kir·ib·an ['entering']. The *n* expresses the in-
tensification of the circumstantial expression, as if you repeat it. You
may strengthen the circumstantial expression by repeating its form.
Thus you say kal·ū kal·ū. It is not permitted to pluralise the circum -
5 stantial expression: it always remains singular, whether the referent of
the circumstantial expression is singular, dual or plural. Thus it is not
permitted to say *kal·ū·lar, nor *ṣur·ā·lar or *kal·ib·lar.

4. Transfer *(intiqāl)* and derivation *(ištiqāq)* belong to the condition
of the circumstantial expression but only in those cases where they are
10 possible. Thus you say sangar kal·dī buġdāy ʾaŋ·luh bul·ub·un for
'Sangar came [being] brown' *(ǧāʾa sangar ʾasmar)*. Buġday means
'wheat' *(qamḥ)*; ʾaŋ means 'colour' *(lawn)*; lū has the meaning
'possessor' *(ḏū)*; bul·ub·un means 'becoming' *(ṣāʾiran)*. For 'Sangar
came [being] long' *(ǧāʾa sangar ṭawīlan)* you say sangar kal·dī ʾuzun
15 bul·ub·an and for 'he came laughing' *(ǧāʾa ḍāḥikan)* [you say] kal·dī
kulā and also kul·ub·an or kul·ub. For 'crying' *(bākiyan)* you say
ʾaġla·y·ū, ʾaġla·y·ub·an or ʾaġla·y·ub. For 'this Sangar is standing'
(hāḏā sangar qāʾim) you say bū sangar ṭur· ṭur·ub·an[90] or ṭur·ub or
ṭur·ū. ṭur·miš may only be used as a predicate and not as a circum -
20 stantial expression.

5. For 'Sangar was beaten in the house' *(ḍuriba sangar fī d-dār)* you
say—in case that he was in the house—sangar ʾur·ul·dī ʾaw·dā. For
'in front of you [sg.]' *(ʾamāmaka)* you say ʾaln·iŋ·dā. For 'Sangar
came, the sun [being] rising' *(ǧāʾa sangar wa-š-šams ṭaliʿa)* you say
25 sangar kal·dī kun·duġ·miš yi·dī, <138> i.e. 'the sun was rising'
(kānat aš-šamsu ṭāliʿa). For 'Sangar came, his face [being] reddening'
(ǧāʾa sangar wa-waǧhuhu muḥmirr) you say sangar kal·dī yuz·ī qizil
bul·ub·an. You may also say yuz·u qizil and by this the circumstantial
expression is meant. <VD 54ᵛ> It is also permitted to say qizar·miš
30 yuz·lū 'having a reddening face' *(ḏū waǧh muḥmirr)*. You use a *r* even
though the base form is a *l* [sc. in qizil 'red']. But when you derive [a
verb] from it, you say qiz·ar·dī, changing the *l* into a *r*. Likewise they
say ṣarġ·ar·dī if something becomes yellow, ʾaġ·ardī if somethings
becomes white, changing the *q* [in ʾaq] into a *ġ*. They say qara·r·dī if
35 something becomes dark and kuk·ar·dī if something becomes blue.
They say yaš·ar·dī if something becomes green and buz·ar·dī if some -
thing becomes dust-grey. If the nominal sentence is negative —e.g.
'Sangar came the sun [being] not rising' *(ǧāʾa sangar wa-mā aš-šams*

---

[90] One would expect ṭuruban instead.

ṭāliʿa)— you may say **kun·duġ·ma·miš** yidī or **kun·duġ·ma·duq**
yi·dī or **yuz·ī qizar·ma·miš yi·dī** or **qizar·ma·duq yi·dī**.

6. For 'Sanğar came laughing' *(ğāʾa sanğar yaḍḥaku)* you say **sanğar**
kal·dī kul·ar. For 'Sanğar, whose father laughs, came' *(ğāʾa sanğar*
yaḍḥaku ʾabūhu) you say **ʾata·s·ī kul·ar**. For 'Sanğar came and he had
laughed' *(ğāʾa sanğar wa-qad ḍaḥika)* you say **sanğar kal·dī kul·dī**. For
'and he had cried' *(qad bakā)* you say **ʾaġla·dī**. For 'and his father had
cried' *(wa-qad bakā ʾabūhu)* you say **ʾata·s·ī ʾaġla·dī**. For 'and he does
not laugh' *(mā yaḍḥaku)* you say **kul·ma·s**. For 'he had not laughed'
(mā ḍaḥika) you say **kul·ma·dī**. In the negative form of **sanğar kal·dī**
kulā, you say **sanğar kal·dī kul·ma·din** or **kul·ma·yin**. This means
'Sanğar came not laughing' *(sanğar ğāʾa ġayru ḍāḥik)*. In the case of
the negative form of **ʾaṭ·nī min·du·m ʾayarla·b** ['I saddled the horse
and mounted'] you use **ʾayarla·ma·din** or **ʾayarla·ma·yin**. The use of
ma·yin with a *y* is Qipčāq.

7. If the circumstantial expression is connected with a direct object,
you say **ʾatmak yi·yū** with the meaning 'eating bread' *(ʾākilan ḥubzan)*,
and for 'riding a horse' *(rākiban farasan)* you say **ʾaṭ bin·ū**. You may
also use **bin·ib·an**,[91] **bin·ib**, **yi·y·ib** or **yi·y·ib·an**. It is permitted to use
any of these.

8. It is permitted to place the circumstantial expression before the
single verb or [before] a verb together with a noun, which is the most
eloquent, whether the circumstantial expression is a sentence, a single
word, a locative or a noun governed by a particle of the genitive, and
whether it is a negation or an affirmation.

9. The circumstantial expression is subject to the rules for either the
undetermined words or the determined words.

10. The determination is as in the examples we have shown. As for
the [use of the circumstantial expression] with an undetermined word:
for 'a man came laughing' *(ğāʾa raǧulun ḍāḥikan)* you say **ʾar kal·dī**
kul·ā. It is permitted to say **kul·ub·an**, and also **kul·ub**, according to
our previous remarks.

11a. If the referent of the circumstantial expression is a relative, the
circumstantial expression does not precede it. <VD 55ʳ> For 'the one
that beats came laughing' *(ğāʾa ḍ-ḍāribu ḍāḥikan)* you say **ʾur·ġan**
kal·dī kul·ā. You may also say **kul·ā kal·dī ʾur·ġan** but not **kul·ā**
ʾur·ġan kal·dī. But it is permitted to say **kal·dī kul·ā ʾur·ġan**.

11b. If the relative is **ʾul kim**, you say **ʾul kim ʾur·dī kul·ā kal·dī**. For
'the one that beat came laughing' *(ğāʾa l-laḍī ḍaraba ḍāḥikan)* you may

---

[91] The first consonant of this word is illegible.

say **kal·dī kulā**, e.g. **kal·dī kul·ā ʾul kim ʾur·dī** or **kul·ā kal·dī·kim**
ʾur·dī. It is not permitted to separate the relative from its clause, nei-
ther with the circumstantial expression nor with the verb. For 'the one
in the house came laughing' *(ǧāʾa l-laḏī fī l-bayti ḍāḥikan)* you may say
5   **ʾaw·da·kī kul·ā kal·dī**, **kal·dī kul·ā ʾaw·da·kī** or **kul·ā kal·dī**
ʾaw·da·kī. *<139>*

40. Chapter on the direct object

1. The [direct] object can either be fully nominal *(ẓāhir)* or pronomi-
nal *(muḍmar)*.
10   2. If the object is fully nominal, in the case of the example 'I ate the
fish' *(ʾakaltu as-samaka)* you would say **bāliq·nī ya·du·m**. Nī is the
marker of the accusative as far as the proper object *(al-mafʿūl bihi aṣ-*
ṣarīḥ) is concerned.
3a. [The verbs] that are transitive to two objects in the language of
15   the Arabs are, in this language, transitive to one of them with **nī**, and
to the other with **ġā** or with **ḳā**. Thus you say **sanǧar·ġā ṭū·nī**[92]
bir·du·m for 'I gave Sanǧar a cloak' *(ʾaʿṭaytu sanǧara ṯawban)*. In the
case of the first [word] which is the first object in the Arabic language
ġā is attached to the base form *(ʿalā l-ʾaṣl)*, and **nī** [is attached] in [the
20   case of] the second [word], which is a second [object] in the Arabic
language, and the opposite is not permitted.[93]
3b. For 'I reckoned Sanǧar [to be] leaving' *(ḥasibtu sanǧara ḫāriǧan)*
you say **sanǧar·nī čiq·miš saǧan·du·m**. You suffix **nī** to the first ob-
ject and leave the second one without [any marker], suffixing neither
25   **nī** nor **ġā** to it. You suffix **ḳā** instead of **ġā** when a palatalised conso-
nant precedes. An example of this is **bay·ḳā ṭu·nī birdu·m** ['I gave the
cloak to the chief'].
4. If it is [a verb] which is transitive with 'to' *(ʾilā)* in the Arabic lan-
guage, whether it is stated or deleted, according to linguistic licence
30   *(tawassuʿ)*, they use **ġā** or **ḳā** in accordance with the preceding remarks
about the velarisation and palatalisation. An example of this is
sanǧar·nī kur·du·m i.e. 'I saw Sanǧar' and **sanǧar·ġa baq·ṭi·m** 'I
looked at Sanǧar' *(naẓartu ʾilā sanǧar)* <VD 55ᵛ> and **bay·ḳā**
bar·du·m i.e. 'I went to the chief' *(ḏahabtu ʾilā l-ʾamīri)*. It is permit-

---

[92] In **ṭunī** one of the two n's appears to be missing (cf. also 41.5 for **san** and **man**).
[93] This same important passage appears slightly differently in İÜ (82ᵛ5-6): "You
attach **ġā** to the second [word] which is the first object in the Arabic language. And you
attach **nī** to the first [word], which is a second [object] in the Arabic language."

ted to put the object at the end in all of these cases, although putting it
at the beginning is the most correct and the most frequent.

    5. If [the object] is pronominal, you say **ʾur·di manī** for 'he beat me'
(ḍarabanī). Its base form is **ʾur·di man·nī**. You delete the first *n*
5   because of the clustering of two similar consonants. We say that it is
the first [*n*] that is deleted, because the second one is used together
with the *y* to indicate a meaning, namely objectivity *(mafʿūliyya)*. For
'he beat us' *(ḍarabanā)* you say **ʾur·di biz·nī**. For 'he beat you [sg.]'
(ḍarabaka) you say **ʾur·dī san·ī**. Its base form is **san·nī** with gemina-
10   tion [of the *n*]. About **san·nī** we can say the same as [we did] about
man·nī. [For] 'he beat you [pl.]' *(ḍarabakum)* [you say] **ʾur·dī siz·nī**.
You say **ʾur·dī ʾa·nī** for 'he beat him' *(ḍarabahu)* and you say **ʾur·di**
ʾan·lar·nī for 'he beat them' *(ḍarabahum)*. In all of these cases it is
permitted to put the object at the beginning, since it is an independent
15   pronoun in their language.

    6. If both the agent and the object are pronominalised, you say
ya·du·m ʾa·nī for 'I ate it' *(ʾakaltuhu)* and you say **bir·du·m ʾa·nī** for
'I gave it' *(ʾaʿṭaytuhu)*. For 'Sangar, I gave him the cloak' *(sangar*
ʾaʿṭaytuhu aṯ-ṯawb) you say **sangar bir·du·m ṭun·ī ʾan·ġā**.

20   7. If the object is an interrogative noun, you say for 'whom did you
[sg.] beat' *(man ḍarabta)* **kim·nī ʾur·du·ŋ**; the most frequent usage is
to put it at the beginning.

    8. If [the object] is a conditional noun, you say **kim·nī ʾur·saŋ**
ʾur·ġa·man for 'whomever you [sg.] beat, I will beat him' *(man*
25   *taḍrubhu ʾaḍrubhu)* and <140> it is obligatory to put the object at the
beginning, because it is a conditional noun.

    9. If it is a relative—as in 'I saw the one I beat with you [sg.]'
(raʾaytu allaḏī ḍarabtu ʿindaka)—you say **ʾul kim ʾur·du·m san·iŋ**
qaṭi·ŋ·dā kur·du·m. As for **ʾul**, it means 'he' *(huwa)* and **kim** means
30   'who' *(man)*; their combination expresses the meaning 'the one that'
(allaḏī). **ʾUr·du·m** means 'I beat' *(ḍarabtu)* and **san·iŋ qaṭi·ŋ·dā**
means 'with you [sg.]' *(ʿindaka)*. The object is governed by *(yutasallaṭ*
ʿalayhi) the verb and it is connected to it *(yataʿallaq bihi)* as we have
demonstrated. [The same goes for] the active participle, the passive
35   participle and the verbal noun.

41. Chapter on the connection of the active
participle with the direct object or the fully nominal agent

1. You say **sangar 'ur·daği·dur sunqur·nī** for 'Sangar is beating
Sunqur tomorrow' *(sangaru ḍārib sunqura ğadan)*.[94] It is permitted [to
say] **sangar sunqur·nī 'ur·daği·dur, 'ur·daği·dur sangar sunqur·nī**
or <VD 56ʳ> **'ur·daği·dur sunqur·nī sangar**. This is also permitted in
the past tense, i.e. when **miš** is attached. [The active participle in the
past tense] is connected to the direct object in the same way as it is
connected in the future tense. All of these constructions are permitted.
You also attach **miš tur** in other constructions.

2. In the case of [the verbs that are] transitive to two [objects], e.g.
in 'Sangar is giving Sunqur a cloak' *(sangaru mu ᶜṭin sunqura ṭawban)*,
you say **sangar bir·daği·dur sunqur·ğā ṭun·nī**. It is permitted to put
it either at the beginning or at the end in the abovementioned con-
structions. You say **sangar ṣagan·daği·dur sunqur·nī 'urū·ṭur·miš**
for 'Sangar is thinking Sunqur is standing' *(sangaru ẓānnun sunqura
qāʾiman)* or just **ṭur·miš** without **'urū**. For 'Sangar is repeatedly beat-
ing Sunqur' *(sangaru ḍarūb sunqur)*, you say in the intensive form
sangar 'ur·agan·dur sunqur·nī. It is permitted to use the construc-
tions that were mentioned in the [Chapter on] the Active Participle
[see above Section 4].

3. If you connect the active participle to the fully nominal agent *(al-
fāᶜil az-zāhir)*, you say **sangar 'ata·s·ī ṭur·miš** for 'Sangar, his father is
standing' *(sangaru qāʾim ʾabūhu)*. You say **bū yuz·ī kurk·u·lū dur** for
'this one's face is beautiful' *(hāḍā ḥasanu [waǧhuhu][95])*. You say **kuz·ī
surma·lū·dur** for 'Sangar, his eye is dyed black' *(sangaru kaḥīlu
ᶜaynuhu)* and **burn·ī 'uzun dur** for 'his nose is long' *(ṭawīl ʾanfuhu)*.
The difference between **'uzun** and **surma·lū** is that **'uzun** is basically
a descriptive [word], whereas **surmā** means 'antimony' *(al-ʾiṭmid)*
which is basically not a description. Therefore **lū** is attached to it, which
has the meaning of 'possessor' *(ḍū)* and this is used as a descriptor
(wuṣifa bihi). [The case of] **kurk** is similar, having the meaning 'beauty'
(ḥusn). It is originally not an adjective, but when they combine it with
lū, with the meaning of 'possessor of' *(ḍū)*, it is used as a descriptor. **Lū**
becomes a conjunctive word *(wuṣla)* for descriptions of generic nouns,
comparable to expressions in the language of the Arabs such as
'possessor of water' *(ḍū māʾ)* and 'possessor of gold' *(ḍū ḍahab)* and so
forth. <141> They do not say *ʾ**uzun lū**, as they do not say *ᶜowner of

---

[94] 'Tomorrow' is not explicitly stated in the Turkic phrase.
[95] Em. of VD *waǧhika*.

long' *(ḍū ṭawīl)* in the Arabic language. The equivalent *(murādif)* of
'beautiful' *(ḥasan)*, a real adjective, is **kukǧak dur**, except that it has
the connotation of affection and beauty with regard to children.

42. Chapter on the connection of the passive participle
to the direct object

1. In the case of the equivalent *(murādif)* of 'I saw a man whose slave
was beaten' *(ra'aytu raǧulan maḍrūban mamlūkuhu)* you say **bir 'ar·nī
kur·du·m 'ur·ul·miš qul·ī**. It is permitted [to say] **kur·du·m bir
'ar·nī 'ur·ul·miš qul·ī**. The adjective cannot precede <VD 56ᵛ> the
described [noun], therefore they do not say **'ur·ul·miš qul·ī 'ar·nī
kur·du·m* nor **bir 'ar·nī 'ur·ul·miš qul·ī kur·du·m*. The verb can
only be placed in the middle or preceding the direct object, but the ad-
jective always follows the described [noun]. For 'I saw a man whose fa-
ther was given a *dirham*' *(ra'aytu raǧulan mu'ṭan 'abūhu dirhaman)*
you say **bir 'ar·nī kur·du·m bir·il·miš 'aṭa·s·i·n·ā bir 'aqǧā**, in
agreement with these constructions.

43. Chapter on the connection of the verbal noun to the direct object

1. For 'Sanǧar's beating [of] Sunqur is harsh' *(ḍarbu sanǧara sunqura
šadīdun)* you say **sanǧar 'ur·maq laq·ī** or **'ur·maq·ī sunqur·nī qāṭī
dur**. It is permitted to say **sunqur·nī sanǧar 'ur·maq·ī qāṭī·dur** or
sanǧar 'ur·maq·ī sunqur·nī qāṭī·dur. All of this is permitted. The di-
rect object precedes the verbal noun either with its agent or without it,
unlike [the verbal noun] in the Arabic language. This is because the
marker of objectivity appears in the object. For 'I liked Sanǧar's beating
Sunqur' *('aḥbabtu ḍarba sanǧar sunqur)* you say **saw·di·m sanǧar
'ur·maq·nī sunqur·nī**. It is permitted [to say] **sanǧar 'ur·maq·nī
sunqur·nī saw·di·m** and **sanǧar 'urmaq·nī saw·di·m sunqur·nī**. It is
not permitted [to say] **sanǧar saw·di·m 'ur·maq·nī sunqur·nī** nor
sunqur·nī saw·di·m sanǧar 'ur·maq·nī.

44. Chapter on the object of reason

1. [The object of reason *(al-maf'ūl lahu)* is, e.g.,] for 'I came to you
[sg.] [out of] love' *(ǧi'tuka maḥabbatan)* you say **san·i saw·mak·
im·dan kal·di·m**. The meaning of **kal·di·m** is 'I came' *(ǧi'tu)*; **saw·
mak·im·dan** means 'from my love' *(min maḥabbatī)*. The meaning of

san·i is 'to you [sg.]' *(laka).*[96] It is permitted to put the verb at the beginning or at the end in this case, whether <142> the motive and the object depend from the same agent, as we showed in the examples, or from two [different] agents. Thus, for 'I came to you [sg.] out of fear

5   of evil from Sangar' *(ği'tuka ḥawfa sangarin aš-šarra)* you say **kal·di·m saġā sangar yaman·dan qurq·maq·in dan**. Its analysis is **kal·di·m** 'I came' *(ği'tu)* and **ṣaġā** 'to you [sg.]'[97] *(laka)* and **yaman·dan** 'from evil' *(min šarrin)* and **qurq·maq·in dan** 'from fear of him' *(min ḥawfihi)*.

10   2. If it [sc. the object] is not related to the motive *('illa)*, as in 'I be-lieved in order to enter heaven' *('āmantu li-duḥūl al-ğanna)*, <VD 57ʳ> you say **kirtun·du·m kir·mak·um 'uġun 'uġmaq·qā**. [And as for] **kirtun·du·m**, it means 'I believed' *('āmantu)*; **kir·mak·um** [means] 'my entering' *(duḥūlī)*; **'uġun** [means] the [particle] *li* that expresses a

15   motive, **'uġmaq·ā** means 'to heaven' *(li-l-ğanna)*. **'Uġmaq** means 'heaven' *(al-ğanna)*; the *q* that is vocalised with *a* and with which the *q* of **'uġmaq** is assimilated indicates the object. And as to **qā** or **ġā**, they form a way to express the meaning 'to' *('ilā)*, namely the goal, while **nī** expresses the proper object. This is why we say **'aqġa·nī bir·du·m**

20   **sangar·ġā**, meaning 'I gave Sangar the *dirham*' *('a'ṭaytu ad-dirhama sangara)* i.e. 'to Sangar' *(li-sangar)*. The giving actually occurred to the *dirham* and Sangar is the one to whom the *dirham* is handed. The equivalent *(naẓīruhu)* of this in Arabic is 'I handed the *dirham* to Zayd' *(dafa'tu ad-dirhama 'ilā zaydin)*. It is not permitted [to say] 'I handed

25   the *dirham* Zayd' *(dafa'tu ad-dirhama zaydan)*.

3. **Kim** also denotes the motivation with the meaning of 'in order to' *(kay)*. You say **'anda·di·m san·ī kim 'āš ya·kā·san** 'I called you [sg.] so that you eat the meal' *(da'awtuka kay ta'kula ṭ-ṭa'āma)*. This is the most correct, viz. that the particle of motivation follows the noun,

30   although it may also follow the verb. Thus you say **kim ya·kā san 'āš** and you may [add] **san·ī 'anda·di·m**.

45. Chapter on the concomitant object

1. [The concomitant object *(al-maf'ūl ma'ahu)* is, e.g.,] 'I sat next to the column' *(ğalastu wa-s-sāriya)* you say **'ulṭur·du·m dirak bilā**.

---

[96] It is interesting to note that according to western interpretation, **san i** is the object of **sawmak imdan** rather than of **kal dim**. 'To you' *(laka)* as such does not explicitly occur in this sentence.

[97] This must be understood as a paraphrase of the direct object with *li*, rather than a direction.

dirak is 'a column' *(as-sāriya)* i.e. 'a pillar' *('amūd)*. [As for] **bilā**, its meaning is 'with' *(maʿa)*. The concomitant object is only used in combination with an explanatory word that is the equivalent of 'with' *(maʿa)*, namely **bilā** *(muṣarraḥan)*. Likewise, the object of reason is
5   always used with the equivalent *(murādif)* of the particle 'from' which denotes the reason *(min as-sabbabiyya)*, as we mentioned in the preceding chapter [see above Section 44]. Here it is permitted to put the verb either at the beginning or at the end.

46. Chapter on the excepted

10 1. [The exception *(al-mustatnā)* is, e.g.,] For 'the people stood up, except Sangar' *(qāma n-nās 'illā sanğar)* you say **kiši·lar ṭur·di·lar maqar sanğar**. It is not permitted to put **maqar sanğar** before **ṭur·di·lar** nor before **kiši·lar**.
     2. In the case of the exception that is made devoid [of governance]
15 *(mufarraĝ)*, you say for 'none stood up except Sangar' *(mā qāma 'illā sanğar)* **ṭur·ma·dī maqar sanğar**; and for 'none stands up except Sangar' *(mā yaqūm 'illā sanğar)* **ṭur·ma·ğay maqar sanğar**. 'I did not eat except the loaf' *(mā 'akaltu 'illā l-ḥubza)* <VD 57ᵛ> **ya·ma·du·m maqar 'atmak**. It is permitted to say **'atmak·nī**.
20   3. The exception occurs both uninterrupted *(mutaṣṣal)*, devoid [of governance], as in the examples given, and interrupted *(munqaṭiʿ)*. <143> Thus you say for 'I saw people except a horse' *('abṣartu n-nāsa 'illā farasan)* **kiši·lar·nī kur·du·m maqar 'aṭ**.⁹⁸ It is not permitted to put the excepted [noun] between the verb and the word from which
25 the exception is made *(al-mustatnā minhu)*. So it is not permitted to say **kiši·lar maqar sanğar ṭur·di·lar**.
     4. An exception may also be made with what is the equivalent of 'other' *(ĝayran)*, i.e. **'uzkā** and **'ayruq**. Thus you say **kiši·lar ṭur·di·lar sanğar·dan 'uzkā** ['people stood up except Sangar'] and **sanğar·dan**
30 **'ayruq**. **Dan** means 'from' *(min)*, as if one wishes to say 'another from Sangar' *(ĝayr min sanğar)*. It is permitted to put **sanğar·dan 'uzkā** and **sanğar·dan 'ayruq** either before the verb, or the word from which the exception is made, or the verb that is devoid [of gover nance]. It also occurs in the interrupted exception. In that case you say **kiši·lar**
35 **kur·du·m 'aṭ·ṭan 'uzkā** ['I saw people except a horse'] or **'aṭ·ṭan 'ayruq**. The base form [of **'aṭ·ṭan**] is **'aṭ·dan**. The *d* is substituted for a *ṭ* which subsequently assimilates with the [preceding] *ṭ*. You may put

---

⁹⁸ This sentence is probably a calque of the Arabic.

the excepted [word] at the beginning in the interrupted exception, in the same way as this is permitted in the uninterrupted exception.

5. If the exception is made from a nominal sentence, like 'the people are your [sg.] brethren except Sangar' *(an-nās ʾiḫwatuka ʾillā* 5 *sangar)*, you say **kiši·lar qarindaš·lar·iŋ·dur maqar sangar**. It is not permitted to put **maqar sangar** at the beginning, as we have mentioned already.

6. If you make an exception with a noun, you say **sangar·dan ʾuzkā** and **sangar·dan ʾayruq**. It is permitted to put it both at the beginning 10 and at the end as we mentioned in [the Section on] the Verbal Sentence [see above Section 24.2b].

47. Chapter on the specification

1. In the language of the *ʾAtrāk* there is no transferred specification *(tamyīz manqūl)*, for they use it with the base form.[99] Thus they say 15 'Sangar's spirit was good' *(ṭābat nafsu sangar)* **kuŋl·ī yaqšī ḇul·dī**. **Kuŋl·ī** means 'his heart' *(qalbuhu)*, **yaqšī** means 'good' *(ṭayyiban)* and **ḇul·dī** means 'he became' *(ṣāra)*. There is no transferred specification derived from the object either. So they say for 'I planted the earth with trees' *(ġarastu l-ʾarḍa šaǧaran)*[100] **dik·tu·m ʾaǧaǧ yir·dā** ['I planted 20 trees in the earth'], in accordance with the base form.

2. As to the non-transferred specification *(ġayr al-manqūl)*, it [consists of] the countable *(al-muqaddarāt)*, and those are the number *(ʿadad)*, the volume *(mukīl)*, the weight *(mawzūn)* and the length *(mamsūḥ)*.

25 3a. The number. You say **ʾuǧ·ar kal·dī** for 'three men came' *(ǧāʾa ṯalāṯatu riǧālin)*. Its base form is **ʾuǧ ʾar**. **ʾUǧ** is 'three' *(ṯalāṯatu)* and **ʾar** is 'man' *(raǧul)*. It is only used in combination with the singular. You say <VD 58ʳ> **bir·ar** 'one man' *(raǧul wāḥid)* and **ʾikī ʾar**, 'two men' *(raǧulāni)*, up to ten. If you say 'eleven [men] came' *(ǧāʾa ʾaḥad* 30 *ʿašar)* [they say] **ʾun bir·ar kal·dī** [and 'twelve men came'] **ʾun·ikī ʾar kal·dī**. Its base form is **ʾun ʾikī** and **ʾun ʾuǧ**, up to nineteen. [They also say] 'twenty men' *(ʿišrūna raǧulan)* **yikirmi·y·ar**; 'twenty one' *(wāḥid wa-ʿišrūna)* **ʾikirmi·bir·ar**; 'twenty two' *(ʾiṯnāni wa-ʿišrūna)* **ʾikirmi· y·ikī ʾar**, and 'thirty men' *(ṯalāṯūna raǧulan)* **ʾutuz·ar**. Its base form is 35 **ʾutuz ʾar** up to a hundred. Thus you say **yuz·ar kal·dī** i.e. 'a hundred

---

[99] What ʾAbū Ḥayyān means to say here is that sentences of the type *ṭāba sangar nafsan*, literally 'Sangar was good in spirit', in which *sangar* is the subject and *nafs* takes the accusative case, do not occur in Turkic.

[100] Probably a calque of *ġarastu š-šaǧara fī l-ʾarḍi* 'I planted the tree in the soil'.

men came' *(miʾa raǧulin ǧāʾa)* and 'two hundred men' *(miʾatā raǧulin)*
ʾiki·yuz·ar; 'three hundred men' ʾuǧ·yuz ʾar; 'four hundred' durt yuz
ʾar up to a thousand. Thus you say min ʾar ['a thousand men']; 'two
thousand' *(ʾalfāni)* ʾikī min ʾar; 'ten thousand men' ʾun min ʾar.

5   <144>
     3b. In the remaining countable objects the specification only occurs
in the singular. If you construct an active participle from the singular
number, you say ʾilki·nǧī 'the first of the sequence' *(al-ʾawwal al-mu-
rattab)*, it means that a second follows after it. ʾiki·nǧī means 'the

10   second' *(aṭ-ṭānī)*; ʾuǧ·unǧī 'the third' *(aṭ-ṭāliṭ)*; durd·unǧī ['the
fourth'] baš·inǧī ['the fifth'] alṭi·nǧī ['the sixth'] yadi·nǧī ['the seven-
th'] sakis·inǧī ['the eighth'] ṭuqs·inǧī ['the ninth'] and ʾun·inǧī ['the
tenth']. The same goes for the compound number *(al-ʿadad al-mu-
rakkab)* in the Arabic language. You say ʾun bir·inǧī ['eleventh'] and

15   ʾun ʾiki·nǧī ['twelfth'] up to twenty. Thus you say yikirmi·nǧī as if
[the speaker] said 'twentieth' *(al-ʿāširūn)* although this is never ex-
pressed that way in Arabic. Similar to this is the status of the rest of the
groups of ten after it, up to a hundred. Thus you say yuz·inǧī. What
comes after a hundred is not actually used. Derivations of the numbers

20   are not used. When they wish to utter this meaning, they repeat the
numeral while adding a *n*. Thus you say bir·in bir·in ['one by one']
ʾiki·n ʾiki·n ['two by two'] up to ʾun ['ten']. [And] thus you say ʾun·in
ʾun·in i.e. 'ten by ten' *(ʿašara ʿašara)*. In [the Chapter on] the conso-
nants of Addition [cf. above Section 19.11h] we have already treated

25   the attachment of the marker of collectivity (and I refrain from repeat-
ing it here).
     3c. The volume. For 'I have one ʾirdabb[101] of wheat' *(ʿindī ʾardabbu
qamḥin)* you say bir ʾirdab buǧday bar qaṭ·im·dā [literally 'near me'].
     3d. The weight. For 'I have one raṭl[102] of oil' *(ʿindī raṭl zaytan)* you

30   say bir baṭman yaǧ·bar qaṭ·im·dā.
     3e. The length. For 'I have a cubit of cloth' *(ʿindī dirāʿ tawban)* you
say bir qarī ṭun bar qaṭ·im·dā. With the meaning of 'I have a foot of
ground' *(ʿindī šibr ʾarḍan)* you say bir qariš yar bar qaṭ·im·dā.
     4. <VD 58ᵛ> In all of these examples the specification does not pre-

35   cede what is specified by it. It is permitted to put the equivalent
(murādif) of the specification at the beginning and to put the specifica-
tion after it as well. Thus you say ʾar ʾuǧ·ar kal·dī ['men, three men

---

[101] One ʾirdabb equals appr. 198 l.
[102] One raṭl equals appr. 3.2 kg.

came'], its meaning is 'of the men, three men came' *(min ar-riǧāl ǧā'a* *ṯalāṯatu riǧālin)*.[103]

48. Chapter on the competing regency

1. Sometimes two verbs are connected to one or more nouns. The noun may stand at the beginning or at the end. Thus, for 'I beat, abused and stabbed Sangar' *(ḍarabtu wa-šatamtu wa-ṭaʿantu sanǧara)*, you say **'ur·du·m suk̲·du·m ṣanǧ·du·m sanǧar·nī.** You may put **sanǧar·nī** either before, or between any of the verbs.

49. Chapter on the annexation

1. The annexation *('iḍāfa)* occurs with particles and without particles.

2. The particles of the annexation are **dan** with the meaning 'from' *(min)* or 'off' *(ʿan)*; **dā** with the meaning 'in' *(fī)*; **daqin** with the meaning 'to' *('ilā)* and 'until' *(ḥattā)*; **bilā** with the meanings 'with' *(al-bā')* and 'together with' *(maʿa)*.

3. They do not have an equivalent *(murādif)* particle to the *k* of <145> comparison, but they use either a noun, like **kibī** or **dakī**, or a verb, like **'uqša·r**, with the meaning 'he looks like' *(yušbihu)*. This verb is [pronounced] with a *q* and some of those who do not speak this language correctly change it into a *ḫ*. The past tense of **'uqša·r** is **'uqša·dī** with the meaning 'he resembled' *('ašbaha)*.

4. The particle of the oath is **'uǧun.**

5. 'On' *(ʿalā)* is expressed with the equivalent *(murādif)* 'on top of' *(fawqa)*. So they say for 'Sangar [sits] on the horse' *(sanǧar ʿalā l-farasi)* **'aṭ 'usti·n·dā·dur sanǧar.**

6. **Niŋ** has the meaning 'of' *(al-lām)*. You say for 'whose are you [sg.]' *('anta li-man)* **san kim·niŋ·san** and for 'this is Sangar's' *(hāḏā li-sanǧar)* **bū sanǧar niŋ dur.** In the language of the Oġuz [they say] **sanǧar·iŋ dur** with the post-palatal *n*[104] alone and with deletion of the first *n*.

7. **Barū** is used in combination with [words denoting] time in the same way as 'since' *(munḏu)* is used in the Arabic language. For 'I have not seen him since yesterday' *(mā ra'aytuhu munḏu 'amsi)* you say **dun·ā kun·dan barū kur·ma·du·m.** **Dunā**[105] is 'yesterday' *('amsi)*;

---

[103] This is not entirely correct, since 'of' is not reflected in the Turkic sentence; the correct translation should rather be 'men, three men came'.

[104] The text reads *kāf ḥayšūmiyya* 'Bedouin *k̲*, which must be an error for *nūn ḥayšūmiyya.*

[105] The ā in **dunā**, probably to be read [düne], is inserted for phonological reasons.

kun is 'day' *(yawm)*; **dan** is 'from' *(min)*; **barū** is 'since' *(mundu)*][106]
[and] **kur·ma·du·m** means 'I did not notice' *(mā ʾabṣartu)*.

8. The particle **ġalī**, too, is used with the meaning 'since' *(mundu)*.
Thus you say **man bar·ġalī** i.e. 'since I went' *(mundu dahabtu)* and
<VD 59ʳ> also **tur·ġalī** 'since I sat down' *(mundu qumtu)*.[107]

9. For 'I came out of the house' *(ǧiʾtu min al-bayti)* you say **ʾaw·dan
kal·di·m**; 'I was in the house' *(kuntu fī l-bayti)* **ʾaw·dā ʾi·du·m**; 'I
walked until the house' *(sirtu ʾilā l-bayti)* **ʾaw·kā daqin yuri·di·m**. ʾAw
is 'the house' *(al-bayt)*, **kā** is used with the word that is pronounced
palatalised and **ġā** is used with the word that is pronounced velarised.
The meaning of both of them is 'to' *(ʾilā)* but they are [also] used
pleonastically *(murdifa)* to **daqin**.[108] 'I walked to Mecca' *(sirtu ḥattā
makka)* **makka·kā daqin yuri·di·m**. For 'from this side to this side the
ground is yours [sg.]' *(min hādā ṭ-ṭaraf ḥattā hādā ṭ-ṭaraf al-ʾardu
laka)* they say **bū ʾuǧ·dan bū ʾuǧ·ġā daqin ġā yir san·iŋ dur** and they
say **san·iŋ dur** with deletion of one of the two *ns*.[109] **Ġā** may be dele-
ted [from **daqin ġā**].

10. 'I went out hunting' *(ḥaraǧtu ʾilā ṣ-ṣaydi)* **ʾav·ġā čiq·tu·m**. **Ġā**
has the meaning 'to' *(ʾilā)*, as we showed and the meaning of the *lām*
of motivation [i.e. the particle *li*]. Thus, for 'I stood up for Sangar'
(qumtu li-sangar) you say **sangar ġā tur·du·m** or you say **sangar
ʾuġun tur·du·m** i.e. 'for Sangar's sake' *(li-ʾaǧli sangar)*.

11. **daqin** is derived from **dak·dī** i.e. 'he reached' *(balaġa)* and 'he
arrived' *(waṣala)*.

12. [You say for] 'I wrote with the pen' *(katabtu bi-l-qalami)* **qalam
bilā yaz·du·m** or **biti·di·m** or **čiz·di·m**. [And for] 'I came together
with Sangar' *(ǧiʾtu maʿa sangar)* **sangar bilā kal·di·m**.

13. These are the particles of annexation that come at the end [of
the word], as we showed, unlike the particles of annexation in the
Arabic language: they can be used with nouns and examples of their
use with overt nouns have been given.

14. As far as the fact that they are used with suffixed personal pro-
nouns is concerned: you say for 'from me' *(minnī)* **man·dan**; 'from us'
(minnā) **biz·dan**; 'from you [sg.]' *(minka)* **san·dan**; 'from you [pl.]'
(minkum) **siz·dan**; <146> 'from him' *(minhu)* **ʾan·dan**; [and for]
'from them' *(minhum)* **ʾan·lar dan**.

---

[106] Alternative reading from İÜ.
[107] Em. for *mundu qaʿadtu*.
[108] In Western analysis this would rather be the other way around, i.e. **daqin** being a
pleonasm to **ġā**.
[109] This would implicate something like **sannin dur**.

15. And [you say for] 'together with me' *(maʿī)* **man·im·bilā**; 'together with us' *(maʿanā)* **biz·iŋ·bilā**; 'together with you [sg.]' *(maʿaka)* **san·iŋ·bilā**; 'together with you [pl.]' *(maʿakum)* **siz·iŋ·bilā**; 'together with him' *(maʿahu)* **ʾa·niŋ·bilā**; [and for] 'together with

5 them' *(maʿahum)* **ʾan·lar·niŋ bilā**. **bilā** has the meaning 'together with' *(maʿa)* as if it were a locative. The equivalent *(murādif)* of 'in' *(fī)*, which is **dā**, is not used with it, unlike **qaṭi** with the meaning 'at' *(ʿinda)*.

50. Chapter on the oath

10 1. When they swear, they say for 'I have sworn I shall not go' *(ḥalaftu mā ʾarūḥu)* **ʾanṭ·iš·ti·m bar·ma·ġa·man** <VD 59ᵛ> if they mean is a future tense. **Bar·mas·man** serves to negate the present tense *(ḥāl)*. For 'I did not go' *(mā ruḥtu)* you say **bar·ma·di·m**. With the meaning 'I swear, I shall not go' *(ʾanā ḥālif mā ʾarūḥ)*[110] [they say] **ʾanṭ ʾiǧ·miš·**

15 **man bar·ma·ġa·man**.

2. **ʾUǧun** is [used] for the motivation and by this the oath in the future is meant. You say **sanġar baš·i ʾuǧun, sanġar kuz·ī ʾuǧun** and **sanġar ǧān·ī ʾuǧun** i.e. 'for his head's sake' *(li-ʾaǧli raʾsihi)*, 'for his eye's sake' *(li-ʾaǧli ʿaynihi)* and 'for his soul's sake' *(li-ʾaǧli rūḥihi)*

20 **bar·ma·ġa·man**.[111]

3. All of these expressions are used as oaths, instead of [real?] oaths *(mawāqiʿ al-qasam)*. The sentence to which the oath is sworn [comes] after [the oath], whether it is affirmative or negative. It may also precede, thus you say **sanġar kuz·ī ʾuǧun sunqur bar·daġi·dur** ['by the

25 eye of Sanġar, Sunqur will come'] and **sunqur bar·daġi·dur sanġar kuz·ī ʾuǧun**.

4. Baylik wrote in his book: 'In our lands we do not use any of these oaths, but one bends one's forefinger, as one does in the case of the number seventy in the counting system of the Copts, without the nail

30 of the thumb touching it. Then one says **bū ʾanṭ** ['this oath'] "...I did not do such-and-such and I did not commit it". One hardly breaks his oath. Its meaning is: if [the oath] is broken, one will go on living like this, with one's back bent.'

---

[110] Literally: 'I am swearing, I will not go'.
[111] This is the verbal complement that was ommited in the three preceding examples.

51. Chapter on the annexation without a particle

1. The element to which is annexed precedes the annexed in this lan-
guage. Thus for 'Sanğar's slave stood up' *(qāma ġulāmu sanğar)* you say
sanğar qul·ī ṭur·dī. It is permitted to put **ṭur·dī** at the beginning.

5   2. Something has already been said about the annexation in the
Chapter on the Determined Word [see above Section 25.4]. If you an-
nex [a word] to an interrogative noun in the Arabic language you say
'whose slave are you' *(ġulāmu man 'anta)*. With the same meaning in
this language [you say] **kim·niŋ qul·ī·san**. Its literal meaning is 'to
10   whom his slave [are] you' *(li-man mamlūkuhu 'anta)*. And for 'whose
slave you [sg.] may beat, I will beat' *(ġulāmu man taḍrub 'aḍrub)* [you
say] **kim·niŋ qul·in 'ur·saŋ 'ur·ğa·man**. **sā** is the particle of condi-
tion and the post-palatal *n* is the pronoun of the second person. Its base
form is **'ur·ur·saŋ** because **'ur·ur** is the imperfect tense and one of
15   the two *r*s is elided to alleviate. For 'the slave of the one in the house is
standing' *(ġulāmu l-laḍī fī d-dāri qā'im)* [you say] **'awda ki·niŋ qul·ī
ṭur·miš dur** <VD 60ʳ> and **kur·kan·im·niŋ qul·ī ṭur·miš dur** for 'the
slave of [the one] I saw is standing' *(ġulāmu l-laḍī ra'aytuhu qā'im)*
<147> and 'the slave of the one I noticed is standing' *(ġulāmu l-laḍī
20   'abṣartuhu qā'im)* **'ul kim kur·du·m 'a·nī qul·ī ṭur·miš dur**.

3. In this language 'some [of]' *(ba'ḍ)* and 'all' *(kull)* belong to the
nouns that in the Arabic language are annexed both formally *(lafẓan)*
and semantically *(ma'nan)*, or else semantically but not formally. The
equivalent *(murādif)* of 'some [of]' *(ba'ḍu)* in this language is **bir
25   'anğā**. Its base form is **bir** with the meaning 'one' *(wāḥid)* and **'anğā**
with the meaning 'like this' *(miṯlu ḍālika)* and these two words express
the meaning of the partitive. If you say in Arabic 'I ate some of the
fish' *('akaltu ba'ḍa s-samaki)*, in this language with the same meaning,
you say **yi·du·m bāliq·niŋ bir·anğa·s·i·n**. **Si** is the indicator *(dalīl)* of
30   the annexation to the third person if the last consonant of the noun is
vocalised. The *n* is the indicator of objectivity *((dalīl al-maf'ūliyya)*.

4. The equivalents *(murādif)* of 'all' *(kull)* in this language are **tayma**
and **barğā**. They are annexed either to an undetermined noun or to a
determined noun.

35   5. If they are annexed to an undetermined noun, with the meaning
'do not eat all of the fish' *(lā ta'kul kulla samakin)* you say **tayma
bāliq·nī ya·mā**. And for 'do not eat the whole fish' *(lā ta'kul kulla s-
samaki)* [you say] **bāliq·nī barğa·s·i·n ya·mā**. **Tayma** is only used as a
general statement about individual nouns *('umūm al-'afrād;)*. **Barğā** is
40   only used as a general statement about collective nouns *('umūm aš-*

šumūl;): the difference between the two constructions is clear. **Tayma**
is Qipčāq; in Oġuz you say **dakmā**. You say for 'I gave each man a
dirham' *('a'ṭaytu kulla raġulin dirhaman)* **dakmā 'ar·kā bir 'aqġā**
bir·di·m; kā, which is attached to it, is the indicator of objectivity and
5   **dakmā** forms a composite with what follows it: **bir** and **'ar**. For this
[reason] the marker of the object appears at the end of the composite.
For 'I took all of it' *('aḫadtu al-kulla)* [You say] **barġa·s·i·nī 'al·du·m.**
'All of it' *(kullahu)* means **barġa·sa·nī;** the *s* [indicates] the annexa-
tion and the *n* the objectivity.
10   6. If you annex to a pronoun, you annex either to a first, or a sec-
ond or a third person.
       7. If you annex to a first person, you say **qul·um ṭur·dī** ['my slave
stood up'] and **qul·um·nī 'ur·du·m** ['I beat my slave']; **qul·umuz**
ṭur·dī and **qul·umuz·nī 'ur·du·q.** [If you annex] to a second person,
15   you say **qul·uŋ ṭur·dī, qul·uŋ·nī 'ur·du·ŋ** and **qul·uŋuz[·nī]**
'ur·du·ŋuz. [If you annex] to a third person, you say **qul·ī ṭur·dī,**
qul·u·nī 'ur·du·m ['I beat his slave'] and **qul·lar·i ṭur·du·lar** i.e.
'their slaves stood up' *(mamālīkuhum qāmū).* **Lar** is attached to the
singular if you annex to a singular; [its] *r* is vocalised with an *i*. This in-
20   dicates <VD 60ᵛ> the plural of the annexed; it may also indicate the
plural of the ones to whom is annexed *(al-muḍāf 'ilayhim).* This can be
deduced from the context *(siyāq al-kalām).* With the meaning 'Sangar,
his slaves are standing' *(sangar mamālīkuhu qā'imūn),* you say **sangar**
niŋ qul·lar·i ṭur·miš·lar dur. If you [mean] 'the Sangars, their slaves
25   are standing' *(as-sanāġir mamālīkuhum qā'imūn),* you say **sangar·lar**
niŋ qul·lar·i ṭur·miš·lar dur.

52. Chapter on the appositions

1. The attribute precedes the qualified noun if it is singular. For 'I saw a
wise man' *('abṣartu raġulan 'āqilan)* you say **kur·du·m bir 'uṣlū 'ar**
30   <148> and **bir ṭur·miš 'ar** for 'a standing man' *(raġulan qā'iman);*
and **'ikī 'uṣlū 'ar kur·du·m** for '[I saw] two wise men' *(raġulayni*
'āqilayni). They satisfy themselves with putting the adjective in the dual
form [and do not need] the dual form of the described noun.
       2. And for 'wise men' *(riġālan 'uqalā')* [you say] **'uṣlū 'ar·an·lar.**
35   Its base form is **'ar·lar**, but they add a *n*. to it. Let us study the motive
for the addition of the *n*. In the case of **'aṭ** you say just **'aṭ·lar** and in
the case of **qul** just **qul·lar.** They satisfy themselves with putting the
described noun in the plural and [and do not need] the plural of the
adjective, i.e. **'uṣlu.** If the adjective is singular and the described noun

is not mentioned, the plural is attached to it. Thus you say ʾuṣlū·lar kur·du·m i.e. 'I saw the sage' *(raʾaytu al-ʿuqalāʾ)*. They may use ʾaran and with this the plural is meant. So they say ʾan·lar ʾar·an ʾuṣlū·lar·dur i.e. 'they are the wise men' *(hum ar-riǧāl al-ʿuqalāʾ)*. For

5   'I saw two standing men' *(raʾaytu raǧulayni qāʾimayni)* they say ʾikī ṭur·miš ʾar kur·du·m and '[I saw] standing men' *(riǧālan qāʾimīn)* ṭur·miš ʾar·an·lar kur·du·m. They suffice with the plural of the described noun and [refrain from using] the plural of the adjective. You do not pluralise, except when the described noun is deleted or when it

10 is used as a predicate, as in kiši·lar ṭur·miš·lar.

   3. If a noun in the genitive case occurs as an attribute to an undetermined word, you say for 'I visited a man from the Turks' *(daḫaltu ʿalā raǧulin min at-Turk)* kir·di·m bir ʾar qaṭ·i·n·ā turk tan. Its base form is qaṭ·i·n·ǧā, but [ǧā] is elided although it is permitted to use it.

15 Its [literal] meaning is 'I entered to with a man from the Turks' *(daḫaltu ʾilā ʿinda raǧulin min at-Turk.).*[112] Tan is in fact identical with dan and it expresses the meaning 'from' *(min)*.

   4. If a locative occurs as attribute to an undetermined noun, you say for 'I visited a man at yours [sg.]' *(daḫaltu ʾilā raǧulin ʿindaka)*

20 kir·di·m bir ʾar qaṭ·in·ǧā. Its base form is 'to with a man' *(ʾilā ʿinda raǧulin)*. If you mean 'I visited the chief' *(daḫaltu ʾilā l-ʾamīri)* you do not use qaṭ·in·ǧā <VD 61ʳ> since it is deleted in their language. Thus you say kir·di·m bay·ḵā.

   5. If a verbal sentence occurs as an attribute, with the meaning 'a

25 man stood up whom I loved' *(qāma raǧulun ʾaḥbabtuhu)* you say ṭur·dī bir ʾar kim saw·du·m ʾa·nī. Kim ties the sentence that occurs as an attribute to the described words that precede it, e.g. 'a man stood up whom I love' *(qāma raǧulun ʾuḥibbuhu)* ṭur·dī bir ʾar kim saw·ar man ʾa·nī.

30 6. If a nominal sentence occurs as an attribute, for 'a man stood up whose father is going' *(qāma raǧulun ʾabūhu ḏāhibun)* you say ṭur·dī bir ʾar kim ʾaṭa·s·ī kit·miš tar. Kim links the sentence that occurs as an attribute with the described [element].

   7. If a conditional sentence occurs as an attribute, for 'a man stood

35 up, if you [sg.] see him you will love him' *(qāma raǧulun ʾin raʾaytahu ʾaḥbabtahu)*, you say ṭur·dī bir ʾar kim kur·ar·saŋ saw ka·san. Kim links the conditional sentence that occurs as an attribute, with the described [element]. It is not permitted to put the noun in the genitive case at the beginning, nor the locative, nor the conditional sentence,

---

[112] The Arabic original is grammatically correct, since ʿinda 'with', a ẓarf, may be preceded by a ḥarf ǧarr

nor any other element when it occurs as an attribute to what is de-
scribed; you must put it at the end, as you do in the Arabic language.

8. If the attribute is either a derived or a non-derived noun, it needs
special analysis a part of which has been dealt with; the rest is yet to be
5   discussed in the description of the determined nouns.

9. If more than one attribute is used, they are put at the beginning
since <149> single adjectives must also be placed at the beginning.
Thus, for 'I saw Sangar the wise, kind writer' *(ra'aytu sangara al-kātiba
al-ʿālima al-karīma)* you say **ǧumart bilk̠ā bitikǧī sangar·nī kur·**
10   **du·m.**

10. Know that there are five kinds of determined nouns, as we men-
tioned in the Chapter on the Determined and Undetermined Nouns:
the suffixed pronoun, the proper name, the indicated word, the relative
and the noun that is annexed to any one of these [see above Section
15   25].

11. As for the suffixed pronoun, it cannot be described and you do
not describe with it. As for the proper name, you do not describe it and
you do not describe with it; but you may describe with a demonstrative
pronoun, a relative or an annexed noun.

20   12. An example of a description with the demonstrative pronoun
with the meaning 'I saw this Sangar' *(ra'aytu sangara hād̠ā)* you say
kur·du·m sangar·nī 'ušbū.[113] The most correct usage is **'ušbū san-
ǧar·nī kur·du·m**, in other words, the demonstrative pronoun pre-
cedes.

25   13. An example of a description with the relative for 'I saw Sangar
who stood up' *(ra'aytu sangara allad̠ī qāma)* you say **kur·du·m san-**
ǧar·nī 'ul kim t̠ur·dū.[114] <VD 61ᵛ> The attribute does not precede
here. If you mean 'I saw Sangar the coming [one]' *(ra'aytu sangar al-*
ǧā'iʾa) you say **kur·du·m sangar·nī kal·kan·nī.** It is also permitted to
30   say **kur·du·m kal·kan·nī sangar·nī.** The first [sentence] is according
[to the rules of] the attribute *(ṣifa)* and in the second [sentence] it be-
haves like an apposition *(badal).* For 'I saw Sangar who is in the house'
(ra'aytu sangara allad̠ī fī d-dāri) you say **kur·du·m 'aw·da·kī san-**
ǧar·nī, according to the [rules of] the substitution. It is permitted to
35   say **kur·du·m sangar·nī 'aw·da·kī**, according to [the rules of] the
attribute *(ṣifa).*

14. An example of the description with the annexed [noun], you say
with the meaning 'I saw Sangar, your [sg.] slave' *(ra'aytu sangara*

---

[113] Although the Arabic version of this sentence is grammatically correct, the
grammaticality of the Turkic sentence is disputable.
[114] idem.

ğulāmaka) **kur·du·m qul·uŋ sangar·nī**. This is the usage with the at-
tribute put at the beginning. You say for 'I saw Sangar, Sunqur's slave'
(raʾaytu sangara mamlūka sunqurin) **sunqur qul·u·nī sangar·nī kur
·du·m**. It is permitted to say **sangar·nī sunqur qul·u·nī kur·du·m**.

5   15. An example of the description of the indicated noun is 'I saw
this wise man' *(raʾaytu hādā l-ʿālima)* **kur·du·m ʾušbū bilka·nī**. The
description stands at the end and the marker of the object appears in
it. It is permitted [to say] **ʾušbu·nī bilka·nī**, as if it were used as an ap-
position *(badal)*. The former is used according to the [basic] construc-
10   tion *(ğiha at-tarkīb)*. Because of this the marker of the object only ap-
pears in the adjective.

16. An example of the description of the relative. With the meaning
of 'I saw the kind coming one' *(raʾaytu al-ğāʾiya al-karīma)* you say
kur·du·m kal·kan·nī ğumart·nī. It is permitted to say **kur·du·m
15   ğumart·nī kal·kan·nī**. And with the meaning 'I saw the coming one,
this one' you say *(raʾaytu l-ğāʾī hādā)* **kur·du·m kal·kan·nī
ʾušbu·nī**.[115]

17. An example of the description of the noun.annexed noun. For
'I saw Sangar's slave, Sunqur's brother' you say *(raʾaytu mamlūka
20   sangarin ʾahā sunqurin)* **sangar qul·u·nī sunqur qarindaš·ī
kur·du·m**, but it is more correct to say **sunqur qarindaš·ī sangar
qul·u·nī kur·du·m**.

53. Chapter on the conjunction

1. You say for 'Sangar and Sunqur stood up' *(qāma sangaru wa-sun-
25   quru)* **sangar sunqur tur·du·lar**. It is permitted [to say] **tur·du·lar
sangar sunqur**. They do not have a word that is literally equivalent
(turādif) to the conjunction 'and' *(wāw al-ʿatf)*. They use the two
nouns, the one following the other <150> without any link.

2. You say for 'Sangar stood up and Sunqur' *(qāma sangar fa-sun-
30   qur)* or 'and then Sunqur' *(tumma sunqur)* **sangar tur·dī ʾan·dan
sunqur**. ʾAn·dan expresses the notion of sequence without regard to
the [distance in] time. It is used instead of 'and' *(al-fāʾ)* and 'then'
(tumma). No particle in their tongue has the special function of either
particle, but they use **ʾan·dan** for the common meaning *(al-qadr al-
35   muštarik baynahimā)* of both of them.

3. <VD 62ʳ> With the meaning 'the people, even Sangar are standing
up' *(an-nās hattā sangar qāʾimūna)* you say **kiši·lar tur·miš·lar dur**

---

[115] These appear to be appositions rather than adjectives.

sangar taqī ṭur·miš dur [Lit.: 'the people are standing, Sangar is standing too']. **Taqī** expresses the meaning of 'even' *(ḥattā)*, but it is not its equivalent *(murādif)*, for its real meaning is 'too' *(ʾayḍan)* which includes 'even' *(ḥattā)*.

4. For 'Sangar or Sunqur stood up' *(qāma sangar ʾaw sunqur qāma)* you say **sangar mū ṭur·dī yā sunqur**. This is the way they use it, the particle **yā** being the equivalent *(murādif)* of 'or' *(ʾaw)*. For 'did Sangar stand up or Sunqur did stand up' *(ʾa qāma sangar ʾam sunqur)* you say **sangar mū ṭur·dī yā sunqur** [Lit.: 'did Sangar stand up or did Sunqur']. They say [also] **sangar mū ṭur·dī yuq sā sunqur**. This expression has come to mean 'did Sangar stand up, or did Sunqur' *(ʾa qāma sangar ʾam sunqur)*.

5. **Yuq** is originally a noun with the meaning 'lack' *(maʿdūm).*[116] This is proved by the fact that it is possible to annex to it. With the meaning 'poor' *(faqīr)* they say **yuq·lū** 'possessor of lack' *(ḏū maʿdūm)*. Annexation is one of the characteristics of nouns. [But] **yuq** came also to be used as the negating [particle] *mā*. As for **sā**, its meaning is 'if' *(ʾin)*, as if [the speaker literally] says 'did Sangar stand up, no if Sunqur' *(ʾa qāma sangar lā ʾin sunqur*. [Sā] corresponds to the meaning 'if Sunqur' *(ʾin sunqur).*[117]

6. You say for 'Sangar stood up, not Sunqur' *(qāma sangar lā sunqur)* **sangar ṭur·dī sunqur ṭur·ma·dī** [lit.: 'Sangar stood up, Sunqur did not stand up']. They do not have an equivalent *(murādif)* for our expression 'not Sunqur' *(lā sunqur)*. For the meaning 'Sangar stood up—not so, it was Sunqur' *(qāma sangar bal sunqur)* you say **sangar ṭur·dī yuq sunqur ṭur·dī**. For 'Sangar did not stand up but Sunqur [did]' *(mā qāma sangar lākin sunqur)* you say **sangar ṭur·ma·dī ʾawat sunqur ṭur·dī** [lit.: 'Sangar did not stand up, yes Sunqur stood up']. And for 'either Sangar, or Sunqur stood up' *(qāma ʾimmā sangar wa-ʾimmā sunqur)* you say **yā sangar ṭur·dī yā sunqur ṭur·dī** as I said in the case of 'or' *(ʾaw)*.

7. For 'I stood up and you' *(qumtu ʾanā wa-ʾanta)* you say **man ṭur·du·m san taqī ṭur·du·ŋ**. **Taqī** reflects the meaning of the conjunction for the plural and it reflects the meaning 'until' of comprisement *(ḥattā bi-t-taḍmīn)*, according to the preceding remarks. For 'I

---

[116] *maʿdūm*, in fact, is a passive participle meaning 'lost, gone'.

[117] The author here refers to the Turkic sentence in the preceding paragraph: **sangar mū ṭur·dī yuq sā sunqur**. Note that he does not consider **sā** a part of the verb (which here perhaps should be reconstructed as **yuq *i·sā** 'not be-COND' = 'if this is not'), but rather an independent particle, the equivalent of the Arabic conditional particles *law* and *ʾin*.

stood up and Sangar' *(qumtu ʾanā wa-sangar)* you say **man ṭur·du·m sangar taqī ṭur·dī** and **sangar ṭur·dī san taqī ṭur·du·ŋ** for 'Sangar stood up and you' *(qāma sangar wa-ʾanta)*.

8. For 'I stood up and then you [did]' *(qumtu ʾanā ṯumma ʾanta)*
5 you say **man ṭur·du·m ʾan·dan san**. And for 'you stood up and [then] I [did]' *(qumta ṯumma ʾanā)* or **san ṭur·du·ŋ ʾan·dan man** ['You stood up and then I did']. <VD 62ᵛ> For 'you stood up and then Sangar [did]' you say **san ṭur·du·ŋ ʾan·dan sangar**. For 'I ate the fish up to its head' *(ʾakaltu as-samaka ḥattā raʾsihā)* you say **bāliq·nī yidu·m**
10 **baš·in ġā daqin ġā**, from which it is permitted to elide **ġā**.

9. **Ġin·ğā** and **ḵin·ğā** are equivalents *(turādif)* of 'until' *(ḥattā)*. They are attached to the last [consonant] of the verb. **Ġin·ğā** is attached to velarised words and **ḵin·ğā** to palatalised words. As for the expression **dakin·ğā** <151>, its base form is **dak kin·ğā**. One of the
15 two *k*s assimilates with the other and is deleted with the purpose of alleviation. The past tense of **dak** is **dak·dī** meaning 'he reached' *(balaġa)*. If **ġin·ğā** or **ḵin·ğā** is joined to the [second] verb, the verb has the form of the imperative, but the meaning is the future. For 'stand here, until I come' *(qum hunā ḥattā ʾagīʾu ʾanā)* you say **ṭur**
20 **mundā man kal·ḵin·ğā**, as if the speaker said 'until come I' *(ḥattā ʾağiʾ ʾanā)*. With the meaning 'you stood until I came' *(qumta ʾanta ḥattā ğītu ʾanā)* they say **ṭur·du·ŋ san man kal·ḵin·ğā**; **ḵin·ğā** and **ġin·ğā** are attached to the verb.

54. Chapter on the corroborative

25 1. The corroborative *(taʾkīd)* may be expressed formally *(lafẓī)*; in that case it consists in the repetition of the word *(ʾiʿāda al-lafẓ)* itself. It occurs with single words and sentences.

2. In the case of single words, it occurs with a noun e.g. **sangar sangar ṭur·dī** ['Sangar Sangar stood up'], with a verb **sangar ṭur·dī**
30 **ṭur·dī** ['Sangar stood up, stood up'] and with a particle, e.g. **ʾawat ʾawat** ['yes, yes']. If the particle is connected to some [element], it is repeated together with it. Thus you say **ʾaw·dā ʾaw·dā ṭur·du·m** ['I stood up in the house, in the house']. When [it is used] in a sentence you say **sangar ṭur·dī sangar ṭur·dī** ['Sangar stood up, Sangar stood
35 up'].

3. The corroborative may also be expressed semantically *(maʿnawī)*. Sometimes it occurs with verbal nouns and sometimes with the restricting nouns *(ʾalfāẓ maḥṣūra)*.

4. Examples of its expression by a verbal noun are ṭur·du·m
ṭur·maq ['I stood a standing up'] and ya·du·m ya·mak ['I ate a
meal'].

5. In the case of the restricting nouns, the corroborative may be used
to emphasise the agent. The verb originates from it and not from
something else. For 'Sangar himself stood up' *(qāma sangar nafsuhu)*
you say **sangar ṭur·dī 'uz·ī**. It is also permitted to say **'uz·i bilā**. The
meaning of **bilā** in this construction is 'together with' *(ma'a)*, as if they
say 'together with himself' *(ma'a nafsihi)* i.e. 'with himself' *(bi-nafsihi)*.
Bilā is also used with the meaning of the *b*, as if the speaker says 'he
stood up with himself' *(qāma bi-nafsihi)*. But they do not build
affirmative sentences with the equivalent *(murādif)* of 'eye' *(al-'ayn)*,
so they do not say *****sangar ṭur·dī kuz·ī** [lit. 'Sangar, his eye stood up']
nor *****kuz·i bilā** [lit. 'with his eye'].[118]

6. Expressions that indicate inclusion with the meaning 'all' *(kull)*
are **barġā** and **qamuġ**. For 'the people all stood up' *(qāma an-nāsu
kulluhum)* you say **kiši·lar ṭur·di·lar barġa·lar** and also **barġa·lar·u**.
<VD 63ʳ> You may also say **qamuġ·lar** or **qamuġ·lar·i** i.e. 'all' *(al-
ġamī'u)* or 'all of them' *(ġamī'uhum)*, respectively. For 'all of this gold
was seized' *(hādā d-dahabu kulluhu qubida)* you say **bū 'altun
barġa·s·ī 'al·an·dī**. **'Al·an·dī** [sic; for 'alin-] means 'was taken'
('uḫida) and with this [word] you express the meaning 'was seized'
(qubida).

7. It is always permitted to emphasise an undetermined word,
whether it is divisible *(mutaġaziya)* or indivisible *(ġayr mutaġaziya)*.
For 'I ate a whole loaf' *('akaltu raġīfan kullahu)* you say **yi·du·m bir
'atmak·nī barġa·s·i·nī**. **Nī** may be elided from the first word [viz.
'atmaknī]. An example of an indivisible word is: 'a man stood up
himself' *(qāma raġulun nafsuhu)* **ṭur·dī bir 'ar 'uz·ī** and **'uz·ī bilā**. It is
obligatory to put the emphasising element after the confirmed ele-
ment, as it is in the Arabic language.

55. Chapter on the [syntactical] substitution

1. The [syntactical] substitution *(badal)* is divided into *totum pro toto
(kull min kull)*, *pars pro toto (ba'd min kull)* and comprehensive
(ištimāl).[119]

---

[118] The Turkic **kuzī** and **kuzī bilā** are calques of the Arabic expressions 'aynuhu 'his
eye', and *bi-'aynihi*, respectively. The latter means 'with his eye' and, as an idiomatic
expression, 'by himself'.
[119] Cf. Owens 1988: 324 'inclusive'.

2. In the case of the first you say 'Sangar, your [sg.] brother stood up' *(qāma sangar ʾaḫūka)* <152> ṭur·dī sangar qarindāš·iŋ.

3. In the case of the second you say 'I ate the loaf half' *(ʾakaltu r-raġīfa niṣfahu)* yi·du·m bir ʾatmak biğuq·nī and [you may also use]

5   yāru·s·nī. Yāru means 'a half' *(aš-šiqq)*. The *s* is is attached to indicate the annexation [sc. to a third person] when the last consonant is vocalised: nī indicates objectivity and the marker of the object appears in the second noun and not in the first.

4. In the case of the third you say 'Sangar, his beating made me hap-

10   py' *(sarranī sangar ḍarbuhu)* saw·in·dur·dī man·ī sangar ʾur·maq·i and 'his knowledge' *(ʿilmuhu)* bil·maq·i. This example may serve as a rule according to which you say saw·in·du·m sangar·dan ʾur·maq dan ['I was happy with Sangar, with [his] beating']. The *ʾAtrāk* also use saw·in·du·m sangar ʾur·maq·i·dan which means 'I was pleased with

15   Sangar's beating' *(surirtu min ḍarbi sangar)*. This concludes the Chapter on the Syntactical Substitution.

5. It seems that the substitution of the comprehensive is not used in their language so that it would be necessary to pay special attention to it.

20   56. Chapter on the condition

1. [Chapter on the condition *(šarṭ)*] and what is connected to it, as far as the rules of the equivalents of 'if not' *(law lā)*, 'if *(law)*, 'when' *(lammā)*, and 'whenever' *(kullamā)* are concerned.

2. The particle of condition is sā. This is its basic function *(al-*

25   *mawḍūʿlahu ʾaṣlan)*. If it corresponds to the conditional 'if *(ʾin)* it is a condition in the future tense. If it expresses what was going to occur because of the occurrence of something else, then it is the equivalent of the hypothetical 'if *(law)*. In that case it is a condition in the past tense. An example of [the condition in] the future tense is 'if Sangar

30   stands up, Sunqur will stand up' *(ʾin qāma sangar qāma sunqur)* sangar ṭur·sā sunqur ṭur·ğay. The apodosis *(ğawāb)* may be put at the beginning, e.g. sunqur ṭur·ğay sangar ṭur·sā. <VD 63ᵛ> The verb and the particle of condition are never separated *(lā yufṣal)*. The apodosis only occurs in the phonetic form of the future tense *(lafẓ al-mus-*

35   *taqbal)*; the conditional sentence only occurs in the phonetic form of the imperative *(lafẓ al-ʾamr)*. By this we do not mean a real imperative; the phonetic form of the imperative is basically used instead of the

phonetic form of the future tense and its apodosis. Thus you say 'if
Zayd stands up, ʿAmr stands up' (ʾin qāma zayd qāma ʿamr).[120]

3. If the verb in the conditional sentence is negated, you say
ṭur·ma·sā sunqur ['if Sunqur does not stand up']. So you insert the
particle of negation before the particle of condition.

4. If the verb agrees with the pronoun of a first person, you say
ṭur·sa·m ʾur·ġay·man i.e. 'if I stand up, I will beat' (ʾin qumtu
ḍarabtu); or with a second person you say ṭur·sa·ŋ ʾur·ġay·san i.e. 'if
you [sg.] stand up, you will beat' (ʾin qumta ḍarabta).

5. If the apodosis (ǧawāb) is a nominal sentence, you say for 'if
Sangar stands up, Sunqur will be standing' (ʾin qāma sangar fa-sunqur
qāʾim) sangar dah ṭur·dī ʾi·sā sunqur ṭur·miš dur ['If Sangar had
stood up too, Sunqur is standing up']. Sā is the particle of condition. ʾi
indicates the meaning 'he was' (kāna). The corresponding meaning is
'if he was stood up' (ʾin kāna qāma). This construction occurs in the
conditional sentence, if the apodosis is a nominal sentence.

6. [Sā] may be used if the verb is negated. In that case you say
sangar ṭur·sā sunqur ṭur·ma·ġay ['if Sangar stands up, Sunqur will
not stand up'].

7. [Sā] may also be used if the verb is an imperative, for 'if Sangar
comes beat him' (ʾin ǧāʾa sangar fa-ʾiḍrabhu) you say sangar kal·sā
ʾur·ġil ʾa·nī. You may also say ʾur ʾa·nī.

8. You may use [sā] if the verb is a prohibition. You say kal·sā
ʾur·ma·ġil ʾa·nī ['if he comes, do not beat him'] and ʾur·ma ʾa·nī. Ġil
is used with the prohibition, the same way it is used with the impera-
tive. Thus you say ya·ma·kil i.e. 'do not eat' (lā taʾkul), the same way
as you say <153> yi·kil i.e. 'eat' (kul).

9a. As an example of sā as an equivalent of the hypothetical 'if'
(law) in the past tense, [you may say] 'if Sangar had stood up, Sunqur
would have stood up' (law qāma sangar qāma sunqur) you say
[sangar][121] ṭur·miš mis·sā sunqur ṭur·miš dī.[122] The base form of dī
is ʾi·dī with the meaning 'he was' (kāna). The vowel [sc. i of ʾi·dī] was

---

[120] The author intends to point out that the conditional suffix sā is attached directly
to the stem which is best visible if the verb has the shape of the imperative form. In the
Arabic tradition, however, there is no concept of 'stem'. Furthermore, in his view the
verb has no inflection, since he considers sā equivalent to the Arabic particle of condi-
tion, rather than a verbal ending.

[121] Alternative from İÜ as emendation for sunqur.

[122] missā is a grammatical element unknown to me. It may be the suffix miš assimi-
lated with sā but this explanation would not satisfactorily account for the form ṭurmiš
missā where one would have to assume two subsequent occurrences of miš (cf. also
Čaykovskaya 1981: 102).

transferred to the silent consonant before it [sc. the *š* in **miš**] that
indicates the active participle. The [overall] meaning has become 'he
was standing' *(kāna qāʾiman)* in the apodosis, therefore the speaker
says **ṭur·miš·dī**. **Ṭur·miš** is the predicate to **ʾi·dī**. As for the condi-
5   tional sentence, it consists of [the words] **sangar ṭur·miš**, in which
ṭur·miš means 'standing' *(qāʾim)*. Furthermore, **sā** is the particle of
condition to which **mis** can be added.

9b. In this case [the first] *s* assimilates with the [second] *s* and thus
results <VD 64ᵛ> **mis·sā**. **Mis** is then added [to **ṭur·miš**] and thus the
10   whole comes to mean 'if Sangar had stood up, Sunqur would have.
stood up' *(law qāma sangar qāma sunqur)*. **Mis** is used as an extra el-
ement *(taʿlīq)* in the past tense. For 'if Sangar beat, Sunqur [would]
beat' *(law ḍaraba sangar ḍaraba sunqur)* you say **sangar ʾur·miš**
mis·sā sunqur ʾur·miš·i·dī.

15   10. Thus are the rules of this chapter. With these two constructions
the condition in the future and past tense is specified. For 'if you [sg.]
had stood up, I would have stood up' *(law qumta qumtu)* you say
ṭur·miš mis·sa ʾi·di·ŋ ṭur·miš·di·m [Properly: 'if you had been
standing, I would have been standing'], and for 'if I had stood up, you
20   [sg.] would have stood up' *(law qumtu qumta)* you say **ṭur·miš mis·sā**
ʾi·di·m ṭur·miš ʾi·di·ŋ.

11. For 'if you [sg.] stand up, I will stand up' *(ʾin qumta qumtu)* you
say **ṭur·sa·ŋ ṭur·ġa·man**; for 'if I stand up, we will stand up' *(ʾin*
ʾaqum naqum) you say **ṭur·sa·m ṭur·ġa·biz**[123]; for 'if you [pl.] stand
25   up, we will stand up' *(ʾin taqūmū naqum)* **ṭur·sa·ŋiz ṭur·ġa·biz**.

12a. For 'whomever you [sg.] beat, I will beat' *(man taḍrib ʾaḍrib)*
you say **kim·nī ʾar·sa·ŋ ʾur·ġa·man**. They use the particle of condi-
tion and the noun in one construction. In the Arabic language the
meaning of the particle of condition is included in the noun [i.e. *man*
30   'whomever']. In this language, however, the meaning of the condition
is separated from the noun and is expressed explicitly with a particle.

12b. For 'whenever you [sg.] visit me, I will honour you' *(kullamā*
ǧiʾtanī ʾakramtuka) you say **tayma kal·du·k·uŋ·ġā ʾaġirla·ġa·man**
san·nī. **Tayma** reflects the meaning 'all' *(kull)* and **kal·du** means 'he
35   came' *(ǧāʾa)*. The *k* in **kal·du·k·uŋ·ġā** is an addition *(zāʾida)*, whereas
the *n* [sc. ŋ] indicates the second person. The *m* of the first person may
be used on the same place. Thus for 'whenever I come' *(kullamā ǧiʾtu)*
you say **kal·du·ku·m·ġā**; 'whenever you [pl.] come' *(kullamā ǧiʾtum)*
kal·du·kuŋuz ġā; 'whenever they come' *(kullamā ǧāʾū)* **kal·du·k·**

---

[123] Reconstruction for **ṭur·sa·ŋ ṭur·ġa·san**.

lar·in·ǧā; 'whenever he comes' *(kullamā ǧāʾa)* kal·du·k·ǧā and kal·du
k·muz·ǧā [sic] 'whenever we come' *(kullamā ǧiʾnā)*. The [pronominal]
markers *(ʿalāmāt)* were placed between the *k* and [ǧā][124] and their
combination conveys the expression for 'whatever' *(mahmā)* in Arabic.
If you use **tayma** with this, it becomes to mean 'whenever' *(kullamā)*,
as if the speaker says 'whatever whenever' *(mahmā kullamā)*.

13. You say **sanǧar ṭur·iǧaq sunqur ṭur·dī** with the meaning 'when
Sanǧar stood up, Sunqur stood up' *(lammā qāma sanǧar qāma sunqur)*.
[Ǧiq][125] means 'at the time' *(ḥīna)*. And for 'when I stood up, Sanǧar
stood up' *(lammā qumtu qāma sanǧar)* you say **man ṭur·iǧaq sanǧar**
ṭur·dī [and for 'when Sanǧar stood up, I stood up' *(lammā qāma*
sanǧar qumtu) you say **sanǧar ṭur·iǧaq ṭur·du·m**][126] and for 'when
you [pl.] stood up, we stood up' *(lammā qumtum qumnā)* you say **siz**
ṭur·iǧaq ṭur·ṭu·q. Or you say [**biz**[127]] **ṭur·ṭu·q**.

14. <154> And for 'when they stood up, we stood up' *(lammā qāmū*
qumnā) you say <VD 64ʳ> **ʾan·lar ṭur·iǧaq ṭur·ṭu·q**; for 'when he
stood up, we stood up' *(lammā qāma qumnā)* you say **ʾul ṭur·iǧaq**
ṭur·du·q.

15. You say **quyaš baṭ·ṭi·ǧi·n·dā kal·ka·man**. [The element] **dā** in-
dicates 'in' *(fī)*, **ǧin** indicates 'at the moment' *(waqta)* and **ǧin·dā** is
equivalent to 'when' *(ʾiḏā)*. They use **quyaš**, i.e. 'beams of the sun'
(šuʿāʿ aš-šams) for 'the sun' *(aš-šams)*. Its meaning is 'I shall come to
you [sg.] when the sun sets' *(ʾaǧīʾuka ʾiḏā ǧābat aš-šams)*. You say
quyaš čiq·ṭi·ǧin·dā[128] **kal·ka·man** i.e. 'I shall come to you [sg.] when
the sun goes away' *(ʾaǧīʾuka ʾiḏā ḥaraǧat aš-šams)*.[129]

16. For 'if it had not been for Sanǧar, Sunqur would have stood up'
(law lā sanǧar qāma sunqur) you say **sanǧar daḵul mis·sā sunqur**
ṭur·miš·i·dī and for 'if it had not been for me, Sanǧar would have
stood up' *(law lā ʾanā qāma sanǧar)* you say **man daḵul mis·sa·m**
sanǧar ṭur·miš·i·dī. Its base form is **mis·sā ʾi·di·m**; but ʾi·dī was
elided and the pronoun has remained as a reference to the elided ele-
ment. The most correct form is the use of the base form without eli-
sion.

17. For 'if it had not been for us, Sanǧar would have stood up' *(law*
lā naḥnu qāma sanǧar) you say **biz daḵul mis·sa y·i·du·k sanǧar**

---

124 Addition from İÜ.
125 Addition from İÜ.
126 Addition from İÜ.
127 Em. İÜ of VD **bir**.
128 Em. İÜ of VD **ḥiq ṭiǧin dā**.
129 In fact, 'to you' is not explicitly stated in the Turkic sentence.

ṭur·miš·i dī; 'if it had not been for you [sg.]' *(law lā 'anta)*—san
dakul mis·sa y·i·du·ŋ; 'if it had not been for you [pl.]' *(law lā
'antum)*—siz dakul missa yidiŋiz;[130] 'if it had not been for him' *(law
lā huwa)*—'ul dakul mis·sa y·i·dī; 'if it had not been for them' *(law lā
5   hum)*—'an·lar dakul mis·sa 'i·di·lar.
        18. The pronouns from which depends the particle [that expresses]
the negation *(al-imtinā')* of existence precede [i.e. lā]. The outcome of
the study of mis·sā is that sā is the particle of condition.

57. Chapter on the quotation

10   1. [The quotation *(al-ḥikāya)* is, e.g.,] 'Sanǧar said Sunqur is standing'
(qāla sanǧar sunqur qā'im), you say sanǧar 'ayit·tī sunqur ṭur·miš.
You quote with the equivalent of what is said. In the imperfect tense
[they use] 'ayd·ur, which form is 'ayit·ur.[131] [In this verb] the t is
substituted for a d; the i together with the y is regarded as too heavy,
15   because of the change of i into u; the whole therefore is alleviated by
silencing the y.[132]
        2. The equivalent of 'he said' *(qāla)* is di·dī in Oǧuz, viz., nā·di·dī
'what did he say' *('ayya šay' qāla)*. In the imperfect tense [they say]
di·r, i.e., the d vocalised with an i.

20   VII

58. Chapter on the particles

1. The particles have already been mentioned here and there in the
preceding chapters about the rules [of their application]. <VD 65ʳ> We
will collect them in this section *(faṣl)*.
25   2. We mention <155> the particle of answer *(ḥarf al-ǧawāb)*, which
is 'awat, with the meaning 'yes' *(na'am)*, and yuq meaning 'no' *(lā)*.
We have already mentioned that yuq is a noun, for which we also have
given our arguments; it means 'lack' *(ma'dūm)*, even though it is used
with the meaning of 'no'. It is suitable for both the meaning 'no' *(lā)*
30   and the meaning 'lack' *(ma'dūm)*.
        3. The particle of negation is a m, vocalised with an a [i.e. ma].

---

[130] Em. of biz dakul missa y·i·du·k.
[131] Em. from ÎÜ for VD 'ayitt·ur.
[132] The argumentation can be schematically summarised as follows: 'ayit·ur →
'ayid·ur → 'ayd·ur.

4. The particle of prohibition is a *m* too, vocalised with an *a* [i.e. **ma**].

5. The particle of interrogation is a *m* vocalised with an *u*, it may be vocalised with an *i* because of the vowel harmony *(li-l-'itbā^c)* [i.e. **mu/mi**].

6. The particles of the imperative are **sun, yim** and **lim**.[133]

7. The particles of the annexation are **dan, dā, kā** and **birlā**. They also use **bilā** and **niŋ**.

8. The particles of conjunction are **'an·dan** and **yā**.

9. The particle of exception is **maqar**.

10. The particle of attention is **'uš**.

11. The particle of condition is **sā**.

12. The particle of both the verbal noun and motivation is **kim. Kim** has the meanings of a conditional noun, an interrogative noun, a relative and a particle of the verbal noun, with the meaning 'that' *('anna)* and the meaning of a particle of motivation with the meaning 'so that' *(kay)*.

13. The particle of wishing *(ḥarf at-tamannā)* is **kaškā**.

14. All of these are independent particles and [independent] words.

15. As to what is added for one of the meanings that occur as part of the root of another word, e.g. the diminutive, the plural, the transitive, or the construction of a passive form, we do not intend to list it in this chapter; since it has already been mentioned.

....

What we intended to put in this book has been completed. Praise be to God alone, His blessings be upon our prophet Muḥammad, his descendants and his companions and may
He grant him salvation.[134]

This book was finished on Monday the fifteenth of the venerable month of *Ša^cbān* in the year seven hundred and thirty five,[135] may God make it end happily.

---

[133] For the third and first person singular and the first person plural, respectively.

[134] İŪ reads: 'Praise be to God, He is my Measure. I am the servant of the Almighty God Aḥmad bn ʿUmar aš-Šāfiʿī may God be kind to him, the representative in the respectable government of the protected city of Lāḏaqiya, on the fourth of *Ǧumādī al-ʾūlā* of the months of the year eight hundred and five [/November 30th, 1402]. Praise be to God alone and His blessings be upon Muḥammad, the prophet of mercy and leader of the nation, on his descendants and his companions, may God grant him salvation. Our Measure is God. — The End'.

[135] This agrees with April 10th, 1335.

This copy was manufactured based on the author's own manuscript. He had written in the epilogue: '[This book] was finished in the night that had started with the morning of Thursday, the thirtieth of the glorious month *Ramaḍān*, in the year seven hundred and twelve,[136] at the

5   *Madrasa* of al-Malik aṣ-Ṣāliḥ in the land of Egypt. Its compiler, 'Abū Ḥayyān Muḥammad bn Yūsuf bn ʿAlī bn Yūsuf ibn Ḥayyān an-Nafazī al-'Andalusī, inhabitant of the land of Egypt, may God protect it, wrote it down with his own hand.'

.-.-.-.-.

---

[136] Ramaḍān 30th, 712 (February 6th, 1313) happened to be a Monday. Ramaḍān 30th, 711 (February 20th, 1312), indeed, was a Wednesday. It may be more than a coincidence that the autograph is said to be finished directly after the fasting.

VERBAL AND NOMINAL PARADIGMS AND SUFFIXES BASED ON DATA FOUND IN *KITĀB AL-ʾIDRĀK.*

VERBAL PARADIGMS

The following overviews provide the verbal and morphological forms as they are encountered in the grammatical parts of *Kitāb al-ʾIdrāk*. Wherever possible, forms were given for both front and back verbal stems, exemplified with the stems **kal**- and **ṭur**-, respectively, and separated by a slash /.

Present Tense
1a sg kal·ur·man
1b   kal·ur·um (Oġuz)
2 sg
3 sg   kal·ur
1 pl   kal·ur·biz
2 pl   kal·ur·siz, -siz·lar
3 pl   kal·ur·lar

Present Tense Interrogative
1 sg   kal·ur·mu·man
2 sg   kal·ur·mu·san
3 sg   kal·ir·mū
1 pl   kal·ur·mu·biz
2 pl   kal·ur·mu·suz
3 pl

Present Tense in stems ending in -lā
and monosyllabic verbs ending in a vowel
3 sg   suz·lā·r - suz·lā·yur; si·r - si·y·ur

Present Tense Conditional
(1 occurrence)
2 sg   kur·ar·sa·ŋ

Present tense Negated
1 sg   kal·ma·s·man
2 sg   kal·ma·s·san
3 sg   kal·ma·s
1 pl   kal·ma·s biz
2 pl   kal·ma·s·siz; -lar·siz; -siz·lar
3 pl   kal·ma·s·lar

Past Tense
1 sg   kal·di·m, -du·m; -du·man
2 sg   kal·di·ŋ
3 sg   kal·dī; kal·du / ṭur·dū, -dī
1 pl   kal·du·k, -·lar -·miz, -biz /ṭur·ṭu·q
2 pl   kal·diŋiz, -di·ŋiz·lar
3 pl   kal·di·lar/ ṭur·di·lār

Past Tense Negated
1 sg   kal·ma·du·m
2 sg   kal·ma·di·ŋ
3 sg   kal·ma·dī
1 pl   kal·ma·du·k, ma·duk·lar
2 pl   kal·ma·di·ŋiz
3 pl   kal·ma·dī·lar

Past Interrogative
1 sg   kal·di·m·mū
2 sg   kal·di·ŋ·mū
3 sg   kal·di·mū
1 pl   kal·du·k·mū
2 pl   kal·di·ŋiz·mū
3 pl   kal·di·lar·mū

Future Tense
1 sg   kal·ka·man / ṭur·ġa·man
2 sg   kal·ka·san, kay·san / ṭur·ġa·san
3 sg   kal·kā, kal·kay / ṭur·ġā, ṭur·ġay
1 pl   kāl·ka·biz / ṭur·ġa·biz
2 pl   kal·kay·siz; -siz·lar
3 pl   kal·kay·lar

Future Tense Negated
1 sg   kal·ma·kay man
2 sg   kal·ma·kay·san
3 sg   kal·ma·kay
1 pl   kal·ma·kay·biz
2 pl   kal·ma·kay·siz
3 pl   kal·ma·kay·lar

Future Tense Interrogative
1 sg   kal·ka·mu·man

2 sg   kal·ka·mu·san

3 sg   kal·ka·mū
1 pl   kal·ka·mu·biz
2 pl   kal·ka·mu·suz
3 pl   kal·ka·lar·mū

Future tense (Oġuz)
1 sg   kal·a·m
       (>kal·a·man)
2 sg   kal·ā·sin
       (neg.: -mi·yā·sin)
3 sg   kal·ā
1 pl   kal·ā·wuz
2 pl
3 pl   kal·a·lar

Imperative

1 sg	kal·kā·yim	
2 sg	kal·ķil / ṭur; ṭur·ġul	
3 sg	kal·sun	
1 pl	kal·kālim (Oġuz?)	
2 pl	/ ṭur·uŋ, -uŋ·uz; -uŋ·siz	
3 pl	kal·sun·lār	

Conditional Verbs

1 sg	ṭur·sa·m
2 sg	ṭur·sa·n, ur·sa·ŋ
3 sg	ṭursā (negated: ṭur·ma·sā)
1 pl	
2 pl	ṭur·saŋiz
3 pl	būl·sa·lar

Prohibition

1sg	kal·ma·kāyim
2sg	kal·mā / bar·mā·mā [1]
3sg	kal·ma·sun
1pl	kal·ma·kālim
2pl	kal·ma·ŋiz, -ma·siz
3pl	kal·ma·sun·lar

Finite 'Konverbial' Form:
'whenever I come' or
'each time I come'

1 sg	kal·duk·um·ğā (Oġuz?)
2 sg	kal·duk·uŋ·ğā
3 sg	kal·duk·ğā (sic)
1 pl	kal·duk·umuz·ğā
2 pl	kal·duk·uŋuz·ğā
3 pl	kal·duk·lar·in·ğā

---

[1] 2sg with extra **mā** has a contemptuous connotation, with the meaning of 'woe unto you' *(waylaka)*. EDT 765b "here you are". 'Abū Ḥayyān gives three examples, all of them on back verbal stems, 'ulṭur-, yaṭ-, bar-.

NOMINAL AND VERBAL SUFFIXES

Pronouns		Possessive Endings		
1 sg	man	1 sg	-(I)m	**qul·um**
2 sg	san	2 sg	-(I)ŋ	**qul·uŋ**
3 sg	ʾūl	3 sg	-(s)I	**qul·ī / ʾaṭa·sī**
1 pl	miz/biz or mizlar/bizlar	1 pl	-mIz	**qul·umuz**
2 pl	siz	2 pl	-ŋIz	**qul·uŋuz**
3 pl	ʾalār/ʾanlar (ʾular, bular marked Oġuz)	3 pl	-lArI	**qul·lar·i**

I Suffixes attached to nominal stems

Plural Suffix

-lAr	ʾanlar 105; baliqlar 106.
-n	ʾaranlar 148.

Comparative

-rAK	comparative suffix; **birğīrak** 108; **barğīraq** 108 **yaman-raq** 107.

Relative

-ki	relative suffix ʾawdākiniŋ 119;

See also verbal suffix -GAn

Case Endings

-nIŋ	genitive **bāliqniŋ** 147.
-Iŋ	genitive Oġuz variant **sanğariŋ** 145.
-GA	dative ʾavġā 145; **bayk̲ā** 139.
-A	dative after possessive ending 3sg ʾaṭasinā 141.
-nI	accusative **bāliqnī** 139.
-n	accusative after possessive ending 3sg. **biranğasin** 147.
-DA	locative ʾawdā 119.
-DAn	ablative ʾawdan 145; (assimilated ʾaṭṭan 143)

Numerals

-(A)GI	collective **bīšakū** 114, ʾalṭaġū 115.
-(A)wI	collective ʾīkawu 115 Final vowel probably possessive

	suffix. EDT 364a.
-In	distributive **birin birin** 144.
-lA-	in collective nouns with 3sg poss suff. **'uǧawlaŋiz** 'the three of you', **'ikalasī** 'both of them', etc.. According to 'Abū Ḥayyān used with possessive suffixes. Not used with numerals above **sakiz** 'eight'. EDT 25b derives it from a Mongolian collective suffix.
-(I)nǧI	ordinal numeral suffix **durdunǧī** 144, **bašinǧī**144.
-mĬš	occurs in the Turkic words for 'sixty' and 'seventy' and is therefore associated with 'the multiplication by ten' **'alṭmiš** 115.

Denominal Verbs

-Ar-	denominal in colours **'aǧardī** 112.
-lA-	denominal verb suff. **baǧašla-** 104; **utrukla-** 115
-nA-	? unknown, only in ṣamas-nadī 113, there is an obvious relationship to ṣamaṣ.

Miscelleaneous:

1. -(A)ǧIK	diminutive *turkmānī* variant **'azaǧuq** 105; **'azǧuq** 105.
2. -GIKAs	diminutive **'at ǧuqas** 105.
3. -GA	idem **'anǧā** 147; **'aqǧā** 114.
4. -KInA	idem. **'atkinā** 112, apparently also suffixed to pronouns and plurals: **'anlarkinā** 105.
-ǧI	occupation **'ašǧī** 113.
-lAyIn	Equational suffix **'arslan layin** 128.
1. -lIK	rank or possession **baylik** 109.
2.	purpose or destination **bitilik** 110; **ṭunluq** 109.
3a.	characteristic or equipment **malluǧ** 106.
3b. -lI/-lIh	idem; probably *turkmānī* variant: **'aŋluh** 137
4a. -lIK	relational suffix **rūmluǧ** 105.
4b. -lI	idem. *turkmānī* variant **rūmlū** 105.

II Suffixes attached to verbal stems

Imperative

-GIl	attached to 2 sg **'urǧil** 152.

Adjectives

-(I)K	with semantic passive connotation **'aǧuq** 107.

Instruments
- G   instrument **'alak̠** 109.
- GI   idem. **'alak̠ū** 109; **biǧqū** 109.

Nouns
-AǧAK   indicates a place **kalaǧak** 109.
-AK   idem **ṭuraq** 123.
-AǧI   indicates permanence of adjective **bar-aǧī** 106
-KA   indicates a characteristic or occupation **bilk̠ā** 106.
-wIK   ? there is one single occurrence of this suffix, which
                Abū Ḥayyān relates to a verbal stem **ǧatlawuk** 116.

Verbal suffixes

a. causative suffixes:
-Ar   causative suffix **ǧiqar-** 110, **kaǧur-** 117.
-DIr   idem. **bildir-** 104 ; **kaǧdur-** 117, **ǧiqtur-** 110.
-zIr–   idem. **'amzur-** 110.
-GIz   idem. **ṭurǧuz-** 110.
-t-   causative suffix after stem ending in vowel **dapalat-**
                110; **yurit-** 110.

b. passive, reflexive and reciprocity:
-(I)š   suffix of reciprocity **ṣanǧiš-** 111.
-Il-   passive suffix **kasil-** 110; **biril-** 107.
1. -n-   passive/reflexive suffix **'ayar lan-** 129; **dapalan-** 107.
2. -n–   medio-passive form **'alin-** 151 **yaṣṭan-** 110.

Gerunds
a. non-finite
-A   a certain state or condition while the action is taking
                place **kulā** 137.
-GAlI   'since' **barǧalī** 145.
-KčA   repeated action (one occurrence: **kaldukǧā** 115).
-GinčA   indicates a parrallel action **kalk̠inǧā** 151.
-(y)U   idem. **kalū** 137; **'aǧlayū** 137.
-mAdIn   without... **'ayarlamadin** 138.
-mAyIn   idem. Qipčaq variant **'ayarlamayin** 138.
-(y)(I)p   expresses a simultaneity or state **'aǧlab** 113; **'ayarlab**
                138; **'aǧlayub** 113.
-(y)IpAn   idem. **'aǧlayuban** 137; **biniban** 138.

-IptIr past tense 'ayarlanubtur 129.

b. finite[2]

-AGAn indicates an intensive active participle baraġan 107.

-(y)IġI permanence of adj. part.biriġī 106; kilaġi 106.

-DAġI future participle bardaġi- 146.

-(i)ġAK According to 'Abū Ḥayyān this suffix indicates a simultaneity, or a parallel action; it probably also expresses an intention to carry out the action; ṭuriġaq 154.

1. -GAn indicates the active participle past tense barġan 112, dapalankan 112.

2. -GAn used for relative phrases in transitive verbs kurkanimniŋ 146.

-DIKImġA expresses repeated action, conjugated. kaldukumġā 153, etc. see verbal paradigms

-DIGIndA consists of -DIK+poss(+n)+LOC the typically Oġuz variant of 1. -GAn; baṭṭiġindā 154.

Infinitive Forms

-m infinitive, also used as verbal noun for one instance of the action barim 108

-mAK infinitive. barmaq 114; (can be used with possessive suffixes sawmakimdan 141).

-mAKlIK idem, renders the infinitive more abstract barmaqliq 114.

-(y)Iš infinitive, also used as verbal noun for one instance of the action yiyiš 134.

Tenses (see also verbal paradigms)

-Ir present tense bašlār 121; binar 128, 'alur 112.

-DI past tense (Oġuz always Dī) 'aġladī 138; biġtī 109, 'aldum 147.

-GA future tense see paradigmata.

-yIr present tense when stem ends in lā suzlāyur 121, or of monosyllabic verbs ending in a vowel siyur 111

-mIssA untraced suffix 'awdāmissa 113; cf. *Durra* 20ʳ 12 (ed. Zajączkowski 80 ḫūš missā)

---

[2] These gerunds are finite in the sense that they can be followed by a pronoun, or possessive.

-mIš	dubitative; inferential **dapalanmiš** 107, **duġmuš** 106.
-mIš yIdI	past perfect **duġmiš yidī** 137.
-mAdIK yIdI	past perfect negated **duġmaduq yidī** 183.
-mAGA	future tense negated 1sg **barmaġaman** 146.
-mAs	Present Tense Negated **kalmas** 113.

Corroborative
-DIr	corroborative element 122.

APPENDIX TWO

LIST OF TURKIC WORDS

The following list is meant in the first place to serve as a reference list of Turkic words that occur in the grammatical parts of *Kitāb al-idrāk li-lisān al-Atrāk* (ed. Caferoğlu 101-154).

A further reason for listing these words is the fact that they thusfar have hardly been included in turcological studies on the 14th century Mamlūk sources, as most manuals and dictionaries only refer to the listing of Turkic in the lexicographical part (ed. Caferoğlu 1931:1-100). However, the material in the gammatical part provides many additional lexical items, including some whose meaning and etymology I have not been able to determine in the available sources. I have aimed at completeness as far as the occurrences and loci of the entries is concerned, although for a given form only one place is cited.

In the material was found a large number of words with **b-**, **d-** and **-ġ-**, most of which cannot be of Qipčaq origin and probably are of Oġuz provenance, e.g., **bunila-**, **dilkū**, **dapi-**, **durdaḳū**. Typically Oġuz is also the verb **ġumal-**.

In first instance reference was made to Abū Ḥayyān's word list (abbreviated as WL; in Caferoğlu pp. 1-100). Further, I have attempted to verify all words in Clauson's (1972) EDT; in those cases wher it does not give sufficient information other sources were consulted, i.e., Codex Comanicus (CC), Doerfer 1976, Räsänen (Räs.), Sewortian (Sew.). Oġuz/Ottoman material was verified in: Derleme Sözlüğü (DS), Redhouse (Redh.), Steuerwald (1972) (Steuerw.), Tarama Sözlüğü (TS), and Söz Derleme Dergisi (SDD). Occasional Persian loan-words were verfied in Steingass.

The transcription system used here is explained in Part One, Chapter One of this book. All entrances are listed according to their respective stems. If a noun or verbal stem as such occurs in the text, the translation is immediately followed by the respective page number in Caferoğlu's edition. References to derivational forms are given only once with indication of the page number on which they occur first.

ʾaqġā little coin 114; WL 17; Doerfer: 506; ʾaqġanī 142.

ʾaqša– to caress WL 17 ʾaqša 104.

ʾar man 112; ʾardur 133; ʾark̲ā 147; ʾarnī 141.

ʾara between 101; WL 11.

ʾaran men properly pl., also used as sg. 112; ʾaranlar 112.

ʾaraslan lion 103.

ʾard back, backside 101; ʾardiŋdadur 122; ʾardinda 135; ʾardiŋdan 135; ʾardimdā 136; ʾardimizdā 136; ʾardiŋdā 136.

ʾar– see ʾur–.

ʾari bee 101; cf. WL 10: ʾarū.

ʾarmaġan gift 122; EDT: a Pers. loan-word in early Oġ..

ʾarqataq ? 102; WL 11: ʾarqatāq 'scull with a long nose'? (ǧamǧamā muṭāwala al-būz).

ʾarslān 122; see ʾaraslan.

ʾāš food 142.

ʾāšġī cook 105; ʾašġī 113.

ʾat flesh, meat 118; also ʾaṭ.

ʾaṭ horse 112; ʾāṭ 129; ʾaṭlar 106; ʾaṭkinā 112; ʾaṭṭan 143.

ʾaṭ– to throw; ʾaṭ 103.

ʾaṭ ġuqas little horse 105; -t is not indicated as back, inferred from the velar suff. WL 15.

ʾaṭa father; ʾaṭasī 113; ʾaṭasinā 141.

ʾatak skirt; edge 101.

ʾatmak bread 102; ʾatmaknī 142.

ʾav game; hunting WL 25; ʾav 'velarised'. ʾavġā 145.

ʾaw tent, house 145; ʾawdāmissa 113; ʾawdadur 119; ʾawdā 119; ʾawdādur 119; ʾawdākilārniŋ 119; ʾawdākiniŋ 119; ʾawdakī 138; ʾawdan 145; ʾawk̲ā 145; ʾawda kiniŋ 146.

ʾawat yes 150.

ʾayakū rib 101; WL 27.

ʾayanla– ?perhaps variant of ʾuyanla.

ʾayarla– to saddle; ʾayarlab 138; ʾayarlamadin 138; ʾayarlamayin 138.

ʾayarlan– to be saddled; ʾayar lanubtur 129.

ʾaybak 108; prop. name comp. of ʾay 'moon' and bak 'chief'.

ʾayduġdī 118; prop. name comp. of ʾay and duġdī.

ʾayit– to make speake; ʾayittī 154.

ʾayittur– to make speak 154; cf. ʾaydur–.

ʾaykanla– see ʾayanla and ʾuyanla. ʾaykanla 104.

ʾaykir– to twist see ʾakir.

ʾayruq different; other 143.

ʾaz few 105; ʾazaġuq 105; ʾazġuq 105.

balṭir calf of the leg 102.

babaġuk eye-apple 102; WL 28: babaġuk ʾinsān al-ʿayn 'person of the eye'.

baġā frog 118.

baġašla– to give; baġašla 104.

bak̲ chief 126; WL 35 bak̲lardur 119; bīk̲ā 129.

baliq fish; baliqlar 106; bāliqliq 114; bāliqnī 139; bāliqniŋ 147; baluq 120; WL 36: baliq.

balkur– to appear; balkurdī 112.

balkū sign 112.

baq– to look at; baqtim 139.

baqiš look, glance 108.

bar there is 132; WL 29.

bar– to go, to leave; baraġī 106; baraġan 108; bardī 111; barġan 112; bariġī 113; barmaq 114; barmaqliq 114; barġil 115; barmā 117; barġālim 121; barġāyim 121; bardum 139; barġalī 145; bardaġidur 146; barmas-

man 146; barmadim 146; bar-
maǧaman 146.

barǧa all; barǧā 101; barǧamiz
106; barǧasinī 147; barǧalar
151; barǧalaru 151; barǧasinī
151; barǧasī 151.

barǧīraq more going 108

barki– ?to be strong WL 30; WL
34: barki 104.

barū since 145.

baš head; baši 146; bašin ǧā 150.

bašawu each five 116; see bīšakū.

bašinǧī fifth 144.

bašla– to begin; bašlār 121; baš-
lāyur 121; bašladī 127.

bašmaq shoe 116; WL 31.

baṭ– to sink; (sun) to set; baṭṭi-
ǧindā 154.

baṭman a unit of weigth 144.

bay chief 108; see baḵ/bak; baynī
128; bayḵā 139.

baylik the rank or possessions of
a baḵ/bay 109.

baza– to ornament; baza 103.

biki variant of kibī

biǧqū knife 109.

biǧuq half; biǧuqnī 152.

biǧ– to cut; biǧtī 109.

bil– to know; bil 127; bildī 128;
bildum 129; bilmaqi 152; (sic).

bilā with 122; see birlā.

bila– to sharpen; biladī 109.

bilakū whetstone 102; WL 35.

bilazik bracelet 102.

bildaraq ?leaf 103; WL 35; *Bulǧa:*
"*yābüldüräk* 3,11… feuillage'
(coll.)." EDT: yapuldurak.

bildirǧin quail 103; Acc. to WL 36
this form is Oǧ., the Qip. form
being buyurǧun.

bildir– to inform; bildir 104;
Transcr. as in İÜ; bildirdum
129.

bilḵā wise man 106.

bin– to mount; binar 128; binib
138; biniban 138; binū 138; see
also min–.

bir one 141; birin birin 144.

bir– to give; biriǧi 106; birǧirak
108; birdī 111; birdum 139;
birdaǧidur 140; birmadī 132.

birakū single 115; Sew.; TS
bireǧü; birawuwu 116.

biril– to be given; barildī 115;
barildaǧī 116; barilmiš 116;
birildī 107; birildaǧī 107; bir-
ilmiš 107; birilkan 112; biril-
kān 116.

biriṇǧī first 144.

biriš act of giving 108; act of giv-
ing.

birlā together with 117 (see bilā)

bisū ?education 101; WL 2 7
('ayanū) derives it from bisla-
'to raise'; also WL 31: bis/bisū
both back.

bīšakū all five 114.

biši– ? 103; see yiši–.

biti anything written, book, let-
ter 101; bitiṇ 133.

biti– write; bitidim 145.

bitikǧi scribe 149; Doerfer 718.

bitilik something to be written
on 110; bitilik 114.

biyiš ? 134 perhaps yiyiš, see yi–.

biz we 113; biziṇbilā 146; bizdan
145; biznī 139.

bizkak cold fever 102; WL 31.

bū this 105; būkinā 105; bular
118; bulār 118; bundā 119;
bunǧuq 105; bunǧuqas 105;
'ibunī 1 2 1 ; bunlarǧuq 105;
bunlarǧuqas 105; bunuṇ 120.

buda– to prune a tree; buda 103;
Although the Ar. orth. indi-
cates that buda- is a front
word, it is prob. back. Initial
d- points to Oǧ..

buǧā bull 101.

buǧdāy wheat 137; buǧday 137; WL 33-4 buǧday but buyday as Qip.; mod. Qipč. lang.—Tat. boday; Nogai/Kklpk biday; Bašk. bïyday; Karaim/ Kumyk buday; Kaz. biyday—buǧday is prob. Oǧ.. cf. Sew. 232-3.

buk- to bend; Buktī 118; prop. name.

bukrī bent 102.

bulǧa- to stir 104.

bulqi- ? to shine WL 36: balqidī; Sewortian: "balqı" bulqi 104.

bunila- ?to reproach cf. WL 36; bunila- with b- is a typically Ott. form; bunila 104.

buqawu handcuff 102; buqawu not on WL 33 (EDT) but on WL 34.

burnī 140; see burun.

bursuq badger 102; WL 29.

burun nose, in certain constructions also 'in front'; 'before' 136.

buṣariq ?dust, haze 102; WL 32; Steuerw. "pusarık neblig".

buzar- to become grey (?ḇ-); buzardī 112; buzardī 138

buzǧuš ?grey bird, probably gerfalcon 116; WL 31.

ḇul- to become; to be; ḇuldī 123; ḇuldī 143; ḇūldīlar 125; ḇulduŋ 124; ḇūlsalar 125; ḇuluban 137; ḇulubun 137.

ḇuz grey 112; WL 30.

čaǧan ?falcon WL 43; čaǧanǧuq 105.

čaǧiltilaq ?dry dirt 103; WL 44'; DS. But Doerfer 1095 "Kinderspielzeug".

čakalan- ?to heal slowly WL 44; čakalan 104; .

čališ ?fight 109; Bulǧa 6,10.

čaštula- ?to spy WL 43; čaštula 104.

čatla- ?to make a cracking sound WL 42; čatladī 116.

čatlawuk (hazel)nut 102; see čatla-. TS "çatlağuç"; Doerfer (1069) "čatlāqūč 'die Früchte des Mastix-Baumes'.".

čī ? 101; WL 41: čī 'fat milk raḥm (?)'; probably EDT: "çig 'moist,'", where WL čī is not given but čik on WL 44.

čïbǧiq sparrow 102; WL 41; Doerfer 1146; Bulǧa: "čypčaq 11,12".

čiǧanaq elbow 102; in İÜ with ǧ; WL 44; Räs.: "čykanak".

čiq- to go out; čiqtum 145; čiqtiǧindā 154.

čiqar- to bring out 'to bring out, send out'." čiqar 110.

čiqtur- to bring out; čiqtur 110.

čiz- to scratch; to draw; čizdim 145.

čumal- to squat on one's heels WL 46. Steuerw. "çömel- sich hinkauern (nach orientalischer Sitte auf den Fersen)"; čumal 104.

čuri- to decay WL 43; 121. čuridī 107; čurimiš 107; čuridaǧī 107.

čuruk rotten 107; Doerfer 1078.

čuruyiǧī decaying 106; cf. čuri-

čuzi- ? çöj - çöz?. čuzi 103.

čuwalduz ?packing needle 103; WL 45.

dabi- ?to dry up dapi 103; perhaps DS "depimek a slightly moist object, to begin to dry up"; exists in several forms; in Oǧ. usually t-/d- (Az. tepi-); Ttü./Ott. d-, in Qip. languages (Tat/Kaz) forms with k- prevail.

dah too 152.

dak– to reach; dak 1 5 1; dakdī 145; dakinǧa 150.

dakā he-goat 117.

dakī until like 145; (see also daqin).

dakmā until 147; see tayma.

dalū mad 117.

damrakū a skin disease 102.

damur iron 116.

dapala– to kill WL 47; d- perh. indication for Oǧ. provenance. *Bulǧa* II 51; Redh. tepele-; dapala 104.

dapalan– to be killed; dapalandaǧī 107; dapalankan 112; dapalanmiš 107.

dapalat– to make kill; dapalat 110.

daqin until like 128; altern. for daqī, which, in its turn, may be an emphatic form of dakī; cf. Kaz. -day; daqin ǧā 145.

dawāquš ostrich 118; compound of dawā 'camel' and quš 'bird'; WL 51.

dakul is not 122; dakulmissā 113.

dik– to plant; diktum 143.

dilkū fox 117.

diltaq ?reason 117; CC: "syltov" Räs. "šiltä".

diluk hole 102; in İÜ Ms. daluk.

dirak column 142.

dirkī portable table 102.

diz knee 116.

di– to say acc. to AḤ this is Oǧ.; recurs on WL 48; didī 154.

duǧ– to be born; duǧmuš 106; duǧmušmidur 131.

dun night; yesterday; budundur 122.

dunā yesterday 145; see tun.

durdakū all four 114; durdawa 116.

durdunǧī fourth 144.

durt four 143.

duš midday 135.

duša– to spread out; duša 133.

dušan– to be spread out; dušan 135.

dutun smoke 116.

farman order 1 0 1; em. for Caferoǧlu farmaz.

firišti angel Telegdi 1938: 309; firištilār 101.

ǧān soul; ǧānī 146.

ǧumart generous 149; Redh. 735.

ʾi– 1 to be; ʾidī 118; ʾidiŋ 123; ʾidiŋiz 123; ʾidum 145.

ʾi– 2 to send; ʾidī 121.

ʾibunī 121; see ʾi- 2. Note b- for m-.

ʾiǧ– to drink; ʾiǧ 103; ʾiǧmišman 146.

ʾiǧakū all three 114; i- perhaps a scribal error for u-.

ʾiǧdur– to make drink; ʾiǧdur 117.

ʾiǧur– to make drink Oǧ. forms in -ur, Qip. languages: Tat. čer-; Kaz. iškiz-; ʾiǧur 110.

ʾīkakū both 114; see also ʾiǧakū, durdakū, bišakū ʾalṭaǧū and yidakū; ʾīkawlasī 1 1 6; ʾīkawu 116; ʾīkawkaŋuz 116.

ʾikala– to file, polish 104; WL 18; Räs. ykyla 86.

ʾikawkamuz both of us 116; see ʾīkakū.

ʾiki two; ʾikiyuzar 143.

ʾikila– Perh. to do something twice; ʾikila 104.

ʾikin ʾikin two by two 144; see ʾiki.

ʾikindī afternoon (prayer) 102.

ʾikinǧī second 144.

ʾilk first 136.

ʾilkarū forwards; before 136.

ʾilkinǧī first 144; see ʾilk.

ʾimdī now 117.

ʾinǧū pearl 102.

ʾiqša– ?to creak WL 17; ʾiqša 104.

ʾira– to rock WL 11; ʾira 103.

ʾirdab measure 144.

ʾišittur– to make hear; ʾišitturdī 116.

ʾišit– to hear; ʾišitk̟il 121.

ʾisillik warmth 102; WL 13.

ʾisirǧā earring 102; WL 13; Räs. syrga; Salar sˊrgẓa WYugur sırqa.

ʾit dog 112; ʾit ǧukas 105; ʾitǧukas 113; ʾitlar 112.

kačā night 135.

kaǧdur– to make pass; kaǧdur 117.

kaǧur– to make pass; kaǧur 117.

kal– to come; kalaǧak 109; kalakan 108; kalalar 129; kalam 130; kalaman 130; kalasan 130; kalā 129; kalāsin 129; kalāwuz 130; kaldilar 112; kaldilarmū 131; kaldim 141; kaldimmū 131; kaldimū 131; kaldiŋ 129; kaldiŋiz 129; kaldiŋizlar 129; kaldiŋizmū 131; kaldiŋmū 131; kaldu 153; kaldukǧā 153; kalduklarin ǧā 153; kaldukumǧā 153; kalkumuzǧā 153; kaldukuŋuz ǧā 153; kaldubiz 130; kalduk 130; kaldukǧā 115; kalduklar 130; kaldukmū 131; kaldukuŋǧā 153; kaldum 130; kalduman 130; kaldumiz 130; kaldī 114; kaldī kulā 137; kalib 113; kaliǧī 113; kalirmū 131; kalkalarmū 131; kalkaman 130; kalkaman 154; kalkaman 154; kalkamuniz 131; kalkamusan 131; kalkamusuz 131; kalkamū 131; kalkan 112; kalkasan 129; kalkay 129. kal-

kaylar 129; kalkaysan 129; kalkaysiz 129; kalkaysizlar 129; kalkā 121; kalkālim 121; kalkāyim 121; kalk̟il 115; kalk̟inǧā 151; kalmadiŋ 130; kalmadiŋiz 130; kalmaduk 131; kalmaduklar 131; kalmadum 131; kalmadī 130; kalmadīlar 130; kalmak 114; kalmakay 130; kalmakay man 131; kalmakaybiz 131; kalmakaylar 130; kalmakaysan 130; kalmakaysiz 130; kalmaklik 114; kalmas 113; kalmas biz 131; kalmaslar 113; kalmaslarsiz 130; kalmasman 131; kalmassan 113; kalmassiz 113; kalmassizlar 130; kalmiyāsin 130; kalsā 152; kalsun 120; kalsunlār 120; kalub 138; kalur 111; kalurbiz 130; kalurlar 129; kalurman 130; kalurmuniz 131; kalurmusan 131; kalurmusuz 131; kalursiz 129; kalursizlar 129; kalū 137; kālkabiz 130.

kamulduruk strap of a saddle 103; CC 151 comuldruc.

kas– to cut; kastī 109; kasdī 110.

kasakū knife 114; see kaskū.

kasaltakī lizard, gecko 103; WL 82.

kasil– to be cut; kasildī 110.

kaskū knife 109.

kaškā would that 113 (Pers.); WL 82.

kaz time 132 (Pers.); WL 81.

kazuǧ patrol 102; WL 82: kuzuǧ 'patrol'.

kībī mould; like 128.

kiǧit crossing place 102; WL 79 kačut; most Oǧ. lang. have a form ending in -t/-d (Az. kečid; Ttü gečit/d), whereas in Qip. lang. -v/-ü seems com-

mon (Tat. kičü; Kum. gečiv; Nogai kešüv). There are also forms in -k (Čaġ. kečik; Tuv. kežig).

kilaġï coming 106; see kal–.

kim 1 who 101; WL 84 kimdur 131; kimniŋ 120; kimnī 132.

kim 2 that, who (relative) 142.

kindin afternoon afternoon prayer 135; see ʾikindin.

kir– to enter; kirib 113; kiriban 137; kirmakum 142.

kirbuk eyelash 102; WL 80.

kiriš– to penetrate; kiriš 104.

kirtun– to believe; kirtundum 142. Cf. *Dīwān* (DK 18). Apparently not in Qip. languages. Redh. kirtin-; TTS I 471 II 460 IV 524).

kišan leg irons for a horse 102; in most Qip. languages Tat./Kom. kišen; Kaz. kisen.

kiši man; person 132; kišilar 132; kišilarḳā 132; kišilarnī 143.

kubkuk intensive blue 113; see kuk.

kuġuġuk small 105; WL 70; Doerfer 1621.

kuk sky; blue 112.

kukar– to become blue; kukardī 112.

kukġak small 141; Doerfer 1664.

kul– to laugh; kulub 137; kul-akan 112; kular 137; kularmišša 113; kulā 125; kulā kaldī 137; kulḳil 115; kulmadin 138; kul-madī 138; kulmas 138; kul-mayin 138; Alternative form of kulmadin; kuluban 137; kuluban 138; kul kul 120.

kulaf rose water 101; WL 83; Telegdi 1938:315. Per. comp. of *gül* ʿroseʿ and *ab* ʿwater.

kulaġ ?laughing 107; Derived

from kul–. Tat. küleč.

kumkuk intensive blue 117; see kubkuk.

kun sun; day 106; kundan 145; kunduġmaduq yidī 138; kun-duġmamiš yidī 138; kunduġ-miš dur 132; bukundur 122.

kunduz daytime 135.

kuprū bridge 102; WL 78: kuprū ʿbridgeʾ; n Oġ. ʿbalanceʾ.

kur– to see; kurdum 112; kursaŋ 119; kurkanimniŋ 147.

kurdur– to make see; to show; kurdur 110; .

kurk shape; beauty 140; .

kurkulū beautiful 133; kurklū dur 119; kurkulū dur 140.

kurkuz– to make see; to show; kurkuz 110.

kurt a kind of tree 101; WL 81.

kustar– ?to show causative form of kör-se- to want to see especially in West Oġ. (Ttü/Az). kustar ḳil 120.

kuz eye; kuzī 140.

kuzak ?great-eyed 102; WL 81; İÜ (37ᵛ14) kuzuk; not in Tat.

kuŋul heart; in WL 85 with -ŋ, three dots underneath the *nün*. (absent in İÜ 39ᵛ6); kuŋlī 143.

Lāġīn properly falcon used as a prop. name; 118; Doerfer (1728) "*lāčīn* ʿWanderfalkeʾ.".

makar/maqar except 126; cf. Steing. *Bulġa* 15,10/11. The use of this word in the Turkic sentences in *Idrāk* matches the Pers. *magar*. In Tur., though, it seems to yield grammatically incorrect phrases.

Makka Mecca; Makkaḳā 145.

malluġ equipped with, owner of *māl*, cattle; malluġdur 106.

man I 113; mandan 145; man-

imbilā 146; manī 139.

maqar except 142; see makar.

min a thousand 143.

min– to make mount see bin–; mindum 138; mindī 112.

mindir– to make mount; mindir 110; mindurdī 110; mindurkay 110.

? miniš way of mounting 109.

mis 153; per. assimilation of miš with the cond. particle sā; see Čaykovskaya 1981 missā idiŋ 153; missā 153.

miz we 130.

muštila– to bring a good message WL 88; Doerfer 1737; muštila 104.

muštuluq good news; gift for bringing good news 102; WL 88; Doerfer 1737.

nā what? 132; nādur 132; nā uǧun 133.

načā how many? 132; WL 89: načā. (mixed); načādur 132.

načuk how? 132 WL 89: načuk. (mixed).

nadan why 132.

naluk why? 132; WL 90, especially Qip..

nasā thing 133; WL 90.

natā how? 132; WL 90; Acc. to AH especially Oǧ.; EDT points to Ott. nete.

nišā why? 132; WL 90; Acc. to AH Oǧ. equivalent for for naluk.

qātī hard; qātī dur 141.

qabarǧuqlū equipped with a blister or a carapace 118.

qabṭurǧā ? 103; WL 67: qubṭurǧā aṣ-ṣawlaqu?

qaǧan when? 124; qaǧandur 131.

qal– to remain; qalmamišdī 106.

qalam pen 145; Ar. *qalam*.

qamuǧ all 151; EDT "Middle Pers. (*hamāg*)"; qamuǧlar 151; qamuǧlari 151.

qamurṣuǧā ant 103; most likely not Oǧ., since in those language qarınǧa or the like is used.

qanī where? 122; see qay.

qapu door, gate 101; WL 68.

qarāǧā goose 114; WL 70; only as qarāǧā qāz *al-lagālag; Bulǧa*: "*qaraža qaz...* 'oie noirâtre'." but Redh. "Somewhat black".

qaraldū shadow or silhouette of a person 102; WL 70; TTS IV karaltı 'shadow'; but Doerfer karantı (2270).

qarar– to become black; qarardī 138.

qarǧū WL 70: qarǧū 'Pers. spear'; but EDT: "a watchtower".

qari 1 old; qarisī 113.

qarī 2 cubit 144; probably not Oǧ. (but cf. Tkm, where it has a different meaning).

qarindaš brother; sister; qarindāšiŋ 152; qarindašlarıŋdur 143.

qariš a span 144.

qarši opposite place; qaršiŋdā 135.

qasṭaliq mussel 102; WL 72.

qaṭi side 146; qatinā 148; qatimdā 144; qatinǧā 148; qatiŋdā 140. Although listed in *'Idrāk* with -i, in view of the poss. endings 3sg. it takes, it must end in -t (as in EDT).

qatirǧā a gale 102; Doerfer 1437. WL 73: qatirǧā aš-šīnī" WL 70 qirbat also šīnī. *Bulǧa* "*qaṭurgan, qāṭurgan* 10,1."

qay which?; what?; qaydā is formed by adding the loc. suff. dā. qaydā 124; qaydādur 131.

qaygū sorrow, grief 112.

qayġur– to be grieved; qayġurdī 112.

qaytur– to turn back (tr.); qaytur 117.

qač how many? 132; WL 69.

qibqizil intensive red 113; see qizil also kumkuk, kubkuk etc.

qil a hair 103.

qilquyruq woodgrouse 103; WL 74; *Bulġa* 12,4. Noun comp. of qil 'hair' and quyruq 'tail', i.e. an animal, a bird, named after its hairlike tail.

qimilda– ? to shiver WL 75; qimilda 104. Redh. 1473. Perhaps conn. with EDT kamıl- to be struck down.

qirb– to shear WL 71 qirpdī (mixed) and WL 70 qirqdī; qirb 104.

qisā short 101; WL 72.

qizar– to become red; qizar 104; qizardī 112; qizarmadiq yidī 138; qizarmamiš yidī 138; qizarmiš 138.

qizil red 112.

quburġū owl 102.

quġa old man; master WL 69; quġaŋ 127; quġasī 113.

qul slave 105; qul ġuqas 105; qulġuqas 113; qulin 146; qulī 119; qulīdur 119; qulīsan 146; qulkinā 105; qullar 106; qullari 147; qullarī 119; qulllariŋ 132; qulum 118; qulumnī 147; qulumuz 119; qulumuznī 147; qulunī 147; quluŋ 118; quluŋnī 147; quluŋuz 118.

qulnaġī in foal 102.

qunrawū bell 102; -ŋ- based on EDT.

qunuqluq visit 122.

qurbaġā frog; toad 102.

qurq– to fear; qurqmaqin dan 142 note spelling with nunation instead of -n.

qurṭul– to be saved WL 70; qurṭul 104.

qušluq early morning 135 probably a typically Oġ. word.

quyaš beams of the sun 154.

quyruq tail 103.

quyunġī ?goldsmith 102; WL 77.

rūmluġ inhabitant of Rūm; Greek 105; rūmlū 106 (Oġ.).

saġa to you 142; see san.

sakisinġī eigthth 144.

sakiz eight 113.

saksan eighty 112.

salkinġak ?swing 103; WL 53. Ttü. salıncak.

san you [sg] 101; sandan 145; sani 141; saniŋbilā 146; saniŋkibī 132; sanī 119; sanudan 136.

sanġar 106; prop. name, deriv. from the verb 'to pierce'. Sanġar is also the name of a famous Selġūq (Oġ.) ruler. cf. Part One, Chapter 2 n. 37. see ṣanġ–. sanġar ġā 145; sanġar lar niŋ 147; sanġar 153; sanġarġā 142; sanġariŋ dur 145; sanġarnī 119.

sanġiš– to stab one another see ṣanġiš–; sanġištī 115.

saw– to love; sawdim 141; sawkaman 125; sawkasan 119; sawkā 125; sawmakimdan 141.

sawindur– to make happy; sawindurdī 152.

sawin– to be glad, happy; sawindum 152.

si– to break (Tr.); sidī 110; sir 111; siyur 111.

sibildaq ?a hairless animal 103; WL 51.

sin– to break (intr.); sindī 110.

sirkā vinegar; nit 102; WL 52.

siz you pl. 113; sizdan 145; siziŋbilā 146; siznī 139.

suḵ– to curse, to stab WL 53; suḵdum 144.

sulṭān sovereign 132.

sumulda– ?to clear one's throat WL 53; sumulda 104.

sunqur falcon 113; Doerfer 1273 *sonqur*. In *Idrāk* used as a prop. name; sunqurġā 140; sunqurnī 140.

supur– to sweep; WL 51; Doerfer N 102 = 2047; also supurkā. supur 104.

supurkā broom 102; cf. supur–.

surmā something rubbed; antimony 140.

surmalū equipped with surmā; surmalūdur 140.

suzla– to speak; suzlab 113; suzladī 115; suzlayub 113; suzlār 121; suzlāyur 121.

ṣabṣarī intensive yellow 113; see sarī.

ṣabṭur– to lead astray; ṣabṭur 104.

ṣaġā to you 126; see san.

ṣaġan– to think see saġin–; ṣaġandaġidur 140; ṣaġandī 118.

ṣaġin– to think see ṣaġan–. ṣaġindī 128.

ṣaġrī hindquarters of a horse 101; WL 58.

ṣal- to send 134.

ṣalin- to be sent 134.

ṣamaṣ obstinacy 113; WL 60. Occurs in W Qazaq dialects (S. Omarbekov, p.c. Feb 1996, and in NYugur - H. Nugteren p.c. Oct 1997).

ṣamaṣnadī ?to be obstinate 113; see: ṣamaṣ.

ṣan– to reckon; ṣandī 128; Oġ. form of saġin–.

ṣanǧ– to pierce; sanǧdum 144; ṣanǧdī 111.

ṣanǧiš fight; combat 125; ṣanǧiš kunī 111; *Bulġa* 6,8.

ṣanǧiš– to stab one another; ṣanǧištī 111.

ṣanqulan– ?to attack by surprise WL 60; ṣanqulan 104.

ṣaqurġa a tick 102.

ṣaraġan ?a kind of plant 102.

ṣarġar– to become yellow; ṣarġardī 138.

ṣarī yellow 113.

ṣarinǧaqan locust 103; WL 57.

ṣarṣil– ?to be wricked WL 57; ṣarṣil 104. Redh. 1154.

ṣarūġā ? 114; This may be identical with ṣarinǧaqan *q.v.*

ṣat– to sell; ṣaṭṭim 117.

ṣiǧqan mouse; rat 117.

ṣina– to test; ṣinadī 112.

ṣinan– to be tested 112; ṣinanmiš 112.

ṣirt ?back 101; but: WL 57: ṣirt 'hill and in Qip. ploughshare'.

ṣirṭalan ?hyena 102; WL 57; *Bulġa* 10,14".

ṣun left WL 59; WL 60; ṣunindā 135.

ṣunrā later 136; WL: 60.

ṣunuŋ later 136.

ṣur– to ask; to suck WL 57; surā 137; ṣurub 113; ṣuruban 137.

ṣū water 101.

šakird worker 102; Pers. WL 55; cf. Telegdi.

šāmluġ from *šām*; the Levant 105; šāmlū 106 (Oġ).

šapšapī ? 102; WL 54: šapšapī *al-mundaliq* 'loose matter'. Lane: *indalaqa* 'to become loose; fall out'.

šimdī now 117; see ʾimdī.

tabran– to move (intr.); tabran 104.

takā he-goat 117; see dakā .

taktur– to make sew; to make plant; takturgil 121. The back suff. cannot be explained.

talū mad 117; see dalū.

tamur iron 116; see damur.

taqī and; also 150.

tayma every 115.

tilkū fox 117; see dilkū.

tiltaq reason 117; see diltaq.

tiz knee 116; see diz.

tizkinla– perh. to brindle a horse also WL 38; tizkinla 104.

tun night; tunlā 136; Bulġa dünlä 13,12.

tuna– to spend the night; tunadī 123.

tuquṣawu all nine 1 1 6 ; see tuquz.

tutun smoke 116; see dutun.

ʔ taġarġuq knapsack 103; cf. WL 64.

ʔ taġil– to disperse WL 64; taġil 104.

tamdur– to cause to drip see tamzur–; tamdur 110.

tamzur– to cause to drip; tamzur 110; see tamdur–; tamzur 110.

tan dawn; tandadur 122; tandā 126 (final y read as ā); tanlā 136.

tarazī scales 101.

tart– to weigh; tart 103.

tašqun overflowing 102; WL 63.

tulī hail; full 101.

tun cloak; garment 110; tūnī 139; tunnī 140.

tunan– to dress oneself; tunandī 110.

tunluq something prepared for clothes 109; but: Bulġa 6,13 "solde, gages".

tuqsaġū all nine 115; see tuquz.

tuqsan ninety 112.

tuqsinġī ninth 144.

tuqurġin a game cf. 103; WL 65: tuqurġin 'a game like chess...". TS "tokurcun... A game that is played with nine stones on a board with twenty four rectangular fields".

tuquz nine 113.

tuqzan ninety 117; see tuqsan .

turaq stopping place; turaqin 123.

turdur– make stop; turdur 110.

turġuz– to make stop; to lift; turġuz 110 .

turuġī stopping 106; see tur–.

turuš (manner of) standing 108.

tur– to stand; tur 121; tur miš lar dur 147; tur mundā 151; tur turuban 137; turdilār 119; turaġġaq 114; turdū 111; turdī 114; turġabiz 153; turġaġaq 109; turġalī 145; turġaman 153; turġasan 153; turġay 152; turġā 121; turġāman 122; turġān 125; turġānlar 125; turġul– 115; turmadaġī 106; turmadī 150; turmaġay 142; turmaq 151; turmasā 152; turmiš 108; turmiš missā 153; turmiš missā idim 153; turmišdur 122; turmišidī 154; tursā 122; tursan 152; tursaniz 153; turtaġīdur 122; turtuq 153; turū 137; turub 137; turubantur 122; turubtur 122; 111; turuŋ 121; turuŋsiz 121; turuŋuz 121; turur.

tut– to hold; tuttī 117.

ʾuġ 1 three 143.

ʾuġ 2 end; ʾuġdan 145; ʾuġġā 145.

ʾuġawkamuz all three of us 116;

ʾuğawkaŋuz 1 1 6 ; ʾuğawlasī 116; ʾuğawu 116.

ʾuğmaq paradise cf. WL 8 ; ʾuğmaqā 142; ʾuğmaqqā 142.

ʾuğun for; for the sake of 133; .

ʾuğungī third 144;

ʾuğurğuq spindel whorl 103; WL 16.

ʾukrat– to teach; n WL 19; ʾukrat 104.

ʾuŋ right; ʾuninā 135.

ʾul that; he, she, it 118; ʾular 118.

ʾūl 119; see ʾul.

ʾul– to die cf. WL 20; ʾuldağibiz 106; ʾuldağī 114.

ʾuldur– to kill; ʾulduruldağī 107; ʾuldurulmiš 107.

ʾultur– to sit; ʾulturdū 111; ʾulturdum 142; ʾulturuğī 106.

ʾulturqağ a place to sit: chair 103; WL 21; SDD: "oturgaç a chair".

ʾulturuš manner of sitting 115; see ʾultur–.

ʾulu great; elder cf. WL 20; ʾulusī 113; ʾulwasī [sic] 120.

ʾulu– to howl; ʾulu 103.

ʾun 1 ten 116; WL 23; ʾunin unin 144.

ʾun 2 front; ʾunā 136.

ʾuningī tenth 144.

ʾunuŋ his 136; see ʾan.

ʾup– to kiss; ʾup 103.

ʾuqit– to make read; ʾuqit 110.

ʾuqša– to resemble; to caress; ʾuqšadī 145; ʾuqšar 145.

ʾur– to hit, strike; ʾarsaŋ 153; ʾurdağidur 1 4 0 ; ʾurağandur 140; ʾurdum 140; ʾurduŋ 132; ʾurduq 147; ʾurğaman 139; ʾurğanī 119; ʾurğaysan 152; ʾurğil 152; ʾurğul 120; ʾurmaq laqī 141; ʾurmaqī 141; ʾurmaqnī 141; ʾursaŋ 139; ʾur 152.

ʾurağat ?woman 102; WL 25: ʾawrat.

ʾūranla– see ʾuranla–.

ʾuranla– to think, suppose WL 12; TS 'oranlamak; ʾuranladī 128; ʾuranladum 128; ʾuranlarman 129; ʾuranladum 128.

ʾurmağak ?spider 102; but: WL 10: ʾurmağak 'acid cheese'.

ʾurt– to cover; ʾurt 1 0 4 ; ʾurttī 110.

ʾurtū cover; blanket 110.

ʾurtun– to cover oneself; ʾurtundī 110.

ʾurū 1 storage pit 101.

ʾurū 2 up 123; ʾurūturmiš 123; ʾurūturur 123.

ʾurul– to be hit, struck; ʾuruldağī 107; ʾuruldī 107; ʾurulğan 112; ʾurulğān 116; ʾurulmiš 107.

ʾuruš fight 108; ʾurušnī 135.

ʾuruš– to hit one another, fight; ʾurušdī 1 1 1 ; ʾurušğā 111; ʾurušur 111.

ʾust upper surface; ʾustindādur 145; ʾustumda 136; ʾustunā 135; ʾustuŋā 135.

ʾušbū '(just) this'; comp. of the particle ʾuš, used to call attention, and bū this 149; WL 14; Eckmann (1966: 114).

ʾuslū equipped with ʾus intelligence, hence intelligent 112; ʾuslūlardur 112.

ʾusur– to fart cf. WL 15; ʾusurmaq 103.

ʾutla– ?to advise; ʾutladī 115; For the meaning "to graze" (as 8 ʾutla-) one would expect a back form.

ʾutrukla– to ly WL 8; Bulğa II 64; ʾutrukladī 115.

ʾutur– to sit cf. ʾultur-; ʾuturuš 109; ʾuturušim 109; ʾuturušī

109.

ʾuṭuṣiz ? 116; In Ms. orig. ʾuẓusiz, but this form is unlikely, since Turkic does not know the sound ẓ. WL 14: ʾuṭū 'in ṭuqsubā it means 'island' and others say ʾaṭrağ'; Bulğa: "aṭuv 5,5". Hence perh. 'without an island'. Caferoğlu: ʾaṭsiz, i.e. 'horseless' 149.

ʾuṭuz thirty 143.

ʾuyanla– to restrain taking into account the pattern fuʿalla that is treated here, the correct vocalisation must be ʾuyanla rather than ʾayanla (conf. in İÜ). WL 26; ʾuyanla 104.

ʾuyanuq awake 102; WL 26; TS "uyanuk informed".

ʾuydin before 136.

ʾuylā noon 135; Bulğa (13,13) "äjlän".

ʾuyukan ? 102.

ʾuyurū valley 102; cf. WL 27.

ʾuz self; ʾuzī 151.

ʾuz– to tear, to break; ʾuz 108; ʾuzdū 110.

ʾuzkā other 102.

ʾuzuk torn, broken 108.

ʾuzul– to be torn, broken; ʾuzuldū 110; ʾuzulmiš 108.

ʾuzum grape 102.

ʾuzun long 133; ʾuzun dur 140.

yā 1 bow 101; yā ğuğas 105.

yā 2 exclamation; or 150.

yāğī bowmaker, archer 105.

yāru half 152; yārusnī 152.

yā– to eat; yā 103; yadiŋ 132; yadum 139; yakā san 142; yamadum 142; yamak 151; yamakil 152; yamā 147; yar 111; yayū 137; yidum 147; yidī 108; yikil 153; yīb 138; yiyiban 138; yiyiš 108; yiyur 111; yiyū 138.

yabyašil intensive green 117; see yamyašil.

yadawu all seven 116; see yïdakū.

yadiŋğī seventh 144.

yağ oil, grease, fat; yağbar 144.

yağirlik saddle-felt 103; deriv. from yağır 'shoulder'; Doerfer 1877.

yağmala– to robb WL 95; Doerfer 1874; yağmala 104.

yak good, better 109.

yakir– to be angry WL 96; İÜ suggests in WL 96; yakridī rather than yakirdī, which is not fully vocalised; yakirib 137.

yaklik superiority 109.

yakrak better 114; Doerfer 1879.

yaldurum glass, lightning 103; WL.

yaltirik 102; in İÜ; em. for biltirik; Bulğa: "jüldiräk 3,15... 'graines du roseau'."

yamaliq what is meant for mending 109; yamāliq 114. cf Doerfer 1895; yamāq.

yaman bad; yamandan 142; yamanraq 108.

yamyašil intensive green 117; see yašil.

yan side yaniŋdā 135; .

yan– to be eaten; yandī 133.

yaqšī good, nice 143; Doerfer 1869.

yar ground, place 144; yir 145; yirdā 143; see also yir.

yaraldiğan creation pass. part. of yarat– (WL 93) 'the created, or creation; yaraldiğan 115.

yaraṣā bat 101; Doerfer 1836; yaraṣa is typically Oğ. (Az./ Ott./ Tkm); Sewortian IV 140.

yaratğan creator; God 115; WL 93. Doerfer 1843. Common

Qip. form yarġanıl-.

yarilġa– (God) to forgive; yarilġa 104.

yaṣṣī flat 117.

yaṣtan– to prop oneself up on a pillow; yaṣtandī 110.

yaṣta– to prop on a pillow; yaṣtadī 110.

yaṣtuq pillow 110.

yaš moist; yašdur 120; yašlārdur 120.

yašar– to become green, moist; yašardī 112.

yašil geen 112.

yat– to be sufficient; yattī 136.

yatmiš seventy 115.

yattī seven 115.

yawuz bad 107.

yaz– 1 to be about to; yazdī 123.

yaz– 2 to write; yazdī 127; yaz- dum 145.

yi– to eat see yā–.

yidakū all seven 114.

yik good 108; see yak.

yikirmi twenty; yikirmī ar 143; 'ikirmibir ar 143; 'ikirmī ikī 143.

yikirmiŋī twentieth 144.

yikrak better 108; see yakrak.

yin– to be eaten Pass. form of

yī–; yindī 112.

yir ground, place see yar.

yiš(i)– to shine; biši could not be found nor in the alternative spelling, suggested in İÜ, i.e. bisi. Hence it may rather be a scribal error for yiši, 'to shine.' WL 94 yišidī 'he shone'.

yiyiš verbal noun of 'to eat' 108.

yū– to wash; yū 103; yudī 133; yur 111; yuyū 137; yuyur 111.

yuq there is no…, no 109; yuq sā 150.

yuqluq poverty 109.

yun– to wash oneself; yundī 133.

yuri– to walk; yuridim 145; yurïyū 137; yūrī 137.

yurit– to make walk; yurit 110.

yuz 1 hundred 101; yuzar 143; yuz 143; WL 93.

yuz 2 face WL 93: yuzī 140; yuzu 138; yuzī 138.

yuzarlik 103; EDT: "the plant rue, *Peganon harmala*" tran- scr. matches İÜ.

yuziŋī hundredth 144.

yuzlū having a … face 138.

ABBREVIATIONS

ABL	ablative	INT	interrogative
ACC	accusative	KONV	Konverb
AOR	aorist	lt	left
APOC	apocopate	m	masculine
ART	article	md	middle
bm	bottom	MG	Margin Grammar
CAUS	causative	NOM	nominative
COND	conditional	NEG	negation
CONJ	conjunctive	OBL	oblique
CORR	corroborative	PART	participle
DAT	dative	PASS	passive
DEF	definite	PAST	past tense
DIM	diminutive	pl/PLUR	plural
f	feminine	POSS	possessive
GEN	genitive	POSTP	postposition
IMP	imperative	PRES	present tense
IMPF	imperfect	rt	right
IND	indicative	sg	singular
INDEF	indefinite	SUBJ	subjunctive
INF	infinitive	TOP	topicalisation
INFER	inferential		

BIBLIOGRAPHY

Primary sources

al-ʾAnbārī, ʾAbū l-Barakāt ʿAbd ar-Raḥmān bn Muḥammad. *al-ʾInṣāf fī Masāʾil al-ḫilāf Bayna n-Naḥwiyyīn al-Baṣrīyīn wa-l-Kūfiyyīn* ed. Muḥ. Muḥyī ad-Dīn ʿAbd al-Ḥamīd. 1952. Cairo: Maṭbaʿa Ḥiǧāzī.

al-ʾAndalusī, ʾAbū Ḥayyān, Muḥammad bn Yūsuf bn ʿAlī. *Irtišāf aḍ-Ḍarab min Lisān al-ʿArab* (3 vol). ed. Muṣṭafā ʾAḥmad an-Nammās. 1984-1987-1989. (Part I: Cairo: Maṭbaʿa an-Našr aḍ-Ḍahabī; Part II and III Cairo: Maṭbaʿa al-Madanī).

——. *Kitāb al-ʾIdrāk li-lisān al-ʾAtrāk* ed. Ahmet Caferoğlu 1931. Istanbul: Evkaf matbaası.

——. *Kitāb al-ʾIdrāk* MS cod. Veli ed-Dīn 2896.

——. *Kitāb al-ʾIdrāk.* MS University of İstanbul cod. Halis Efendi 6597.

——. *Manhaǧ as-Sālik li-ʾAlfiya Ibn Mālik.* ed. Sidney Glazer. 1947. Washington.

——. *Taḏkira an-Nuḥāt* ed. ʿAfīf ar-Raḥmān 1986. Beirut: Muʾassasa ar-Risāla.

——. *al-Baḥr al-Muḥīṭ* (8 vol.) ed. ʿĀdil ʾAḥmad ʿAbd al-Mawǧūd and ʿAlī Muḥammad Muʿawwaḍ. Beirut: Dār al-Kutub al-ʿIlmiyya 1993.

al-Bayḍāwī, Nāṣir ad-Dīn. *Tafsīr ʾAnwār at-tanzīl wa-ʾAsrār at-Taʾwīl.* al-Maṭbaʿa al-Yamaniyya. 1320/1902 (2 vols.).

al-Kāšǧarī, Maḥmūd Ibn al-Ḥusayn Ibn Muḥammad. *Dīwān Luġāt at-Turk* facsimile edition of the MS; Turkish Ministry of Culture. 1990 (see also Dankoff & Kelly 1982-5).

al-Qārīʾ, al-Minaḥ = al-Qārīʾ, ʿAlī bn Sulṭān, *al-Minaḥ al-fikriyya ʿalā matn al-Ǧazariyya.* Egypt (Cairo?): al-Maṭbaʿa al-Yamaniyya, 1322/1904.al-Fārābī, ʾAbū ʾIsḥāq bn ʾIbrāhīm *Dīwān al-ʾAdab* ed. ʾAḥmad Muḫtār ʿUmar 1974-78 (4 vol.) Cairo: al-Hayʾa al-ʿĀmma.

Ibn Ǧinnī, ʾAbū l-Fatḥ ʿUṯmān. *Sirr Ṣināʿa al-ʾIʿrāb* ed. Ḥasan Hindāwī. (1985) Damascus: Dār al-Qalam.

——. *al-Ḫaṣāʾiṣ.* ed. Muḥammad ʿAlī an-Naǧǧār 1371/1952-56. Cairo: Dār al-Kutub al-Miṣriyya (3 vols).

al-Ḥalīl = ʾAbū ʿAbd ar-Raḥmān al-Ḥalīl bn ʾAḥmad al-Farāhidī. *Kitāb al-ʿAyn* I. Ed. al-Maḫzūmī and as-Sāmarrāʾī. Beirut: al-ʾAʿlamī.

Ibn Kamāl Pāšā *Risāla fī Taḥqīq Taʿrīb al-Kalima al-ʾA ǧamiyya.* Ed. Mohammed Sawaie. Institut Français de Damas 1991.

Ibn Manẓūr, ʾAbū l-Faḍl. *Lisān al-ʿArab* (n.d.) Beirut: Dār Ṣādir (15 vols.).

Ibn al-Muhannā, Ǧamāl ad-Dīn *Kitāb Ḥilya al-ʾInsān wa-Ḥalaba al-Lisān.* ed. Kilisli Muʿallim Rifʿat (Bilge) 1921.

Ibn as-Sarrāǧ, ʾAbū Bakr Muḥammad bn Sahl. *Al-ʾUṣūl fī n-Naḥw* ed. ʿAbd al-Ḥusayn al-Fatlī (3 vols) 1985. Beirut: Muʾassasa ar-Risāla.

Ibn Sīdā = ʾAbū Ḥasan ʿAlī bn ʾIsmāʿīl. *al-Muḥaṣṣaṣ* ed. Laǧna ʾiḥyāʾ at-turāt al-ʿarabī. (n.d.) Beirut: Dār al-ʾĀfāq al-Ǧadīda.

Ibn as-Sikkīt *ʾIbdāl* = Ibn as-Sikkīt, ʾAbū Yūsuf Yaʿqūb bn ʾIsḥāq bn Yūsuf, *Kitāb al-Qalb wa-l-ʾIbdāl.* Ed. by Šaraf, Husayn Muḥammad. Cairo: al-Hayʾa al-ʿĀmma li-šuʾūn al-maṭabiʿ al-ʾamīriyya (1 vol), 1978.

Ibn Sīnā *ʾAsbāb Ḥudūt al-Ḥurūf* ed. Muḥibb ad-Dīn al-Ḥaṭīb 1332/1913/4 Cairo: Maṭbaʿa al-Muʾayyid.

Tarǧumān = Kitāb Tarǧumān Turkī wa-ʿarabī wa-muǧulī in Houtsma 1889.

Lisān see Ibn Manẓūr.

al-Qawānīn al-Kullīya li-Ḍabṭ al-Luǧa at-Turkiyya ed. Muʿallim Rifʿat 1928. İstanbul: Evkaf.

ar-Rāzī, Faḫr ad-Dīn. *at-Tafsīr al-Kabīr.* ed. Muḥy ad-Dīn ʿAbd al-Ḥamīd (11 vol.). Cairo: al-Maṭbaʿa al-Miṣriyya. 1933-8.

Ṣāliḥ, Ibn Muḥammad Mollāh *aš-Šudūr aḏ-Ḏahabiya wa-l-Qiṭaʿ al-ʾAhmadīya fī l-Luǧa at-Turkīya* Bibliothèque nationale, Paris, Supplément Arabe No. 4333.

Sībawayh, ʾAbū Bišr ʿAmr. *al-Kitāb* II ed. Būlāq - Egypt 1900/1317.

aš-Širbīnī, Sayyid Muḥammad *Kitāb nūr as-Saǧiyya fī Ḥall ʾAlfāẓ al-ʾAǧūrrūmiyya* ed. and transl. by Michael Carter 1981. Amsterdam: John Benjamins.

al-Kitāb at-Tuḥfa aḏ-Ḏakiyya fī l-luǧa at-turkiyya. Facsimile edition by T. Halasi Kun in 'La langue des Kiptchaks d'après un manuscrit arabe d'Istanboul' Budapest: Bibliotheca orientalis Hungarica, 1942.

az-Zaǧǧāǧī, ʾAbū l-Qāsim ʿAbd ar-Rahmān bn ʾIsḥāq *al-ʾĪḍāḥ fī ʿIlal an-Naḥw* ed. Māzin al-Mubārak 1959 Cairo: Maṭbaʿa al-Madanī.

az-Zamaḫšarī, ʾAbū l-Qāsim Maḥmūd bn ʿUmar *Kitāb al-Mufaṣṣal Fī n-Naḥw.* Ed. Jens Peter Broch 1879.

al-Wāfī see aṣ-Ṣafadī

Secondary sources

Adang, Camilla. 1993. *Muslim writers on Judaism and the Hebrew Bible from Ibn Rabban to Ibn Hazm.* Diss. Univ. Nijmegen.

Alpay, Gönül. 1973. *Aḥmad-i Dāʾī and his Çengnāme. An old Ottoman Mesnevī.* Sources of Oriental Languages and Literatures, Harvard Printing office.

Atalay, Besim. I '«Et-tuhfet-üz-zekiyye» Çevirmesi'nin tenkidi dolayısiyle' in *Ankara Üniversitesi Dil ve Coğrafya Fakültesi Dergisi* (VI) 87-126.

——. 1949. *Eş-Şüzūr-üz-Zehebiyye vel-Kıtai ʾl-Ahmediyye fil-Lūgat-it-Türkiyye* (Turkish transl. of *aš-Šudūr aḏ-Ḏahabīya*) Istanbul.

Auroux, Sylvain and Queixalos, F. 1984. 'Présentation. Pour une histoire de la linguistique amérindienne en France' in *Amerindia* (6) 1-20.

Ayoub, Georgine. 1991. 'La forme du sens: le cas du nom et le mode du verbe' in Dévényi and Iványi (1991). 39-87.

Ayoub, Georgine and Georges Bohas. 1983. 'Les Grammairiens Arabes, la Phrase Nominale et le Bon Sens' in Versteegh, Koerner and Niederehe (1983) 31-48.

Baalbaki, Ramzi. 1988. 'A Contribution to the Study of technical terms in Early Arabic Grammar—the term 'aṣl in Sībawayhi's Kitāb', in Irvine, Serjeant and Rex Smith (eds.) A Miscellany of Middle Eastern Articles - In Memoriam Thomas Muir Johnstone (1924-83), Harlow: Longmann. 163-177.

Babby, Leonard H. 1986 'The locus of case assignment and the direction of percolation' in Case in Slavic, Brecht and Levine (eds.). Columbus (Ohio): Slavica Publishers. 170-219.

——. 1991. "Noncanonical configurational case assignment strategies" in Cornell Working Papers in Linguistics (9) 1-55.

Babby, Leonard H. and Freidin, R. 1984. 'On the interaction of Lexical and Syntactic Properties: Case structure in Russian', in W. Harbert (ed.) Cornell Working Papers in Linguistics (6) 71-103.

Bakalla, Muḥammad Ḥ. 1982. Ibn Jinni: An early Arab Muslim Phonetician. London and Taipei.

——. 1983. 'The Treatment of Nasal Elements by Early Arab and Muslim Phoneticians' in Versteegh, Koerner and Niederehe (1983) 49-69.

Baskakov, N.A. 1988. Istoriko-tipologičeskaja fonologija tjurkskix jazykov. Moskva: Nauka.

Bazin, Louis. 1959. 'Structures et tendances communes des langues turques' in Fundamenta I. 11-21.

Bergsträsser, G. 1921. 'Das Vorbild von Kāšġarī's dīwān luġāt at-turk' Orientalistische Literaturzeitung (21) 154-5.

Bernards, Monique. 1993. Establishing a reputation: The Reception of the Kitāb Sībawayh. Diss. Univ. Nijmegen.

Berta, Arpad. 1989. Lautgeschichte der tatarischen Dialekte. Szeged: Attila József University.

Biesterfeldt, Hans-Hinrich. 1990. 'Ibn Farīġūn's Chapter on Arabic Grammar in his Compendium of the Sciences' in SHAG II. 49-56.

Blaškovič. J. 1961. Arabische, türkische und persische Handschriften in der Üniversitätsbibliothek zu Bratislava.

Blau, Joshua. 1989. 'Two Studies of Sībawayhi's Kitāb I: The Shift of Arabic Sibilants' in Jerusalem Studies in Arabic and Islam (12) 1-15.

Blochet, Edgar. 1932. Catalogue des Mss Turcs, Bibliothèque Nationale. Paris.

Bobzin, Hartmut. 1983 'Zum Begriff der "Valenz" des Verbums in der arabischen Nationalgrammatik' in SHAG I, 93-108.

Bodrogligeti, 1971. A. 'A Grammar of Mameluke-Kipchak' in Studia Turcica (Budapest). 89-102.

Bohas, Georges. 1982. Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie d'après des grammairiens 'tardifs'. Thèse d'état, Université de Paris III (1979).

Bohas, Georges and Jean-Patrick Guillaume. 1984. Étude des théories des grammairiens arabes. I Morphologie et phonologie. Damascus: Institut Français de Damas.

Bohas, Georges, Jean-Patrick Guillaume and Djamel Eddin Kouloughli. 1990. The Arabic Linguistic Tradition. London: Routledge.

Boyce, Mary. 1968. 'Middle Persian Literature' in *Handbuch der Orientalistik* I Abt. IV Bd. 2 Abschn. 32-65.

Caferoğlu, Ahmet. 1931. *Kitâb al-İdrâk li-lisân al-Atrâk*. Istanbul: Evkaf Matbaası.

Čajkovskaja, Angelina Ivanova. 1981. *Tjurkskaja Grammatika v Arabojazyčnix filologičeskix traktatax xiii - xiv Vekov (glagol')*. Taškent: Izd. Fan.

Cantineau, Jean. 1960. *Cours de phonétique Arabe*. Klincksiec: Paris (repr. of 1941).

Carter, Michael. 1981. *Arab linguistics. An Introductory classical text with translation and notes*. Amsterdam: John Benjamins [Edition of Širbīnī's *Kitāb nūr as-Saġiyya fī Ḥall 'Alfāẓ al-'Aġurrūmiyya*].

———. 1991. 'Elision' in Dévényi and Iványi (1991). 121-133.

CC = *Codex Comanicus* see Grønbech 1942.

Celentano, Guiseppe. 1979. 'Due scritti medici di al-Kindī' in *Annali* (39) 1 Suppl. no. 18 agli. 47-75.

Clauson, Gerard. 1962. *Turkish and Mongolian Studies*. London: The Royal Asiatic Society of Great Britain and Ireland.

———. 1972. *An Etymological Dictionary of Pre-Thirteenth Century Turkish*. Oxford: Clarendon Press.

Comrie, Bernard. 1983. *Language Universals and Linguistic Typology*. Oxford: Blackwell [corr. repr. of 1981].

Corriente, Federico. 1978 'Los Fonemas /p/, /č/ y /g/ en Árabe Hispánico' in *Vox Romanica* (37) 214-218.

Dankoff, Robert. 1972. 'Kāšgarī on the Tribal and Kinship Organization of the Turks' in *Archivum Ottomanicum* (4) 23-43.

Dankoff, Robert & Kelly, James. 1982-5. *Compendium of Turkic dialects*. Translation of Maḥmūd al-Kāšġarī's *Dīwān Luġat at-Turk*. 3 vol.: I (1982): MS 2-304; II (1984): MS 305-638; III (1985) indices to vol. I and II. Harvard: University Press.

Deny, Jean. 1921. *Grammaire de la langue turque*. Paris.

Dévényi, Kinga and Iványi, Tamás (eds). 1991. *Proceedings of the Colloquium on Arabic Grammar Budapest, September 1-7 1991*. The Arabist. Budapest Studies in Arabic (3 - 4).

Devereux, Robert. 1964. 'Judgment of Two Languages. *Muḥākamat al-Lughatain* by Mīr ʿAlī Shir Nawāʾī. Introduction, Translation and Notes.' in *Muslim World* (54) 270-87 (continued in 1965 *Muslim World* [55] 28-45).

Diem, Werner. 1976. 'Some glimpses at the Rise and Early Development of the Arabic Orthography' in *Orientalia* (5) 251-61.

Ditters, Everhard. 1992. 'De "zij" van Sibawayh' ['On "them" in Sibawayh'] in *Elf Wijzen van Interpreteren. Essays over het lezen van teksten uit het islamitisch cultuurgebied*. Nijmegen: Instituut voor Talen en Culturen van het Midden Oosten. 36-54.

Doerfer, Gerhard. 1959. 'Das Krimtatarische' in *Fundamenta* I 369 - 390.

Doerfer, Gerhard. 1965 and 1967. *Türkische und mongolische Elemente im Neupersischen. II Türkische Elemente, alif-tā', II jīm-kāf*. Wiesbaden: Harrassowitz.

———. 1976. 'Woher Stammte Ibn Muhannā?' *Archäologische Mitteilungen aus Iran* (9) 243-251.

Dozy, R.P.A. 1851. *Catalogus Codicum Orientalium Bibliothecae Academicae Lugduno Batavae* Leiden: Brill.

EDT see Clauson (1972).

EI[1] *Encyclopaedia of Islam.* 1908-1934. Leiden: E.J. Brill

EI[2] *Encyclopaedia of Islam.* 1960-. New edition prepared by a number of leading Orientalists. Leiden: E.J. Brill.

Ermers, Robert. 1992. 'Abū Ḥayyān al-Andalusī en zijn grammatica van het Turks: Een Introductie' in *Sharqiyyât* (2) 302-323.

——. 1994. Review of "Risāla fī tahqīq taʿrīb al-kalima al-aʿgamīya by Ibn Kamâl Pasha", Sawaie, Mohammed, 1991, Damas: Institut Français, in *JAOS* (114/2) 294-295.

Ermers, Robert and Michel Limpens. 1995. Review of "Sībawayhi the Phonologist", al-Nassir, A.A., in *Bibliotheca Orientalis* 52 5/6.853-5.

Fazylov, E.I. 1976. 'Zamečanija o rukopisi i jazyke "at-Tuḥfa"' in *Turcologica* (50) 334-340.

Fazylov, E.I. and M.T. Zijajevaja. 1978. *Izkannyj dar Tjurkskomu Jazyku. Grammatičeskij traktat XIV veka na arabskom jazyke* [Russian translation of *at-Tuḥfa az-Zakiyya*]. Taškent: Fan.

Fischer, Wolfdietrich. 1989. 'Zur Herkunft des Grammatischen Terminus *ḥarf* in *Jerusalem Studies in Arabic and Islam* (12) 125-145.

Fleisch, Henri. 1961. *Traité de Philologie Arabe I, Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.

Flemming, Barbara. 1968. *Verzeichnis der Orientalischen Handschriften in Deutschland - Türkische Handschriften* XIII 1. Wiesbaden: Steiner.

——. 1968. 'Ein alter Irrtum bei der chronologischen Einordnung des *Tarǧumān Turkī wa-ʿArabī wa-Muǧalī*' in *Der Islam* (44) 226-229.

——. 1977a 'Zum Stand der Mamlūk-türkischen Forschung' in *Vorträge des XIX. deutschen Orientalistentages vom 28. September bis 4. Oktober 1975 in Freiburg im Breisgau.* ZDMG Suppl. III 2. 1155-64.

——. 1977b 'Literary activities in Mamluk Halls and Barracks' in *Gaston Wiet Memorial Volume* Jerusalem. 249-286.

Fundamenta I = *Philologica Turcicae Fundamenta* I. 1959. Louis Bazin e.a. (eds.). Wiesbaden: Steiner.

Gabain, Annemarie von. 1959. 'Das Alttürkische.' in *Fundamenta* I.

——. 1959. 'Die Sprache des Codex Comanicus' in *Fundamenta* I: 46-73

Garipov, Talmas M. 1979. *Kypčakskije jazyki uralo-povolžya* Moscow: Izdatel'stvo Nauka.

Gibb, E. 1900-09. A *History of Ottoman Poetry* (6 vols) London.

Götz, Manfred. 1979. *Verzeichnis der Orientalischen Handschriften in Deutschland - Türkische Handschrifte,* 4. Wiesbaden: Steiner.

Grotzfeld, Heinz. 1965. *Syrisch-Arabische Grammatik: Dialekt von Damaskus.* Wiesbaden: Harrassowitz.

Grønbech, Kaare. 1942. *Komanisches Wörterbuch. Türkischer Wortindex zu Codex Comanicus.* København: Einar Munksgaard.

Grünert, Max. 1912. 'Al-'Išmâm' in *Wiener Zeitschrift für die Kunde des Morgenlandes* (26) 222-36.

Gully, Adrian. 1994. 'Synonymy or Not Synonymy: That is the Question. The case of the Particle in Medival Arabic' in *ZAL* (27) 36-46.

Guzjev, V.G. 1990. *Očerki po Teorii Tjurkskogo Slovoizmenija: Glagol'* *(na Materiale staroanatolijsko-tjurkskogo jazyka)*. Leningrad: Izd. Leningradskogo Universiteta.

Haarmann, Ulrich. 1988. 'Arabic in Speech, Turkish in Lineage: Mamluks and their Sons in the Intellectual Life of Fourteenth-Century Egypt and Syria' in *Journal of Semitic Studies* (33) I 81-114.

Ḥadīṭī, Ḥadīǧa. 1967. *'Abū Ḥayyān an-Naḥwī*. Baġdād: Maktaba an-Nahḍa.

Halasi Kun, Tibor. 1947. 'Philologica I' in *Ankara Üniversitesi Dil ve Coğrafya Fakültesi Dergisi* (5) 1-37.

——. 1948. 'Philologica II' in *Ankara Üniversitesi Dil ve Coğrafya Fakültesi Dergisi* (6) 415-65.

Hanser, Oskar. 1977. *Turkmen Manual. A descriptive Grammar of contemporary literary Turkmen*. Wien: Verlag des Verbandes der Wissenschaftlichen Gesellschaften Österreichs.

Hattori, Shiro. 1979-80. 'The place of Tatar among the Turkic languages' in *Eucharisterion: Essays presented to O. Pritsak* = Harvard Ukrainian Studies (3/4) 383-93.

Haywood, John. 1965. *Arabic Lexicography*. Leiden: Brill.

Hegazi, Mahmoud M. 1971. *Abū Saʿīd as-Sīrāfī der Sībawayhi-Kommentator als Grammatiker*. PhD thesis. Univ. of Munich.

Henning, W.B. 1958. 'Mitteliranisch' in *Handbuch der Orientalistik* Bd IV Abschn. 1.

Hjelmslev, Louis. 1972. *La catégorie des Cas*. [Second edition of the original of 1935 (ed. by E. Coseriu) München: Fink].

Houtsma, Martinus. 1889. '*Tarǧumān Turkī wa-ʿArabī wa-Muġulī*. Ein türkisch-arabisches Glossar nach einer leidener Handschrift.' (Introduction and edition) Leiden.

İA *İslam Ansiklopedisi.*

Iskandar, Albert Z. 1981. "A Doctor's Book on Zoology. Al-Marwazī ñabāʿī' al-Ḥayāwān (Nature of Animals) Re-assessed" in *Oriens* (27/28) 266-312.

İzbudak, Velet and Kilisli, Muʿallim Rifʿat. 1936. *El-İdrāk Ḥāṣiyesi*. İstanbul: TDK Devlet Basımevi.

Kästner, Hartmut. 1981. *Phonetik und Phonologie des Modernen Hocharabisch*. Leipzig: Enzyklopädie.

Karamanlıoğlu, Ali Fehmi. 1962. 'Kıpçaklar ve Kıpçak Türkçesi' in *Türk Dili ve Edebiyatı Dergisi* (12) 175-184.

Karatay, Fehmi E. 1961. *Topkapı Sarayı Müzesi Kütüphanesi Katalogu* Istanbul. Topkapı Sarayı Müzesi. 2 Vols.

Kelly, James. 1971. 'On defining *Dhū ath-Thalāthah* and *Dhū al-Arbaʿah*' in *Journal of the American Oriental Society* (91.1) 132-136.

——. 1972. 'Remarks on Kāšgarī's Phonology. I. Linguistic terminology.' in *UAJ* (44) 191-193.

——. 1973. 'Remarks on Kāšgarī's Phonology. II. Orthography.' in *UAJ* (45) 144-162.

——. 1976. 'Remarks on Kāšgarī's Phonology. III. Structure of the Diwan.' in *UAJ* (48) 151-158.

——. 1980. 'A closer look at the *Dīwān al-Adab*' in *Eucharisterion. Essays presented to Omeljan Pritsak on his sixtieth birthday*. ed. Ševčenko and Sysyn; vol. III/IV (1979-80; part 2) 497-506.

Kerslake, Celia J. 1994. 'Two Ottoman Turkish Grammars of the Tanẓīmāt Period' in *Proceedings of the VII. Conference of the CIEPO held in Pécs, Hungary, in 1986*. Ankara: Türk Tarih Kurumu. 133-168.

Kinberg, Naphtali. 1987. 'The Concepts of Elevation and Depression in Medieval Arabic Phonetic Theory' in *ZAL* (17) 7-20.

Lane, Edward William. *An Arabic English Lexicon, derived from the best and the most copious Eastern sources....* Book I in eight parts [all published]. Londen & Edinburgh: Williams and Norgate, 1863-93. (Repr., Beirut: Librairie du Liban, 1980).

Lăzărescu-Zobian, Maria M. 1982. *Kipchak Turkic Loanwords in Rumanian*. PhD. Diss. Columbia University.

Levin, Aryeh. 1987. 'The Views of the Arab Grammarians on the Classification and Syntactic Function of Prepositions' in *Jerusalem Studies in Arabic and Islam* (10) 342-367.

——. 1994. 'Sībawayhi's attitude to the spoken language' in *Jerusalem Studies in Arabic and Islam* (17) 204-243.

Lewis, Bernard. 1990. *Race and Slavery in the Middle East. An historical Enquiry*. New York - Oxford: Oxford University Press.

Lewis, Geoffrey L. 1984. *Turkish Grammar*. Oxford: Clarendon Press. (corr. repr. of 1967)

Lyons, John. 1971. *Introduction to Theoretical Linguistics*. Cambridge: University Press.

Mahmutova, L.T. (1982) 'Tatarskiy jazyk v ego otnošenii k drevnepis'mennomu pamjatniku Codex Comanicus po dannym leksiki. Kratki analiz i priloženie' in *Issledovanija po istoričeskoj dialektologii tatarskogo jazyka. Materialy i issledovanija* 2. Kazan'. 68-154

Mansuroğlu, Mecdut. 1954. 'The Rise and Development of Written Turkish in Anatolia' in *Oriens* (7) 250-264.

Meier, Fritz. 1981. 'Aussprachefragen des älteren Neupersisch' in *Oriens* (27-28) 70-176.

Menges, Karl. 1966. 'Report on the Second Excursion to Taškent for Research in Čagataj Manuscripts' in *Central Asian Journal* (11) 131-

——. 1968. *The Turkic Languages and Peoples. An introduction to Turkic Studies*. Wiesbaden: Harrassowitz.

Miquel, André. 1967-1988. *La géographie humaine du monde musulman jusqu'au milieu du 11e siè cle*. Paris - La Haye: Mouton (4 vols).

Mitchell, Terence F. 1993. *Pronouncing Arabic 2*. Oxford: Clarendon Press.

Moldabekov, K. 1992. *Govorim po-kazaxski*. Alma-Ata: Ana Tili.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih* Diss. Universität München.

Musaev, K.M. 1964. Karaimsko - Russkiy Slovar'.

Muskens, G.L. n.d. *Latijnse Grammatica* (rev. by J. Ysebaert). 6th imp. Maastricht: Schenk.

Nassir, ʿAbd al-Munʿim ʿAbd al-ʾAmīr. 1993. *Sibawayh the phonologist. A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in his Treatise Al-Kitab.* London and New York: Kegan Paul.

Nissman, David. 1969. *A study of a Language and Orthography of the Mamlukes of Egypt Based on the Kitāb Bajṭarati ʾl-wāḍiḥ.* PhD. Diss. Columbia University.

Owens, Jonathan. 1988. *The Foundations of Grammar.* Amsterdam: Benjamins.

———. 1990. *Early Arabic Grammatical Theory. Heterogeneity and Standardization.* Amsterdam: Benjamins.

Öztopçu, Kurtuluş. 1990. 'Kitāb fī ʿIlm an-Nuššāb' (in transcription) in *Rocznik Orientalistyczny* (47 I) 5-63.

Poppe, Nicholas. 1963. *Tatar Manual Descriptive Grammar and texts.* Bloomington: Indiana University Publications. Uralic and Altaic Series no. 25.

Pritsak, Omeljan. 1957. Review of Zajączkowski's edition of *Bulġa* in *Der Islam* (32) 362-5.

———. 1959. 'Das Kipčakische' in *Fundamenta* I. 79-84.

Qudayberdy-uly, Šakarim [1858-1931]. 1990. *Rodoslovnaya Tjurkov, Kirgizov, Kazaḫov i ḫanekix dinastij* Alma-Ata: Jazušy (transl. from Kazaḫ by B. Kairbekov).

Redhouse, James W. 1978 [1890] *A Turkish and English Lexicon Shewing in English the Significations of the Turkish Terms.* Istanbul: Çağrı Yayınları (Reprint of the 1890 Ed., Constantinople)

Rieu, Charles. 1888. *Catalogue of the Turkish MSS in the British Museum.* London.

RO *Rocznik Orientalistyczny.*

Robakidze, N.I. 1986. "Interpretatsiya termina 'al-ḳalb' u starykh arabskikh grammatistov" in *Semitologicheskiye Shtudii,* 13/3.42-44.

Robins, Robert Henry. 1993. The Byzantine Grammarians: Their place in history. Berlin and New York: Mouton de Gruyter.

Roman, André. 1983. *Étude de la Phonologie et de la Morphologie de la Koine Arabe* (2 vols) Univ. de Provence.

Rossi, Ettore. 1935. *Elenco dei Manuscritti Turchi della Biblioteca Vaticana.* Cittṭ del Vaticano.

Roux, Jean-Paul. 1984. *La Religion des Turcs et des Mongols.* Paris: Payot.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor.

Saad, George. 1975. *Transitivity, Causation and the Derivation of Passives in Arabic.* Diss. Univ. of Texas, Austin.

Sawaie, Mohammed (Ed.). 1991 see Ibn Kamāl Pāšā.

SDD = *Türkiye'de halk ağzından Söz Derleme Dergisi.* 1939-1957. Türk Dil Kurumu. Istanbul: Maarif.

Šemiyeva, A. and I. Eḥmetov. 1955. *Uygur Tili Grammatikisi I. Ëismi Fonetika hem Morfologija.* Alma-Ata: Kazak SSR Maarip Ministrligi (Ministry of Sciences).

Sertkaya, Osman. 1973. *Osmanlı şâirlerinin Çağatayca şiirleri iii. Uygur harfleriyle yazılmış bazı manzum parçalar i.* Istanbul: Edebiyat Fakültesi.

———. 1975. *Uygur harfleriyle yazılmış bazı manzum parçalar ii.* Istanbul: Edebiyat Fakültesi .

Sewortian, E.V. 1974-1978-1980. *Etimologicheskiy slovar' t'urkskich jazykov. 1 Vowels; 2 V, G, D; 3 Z, Zh.* Moscow: Institute of Linguistics.

SHAG II *Studies in the history of Arabic Grammar. Proceedings of the 2nd Symposium on Arabic Grammar, held at Nijmegen 27 April - 1 May 1987,* ed. Michael G. Carter and Kees Versteegh. Amsterdam: Benjamins (= *Studies in the History of the Language Sciences* 56).

Sinor, Denis. 1969. *Inner Asia. History, Civilization, Languages. A syllabus.* Indiana University. The Hague: Mouton & Co.

SKJa. = *Sovremennyj Kazaxskij Jazyk, Fonetika i morfologiya.* 1962 Alma-Ata: Izd. Akademii Nauk Kazaxskoy SSR.

Sohrweide, Hanna. 1974. *Verzeichnis der Orientalischen Handschriften in Deutschland - Türkische Handschriften* XIII 3. Wiesbaden: Steiner.

Steingass, Francis Joseph. 1947. *A Comprehensive Persian-English Dictionary.* London: Paul, Trench, Trubner.

Steuerwald, Karl. 1988. *Türkisch-Deutsches Wörterbuch.* Wiesbaden: Harrassowitz.

Süssheim, K. "ʿAlī" in *İslam Ansiklopedisi* (I 304-306).

Taha, Zeinab Ahmed. 1995. *Issues of Syntax and Semantics: A Comparative Study of Sibawayhi, al-Mubarrad, and Ibn as-Sarraaj.* PhD. Diss. Georgetown University; Washington D.C.

Talmon, Rafael. 1984. 'at-Tafsīr an-Naḥwī qabla Kitāb Sībawayhī' -Dirāsa fī Taʾrīḫ al-Muṣṭalaḥ an-Naḥwī al-ʿArabī' in *al-Karmil* (5) 37-53.

——. 1997. *Arabic Grammar in its Formative Age. Kitāb al-ʿAyn and its attribution to Ḫalīl b. Aḥmad.* Leiden, New York, Köln: Brill.

Tarama Sözlüğü. 1963-1977. Ankara: Türk Dil Kurumu.

Telegdi, Szigmund. 1938. 'Eine türkische Grammatik in arabischer Sprache aus dem XV. Jhdt.' in *Körösi Csoma Archivum,* Ergänzungsband H.3. Budapest. 282-326.

Theorie Analyses. Analyses, theorie: Études Arabes Université de Paris. Departement d'Arabe - Vincennes.

Tomanov, M. 1965. 'Mahmud Kashgarī' in *Central Asian Review* (13; 4) 330-338.

Troupeau, Gérard. 1976. *Lexique-Index du Kitāb de Sībawayhi.* Paris: Klincksieck.

——. 1989. 'Voyelles et Semi-voyelles dans le *Kitāb* de Sībawayhī' in *Jerusalem Studies in Arabic and Islam* (12) 31-39.

TS = *Tarama Sözlüğü* Türk Tarih Kurumu Basımevi - Ankara 1965 - 1981. 8 vol.

TÜYATOK = Türkiye Yazmaları Toplu Katalogu [The union catalogue of manuscripts in Turkey]. Ministry of Culture. Ankara : Türk Tarih Kurumu Basimevi (1979-).

UAJ *Ural-Altaische Jahrbücher.*

Ülkütaşır, M.Ş. 1946. *Kaşgarlı Mahmut.* Istanbul.

Vajda, Edward J. 1994. 'Kazakh Phonology' in *Opuscula Altaica. Essays presented in honor of Henry Schwarz,* ed. by E. Kaplan & D. Whisenhunt. Washington: Western Washington. 603-50

Versteegh, Kees. 1977. *Greek Elements in Arabic Linguistic Thinking.* Leiden: Brill.

Versteegh, Kees. 1978. 'The Arabic terminology of syntactic place' in *Arabica* (25) 261-281.

—— 1980. 'The origin of the term *qiyās* in Arabic grammar' in *ZAL* (4) 7-30.

——. 1981a. 'A Dissenting Grammarian. Quṭrub on Declension' in *Historiographia Linguistica* (8) 403-429.

——. 1981b. 'La conception des "temps" du verbe' in *Theorie Analyses.* 47-68.

——. 1985. 'The Development of Argumentation in Arabic Grammar - The declension of the dual and plural' in *ZAL* (15) 152-173.

——. 1990. 'Freedom of the Speaker? The Term *Ittisā ʿ* and Related Notions in Arabic Grammar' in *SHAG* II. 281-293.

——. 1994. 'The notion of "underlying levels" in the Arabic grammatical tradition' in *Historiographica Linguistica* (23.3) 271-296.

——. 1995. *The Explanation of Linguistic Causes. Az-Zaǧǧāǧī's theory of grammar. Introduction, translation, commentary.* Amsterdam and Philadelphia: J. Benjamins.

——. (forthcoming) 'Maʿnā in the Arabic tradition' in *The emergence of semantics in four linguistic traditions, sanscrit, Hebrew, Greek, Arabic.* Wout van Bekkum et. al.. (eds.) Amsterdam and Philadelphia: Benjamins.

Versteegh, Kees, and Schippers, Arie. 1987. *Het Arabisch. Norm en Realiteit.* Muiderberg: Coutinho.

Versteegh, Kees, Konrad Koerner and Hans Niederehe (eds). 1983. *The History of Linguistics in the Middle East.* Amsterdam: Benjamins.

VOHD = *Verzeichnis der Orientalischen Handschriften in Deutschland; Türkische Handschriften* III 1-5.

Waetzoldt, Irene. 1978. *Zu den osmanischen Verbformen des 16. Jahrhunderts nach dem Mecmu-i Menazil des Matraqçi Nasuh.* Freiburg im Breisgau: Schwarz (Islamkundliche Untersuchungen; Bd. 48).

Wal, Marijke van der. 1992. *Geschiedenis van het Nederlands.* Utrecht: Het Spectrum.

Waugh, L. and M. Halle (Eds.). 1984. *Roman Jakobson: Russian and Slavic Grammar (Studies 1931-1981).* Berlin: Mouton.

Weiss, B. 1976 'A Theory of the Parts of Speech in Arabic (Noun, Verb and Particle): a Study in ʿIlm al-Waḍ ʿ' in *Arabica* (23) 23-36.

Wild, Stefan. 1965. *Das Kitāb al-ʿAyn und die Arabische Lexikographie.* Wiesbaden: Harrassowitz.

Wittek, Paul. 1928. 'Neuere Wissenschaftliche Literatur in osmanisch-türkischer Sprache. I.' in *Orientalistische Literaturzeitung* (3) 172-178.

Wright, W. 1986. *A Grammar of the Arabic language.* Translated from the German of Caspari. Cambridge: University Press. (2 vols) (First edited in 1859/1862).

Younes, Munther A. 1994. "On emphasis and /r/ in Arabic". In: Mushira Eid, Vicente Cantarino and Keith Walters (Eds.), *Perspectives in Arabic Linguistics VI.* Amsterdam and Philadelphia: J. Benjamins. 215-235.

Zajączkowski, Ananiasz. 1954. *Slownik Arabsko-Kipczacki z okresu Państwa Mameluckiego* Bulġat al-Muštāq fī luġat at-Turk wa-l-qifžāq. II *Verba.* Warszawa.

——. 1958. *Vocabulaire Arabe-Kiptchak de l'époque de l'État Mamelouk* Bulġat al-Muštāq fī luġat at-Turk wa-l-qifžāq. *I Le Nom.* Warszawa.

Zajączkowski, Ananiasz. 1965a. 'Chapitres Choisis du Vocabulaire arabe-kiptchak *ad-Durra al-muḍīʾa fī l-luġa t-turkiyya* I' in *RO* (29) I 39-98.

———. 1965b. 'Chapitres Choisis du Vocabulaire arabe-kiptchak *ad-Durra al-muḍīʾa fī l-luġa t-turkiyya* II' in *RO* (29) II 67-116.

———. 1968. 'Material kolokwialny arabsko-kipczacki w Słowniku «ad-Durra al-muḍīʾa fī l-luġa t-turkiyya»' in *RO* (31) II 71ff.

———. 1969. 'Chapitres Choisis du Vocabulaire arabe-kiptchak *ad-Durra al-muḍīʾa fī l-luġa t-turkiyya* III' in *RO* (32).

ZAL *Zeitschrift für arabische Linguistik.* Wiesbaden: Harrassowitz.

ZDMG *Zeitschrift der deutschen morgenländischen Gesellschaft.* Wiesbaden.

A. INDEX OF SUBJECTS

B. INDEX OF ARABIC TERMS

STUDIES IN SEMITIC
LANGUAGES AND LINGUISTICS

and the Subsequent Reception of the *Kitāb*. 1997. ISBN 90 04 10595 6

24. Belnap, R.K. and N. Haeri. *Structuralist Studies in Arabic Linguistics.* Charles A. Ferguson's Papers, 1954-1994. 1997. ISBN 90 04 10511 5

25. Talmon R. *Arabic Grammar in its Formative Age. Kitāb al-'Ayn* and its Attribution to Ḥalīl b. Aḥmad. 1997. ISBN 90 04 10812 2

26. Testen, D.D. *Parallels in Semitic Linguistics.* The Development of Arabic la- and Related Semitic Particles. 1998. ISBN 90 04 10973 0

27. Bolozky, S. *Measuring Productivity in Word Formation.* The Case of Israeli Hebrew. 1999. ISBN 90 04 11252 9

28. Ermers, R. *Arabic Grammars of Turkic. The Arabic Linguistic Model Applied to Foreign Languages & Translation of 'Abū Hayyān al-'Andalusī's* Kitāb al-'Idrāk li-Lisān al-'Atrāk. 1999. ISBN 90 04 113061

29. Rabin, Ch. *The Development of the Syntax of Post-Biblical Hebrew.* 1999. ISBN 90 04 11433 5