A. RÓNA-TAS

# AN INTRODUCTION TO TURKOLOGY

*studia uralo-altaica* 33

*studia uralo-altaica*

*33*

Redigunt
P. HAJDÚ
T. MIKOLA
A. RÓNA-TAS

All orders should be addressed to John Benjamins, Amsterdam, Holland.

UNIVERSITAS SZEGEDIENSIS DE ATTILA JÓZSEF NOMINATA

A. RÓNA-TAS

# AN INTRODUCTION TO TURKOLOGY

SZEGED
1991

Editionis curam agit

KLÁRA SZŐNYI-SÁNDOR

# PREFACE

The contents of this introduction had a long conception period. I tried to re- and reshape it since I begun teaching Turkology at the Attila József University, Szeged, Hungary, in 1974. I first outlined the basic contours of this variant in Bonn in the academic year 1982/83, when I tought there as a visiting professor in the *Zentralasiatisches Seminar*. Coming back to Szeged I changed the language from German to English and used the draft for graduate courses. In the meantime I learnt that several of my colleagues were preparing similar introductions. I almost left my manuscript unfinished when I met some of these colleagues and realized that our approaches were basically different. I became convinced that our respective introductions would fit into a greater framework complementing each other. Thus I set myself to finish the work. Unfortunetely other inevitable duties have hindered me to complete the entire work. This volume is only the first of a series but does not bear this numeral because I am in this respect superstitious. The second volume is practically ready. It will contain the introduction to the sources in the Manichean, Sogdian, Uighur and Arabic scripts. These systems of writing were used to render Old Turkic texts. In a further part I intend to deal with those writing systems in which we find Old Turkic words, names and isolated phrases, such as Chinese, Pahlavi, Georgian, Armenian, Greek, Latin, Cyrillic, Hebrew. In an Appendix I shall summerize our knowledge on the inscriptions written with the East European "Runic" script(s) wich I consider practically undeciphered. If space will allow I plan to add Syriac and Phagspaal though Turkic texts written in these systems pertain to the Middle Turkic period. Thus volumes I and II will offer a kind of graphematic and ortographic analyzis of Old Turkic. In Volume III all loanwords from Old Turkic into non Turkic languages will be dealt with. Here the earliest Turkic layers in Mongolian and in the Uralic languages, among them of course Hungarian, will have our main attention.

I hope very much that the three volumes will lay a solid foundation for a phonetic and a phonological analyzis of Old Turkic.

Compiling this book I have tried to write for graduate level students of Turkology. I hope, however, that it will be useful for undergraduates, as well, presupposing a well qualified instructor. Perhaps some parts will be useful for non Turkologists in the adjacent fields.

I would like to offer my sincere thanks to all my colleagues in Szeged who helped me with advises and corrections. My special thanks are due to Klára Szőnyi–Sándor who took care of editing the text and to Etelka Szőnyi for drawing the tables.

# CONTENTS

PART ONE–GENERAL INTRODUCTION ..... 7
THE TURKS AND TURKOLOGY ..... 9
TURKIC AND THE OTHER LANGUAGES ..... 14
Turkic and the Altaic languages ..... 14
Turkic and Uralic ..... 17
Turkic and Indo-European ..... 18
Turkic and the Ketic languages ..... 18
Turkic and the so called Nostratic language family ..... 19
THE PERIODIZATION OF TURKIC LINGUISTIC HISTORY ..... 21
General considerations on linguistic periodization ..... 21
The main periods of Turkic linguistic history ..... 24
THE CLASSIFICATION OF THE TURKIC LANGUAGES ..... 31
A SHORT HISTORY OF THE TURKIC PEOPLES ..... 34
THE MONGOLS. HISTORY, PEOPLES AND LANGUAGES ..... 41
THE MANCHU-TUNGUZIANS ..... 44
THE WORLD RELIGIONS AND THE ANCIENT TURKS ..... 46
Shamanism and Tengrism ..... 46
The spread of Buddhism to Central Asia and the Turks ..... 47
The spread of Manicheism to Central Asia and the Turks ..... 50
The spread of Islam and Ancient Türks ..... 50

PART TWO–INTRODUCTION TO THE STUDY OF THE OLD TURKIC LANGUAGES IN DIFFERENT SCRIPTS ..... 53
THE SOURCES OF THE HISTORY OF THE OLD TURKIC LANGUAGES IN DIFFERENT SCRIPTS ..... 55
THE SOURCES IN RUNIC SCRIPT ..... 55
The origin of the East Turkic Runic script ..... 57
The orthographical rules and the graphotactics ..... 58
Vowels ..... 58
Consonants ..... 60
THE SOURCES IN BRAHMI SCRIPT ..... 63
Historical background ..... 63
Tocharian and Tocharian Brāhmī ..... 63
Northern Saka or Tumshuquese ..... 65
Turkic texts in Northern Brāhmī ..... 67
Earlier Studies ..... 67
The texts ..... 69
The paleographical origins ..... 70
The new signs ..... 71
The transcription ..... 74
General remarks ..... 74
Graphotactics and transcription ..... 74
The representation of the Turkic vowels ..... 77
The representation of the Turkic consonants ..... 82
Khotanese Saka and Khotanese Brāhmī ..... 85

Turkic in Khotanese Brāhmī ..... 88
The representation of the Turkic vowels ..... 89
The representation of the Turkic consonants ..... 90
THE SOURCES IN TIBETAN SCRIPT ..... 92
Historical background ..... 92
Old Tibetan literature as a source of the Turkic linguistic history ..... 92
The Tibetan alphabet and Tibetan orthography ..... 93
The most important tendencies in the history of Tibetan ..... 95
Turkic texts in Tibetan Script ..... 95
The transcription and representation of Turkic vowels ..... 97
The vowels in initial position ..... 97
Vowels in postconsonantal position ..... 99
The functions of the *a-churi* ..... 100
The transcription and representation of the Turkic consonants ..... 102
Consonants in initial position ..... 102
Consonants in non initial position ..... 104
On the date of Pelliot tibetain 1292 ..... 107
Some remarks on the Turkic language of Pelliot tibetain 1292 ..... 107
Some special traits of other Turkic texts written in Tibetan ..... 109
The Turkic material in the translated texts: Pelliot tibetain 1283 ..... 109
Turkic names and titles in Old Tibetan texts ..... 110

TABLES ..... 111
Table I (The inventory of the graphemes of the East Turkic Runic Script) ..... 111
Table II (The Tocharian Brāhmī alphabet according to Krause and Thomas 1960) ..... 112
Table III (The graphic representation of the Turkic vowels in Tibetan) ..... 113
Table IV (The graphic representation of Turkic consonants in Northern Brāhmī ..... 114
Table V (The graphic representations of the vowels in Khotanese Brāhmī) ..... 115
Table VI (The graphic representations of the consonants in Khotanese Brāhmī ..... 116
Table VII (The Tibetan alphabet) ..... 117

REFERENCES ..... 118

APPENDICES ..... 129
Appendix I (The first attempt of Hitch to identify the new Tumshuquese letters after Hitch 1981) ..... 131
Appendix II (The Indian alphabets from Sander 1968) ..... 132
Appendix III (Text G, lines 10–18 from Gabain 1954 – „Tafel II") ..... 152
Appendix IV (The *Turkic Khotanese Wordlist* from Bailey 1974) ..... 156
Appendix V (The fragment of the Turkic *Sekiz Yükmek*) ..... 169

# PART ONE
# GENERAL INTRODUCTION

# THE TURKS AND TURKOLOGY

Turkology is a branch of science which deals with the Turks. In the broadest sense it is interested in any fact which is connected with the Turks, but in a narrower sense it deals with their language, history, literature and other cultural activities. In this book the term *Turkology* will be used for one aspect of these studies, namely the linguistic one. Accordingly Turkology will be used as a term denoting the study of the Turkic languages, their history and present state. We shall interested ourselves in the history of the Turkic languages with the aim of reconstructing this history, but also of gaining data on the history of the people and their culture. In the second part of this book we shall also try to give a picture of the present Turkic languages with the aim of providing the essentials for their comparison.

Before we go ahead we have to define the term *Turk*. In the usage of the present times the name *Turk* is used to denote two different groups of people. In a narrower sense *Turk* is the name of the Turkish speaking population of the Republic of Turkey. In the following we shall call their language *Turkish*. In some books this Turkish language is called *Ottoman Turkish, Turkish of Turkey* and so on. We shall use *Ottoman Turkish* for the old literary language of the Ottoman Empire. During the long history of the Ottoman Empire there existed also a spoken language with many dialects, these spoken idioms we shall quote as *Osmanli*. About the periodization of these idioms and of their relative chronology we shall speak later.

In a broader sense Turk denotes any people which speak a language which belongs to the Turkic group of languages. For the languages spoken by the Turkic people we shall use the generic term *Turkic*. The contrast *Turkish* and *Turkic* is artificial and gained recognition only in the last few decades, it has been modelled after the opposition *German* for the German language and *Germanic* for a larger group of languages to which belong also English, the Scandinavian languages Norvegian, Swedish, Danish and also Flemish and the language of the Netherlands. A few English scholars oppose the usage of *Turkic*, they suggest using *Turkish* for the whole language family and *Republican Turkish* for the Turkish of Turkey. The other European languages also have problems with this distinction. In German *Türkisch* is used both for the smaller and the larger group, recently *Türkei Türkisch* is appearing for the smaller group, *Osmantürkisch* is used for the old literary language. In French the name *turc* denotes both the smaller and the greater group, but also here *turc* as well as *turque du Turquie* is used. In Russian until the thirties of this century *tureckiy* was the term of all languages, but since then a new differentiation arose, and now *tureckiy* denotes only Turkish and *turkskiy* is the term denoting Turkic.

In this book *Turkology* will be used for the study of the Turkic languages independent of when they were spoken.

We shall speak about historical Turkic languages spoken in the past and present or modern Turkic languages spoken recently.

There does exist a considerable uncertainty about the name *Türk*. Not only its etymology is debated but also its early pronunciation and usage. Some authors claim that the oldest fixed form of this name is not *Türk* but *Türkü*. This opinion is based on the following facts: In the East Turkic Runic inscriptions we find the word written with letters *TWRK*, where the last letter *K* is written with a special sign which occurs only before and after *ö* or *ü*. This would entitle us to read *türkü, türük, türkö, türök*. It is however almost a rule without exception that in the great inscriptions final vowels are written in full, i.e. with a separate sign, therefore only the forms *türük* and *türök* would be plausible. "Would be" in the previous sentence refers to the fact, that all this is true only in the case of non-reduced vowels. So this written form as well reflect a reduced labial in word final position. In the Chinese sources we find two hieroglyphs, which render the pronunciation of the name of the Turks. The first had to be read in Middle Chinese as *t'uɑt* or *d'uɑt* (Karlgren No. 489a), and is pronounced in the present Chinese language of Beijing as *Tu* (*t'u*), the second was pronounced as *kiwat* (Karlgren 301c) and is pronounced in modern Pekingese as *que* (*chüe, küe*). Later on we shall deal in a more detailed way with the Chinese transcription of Turkic, here it will suffice if we state that the present pronunciation of these hieroglyphs is *Tuchüe*. The Middle Chinese rendered something as **türküt*, or *türkür*, but the second form makes no sense. The first is not the basic form of the name, but reflects a non Turkic form with a suffix. It has been suggested that the final *-t* is the sign of the plural, and in fact some ethnonyms are frequently used in the plural, as we would say "the Turks". This plural can be either Sogdian or Mongolian, but not Turkic. The Mongolian plural is in fact *-d* and can be used only for words ending in *-n, -l, -r*, thus a form *Türküd* would be possible only if the singular would have been *Türkün, Türkül* or *Türkür*, none of which is anywhere recorded. The plural in *-t* in Sogdian is not dependent on the final, so it is very likely that the Chinese form reflects a Sogdian *twrkwt* as Harmatta (1972) suggested. This is also historically very plausible, since as we shall see that the Sogdians played an important role in Central Asia at the time of the appearance of the Turks. And in fact in the recently discovered Inscription of Bugut which is written in Sogdian we find the form *trkwt*. The Sogdian form **türküt* tells us nothing about the final vowel of the name, because if the original name was *Türk*, a connecting vowel *-ü-* would have had to be added because the cluster *rkt* in the final was unpronounceable. According to Bailey (1985 102) the ending would be an Iranian adjective suffix *-kut*, thus **türküt* would be in fact an adjective and not a noun (more precisely

*türk-kut*). This possibility is connected with the etymology of the ethnonym. Some authors quote the Tibetan name of the Turks. This occurs in the early Tibetan sources as *dru-gu*. The Tibetan could have written also *drug* (in some sources it occurs, but seems to be secondary) but it could not write *turk* or *durk*, because it had no final *-rk*. The Tibetan form is however not a direct reflection of the name. It is of Khotanese origin. In Khotan the inhabitants spoke an East Iranian language until the end of the 10th century which was called Saka. In this language *ttrukä* is the original form but in later texts *tturka* is also found (see Bailey 1985). The initial double *tt* denotes only a simple stop *t-* and the final *-ä* was not pronounced in later Khotanese Saka. The first form thus reflects a name *truk*, perhaps earlier *truka*, the second reflects *Türk*. The Khotanese form *truk* seems to be of great age, because the Tibetan form can be only considered as the reflection of *truk* and not of *Turk*. The latter would have been rendered as *dur-gu* or *dur-ku*, as e.g. the name *Türgesh* is rendered as *dur-gyis*. Considering the Khotanese *Truk(a)* we have two options. Either this is the original name or this is a secondary development in Saka. If we have to do with the first possibility we are confronted with a name of non Turkic origin which has been transferred to the Turks who adapted this name *Truk* to their phonetical system as *Türk*, a case which is very frequent with ethnical names, let us only refer to the name of the Russians or the Hungarians. In this case all etymologies which later connected the ethnonym *Türk* with Turkic words would be late folk etymologies. This is also very frequent with ethnonyms. In the second case *Truk(a)* would be a Khotanese Saka development and would render an earlier *Türk* with perhaps a vowel at the end of uncertain quality. Even in this case we cannot be certain whether the name itself is of Turkic origin, since the Sakas could have learnt it from other people, or from the Turks themselves but even in the latter case it is not necessarily a Turkic word, they could have given themselves a name of non Turkic origin as do the Russians, or the French, (as we know *Rus* is not a Slavic nor *Frank* a Romance word).

Most scholars, however, opt for the Turkic origin of the name *Türk*. The opinions about the Turkic etymon of the name *Türk* can be divided into two groups. To the first pertain those who depart from the adjective *türk*, the others try to find other words. To the latter pertain scholars who derive the name from a suffixed form of Turkic *törü-* 'to come into existence, to be created'. This idea is based on the fact that many ethnic names are connected with the concept of 'being'. This suggestion has however to be refuted, because the vowel of the first syllable is safely everywhere *-ü-* or its regular reflex. Some tried to combine the name *Türk* with the noun *töre* 'place of highly esteemed guests in the yurt' or *törü* 'customary law', both ideas have to be refused because of phonetical and also semantical difficulties. An idea that the basic word was not *türk* but *tür*, to which suffixes as *kü* and *n* or *t* have been attached is

also not acceptable. The only Turkic word which has a chance of being connected with the ethnonym is *türk*. But also here we are confronted with different opinions and confusing facts. The word occurs in the Old Turkic texts, but as we shall see its exact meaning in these text is also debated. Then it occurs in the Divān of Mahmud al-Kāshgari, where we read the following: *"TRK ' name of a city in the country of the Turks', TURK 'Name of the son of Noah, God's blessing be upon him. This is the name by which God called the sons of Turk son of Noah, just as 'Man' is the name of Adam, peace be upon him, in the following verse [citations from the Koran follow...], ...it is a collective noun, since there is no one who can be excluded from this singular. In the same way 'Türk' is the name of the son of Noah, in the singular; but when it refers to his sons it is a collective - like the word 'human', it is used for singular or plural. Likewise, 'Rum' is the name of Rum son of Esau son of Isaac, God's blessings be upon him, and also his sons were called by that name. I state that at-Turk is the name given by God. This is on the authority of the venerable Shayk and Imam, al-Husayn Ibn Khalaf al-Kāshgari, who was told by Ibn al-Gharqi, who said...."* and follows the praise of the Turks, also a poem and then it is continued: *"The singular is TURKu as well as the plural. kim sän 'who are you?'; türk män 'I am a Turk'; türk süsi atlandi 'The troops of the Turks mounted"*. The point is here that Kashgari joining the Mohammedan tradition connects the ancestor of the Turks with Noah, and further he has to make clear to the Arabs that the word *Türk* has no special plural form as most of the Arabic ethnonyms have. (Later the form *terak* was used). He points to the fact that this is a collective noun as well, it should not to be suffixed e.g. with *-ler*, so you have to say *türk süsi atlandi* and not **türkler süsi atlandi*, or *atlandilar*. Here the only problematical place is where Kāshgari writes once the singular as *TURKu*. In the given place the word is written with two dammas, and until the last edition of Dankoff and Kelly (1982, 274) it was either neglected or uncertain. Now according to Dankoff the second *u* has been written by a later hand and does not pertain to the original. After this item follows a third one: *TURK* a particle of time indicating 'the mid-point of the ripening of a fruit'. Thus *türk üzüm ödi* 'the mid-time of ripening grapes', *türk quyaš ödi* 'mid time (i.e. noon)', *türk yigit* 'a young man at the prime of his youth' (all texts in the translation of Dankoff op.cit.). The only modern spoken Turkic language where this word seems to survive is Kirghiz, where the word *türk* denotes 'big, fat, about sheep; ob ovcah krupnyj i žirnyj' (Yudahin). In Ottoman we find *türk* as 'young, beautiful' but it is uncertain whether this meaning is a secondary development of the ethnonym or the same word as in Kirghiz. This ethnonym has in many languages the meaning 'young, beautiful (people)' and the opposite 'terrifying, coarse (people)', as e.g. in Persian or in French, German. Thus Clauson (1962) seems to be right when he writes that in none of the Turkic languages of the Middle or Modern period

does the word mean 'strength'. He also concluded that the idea that *Türk* should be derived from a Turkic word for 'strength', as suggested by Németh (1926) and accepted by most scholars should be abandoned. The question of the word *türk* in the Old Turkic texts is, however another case. In some texts it stands alone and it is not always clear whether it denotes an ethnic group or is a qualifier, as e.g. in *türk bodun* 'Türk people'. In some other texts it is the second part of a so called hendiadys, a binomial expression where both parts of the construction have the same, or almost the same meaning. Where *türk* stands as part of a binomial expression it is always *erk türk*. The word *erk* has the basic meaning in these texts 'power, strength, will, free will'. It had a disyllabic form *erik* from which the Hungarian word *erő* 'strength' can be derived. It was borrowed also by Mongolian in the form *erke* and has there the meaning 'power, authority'. The word *türk* had to have the same meaning as, or a semantic field very near to *erk*. Taking into account the data of Kāshgari the noun *türk* had to denote the strength and authority reached by somebody when he was at his biological maturity, a similar meaning was also used for animals. It is unlikely that this should have been the name of an ethnic group, but it could have become the *epiteton ornans* of a ruling clan 'the mighty, those who have the authority, the strength, the power', and from this it became first the name of a tribe ruled by this clan and later the name of all people speaking the same language or living in the same Empire.

This solution does not exclude the one mentioned earlier, namely that the name *Türk* is finally of non Turkic origin. It may have been of foreign, non Turkic origin, but later when it became the name of the ruling Turkic clan it was reinterpreted as a Turkic name and used consciously as such. In 551 A.D. a coalition of tribes under the leading tribe of the Turks founded a new Empire in Inner Asia and later the name became the name of this Khanate and the name of the people who pertained to it. Even later, after the dissolution of the first Turkic Kanate the name became a generic name for all groups of people speaking a language related. According to the Chinese Annals the Turks got their name from the Golden Mountain which had the form of a helmet, and a helmet is called in their language **türküt*. No such or similar word is hitherto known from any Central Asiatic language, with the exception of Khotanese Saka where we find *tturakä* 'cover', in Turkic *qapqaγ*. (On the earlier history of the question without the last Saka data see Doerfer TMEN II 483-495, Clauson 1962.)

We see that the name *Turk* had its own history before it became a term in European scholarship. Turkology in the narrower sense used in this book deals with the languages of the Turks.

# TURKIC AND THE OTHER LANGUAGES

*Turkic and the Altaic languages*

*Altaic* is the name of a language family. *Language family* is a term used for a special type of a group of languages. To understand what exactly a language family is we have to know the difference between such terms as *linguistic identity, linguistic similarity* and *linguistic correspondence*. Two linguistic forms are identical if and only if their phonetic shape and semantics are the same. Similar are two units of a language when they are not identical but their form and/or their meaning resemble each other. Two linguistic units (words, phrases, structures etc.) correspond to each other if their difference can be described by stating regularities. Languages cannot be grouped according to their linguistic identities because in no two languages do identical features exist, at best features are only similar. Even in cases of formal (phonetic) and semantic identity their usage, i.e. the rules for their use in the different languages are different.

To a language family belong languages which have the same genetic affinity, or with other words which belong to the same group of languages where the group is defined by the genetic principle. The genetic principle says that two or more languages belongs to the same genetic group of languages if and only if their basic subsystems can be followed back with help of tracing regular changes to a common language. The theory of linguistic relationship needs further elucidation but we cannot go here into details (see for details Róna-Tas 1978). In the genetically related languages we find regular correspondences but neither identity nor similarity is needed to prove genetic relationship. Therefore neither identities nor similarities tell us anything about the genetic relationship of languages.

Similarity between languages can be due to several factors of which the most important are the following: 1. chance, 2. typological affinity, 3. convergency from independent different points of departure, 4. historical causes. The historical causes can be: a. genetical relationship, b. areal contacts, c. the impact of a language from outside, which may act as a superstrate or a substrate. As we see similarities can be due to many factors. For the linguist those features are of interest which are due not to superficial similarities but to regular correspondence. Chance can never cause regular correspondence, but all other factors can. Therefore even regular correspondences between two languages do not necessarily say anything about their genetic relationship. The genetic relationship is a very special kind of the possible grouping of languages. Why has the genetical grouping a special status in linguistic science?

Languages can be grouped by criteria which can be arbitrary or non arbitrary, exhaustive or non exhaustive and unambiguous or not unambiguous.

If a criterion is non arbitrary this means that this grouping is the only possible. If a grouping is non exhaustive this means that there may exist a language which belongs to none of the groups, and if a grouping is not unambiguous this means that a language may belong at the same time to two or more groups (see Greenberg 1957). How we can group languages? In many ways. The simplest way is to group languages according to their geographical distribution, as "languages of Africa", "Baltic languages" (languages spoken in the Baltic and not the Baltic branch of the Indo-European languages, thus including also Estonian not only Lituanian and Latvian), "languages of Inner Asia". Such a grouping is arbitrary, because it is arbitrary where we draw the borders, it may be exhaustive but it is not unambiguous, any of the languages may belong to another geographical grouping, e.g. the languages of Inner Asia to the languages of Asia. Languages can be grouped according to their typological character. We can speak e.g. of languages which have vowel harmony, languages which have the sequence Subject Object Predicate, languages which have no grammatical gender. Such groupings are arbitrary, because any such features or their combination can used for grouping. Then typological grouping is exhaustive, because if we have two groups as e.g. the languages with vowel harmony and the languages without it, any of the languages will belong to one of the two groups, and typological grouping is unambiguous because there will be no language which could belong to both of the groups at the same time. Genetic grouping is a way of grouping which is non arbitrary. There do not exist two groupings made according to the genetic principle which could be equally true, any language pertains to one and only to one genetic group.

| | non arbitrary | exhaustive | unambiguous |
|---|---|---|---|
| areal | - | - | - |
| typological | - | + | + |
| genetic | + | + | + |

*TABLE I.*

The grouping according to the genetic principle has a second advantage. If properly used it can offer important insights into the history of the people who spoke it.

Now we can turn to the term *Altaic linguistic family*. The name *Altaic* comes from the Altai mountains in Central Asia, once supposed to be the original homeland of the people who spoke the Altaic languages. The term *Altaic languages* has changed its content during the last two hundred years. In the

middle of the 19th century some scholars (as e.g. W. Schott) used the term *Altaic languages* for a group which we would call nowadays *Uralic and Altaic languages*. From the end of the last century *Altaic languages* denoted the Turkic, the Mongolian and the Manchu-Tunguzian languages. Later on some scholars suggested that Korean should be also called Altaic, and finally there are some scholars who claim that Japanese is one of the Altaic languages.

The study of the history of the Korean and the Japanese languages in respect of a comparison of these languages with other Altaic languages has not reached a scholarly level which would enable us to use these data for Turkological purposes. If such a genetic relationship ever existed it is so remote that with our present data and methods the output of the comparison of Korean and/or Japanese with the Turkic languages can be ignored. Neither Korean nor Japanese ever had any traceable linguistic contacts with the Turkic languages thus, unlike Mongolian and Manchu-Tunguzian their study does not help the reconstruction of later phases in the history of the Turkic languages. For a long time in the 19th century the genetic relationship of Turkic, Mongolian and Manchu-Tunguzian was not a hypothesis which had to be proved, but evidence which had to be demonstrated. The claim that the Altaic languages are genetically related became the object of serious scholarship with the works of the great Finnish scholar G. K. Ramstedt. His views were further elaborated by N. Poppe and some other scholars. According to them it can be demonstrated by sound linguistic methods that the three groups, Turkic, Mongolian and Manchu-Tunguzian are genetically related. Ramstedt in his later works included also Korean in the Altaic linguistic family. The views of Ramstedt and Poppe slowly became generally accepted in the first half of this century. A few scholars remained sceptic, others formulated their views more cautiously, but by and large until the middle of this century the genetic relationship of Turkic to Mongolian and Manchu-Tunguzian was considered as proved. Beginning with the fifties a few scholars tried to challenge and refute the arguments of Ramstedt and Poppe. Among them Clauson, Doerfer, Shcherbak and the author of this Introduction could be mentioned. (For my arguments see in detail Róna-Tas 1974b). As a conclusion of the "Debate of the Altaists and Anti-Altaists" as this discussion is usually called, two conclusions have been unanimously accepted. The one is that the Turkic and the Mongolian languages on the one hand and the Mongolian and Manchu-Tunguzian languages on the other had long lasting and early contacts, which caused several layers of loanwords. Whatever the case with the genetic relationship of the three groups may be, first these later layers have to be separated and only what remains can be considered as the stock of a possible common Proto Altaic language. The discussion goes on only about what we have to consider as the earliest and therefore (if such existed) as the common stock. The other conclusion was that the similarities

and correspondences among the three branches can be due to several factors. Some turned out to be secondary typological convergencies others are due to areal contacts. But in any case, it became clear that the most essential problems of the history of the Turkic languages cannot be solved without the profound knowledge and analysis of the Mongolian and in some respects of the Manchu-Tunguzian linguistic history. The opinion of the author of this Introduction is that those linguistic correspondences which have been quoted by Ramstedt, Poppe and their followers as arguments in favour of the genetic affinity of the Altaic languages cannot be accepted as such. They witness early contacts and are loanwords. Nevertheless after having separated these very old layers, the remaining very thin layer may pertain to a common Altaic proto-language. We shall return to some of the questions of detail and demonstrate them by the help of examples.

*Turkic and Uralic*

Two large groups belong to the Uralic linguistic family: the Finno-Ugric and the Samoyed languages. For a long time in the 19th century scholars believed that the Uralic and the Altaic languages were genetically related. This was categorically denied by Ramstedt, but a pupil of Ramstedt Räsänen was an adherent of this view, and in a few papers Poppe also returned to this opinion. Some Uralists, e.g. B. Collinder, also argued in favour of the relationship of Uralic and Altaic (see the discussion in JSFOu 1983, for my views see there 235-251). A very special view has been expressed by J. Németh. Németh did not speak about linguistic relationship among Uralic and Altaic but about special "relation-like old contacts" between Turkic and Finno-Ugric (see Németh's view in Németh 1942-47 and my analysis of his views in Róna-Tas 1983c). Another formulation can be met in the works of D. Sinor. According to him the respective languages have to be placed as follows:

Samoyed Tunguz

Finno-Ugric Mongol

Turkic

This means that Turkic had special early contacts with Mongolian and Finno-Ugric, Finno-Ugric with Turkic and Samoyed, Samoyed with Finno-Ugric and Tunguz, Mongol with Tunguz and Turkic (see now in detail Sinor 1988). This hypothesis would locate the five proto-languages in an early area where the contacts were circular but not diametral.

There is no doubt that the Turkic languages had very early contacts with both the Proto-Samoyed (see Róna-Tas 1980) and the Finno-Ugric languages. For reconstruction of the Turkic linguistic history very important are the early loanwords from an r-Turkic language in Proto Samoyed. The continuous contacts of Turkic languages with the Finno-Ugric languages offer a great help in reconstructing Turkic. In fact the earliest monuments which we can reconstruct for the history of the Turkic languages are the loans in Mongolian, but very early are also those borrowed by Proto Samoyed, Ugric and Hungarian (see the detailed study in Róna-Tas 1988b).

*Turkic and Indo-European*

The Turkic languages had very early and from then on continuous contacts with the Iranian group of the Indo-European languages, to which belong not only Persian, but also several other languages, such as Saka, Sogdian, Khvarazmian etc. But since the end of the last century scholars have also dealt with the possibility that Turkic or Altaic had early contacts with the not yet differentiated Indo-European or with very early Indo-European languages other then Iranian or even Indo-Iranian. Ramstedt (1946-47) and Németh (1942-47) expressed the view that some very early migrating cultural words reached the Altaic (Ramstedt) or Turkic (Németh) languages. In some cases the direction of borrowing is uncertain. Németh called attention to a few early Tocharian loanwords in Turkic and I suggested very hypothetically a few more parallels (Róna-Tas 1974a, on which now see the critical evaluation of Ivanov 1988). Turkic parallels with early Indo-European forms have been dealt with by Gamkrelidze and Ivanov in their recently published book (Gamkrelidze and Ivanov 1984) and I have discussed the problem recently (Róna-Tas 1990). The result is that in Turkic but not in Mongolian or in Manchu-Tunguzian a small group of early migrating cultural words can be found which do not seem to result from early language contact. The oldest form of Turkic was spoken in an area where these migrating words could easily reach it. Though I am even more sceptical about the validity of my own Turkic-Tocharian comparisons as Ivanov seems to be, some traces of such a contact can be supposed. Another source of the migrating cultural words quoted above could come from the language group which Gamkrelidze and Ivanov call Old European, i.e. the common ancestor of Balto-Slavic, Illyrian, Italo-Celtic and Germanic.

*Turkic and the Ketic languages*

The group of Northern Asiatic languages which do not belong to any of the known linguistic families and are most probably not even related to each other are called Paleoasiatic languages. Among them only one is of interest to Turkological studies, this is *Ketic* or as it was earlier called by some authors

*Yenisei Ostyak* (not related to the Finno-Ugrian Ostyak language, and this name was later abandoned). Ketic itself is a group of languages, some of which died out only in the 19th or even 20th centuries. Besides Ketic proper such local variants as Arin, Assan, Kot, Pumpokol were recorded in the 18th century (see Dul'zon 1961). These languages were or are still spoken around Krasnojarsk, North to the area where the present Khakass language is spoken. It has been suggested that the Ketic language earlier played a greater role, some even thought that the Asiatic Huns or the Hsiung-nu people spoke in this language (see Ligeti 1950, Pulleyblank 1961-1962). From the middle of this century Ketic studies got a new impetus (on earlier studies see Jakobson; Hüttle-Worth and Beebe 1957, Vdovin-Tereshchenko 1959), but we are far from having a reliable material for early, or reconstructed Ketic. Nevertheless a few early Turko-Ketic parallels show that further investigations promise to be fruitful (some parallels are dealt with in my unpublished dissertation Róna-Tas 1970).

*Turkic and the so called Nostratic language family*

The name *Nostratic* was coined by H. Pedersen (1903) from the Latin *noster* 'our'. The hypothesis has been revived and placed on new foundations by V. M. Illich-Svitych and his school. According to this hypothesis there existed a common proto-language of the following six great linguistic families: Altaic, Uralic, Indo-European, Dravidan Semito-Hamitic and Kartvelian. Dravidan is the common name of the group the languages which had been spoken on the Indian subcontinent before the Aryan Indic people occupied it, and many of the present day languages in India and Sri Lanka are descendants of Dravidan, the most well known of which is Tamil. Kartvelian is the name of a group of Caucasian languages of which the most important is Georgian (many of the Caucasian languages are not related to Kartvelian). It is the merit of Illich-Svitych that he observed that of these six linguistic families the most problematical is Altaic, and its members (that is Turkic, Mongolian and Manchu-Tunguzian) are farther from each other than any other two members of any other linguistic families (Illich-Svitych 1971, 69). Nevertheless Illich-Svitych and his followers claim that the six language families are genetically related. The methodological problem with this and similar claims is that we do not have the adequate methods to investigate the case. We have to work with rules and regularities which can be observed in cases of or abstracted from extant languages. The functions of the language and the way it changed were essentially formed in the Neolithic Age. It needed a certain size of group of speakers who lived together constantly, a certain amount of steadiness in the migration of the people and it is not a mere chance that all known language families can be followed at best to the Neolithic Age. The human language is of course much older than the Neolithic Age, but known languages do not go beyond the beginnings of the Neolithic

Age, and therefore it is theoretically futile to research any proto-language which must have existed earlier than the Neolithic Age as is the case with Nostratic. A second methodological problem of the Nostratic school is that the reliability of any reconstruction has an inverted probability to the number of the asterisked, hypothetical forms which are needed to bridge over the gap between the known data and the proto-language. Those who beleive in the Nostratic theory need a long chain of hypothetical forms to reach their final reconstruction. If there existed an Altaic proto-language this had to be extant earlier than 5.000 B.C. and the Nostratic proto-language, if it ever existed, must have existed many thousand years earlier. We know nothing about eventual changes during these many thousands of years, nor about foreign influences etc. The Nostratic hypothesis does not help in Turkological studies.

To sum up: Early Turkic had very close contacts with Mongolian, important early contacts with Proto Samoyed, with the earliest Finno-Ugric languages, with Tocharian and Old European, later with Iranian, and perhaps with early Ketic. Later several Turkic languages came into contact with other languages as well. All these linguistic contacts help us to reconstruct the history of the Turkic language(s).

# THE PERIODIZATION OF TURKIC LINGUISTIC HISTORY

*General considerations on linguistic periodization*

Periodization in historical linguistics is a practical device. Since language changes continually, and there are no leaps and bounds in the stream of linguistic history, we cannot accept any periodization based on the specific internal features of the language, which does not mean that we cannot characterize periods in them. The segmentation of the bundle of changes we give depends on our concept of language situation, on our methods of linguistic reconstruction and the character of the sources at our disposal. By language situation we understand the language as a kind of action and its social environment. The periodization of linguistic history is practical insofar as it serves as a short cut instead of a cumbersome description of chronological relations and helps to treat chronologically different data in an easier way.

Since most of the past stages of linguistic history appear as a result of linguistic reconstruction we have to mention briefly the discussion about the status of reconstructed languages. Some scholars claim that reconstructed languages are identical with the living languages of the respective period, other deny the reality of the reconstructed languages. We can accept neither of the two extremities. Reconstruction is an approach which always tries to approximate the natural language of a past period, the grade of approximation may be greater or less depending on many circumstances. But it is a natural language which is or should be reflected by the approximation. Thus we may know fewer rules of a reconstructed language than existed but a reconstructed language has no rules which did not exist in fact, save they are more simple. This already had been clearly formulated by Bloomfield (1933, 302-303) who distinguished occurent and reconstructed languages. The latter has to try to reflect the former as closely as possible, but, by definition, it can never be identical with it.

In most cases reconstruction is not devoid of contradictions. Some of these contradictions are simply due to the weakness of our methods or to insufficient data. Other contradictions can be eliminated by assuming or recognizing secondary developments. But most reconstructions have some inherent contradictions, and many attempts have been made to master this problem. One of such attempts was the famous wave theory by J. Schmidt, (1872) who supposed that those contradictions which were inevitable in the family-tree model of Schleicher (1871) can be solved by supposing that some changes did not occur along the lines of the family tree but spread as waves over the territory inhabited by the speakers. Trubetzkoy (1933) tried to overcome these contradictions by supposing that Indo-European developed from different languages through an areal, secondary process, ("they became Indo-European languages"). He assumed that what can be reconstructed without contradictions originated

from a secondarily converged areal group of languages, and the contradictions are due to the fact that originally the languages pertained to different genetic groups. A third solution has been suggested by the Turkologue E. V. Sevortyan (1971). According to Sevortyan it pertained to the nature of past languages that they had contradictions. Forms of the Turkic verb word for 'to sing' can be reduced to two old forms *et-* and *öt-*, but these two cannot be reduced to one, that is, we don't know which of the two or a third is the earlier. According to Sevortyan such "doublets" pertain to a stage of the history of the Turkic language which he called Common Turkic (obščee tjurkskoe sostojanie). This solution, in the form Sevortyan suggested, is likewise unacceptable. In the case of *et- ~ öt-* Sevortjan speaks of the "alternation" of *e* with *ö*. This is an abuse of the term "alternation". The term "alternation" can be used only in cases of free variants within a dialect or idiom. Free variants are forms which can be used by one and the same speaker whithout changing the meaning or grammatical function of the word (or morpheme). This was however not the case with the forms *et- ~ öt-*. There is no Turkic language where these two forms freely interchange and we have no reason to suppose that it ever existed. If we suppose that *ö* alternates with *e*, *e* alternates with *i*, *i* alternates with *ï*, *ï* alternates with *u*, as Sevortyan supposed, then we reconstruct a language where practically all vowels can be freely used instead of one another. This would mean the neutralization of their phonemic opposition (labial for illabial, and front for back having been explicitly suggested by Sevortyan, 1971 8).

Doerfer suggested (1971, more coutiously 1975-76) that if we have "contradictory" correspondence what we have to do is to reconstruct the different prototypes. For instance if we find:

| Turkic | Chuvash |
|---|---|
| *ār-* 'to get tired' | *ïr-* |
| *qāz* 'goose' | *χur* |

we have to reconstruct different proto-phonemes, say $a_1$ and $a_2$ for Proto Turkic. No doubt in some cases this procedure is justified. But let us extend this series to Yakut:

| Turkic | Chuvash | Yakut |
|---|---|---|
| *ār-* 'to get tired' | *ïr-* | *ïr-* |
| *qāz* 'goose' | *χur* | *χas* |
| *āy* 'moon' | *uyaχ* | *ïy* |
| *yār-* 'to split up' | *śur-* | *ïr-* |
| *qān* 'blood' | *ɣun* < **χiun* | *χan* |

In this case we would have to reconstruct five proto-phonemes. Now if we add that the Tuvanian *ay* corresponds to the Yakut *ÿy* 'moon', while both Tuvanian *χĭrĭn* and Chuvash *χĭram* correspond to the Yakut *χarĭn* 'belly', we find that we have either to give up, or to reconstruct as many proto-phonemes as there are "contradictions". While Sevortyan's approach menaces with oversimplification, the way of Doerfer, if this were the only possible solution, leads to overcomplication (in 1971 he suggested 30 vowel phonemes for Proto Turkic, but in the system lie even more).

If we distinguish proto-languages and occurrent languages, it will be clear that any proto-language projects its findings on a synchronous screen, although the features themselves existed in a number of different places and at various times. This is the first reason why we see "contradictions" on the screen. I propose to call the earliest stage of the history of the Turkic languages *Ancient Turkic* (1970, in 1982 I used the term *Ancient Turkish* but *Turkic* is the more appropriate one), which has to be divided into two periods: Early Ancient Turkic (EAT) and Late Ancient Turkic (LAT). EAT lasted from the dissolution of the Altaic unity (or from the end of Pre-Turkic) until the appearance of those dialects which later became the respective nuclei of the various Turkic languages and language groups. In the EAT period, there was what might be called a heterogeneous linguistic unity. There was one language, spoken with a number of local differences. The differences, like the groups speaking them, were unstable. The various groups understood each other, and had contacts of varying degrees of intensity. Along with the historical changes taking place, the language situation slowly changed, too. Some groups became more stably connected and slowly dialects appeared. With the appearance of dialects, LAT took shape. The dialectal features appeared - as in all languages - as isoglosses. Isoglosses are imagined lines which connect in the geographical space features which pertain together. Many of the isoglosses coincided and formed bundles, but some did not, they simply crossed the others. Unless we keep in mind this finding of modern dialectology, we will be unable to understand the problems connected with the reconstruction of Proto-Turkic. Not only isoglosses can cross each other, also some single features may do so. In a certain case the distribution of a phonological feature is marked by an isogloss, on the one hand there is a feature X on the other hand we find feature Y. Most of the words follow this distribution but sometimes we find "undisciplined" words which do not "respect" the main phonological border and cross it. This happens in many cases with living dialects. When those speaking these dialects went their separate ways for historical reasons, they took with them also the "undisciplined" words. If we do find them in later, already independent languages we find apparent "contradictions". Thus one reason for the "contradiction" is our system of backward projection on a synchron screen. The second is that in fact the spatial

distribution of the features cannot always be clearly demarcated. There is also a third reason for these "contradictions". Language never changes in an instant. Because of several reasons there develops a tendency to change a feature X in a given language to feature Y. This tendency affects most of the words in the given category. Sometimes it affects all members of the given category in other cases only most of them. But even those members which are affected do not change at once. Frequency of use, more stable occurrence in frequently used expressions can delay the change. If the tendency to change is just starting when the speakers of the various dialects go their separate ways, it may happen that the change takes place with varying intensity in the different dialects. The shortening of the primary long vowels is a tendency which can be observed in all Turkic languages. The speed with which this change took place, however, differed greatly. We have reason to suppose that LAT was a language where the rules governing the lives of the dialects were effective.

If we exclude those "contradictions" which are due to our lack of knowledge, all other "contradictions" can be solved if we suppose that the Ancient Turkic language was a living, occurent language, which had chronological and dialectal dimensions and differences, and our data contradict each other because they reflect either chronological or dialectal differences, isoglosses etc. in this Ancient Turkic language. It is not always the task to ignore these problems, our task is to reconstruct the complicated system as it was. No doubt, such a reconstruction can never be perfect, but it has to try to attain a maximum of approximation.

*The main periods of Turkic linguistic history*

In my 1982 paper I suggested the following periodization of the Turkic linguistic history:

Altaic/Preturkic
Ancient Turkic
    Early Ancient Turkic
    Late Ancient Turkic
Old Turkic
    Early Old Turkic
    Late Old Turkic
Middle Turkic
    Early Middle Turkic
    Late Middle Turkic
New Turkic
    Early New Turkic
    Late New Turkic
Modern Turkic

*Altaic* is a term used for the common language presumably spoken by the common ancestors of the Turks, Mongols and Manchu-Tunguzians, and probably by other groups as well. It has to be stressed that this language, too, must be thought of as having its own long history, its territorial variants and areal subgroups. As we have seen there is a very important theoretical border over which we cannot follow the history of these languages backward: the beginning of the Neolithic Age. As we have stated above in the Neolithic Age the nature of the human society, the system of interhuman communication drastically changed. With our present methods we cannot reconstruct the languages spoken earlier, because in the Mesolithic and Paleolithic Ages the size of the extant human groups, their migration and their communication system were totally different from those in the Neolithic Age. The instability of the communicative system in ages earlier than the Neolithic Age makes it theoretically impossible to reconstruct there any detectable rule or tendency in language system and language change. The Neolithic Age begun about 5.000 B.C. in the highly cultivated areas of Egypt, Mesopotamia, India and China and spread over Eurasia within a thousand years. Since the Altaic languages were relatively far from these high early cultures (see the Indo-European contacts above p.15.), the limitation of the reconstruction can be placed at the middle or the end of the fifth millennium, that is 4.500-4.000 B.C. As we shall see, the first separation of the Turkic languages must be placed in the first centuries B.C. That means that the Ancient Turkic period ended at this time. For the development and formation of Late Ancient Turkic we have to suppose a long period, which was preceded by Early Ancient Turkic. We do not know how long Early Ancient Turkic lasted and thus we do not know when it begun. As a working hypothesis we can assume that the Altaic language may have existed between 4.000 and 3.000 or so. According to Ramstedt (1957 15) the Altaic unity was dissolved around 4.000 B.C., Ligeti (1953 358) joined this opinion, but later (1963 384) he put it between 3.000 and 2.000 B.C.

Unlike those who deny that such a language existed, I admit the possibility of its occurrence; but unlike those who take its existence for granted, I consider the common Altaic language only as an unproven hypothetical model. In any case the bulk of the famous and much debated Chuvash-Mongolian parallels do not pertain to the common Altaic language, most of them are Turkic loanwords in Mongolian. For those who deny the existence of the Altaic community, the neutral term *Pre-Turkic* can be recommended (cf. German *Altaisch*, *Prä-Türkisch*, Russian *altajskij*, *dotjurkskij*).

Ancient Turkic denotes the stage following the formation of the separate Turkic language. Most probably this happened as a result of the separation of its speakers from the speakers of the other Altaic languages. The areal consolidation certainly lasted a long time. In Early Ancient Turkic we can assume

local variants, but these had no direct contacts with the dialects which emerged in the Late Turkic period. The beginning of the Late Turkic period was marked by the formation of those Turkic dialects which later became the basis for the various groups and single languages. In my earlier papers I supposed that the appearance of Rhotacism and Lambdaism may have been perhaps the first of the dialectal peculiarities which separated the *Chuvash-Bulgharian* group (ChBT) from the other, which we usally call *Standard Turkic* (ST). This was not necessarily so. For the appearance of the Rhotacism we have a definite date, the appearance of the stirrup.

The word for 'stirrup' is common to all Turkic languages including the so called Chuvash group. In each of the languages and its earlier monuments it has this meaning and only this meaning (some minor extensions of the meaning are late secondary developments). The earliest form which can be reconstructed for the present-day Chuvash *yărana* 'stirrup' is **irä ä*. In case of Standard Turkic languages we can reconstruct two prototypes: **üze gü* for the Oghuz, Kipchak and Turkestan Turkic languages, and **izeŋge* for Baraba, Khakass, Tuvanian, Yakut and Yellow Uighur. Since we find Rhotacism in this word, Rhotacism must have appeared after the time when the Turks became acquainted with the stirrup. The picture is also otherwise very instructive. We find two isoglosses. One of the isoglosses separated those Turkic languages which had an *r* in place of the *z*. The other isogloss separated those Turkic languages which had the *i-e-e* vocalism from those which had the *ü-e-ü* vocalism:

| | |
|---|---|
| *iräŋä* | *izeŋe* |
| | *üzeŋü* |

*TABLE II.*

As we see on the one hand the Chuvash belongs to a separate group, but on the other it has a common isogloss with the Siberian languages (Yellow Uighur migrated to the south in the 9th century A.D.). This is a situation which is typical for dialects. Can we reconstruct a common protoform for these three different reconstructions? I have dealt with the word in Róna-Tas 1973 and 1982a, 120-122. Recently Ş. Tekin quoted an Uighur word: *izeŋgülük* from the Hami version of the Maitrisimit (ed. Geng Shimin) where it occurs in the form *izengülükinte* with the meaning 'on his foot-sole'. According to Ş. Tekin, who quotes also earlier suggestions for the etymology, this would be a derivation from the basic word *iz* 'footprint, trace, track' > *izeŋgülük* 'sole' > 'stirrup', from this later *izeŋgü* > *üzeŋgü*. I think *izeŋgülük* is the middle part of the sole

where the stirrup was placed, and thus we have only the form *izeŋgü* to hand which is the earlier form of Yellow Uighur *ezenkü* (read *ezengü* Tenishev 1976 179), *ezengī, ezengo, ezengo* (Malov 1957 25), and which reflects an earlier form than the *üzengü* in the Qutadgu Bilig (ed. Arat, 1947, l. 6110), the earliest occurrence known until now (1069 A.D.). The vocalism *i-e-ü* may be the original one, because only from this can we understand the change into *i-e-e* on one hand and *ü-e-ü* on the other, the latter under the influence of *üze* 'above' (*i-e-e* < *i-e-ü* > *ü-e-ü*).

Thus in our case none of the two types of vocalism reflects the original form, we are confronted with the latest layer of Ancient Turkic and with the help of the Uighur data we can take one step backwards and reconstruct a third type of vocalism. Some possibilities remain, however, open. Without deciding now and here the question of Rhotacism, i.e. whether *z* or *r* is the original sound in this word, we can suppose the following changes:

| | | |
|---|---|---|
| | | *üzeŋgü* |
| | *izeŋgü* | |
| **iXeŋgü* | | *izeŋge* |
| | *ireŋgü* | *iränä* |

or

| | | |
|---|---|---|
| | | *iränä* |
| | *iXeŋge* | *izeŋge* |
| **iXeŋgü* | | |
| | *üXeŋgü* | *üzeŋgü* |

*TABLE III.*

In the first case the opposition of *r:z* was the first, and the changes in vocalism occurred later. In the second case the opposition *r:z* and the changes in vocalism may have occurred contemporaneously or the change in vocalism may have occurred even earlier.

In my 1982 paper I briefly dealt with the chronology of the stirrup in Eurasia and concluded that though we have no clear archaeological data before the 4th century A.D. we can date the object back to the first centuries B.C., (but definitively not earlier) because the first stirrups were made of rope or wood, which have not been preserved, but left their traces in pictures, techniques and semantics. The changes shown in Table III must have occured after the first stirrups appeared among the Ancient Turks. This means that Rhota-

cism in any case and very probably also other changes, (as in our word the change in vocalism) occurred after the first centuries B.C.

The chronology of the appearance of Rhotacism, and with it the dialects of the Chuvash type, can be corroborated by the earliest Turkic loanwords in Samoyed. Donner (1924) has already discussed the earliest layer, but owing to the open question of whether Proto-Turkic had *r or z*, no conclusions on the chronology could be drawn. The history of the studies can be read in Janhunen 1977 and I summed up my results in Róna-Tas 1988b. Such early loanwords as Proto-Samoyed *yür* 'hundred', <- Late Ancient Bulghar-Chuvash *yür* cf. Late Standard Turkic *yüz* leave no doubt that the word was borrowed before the separation of Northern and Southern Samoyed, or the three groups Northern, South-Eastern and Selkup as I am now inclined to suppose. The separation of the main Samoyed groups can be safely dated to the first centuries A.D. (Hajdú 1978 348-349).

In 1974 I suggested investigating a series of Turkic-Tocharian parallels because they seemed to offer important conclusions for the history of the Ancient Turkic dialects. The problem has been recently discussed by Gamkrelidze and Ivanov (1984), by Ivanov (1988) and myself (Róna-Tas 1990). Such words as Toch. B *yasa* > *yäs* > Ancient Turkic *yez* > Ch **jer* ( -> Moksha Mordvin *serä*; -> Mongolian *jer*) seem to indicate that in such words the *z* was original, and this *z* was a substitution for foreign final *-s*. The appearance of the dialects with Rhotacism has to be put after the borrowing of these Old Tocharian loanwords (Old Turkic had also later loanwords from Tocharian, they will be dealt with separately). The great migration of the Tocharians to the East begun about the first centuries of the second millennium and they reached the Chinese border in the first half of the first millennium. We don't yet know where the Ancient Turks contacted the Tocharians, but the Old Tocharian - Ancient Turkic contacts cannot be later than the first half of the first millennium B.C. Since these words appear in all Turkic groups they must have appeared in Ancient Turkic before the final development of the Ancient Turkic dialects. Thus we can conclude that we can approximately date the beginning of the Late Ancient Turkic period to the middle of the first millennium B.C., and the emerging of the opposition of *z:r*, i.e. Rhotacism to the beginnings of the first century A.D.

Old Turkic begun with the separation, formation and consolidation of the independent Turkic languages. This was a long historical process. By "independent Turkic languages" I do not mean the ancestors of the present Turkic languages. The first two groups which emerged were the Bulghar-Chuvash and the Standard Turkic languages, some later independent languages were present only as dialects, others such as e.g. Yakut may have began their separate life within the ST group but earlier than the others. Such groups as Old Oghuz, Old

Kipchak were intermediate forms. Old Turkic has to be divided into two subperiods. Early Old Turkic lasted until the formation of the great Western and Eastern Turkic Empires. In the West this happened in the fifth, in the East in the middle of the sixth century. Late Old Turkic can be divided into three subperiods. LOT I is a period which lasted until the formation of the Second Turkic Khanate (551 A.D.). This is the time when we already have source material on different Turkic groups and their language, but not yet Turkic texts written by the Turks themselves (at least not texts deciphered). The most important event was the formation of a kind of *koiné* and a literary language in the Eastern Khanate and this was used to write texts in the Second Eastern Turkic Khanate with the so called Eastern "Runic" script and then with other scripts. The appearance of the earliest Turkic texts in the first decades of the 8th century marks the beginning of the period LOT II. The beginning of the Arabo-Persian influence was caused by the impact of the Mohammedan expansion, and this marks the beginnings of LOT III. The Mohammedan influence did not reach the northern and the eastern groups of the Turks and in Central Asia the Buddhist religion and also other religions continued their activities. The Old Turkic period ended with the Mongolian invasion.

**Middle Turkic** begins with the Mongolian invasion which spread over the Turks in the first decades of the thirteenth century. It brought a considerable rearrangement of the linguistic situation both in respect of the interrelationship of the Turkic languages to each other and in respect of the impact of the Mongolian language on the Turkic languages. The gradual formation of the literary languages and the formation of those language groups which later became the independent languages of today were events which occurred at different times in the various parts and regions of the Turkic world. In the East the *Khvarazmian* literary language was followed by the *Chaghatay* literary language (both were called by Soviet scholars *Old Üzbek*), several Kipchak groups tried to form their own literary languages like the *Mameluk Turks*, the *Cumans* or the Kipchak groups of the Volga region. This is the period when the Ottoman literary language emerged and took shape, but also other Oghuz groups tried to stabilize their common language. Far from the Muslim influence Buddhist Turks preserved their literary language which slowly changed during the Mongolian rule. Thus I consider the Uighur texts of the Mongolian period as Middle Turkic in contrast to Annamarie von Gabain, who calls all Uighur texts irrespective of their age *Alttürkisch*. Since I define Late Middle Turkic the subperiod of the full development of the literary languages and their mutual effect with the spoken language the beginning of the Late Middle Turkic period has to be set differently for each separate language. (cf. German *Mitteltürkisch*, Russian *srednetjurkskij*).

New Turkic began with the conclusive formation of the present Turkic languages. Its first period is in most cases marked by the struggle between the old literary languages and the spoken ones, while in the later period purism and several dialects influenced the emerging standard language (cf. German *Neutürkisch*, Russian *novotjurkskij*).

Modern Turkic is any Turkic language whose synchronic structure and dialects can be investigated (cf. German *modernes Türkisch*, Russian *sovremennyj tjurkskij*).

We have to add some comments to the above periodization. As we have seen the criteria for the lingusitic periodization were extralinguistic and took into account several aspects: the type of interrelationship between the language units (local groups, dialects, languages etc.), the historical events which influenced what I call the linguistic situation, i.e. the sociolinguistic setting of communication, and the types of sources available for reconstructing the system of the period. The terms *Ancient*, *Old* and *Middle* are of relative significance. All are terms for occurrent languages. Of course we can also use the term *Proto-* as in *Proto-Oghuz* or *Proto-Ottoman* but in this case the reconstruction or the reconstructed is emphasised.

To sum up, the chronological framework can be roughly given as follows (the years were rounded up to make them easy to learn):

Altaic 4.000-3.000
Ancient Turkic
    Early AT 3.000-500 B.C.
    Late AT 500 B.C.-400 A.D.
Old Turkic
    Early OT 400-550
    Late OT I. 550-700
    Late OT II. 700-1000
    Late OT III. 1000-1200
Middle Turkic 1200-

All further chronological borders have to be fixed separately for each language.

# THE CLASSIFICATION OF THE TURKIC LANGUAGES

There are many aspects according to which the present Turkic languages can be classified and grouped. The modern classifications of the Turkic languages go back to an important paper of Samojlovič (1922). Of the more recent classifications those of Baskakov (1952, 1962), Benzing (1953, on which see Rásonyi 1957 315), Räsänen (1949, 1953), Poppe (1965), and Menges (1968) can be mentioned. In the following we shall put forward a classification which follows two different aspects. Where it is possible we prefer the genetical and historical principles, and follow only on the second place geographical aspects. In other cases we have had to choose first the geographical aspect and only then, if at all, the genetical and historical. Where it is possible and/or important I also give the names of the dialects. We shall begin with the classification of the modern Turkic languages:

I. CHUVASH
- D: Viryal, Anatri

II. KIPCHAK or NORTHWESTERN BRANCH
1. Northern or Volga Kipchak
   - Kazan Tatar, D: Central, Siberian
   - Misher Tatar
   - Bashkir
2. Eastern or Aral-Caspian
   - Kirghiz (earlier Kara Kirghiz)
   - Kazak (earlier Kirghiz)
   - Karakalpak
   - Nogai
3. Western or Pontic-Caspian
   - Kumük
   - Karachai-Balkar, D: Karachai, Balkar
   - Crimean Tatar (On the Crimea Turkish and Nogai were also spoken), Dobrujean Tatar, Urum or Greek Tatar
   - Karaim, D: Trocki, Luck

III. OGHUZ or SOUTHWESTERN BRANCH

Türkmen
Khorasan
Azeri (Azerbayjani)
Turkish or Ottoman, Osmanli, D: Anatolian, Rumelian
Gagauz (earlier a dialect of Turkish)

IV. KHALAJ or SOUTHERN BRANCH

V. TURKESTANI or EASTERN BRANCH

Özbeg
New Uighur
Turki (also Taranchi), D: according to the oases as Kashgar, Kucha, Khotan, Turfan Yarkend, Lobnor
Salar
Hoton
Yellow Uighur

VI. SIBERIAN or NORTHERN BRANCH

Siberian Tatar, D: Tobol-Irtish, Baraba, Tomsk
Altai (earlier Oirot), D: Kumandu, Tuba-kizhi, Lebed, Teleut, Tölös, Upper Bij, Urjankhai (Telengut)
Shor, D: Kondom, Mras, Aladag
Chulim, D: Küerik
Hakass, D: Abakan, Kizil, Koibal, Kacha, Kandakov, Mador (Matar), Motor (Matir), Sagai, Salba
Tuva
Tofalar, Karagass (earlier a dialect of Tuva)

VII. YAKUT

D: Dolgan.

The classification of the Middle Turkic languages is less elaborated. All or most of the Modern Turkic languages had their fore-runners in the Late Middle Turkic period, some also in Early Middle Turkic. But in most cases we are confronted with literary languages, which hide or overshadow the spoken languages. These literary languages were used in several centres and also beyond these centres. They were used by people speaking different Turkic languages. The influence of the spoken idioms on the literary languages may have

been different. Thus the following literary languages are not from the same level as regards the classification:

Uighur
Eastern Literary Middle Turkic
    Chwarasmian
    Chagatay (both also called Old Üzbek in the Soviet Union)
Volga Turki
Volga Bulgharian
Middle Azeri
Middle Turkmen
Middle Ottoman
Cuman
Mameluk Kipchak
Armeno-Kipchak

# A SHORT HISTORY OF THE TURKIC PEOPLES

The chronology of the early archeological periods in the steppe region of Eurasia is much debated and far from being exact. The Southwestern parts of this region had close contacts with Persia and Mesopotamia and therefore here the dating is more sure and exact than in regions more to the East and North. The end of the Neolithic Age and the short subsequent Copper Age, the beginning of the Bronze Age is dated in the Carpathian basin to around 1.900 B.C. The appearence of the first, imported bronze tools in the North of Caucasia can be dated to the 3rd milleneum, but it was only in the second half of the 2nd milleneum that the local bronze production began and the copper and arsenic alloy was replaced by a copper and tin alloy. The spread of iron was faster and can be dated to the 6th century B.C.

The first people known to Greek and Persian sources in the territory which is now South Russia were the Khimmerians. They can be traced from the middle of the 2nd milleneum. In the 8th century B.C. they were pushed towards Mesopotamia by the newcomers the Scythians. The various Scythian groups lived in the area from about 750 B.C. until 200 B.C. and fought with the Persians and Greeks. As it is known Alexander the Great (336-323) conquered not only Persia but reached Central Asia. He conquered Samarkand and Baktria in 329 B.C., that is the regions which were earlier Sogdiana and later Baktria. On the steppe the Scythians gave way gradually to several new groups mostly known under the name Sarmatians. As far as we can conclude from the personal names used by members of these people the Khimmerians, the Scythians and the Sarmatians spoke Indo-European, the latter two most probably Iranian languages.

Chinese sources mention the Hsiung-nu as early as in the 9th-8th centuries B.C., but their importance grew considerably only at the beginning of the 2nd century. The founder of the great Hsiung-nu Empire was T'ou-man, who was killed by his son Mao-tun in 209 B.C. The Hsiung-nu Empire reached its greatest power under Mao-tun. He conquered many people living to the North and the West of China among them people which later played a role in the formation of the Altaic and more especially the Turkic people. Among the conquered people the Tung-hu, the Wu-huan, the Ting-ling, the Hu-kie, the Wu-sun and the Yue-chi or as they were later called, the Tocharians should be mentioned. While we know with more or less certainty that the Scythians and the Sarmatians spoke an Iranian language, the language of the Hiung-nu is much debated. Some scholars are inclined to suppose that they spoke a Turkic or an other Altaic language, others assume that they spoke a language close to

the Paleoasiatic Ketic or a language which later became totally extinct. It seems that the question cannot be solved with certainty. On the other hand we have to keep in mind that the Hsiung-nu Empire was not linguistically uniform and though we do not know exactly what language was spoken by the ruling groups, it is sure that some Turkic and Mongolian and perhaps also Manchu-Tunguzian people lived under Hsiung-nu rule. Thus we can conclude that the history of the Turkic people begun with the Hsiung-nu even if we are sceptical about their exact lingustic affiliation.

The history of the Hsiung-nu is known from the Chinese sources. After a long period of the Chou dynasty (1122-255) the Ch'in dynasty united China and the great Emperor Ch'in Shi huang-ti declared himself as the first universal Emperor. The dynasty's rule (255-206) was marked by the struggle against the Hsiung-nu which continued under the rule of the Early or Former Han dynasty (206 B.C. - 25 A.D.). The Great Wall was built against the Hsiung-nu, then the military system of China was reorganized and they slowly learned to master the problem. In the first century B.C. dissent among the Hsiung-nu leaders weakened their power and the Shan-yü, the Hsiung-nu Emperor asked for Chinese protection against his rebellious brother Chih-chih (56-36 B.C.), who under the pressure of the united Hsiung-nu - Chinese attacks moved westwards and reached the region of the Ili river. In 9 A.D. a rebel Chinese leader Wang Mang crushed the power of the Han dynasty and caused great domestic problems in China. The Later or Eastern Han dynasty could reorganize the country only about 25 A.D. and ruled then until 220. The Hsiung-nu reestablished for the time being their own power benefiting from the domestic problems of China, but soon after the restoration of order their weakness came to light. The Shan-yü P'u-nu (48-83) was unable to keep the Empire under his sword and the Hsiung-nu Empire fell to pieces. Some of the tribes settled in Northern China and as with other barbarians later, they founded local dynasties and slowly or rapidly became Sinicized. In 311 A.D. the Hsiung-nu captured the capital of Lo-yang and the southern Hsiung-nu chief proclaimed himself Emperor, an event which proved to be of interest.

While some of the Hsiung-nu played a role in the history of China others moved to the West. At the end of the first century A.D. they conquered what is now Kazakstan and they lived there from about 100 until 350 A.D.

In the 370s a new people appeared on the borders of Europe. They crossed the Volga river and in 375 they attacked the Eastern Goths who lived then in what is now southern Russia. These Goths mixed with the Huns and some of them remained on the Crimean peninsula, where we can follow their life until the 18th century. The Gotic language was last recorded by Ogier Ghislain Busbecq in 1589. The Huns reached the eclipse of their power under Attila (441-453), whose central territory was the Carpathian Basin. He fought

against the Roman Empire, and after the battle of Catalanaum (451) he retired to the East. After his death the Hunnic Empire crashed and disintegrated.

Since the 18th century the question has been debated whether the Huns and the Hsiung-nus are the same people or not. There does exist an early Sogdian letter which relates the destruction of Lo-yang the Chinese capital by a people called Hun. Since the Chinese sources refer to the same event but mention the Hsiung-nu the equation of the two names seems to be sure. The date of this and other Ancient Sogdian letters has been questioned but not the indentity of the two names. The identification of the two names does not necessarily mean the identity of the two people or their language. Some continuity, however, may have existed. The language of the European Huns has also been the subject of discussions. Unfortunately we dispose only of three common words which are said to be Hunnic and these are surely not Turkic. They are: *strava* 'funeral repast' (Jordanes, Getica 49,258), *med-os* 'a drink' (Priskos E1 131.12) and *kamo-n* 'a drink' (the *-n* is the Greek accusative, Priskos E1, 131.14). We know of a few names and titles, none of them is absoltely clear, but some allow a Turkic interpretation. As far as we can judge from analoguous situations the Hunnic Empire was not homogeneous and many languages were spoken in it, among them surely Gothic (on the details see Doerfer 1973).

In between on the Eastern ends of the steppe a new people the **Hsien-pi** attacked the Northern Hsiung-nu. Around 155 A.D. they defeated the Northern Hsiung-nu, and among these Hsien-pi tribes we have to suppose Mongolians as well. Some of the Hsien-pi tribes migrated to the Chinese border, the Tu-yü-huns reached the borders of Tibet and the Tibetans called them A-zha. A mighty Hsien-pi tribe the **Tabgach** conquered Northern China and there founded the Wei dynasty (386-538). Some times later the Turks called China by this very name. In the fifth century the **Zhuan-zhuan** (400-551) founded a great Empire on the territory which is now Mongolia.

In the West after the dissolution of the Hunnic Empire newcomers appeared North of Caucasia. The various **Oghur** tribes migrated in the steps of the Huns and came from the territory which is now Kazakstan. In the middle of the 5th century they supposedly absorbed the various parts of the destroyed Hunnic Federation. The Byzantine sources relate after 463 about these groups which, or at least most of which spoke Turkic languages. **Kutrighurs, Saraghurs,** Sabirs and other tribes fought each other.

In the Zhuan-zhuan Empire tribes rebelled and in the subduing of the revolt the **Turks**, earlier said to be the blacksmiths of the Zhuan-zhuan offered great help. After the victory of the Zhuan-zhuan the Turk chief asked for the hand of the daughter of the Zhuan-zhuan ruler, who refused to consider the chief as an equal partner. The Turks revolted and defeated the Zhuan-zhuan. After their victory they founded the **First Turkic Khanate** (551). They soon

extended their power over other tribes as well, they appeared at the border of Persia, sent ambassadors to Byzantium and subdued the various Turkic tribes in East Europe. After the death of Bumin khan (552), the founder of the Empire, within a decade a great Turkic Empire faced its neighbours. But soon the Empire broke into two parts, the Western and the Eastern Turkic Khanates. The **Western Turkic** ruler Ishtemi khan sent his first embassy led by the Sogdian merchant Maniakh to Constantinople in 567 and a long period of wars and contacts followed both with the Romans and Persia. The **Eastern Turks** annually invaded the Chinese borders. The new dynasty, the T'ang (618-930), only slowly organized their defence but then moved to attack. They made use of the dissent among the various Turkic groups and in the middle of the 7th century one after the other Turkic groups were forced to acknowledge the sovereignty of the Chinese Emperor.

In 555 a new people appeared North of Caucasia. They called themselves **Avars** and were also known by the name **Varhuns**. They rapidly moved to the West where they occupied in 567 the Carpathian Basin which the Longobards had left for Northern Italy, the present-day Lombardy. Many scholars think that the Zhuan-zhuans disappearing around 552 and the Avars were the same people. There were surely also other groups among them, including some tribes of the **Ephtalites**, called also the **White Huns**.

North of Caucasia the Khazars began to build their new empire. In 650 they defeated the **Onoghur Bulgharian** Kingdom of Kuvrat. The people of the Bulghar federation moved in various directions. One group of them reached the Lower Danube where they founded around 683 the Danubian Bulghar State. Other groups moved to the Carpathian Basin, the Balkans and even Italy. A larger group remained under the rule of the Khazars and slowly moved to the North along the Volga river. They became later the Volga Bulghars.

In 682 a small group of Turks under Chinese rule led by the later Elterish khan rebelled. They soon united most of the Turkic groups and founded the **Second Turkic Khanate**. In the First Turkic Khanate they had used the Sogdian language and script for chancellary purposes and inscriptions. This was changed in the second Khanate and we dispose of inscriptions written first in history in Turkic language and in the so called Runic script. The second Turkic Khanate was defeated by a coalition of three tribes, the **Uighurs**, the **Qarluqs** and the **Basmils**. But the Uighurs soon got rid of their allies and under their hegemony was founded the **Uighur Khanate**. The consolidation of their rule lasted several years and was final only in the late fifties of the 8th century.

A mighty rebel most probably of ultimately Sogdian origin An Lu-shan menaced the rule of the T'ang Emperors. The Uighurs came to help the Emperor and in 762 they sacked Lo-yang the capital. The Uighur ruler met here the Manichean priests who converted him, but also Buddhist and Nestorian

missionaries reached the Uighurs and other Turks and tried to convert them. The result was that several scripts were used for translations of the holy scripts of these religions into Turkic. The Uighur society slowly transformed under the influence of the clergy. The Kirghiz put an end to the flourishing Empire, but the Uighurs moved to the South. One group founded around Kocho a new kingdom which lasted until the late 13th century. Another group reached the Kansu region of China and reorganized itself. The became known under the name Yellow Uighur as they are still called today.

The Kirghiz were conquered by the Mongolian-speaking Khitays who soon founded in North China the Liao dynasty (907-1125). One part of the Khitays or Khitans beaten by the united forces of the Chinese and the newcomer Juchens, early realtives of the Manchus, migrated to the west and became known under the name Kara Kitay or Black Kitay. They reached the region around Khwarazm and slowly became Turkicized.

In the West after the victory over the Bulghars, the Khazars strengthened their Empire. Soon they came into contact with the Arabs, who conquered Persia in the 7th century (642 battle of Ktesiphon), begun their fights with the Khazars and gradually reached the Talas valley. Here in 751 was one of the most important battles of the High Middle Ages, the battle between the Arabs and the Chinese with allies on both sides. The Chinese stopped the Arabs. The attacks of the Arabs menaced Europe from three directions. From the Iberian peninsula towards France, against the Byzantine Empire and through Caucasia against the Khazars. In the first years of the 8th century their pressure grew. In 717 theiy stood around Constantinople, in 721 they fought against the Franks at Toulouse and only in 732 could the Franks stop them at the battle of Poitiers. The third front was directed against the Khazars. Already in the 7th century the Khazars and the Arabs fought each other without clear success on either side. In 730 the Khazars defeated the Arabs and in 737 the Arabs launched a major attack on the Khazars. The Khazar khan fled from the capital and withdrew northeast along the Volga. The Arabs, under the leadership of Marwan followed him, and finally forced him to accept Islam. After this had happened, the Arabs withdrew. The Khazars recovered and reorganized their state. The Jewish religion strengthened its influence in the Khazar state already in the 8th century but the famous conversion of the Khan and the leading class took place only around 860. The long Pax Khazarica between 740 and 960 was only disturbed by the Pechenegs, the Hungarians and the revolt of some Khazar groups. The latter called Kabars joined the Hungarians, who formerly lived in the Khazar Empire and who gradually became independent. Their slow move from Levedia to Etelküzü and then to the Carpathian Basin was caused by the attack of the Pechenegs and the pressure on the part of the Khazars. As is known the Hun-

garians appeared at the Lower Danube in 839, and invaded the Carpathian Basin in 862 but finally settled there only in 895.

Those Bulghars who remained in the Khazar Empire also tried to loosen their ties with the Khazars. They slowly moved to the North. Around 800 they reached the territory where now we find the village Bolshie Tarkhany and occupied the southern bank of the Kama river in the first years of the 10th century, that is a hundred years later. The **Volga Bulghars** founded their important state which converted to Islam. Ibn Fadlan the member of the Caliph's embassy visited them in 921-22 and described a flourishing life.

In Western Central Asia several Turkic groups fought each other. In the middle of the 8th century a Turkic ruling clan of Türgesh origin, the **Turk Shahi** or Kabul Shah dynasty established itself in Gandhara. They were Buddhists and one of their rulers bore the title From Kesar, the "Roman Caesar" between 738 and 745. They had matrimonial ties with Khotan allies of China and Tibet. In 809 the Arabs defeated the Turk Shahi dynasty and they converted to Islam.

In the region of Kunduz and Herat, called also historically Tokharistan, the **Yabgus** of Western Turkic origin ruled. They were Buddhists and their rule lasted until 759.

The **Karluks**, who were allies of the Uighurs when they defeated the Turks, moved to the West after the final victory of the Uighurs. They were joined by the **Chighils** and **Yaghmas**. The ruler of the confederation bore the title Yabgu. After the end of the Uighur Empire in 840 the Yabgu took the title Khan, or Kara Khan. Islam slowly gained ground in the Empire, while the **Karakhanids** fought against the Ghaznavids and in alliance with or as mercenaries of several Mohammedan leaders. Satuq Bughra khan converted to Islam, and after his death (955) his son Baytash organized a strong Islamic reign. The Karakhanids extended their power over Bukhara, Samarkand and in the 11th century they had a high culture with literature in Turkic language. At the end of the 11th century they fell into rival sections and were defeated first by the Seljuks and then by the Kara Kitays and the Khwarazmshahs.

In 977 the local Turkic people of Ghazna elected a certain Sebük Tegin as their governor. He was the founder of the **Ghaznavid** dynasty which was only nominally a souzerain of the Kaliph. The majority of the population over which the Turkic clan ruled was Persian, but important Turkic contingents served as a backbone for the military. In the first half of the 11th century the Ghaznavids led predatory campaigns against North India and played an important role in the history of Persia. The Seljuk put an end to their flourishing.

A large group of Oghuz tribes was organized by Seljuk who turned to Islam with his family around 1000. First in alliance with, then against the Karakhanids and Ghaznavids the **Seljukids** estabilished their power first in what is

now West Persia. In the 40s of the 11th century the Seljukid ruler Togril chose Nishapur as his capital. In 1049-50 he even occupied Baghdad which he held until 1058. The next Seljukid ruler Alp Arslan defeated the Byzantine ruler Romanos in 1071 at Manzikert and founded the Seljuk Sultanate of Ikonion or Rum which gained power in Eastern Anatolia and Iraq. The Seljuks menaced the existence of Byzantium and the Holy Land, so their attacks were the first cause of the Crusades. In the 12th century after a short flourishing period the Seljukid Empire fell into rival sections, and the renewed attempts at the restoration of the central power did not last long. In 1194 the **Khwarazmshah** put a final end to Seljukid power.

Along the rivers Tobol and Irtish lived the federation of the **Kimek** tribes and among them we hear in the 9th century for the first time about the **Kipchaks**. After the defeat of the Khazar Empire and following the Pechenegs the Kipchak tribes appear in the territory North to the Black sea. The Kipchaks moving in South Russia were named by the Russians Polovcy, by the Western sources Cuman, or Falben by the Hungarian sources Kun.

In 1196 when the Mongolian tribes elected Temüjin as their khan and he chose the name Chingis, the Kara Kitays, the Kwarasmshah, the Volga Bulghars and the Kipchaks formed the greatest Turkic power in the West. Some smaller groups such as the Kereits, the Öngüts and others lived in the East. South of the Mongols were the Uighurs of Kocho and in the Siberian regions lived various other smaller groups.

# THE MONGOLS. HISTORY, PEOPLES AND LANGUAGES

In Chinese sources we meet the **Tung-hu**, originally the designation of a group of barbarian, non Chinese people. Later it became the designation of a certain group, and among the Tung-hu two main federations were recorded, the **Wu-han** and the **Sien-pi**. Later the Sien-pi federation was divided and the Southern Sien-pi groups were the **Tu-yü-hun**, the **Kitay** and the **Tatabi** and the Northern the **Shi-wei**. Among the Shi-wei a tribe with the name **Mongol** was recorded. Among the above groups also the **T'o-pa** or **Tabgach** has to be counted, and we have some data that all these groups spoke Old Mongolian languages or at least were ruled by Mongolian speaking groups.

After the disappearance of the Kitays, at the end of the 12th century, around 1196 when Temujin was elected as Khan of the Mongols, the following people lived in the neighbourhood: around the Yenisei the Kirghiz, around Kobdo, present West Mongolia, the Naimans, on the Orkhon and Selengga river the Kereits (Nestorians, perhaps they can be connected with the famous Kingdom of Presbyter John), the Merkits South of the Baikal, the Önggüts North of the Huang-ho, the Tatars (known since the 8th century, or even earlier), the Oirats (West of the Baikal). The Mongols themselves occupied the territory around the Onon river.

Having subdued the related and neighbouring people Chingis went first against the Tanguts (1209), then led his troops against the Jurchens (1211-1215). Then he turned to the West, Khwarazm fall in 1219, in 1221 the commander Süböetey launched his first great campaign. In the battle at the Khalkha River in 1222 the Mongols defeated the united troops of the Kievian Rus and the Comans. Chingis turned against the Tangut cities and died in 1227. In 1229 the great Kurultay elected Ögödey as great khan, and he moved his capital to Karakorum. In 1232-34 they finally defeated the Jurchens, in 1236 they occupied Korea. 1235-36 the Mongols attacked the Volga Bulghars. In 1237 the Comans move westwards and 1239 they fled to Hungary. In 1238 the Mongols defeated Suzdal, Kiev, Chernigov and Halich. In 1241 they crashed the united German and Polish forces at Walstatt and invaded Hungary. In the battle of Muhi (11th April, 1241) they defeated the forces of the Hungarian king Béla IV. Ögödey died 1241 and the Mongolian troops left Hungary. Under the Khan **Kubilay** they conquered the whole of China and founded the Yüan dynasty (1280-1368). Twice they tried to invade Japan (1274, 1281), and in 1293 they reached Java. At the end of the 14th century the greatest Empire in world history was united under the Mongols.

The house of **Chagatay** ruled in Central Asia, the house of **Hūlegū** in Persia and the house of **Jochi-Batu** West of Urals, the latter became later the **Golden Horde.**

Nowadays the Mongols live in the Chinese People's Republic, the Mongolian People's Republic and in the Soviet Union where we find the Buriat Autonomous Republic (Ulan Ude is its capital) and the Kalmūk Autonomous Republic (with its capital Elista on the Volga).

The present Mongolian languages are grouped in two branches, the archaic and the non archaic languages. We can divide the non archaic branch into a **central**, a **northern** and a **western** group.

The most important of the non-archaic languages is **Khalkha**, the official language of the Mongolian People's Republic (capital: Ulan Bator). It is spoken by about one million inhabitants and can be divided into several dialects. They use the Cyrillic script introduced in place of Uighur-Mongolian in 1941. Closely related to the Khalkha is the **Chakhar** language, mostly spoken in China in the Inner Mongolian Autonomous Republic (capital Köke Kota or Khökh Khot), together with many other Mongolian languages and dialects such as Üjümchin, Ordos, etc. The Mongols of China use the Old Uighur Mongolian script. These languages pertain to the Central Group.

In the North, in the Buryat Autonomous Republic and the adjacent areas various **Buryat** dialects are spoken by about 400.000 people. The literary language written with the Cyrillic script is based on the Khori dialect.

The Western group is also called **Oyrat.** Many of the Oyrat groups live in West Mongolia, i.e. the Western part of the Mongolian People's Republic. The second group lives in China in Dzungaria and around the Alashan mountain-range. The third group migrated in the 17th century to Europe and lives now in the Kalmuck Autonomous Republic.

The Archaic Branch of the Mongolian languages can be divided into three groups. In the Northeastern territory of what earlier was called Manchuria around Khaylar and Tsitsikar, now Heilungkiang and in the Hulunbuyir Aymak of Inner Mongolia the **Dakhur** or Dagur language is spoken. It has several dialects some of them preserving the initial *h-* of the Middle Mongolian language. It has been suggested that the Dagurs are the descendants of the Khitays. If this is true they were strongly Mongolized, that is their earlier language which was of an Old Mongolian type and of another branch later came under the influence of Middle Mongolian.

The **Monguors** live in the province Chinghai, they and the following groups moved to their recent dwelling places under the Yüan dynasty, but then after the fall of the Yüan-dynasty (1368) they remained there as border guards of the Chinese Emperors. The Monguors speak a very archaic Mongolian language which developed after the 14th century under Tibetan and Chinese

influence. The Tung-hsiang language is spoken in Kansu province and the Pao-an language both in Kansu and Tsing-hai provinces. In Kansu we also find the Mongolian-speaking Yellow Uighurs or Shera Yöghurs. This group lives in a kind of symbiosis with the Turkic speaking Yellow Uighurs: the Sira Yöghurs.

The third group of the Archaic Branch consist of several Mogol dialects spoken in Afghanistan.

# THE MANCHU-TUNGUZIANS

*Manchu-Tunguzian* is the name of a genetically related linguistic group of languages. In Russian linguistic literature this group is called *tunguso-man'-chdžurskij* in an attempt to stress the greater importance of the Tunguzians and the Tunguz branch most of whom live in the Soviet Union. The earliest Manchu-Tunguzian language attested in written sources is **Jurchen**, the language of a tribal confederation which came to power in North China in 1115 under the name Chin (Kin) or Golden Dynasty (also Nü-chen). It was crushed by Chingis Khan and his successors (1234) but important groups lived and spoke their language not only during the Sung and Yüan dynasties but also during the Ming dynasty. There exist early inscriptions, a Jurchen-Chinese glossary and Sino-Jurchen texts. In the first half of the 17th century the **Manchu tribal federation** came to power and conquered China. In 1644 they founded the Ch'ing dynasty, the last in the history of China; the Republic was declared in 1912. Manchu became a literary language written in Uighur-Mongolian script which was later adapted to the special needs of Manchu. The spoken form of Manchu is called **Shibo** or Shibe.

The Manchu-Tunguzian languages are divided into two branches. The Northern Branch is called Tunguzian in the broader sense, and the Southern also Manchu. We dispose now of an excellent comparative dictionary of the Manchu-Tunguzian languages compiled under the direction of V. I. Cincius (1975) and therefore I give here the classification of Cincius adding the names of peoples and languages used in literature other than Russian.

NORTHERN OR TUNGUZIAN BRANCH
- Evenki or proper Tunguzian
- Even or Lamut
- Solon
- Negidal

SOUTHERN OR MANCHU BRANCH
- Manchu (spoken form Shibo or Shibe)
- Nanai or Gold
- Olcha or Ulcha
- Orok
- Udihe
- Oroch

In a paper published in 1974 Ikegami gave the following grouping: I. Evenki, Solon, Negidal, Lamut, II. Udihe, Oroch, III. Nanay, Ulcha, Orok, IV. [Jurchen], Manchu. Essential is here, that Udihe and Oroch pertain to a transitory group, while between III. and IV. the difference is more of chronological nature. Based on the *Comparative Dictionary* of Cincius in 1978 Doerfer revised the earlier classifications. He tried to quantify the differences and came to the following result:

I. NORTHERN BRANCH
1. NE Group: Lamut, Arman;
2. NW Group: Evenki, Negidal, Solon;

II. CENTRAL BRANCH
1. CE group: Oroch, Udihe
2. CW group: Kili, Nanay, Ulcha, Orok (Uilta)

III. SOUTHERN BRANCH
Jurchen, Manchu.

Two dialects have been redefined as languages: *Arman* (earlier a Lamut dialect) and *Kili* or the Dialect of Kur-Urmi, which was considered earlier as a Nanay dialect.

# THE WORLD RELIGIONS AND THE ANCIENT TURKS

*Shamanism and Tengrism*

In order to understand the sources on the Old Turkic languages we have to give a short overview on the spread of the various religions to the Ancient Turks, because most of the material which has remained for posterity was written for religious purposes and by monks and priests of various religions. Several kinds of Turkic religious literary languages appeared which cannot be separated from the historical, linguistic and cultural background of the other peoples who played a role in the missionary activities.

Among the old Turks we can find two types of religions or religious complexes. One was shamanism. In the definition of shamanism I accept the views of U. Johansen (1987, 8-9) which I quote in the original German:

"Dieser Terminus, der wahrscheinlich aus einer tungusischen Sprache von den Europäern übernommen wurde, wird auf Träger religiöser Funktionen angewendet, die

1. sich bewusst in Trance versetzen können, das heisst in einen veränderten Bewusstseinszustand, in dem sie zumindest auf auditive und visuelle Reize ihrer realen Umwelt in vermindertem Masse reagieren,
2. die Fähigkeit hierzu in einem Berufungserlebnis und einer Zeit psychischer Krisen erwerben,
3. in diesem Bewusstseinszustand die Verbindung mit vorgestellten, vom naturwissenschaftlichen Standpunkt gesehen, nicht existierten Wesen aufzunehmen glauben, wobei die Vorstellung durch die Religion, in der sie wirken, bestimmt sind,
4. den religiös motivierten Zustand veränderten Bewusstseins in der Regel im Interesse und in Übereinstimmung mit ihrer Gesamt-Gesellschaft herbeiführen, in der sie somit als religiöse Interpreten wirken und ihr ein Gefühl der Sicherheit gegenüber dem Transzendenten geben,
5. auch in den unmittelbar sichtbaren Formen ihres Wirkens - in Ritualkleidung, Ablauf der religiösen Handlungen oder Ausgestaltung ihres Ortes etwa - Träger einer Tradition sind."

This is a very proper definition of the shaman and Shamanism is the system of religious beliefs in which the central role is played exclusively by the shaman. Of course Shamanism did not disappear in other religions, and the shaman, though he or she lost the central role, may have an important function. In these cases we speak about shamanistic traditions, functions, etc.

Among the Old Turks there existed an other type of religion, which was called by J.-P. Roux (1956, 1962 etc.) Tengrism. In this religion there existed a central God which was identified with heaven (Turkic: *Kök Täŋri*), who was a creator but not an *ex nihilo* creator. The ruler, the king was considered a sacral figure, this sacral king embodied the fortune of the society, if his forces diminished he was sacrified (see Róna-Tas 1987a). Shamanism became an important part of Tengrism.

*The spread of Buddhism to Central Asia and the Turks*

A very good and comprehensive overview is given by R. E. Emmerick in *The Encyclopedia of Religion* (ed. M. Eliade 1987, vol 2 400-404). The spread of Buddhism from India began in the time of the king Ashoka (traditionally 271-231) under whom missionary activities were encouraged and reached Bactria, Gandhāra, that is the present Afghanistan, where after the military expedition of Alexander the Great (327-325) a lively Greek colony was founded. In the middle of the first century A.D. the Yüe-chi tribes who had migrated earlier from the Chinese borders through Gandhāra arrived and, there they adopted the local East Iranian language of Bactria and invaded Northern India where they founded the Kushān dynasty. Under the Kushān ruler Kanishka I (first to second century A.D.) the missionary activities were extended and several people converted to Buddhism, among them Greeks living in Bactria and Gandhāra, and several Iranian groups. The famous Graeco-Bactrian culture brought a turning point in the history of Buddhism, the Greek influence was in some respects essential, as e.g. the depiction of the human form of the Buddha, who had been earlier represented only by symbols (footprint, wheel). It seems to be very likely that the Mahāyāna school of Buddhism arose in the cosmopolitan environment of Gandhāra. Under the Kushān dynasty the influence of Buddhism reached perhaps Khwarazm and Sogdiana, but it was not of central importance. The language of the administration in the Kushān Empire was the so-called Gandhāri Prakrit written in Kharoshthi script. The influence of the Kushān Empire reached China and administrative documents written in the so-called Gandhāri Prakrit in Kharoshthi script have been found in the Kingdom Kroraina (east from Khotan), around the Lob-nor and even in Lo-yang, the Chinese capital. This Gandhāri language was used by the Hinayāna Buddhist sect, the Dharmaguptakas, who were among those who made translations of Buddhist texts into Chinese at the beginning of the 5th century. We know from the Chinese traveller Fa-hsien that in 400 there were more than four thousand Hinayāna monks in Shan-shan (Kroraina). The Kharoshthi texts found in the Niya-valley, the "Niya documents" pertained to the Shan-shan kingdom, but there exist some Buddhist texts which may be dated to the second century, and were written in Gandhāri, but found in the vicinity of Khotan. The influence of

the Gandhāri Hīnayāna Buddhists on the development of Buddhism in Khotan is documented by the oldest strata of Buddhist terminology in Khotanese which shows the impact of Gandhāri.

By the Kushān period monks of another Buddhist Hīnayāna sect, the Sarvāstivāda, were spreading throughout Central Asia, taking with them palm-leaf manuscripts written in Buddhist Sanskrit. Paleographic research (see Sander 1968, 1983) has established a close connection between the monasteries in Bāmiyān and Gilgit (Afghanistan and Kashmir) on the one hand and those in Eastern Turkestan on the other.

Buddhism reached Khotan around the second century A.D., and it was a well-established center of Mahāyāna studies in the third century. Khotanese monks served as translators of Sanskrit texts into Chinese. The Buddhist scriptures were first translated into Chinese from Gandhāri and thereafter from Sanskrit in most cases by Central Asiatic monks themselves speaking some of the Iranian languages.

Kashghar and its surroundings were converted by Saka groups to Buddhism in about 100 A.D., a date when Hinayāna prevailed and this was then spread to the cities of the northern route across the Takla Makan. Saka groups converted the region around Tumshuq whose monastery is thought to date from the fourth or fifth century. They used, as we shall see later, not the Brāhmi script of the Khotanese Saka monks but the Brāhmi used among the Tocharians. The Buddhist culture of Tumshuq was destroyed by the Mohammedan Karakhanids in the 10th century.

It is uncertain when Buddhism first reached the Tocharians, but it surely flourished there in the fourth century in the time of the famous translator Kumārajiva (344-413) the son of an Indian father and a Tocharian mother, who was the sister of the king of Kucha. Kumārajiva first practised Hīnayāna, but later turned to Mahāyāna. He translated many Hinayāna and Mahāyāna texts into Chinese, and it was he who introduced into China the Mādhyamika philosophy. We have reliable sources on the fact that both Hīnayāna and Mahāyāna Buddhism was present among the Tocharians.

In the spread of Buddhism a great role was played by the Sogdians, whose network of commercial houses and trading colonies we find all over Central Asia. They used the Silk Road and took with them not only Buddhism but also Manicheism and Nestorianism. Sogdians were among the early translators from Sanskrit into Chinese, and later Sogdians translated Chinese Buddhist literature into Sogdian. The key role of the Sogdians was appreciated by the Turks, who from their first appearance engaged them in their chancellary and diplomatic activities and thus inevitably came under the influence of the Sogdian Buddhists. This is well demonstrated in the Bughut inscription of the First Turkic Khanate, (around 581), which was written for the Turkic Emperor, but

in Sogdian language and in Sogdian script. In this very important inscription, the Buddhist community is mentioned (on the inscription see Klashtornyj-Livshic 1972).

In and around Tun-huang (Dunhuang) the activity of Buddhist monks may have begun in the first century A.D., but the first attested texts pertain to the third century. In the middle of the fourth century they began to build the complex of caves of which the famous Ch'ien-fo-tung (Qianfodong) or Caves of the Thousand Buddhas are the most known. There the Hungarian scholar Sir Aurel Stein, following the instructions of his compatriot geographer L. Lóczy, discovered the Hidden Library, which contained thousands of manuscripts in many of the Central Asiatic languages. The study of the Hidden Library became in the last years a separate branch of scholarship. A good, unpublished introduction is that of G. Uray (Manuscript, 1988 with a very rich bibliography), and he refers to the school of Fujieda who set as its aim to reconstruct the Hidden Library and study it as an entire complex (see Fujieda 1966, 1981). See on the Cave and its chronology also my brief note (Róna-Tas 1968).

One of the earliest scholars working in Tun-huang was Dharmaraksa (Chinese Fa-hu), who was born in Tun-huang around 230 A.D. He travelled in all parts of Central Asia and aquired many of its languages. He collaborated with Indians, Tocharians, with a Yüe-chi, a Khotanese and probably also a Sogdian. It was one of his Chinese disciples, Fa-cheng, who in about 280 founded the large monastery in Tun-huang.

As we have seen Buddhism appeared already at the court of the rulers of the first Turk Khanate. We know of a Chinese temple-inscription which was written in the fifties of the 6th century under the Western Wei dynasty in which the devotion of Muhan Khan the second son of the founder of the Khanate Bumin Khan, to Buddhism was mentioned (see A. von Gabain 1954). It is known that a Buddhist monk from Kāpishi (Gandhāra, Begram) the famous Jinagupta (528-605) who had spent some time in Khotan, taught Buddhism at the court of the Eastern Turks (cca 574-584) during the rule of Muhan khan's successor Tapar kagan.

From the time of the second Turkic Khanate and the Uighur Empire we have very scanty data on Buddhism among the Turks: we know that the Uighur khan converted to Manicheism in 762, but after the defeat of the Uighurs by the Kirghiz and after their move to the South, as we have seen above, they became devoted Buddhists.

Khotan was occupied by Islam around 1000, the Uighurs of Kocho were reached by the same destiny at the end of the 15th century.

*The spread of Manicheism to Central Asia and the Turks*

The Manicheism founded by Māni (217-276) in Persia was a syncretic religion from elements of the Christian religion, Zoroastrism and Buddhism. Its philosophy was gnostic, and built on a dual principle. The World consists of two main principles: Light and Darkness, the Spiritual and the Material principle. Māni was first supported by the Sassanid Emperor Shāpūr I (242-273), but later disgraced because of the opposition of the Zoroastrian priesthood. Under Bahram I (274-277) Māni was imprisoned, where he died (276) and his pupils were persecuted. They spread over the world. Manichean priests came to Sogdiana and converted considerably large groups of Sogdians. The Sogdian Manichean groups were instrumental in the further spread of the religion. In earlier times they used the Persian and the Parthian languages (the latter was the language of the Arsacid Emperors who were replaced by the Sassanids) for religious purposes, but later the Manichean texts were translated into Sogdian as well. The Sogdians became engaged in missionary activities even in China. Manicheism came in Central Asia under the influence of Buddhism which itself could not avoid the Manichean influence. When in 762 the Uighur Emperor visited the Chinese capital after having helped the Chinese Emperor against the rebellion of An Lu-shan (755-757), he came into contact with Manichean priests, who converted him to Manicheism. After returning to the Uighur homeland, an important group of Turks followed the Khan and joined the Manichean religion. This group begun to write in the socalled Manichean script first translations and later original works in Turkic. After a time Manichean texts were also written in the Sogdian-Uighur script.

*The spread of Islam and the Ancient Turks*

Islam (from Arabic *islām* 'submission to the will of God') the religion of the Muslims (Arabic *muslim* 'those who submitted to the will of God', both words from the root *SLM*), was founded by Muhammed who left Mecca for Medina in 622 this is the Hijra, the year of the beginning of the Muslim era. Muhammed died in 632 and was followed by four elected Caliphs in whose hands both the religious and the political power was united. The unprecedented expansion of the Arabs reached Damascus in 635, they captured Jerusalem in 638, and Persia between 636 (the battle of Ktesiphon) and 642 (the battle of Nihavend). The group around Osman (644-656) helped his family, the Omayyads, to power against the son-in-law of Muhammed, Ali. The latter was backed by the Shiites. Ali was murdered and the Shiites got into a bitter struggle with the Sunnites, the partisans of the Omayyads, who acknowledged the Sunna as a part of the holy tradition and the law. The Omayyads ruled until 750. Under Muaviyya (661-680) they conquered Samarkand, Bukhara and Kabul. During the reign of Valid I (705-715) the Arabs conquered Transoxania and reached

the Indus Valley. They tried to encircle from the East the European states and attacked through Caucasia the Khazars. The most important campaign of the Arabs was, as we pointed out above, in 737 when they provisionally forced the Khazar king to convert to the Islam (cf. the Arabo-Frank battles of Toulouse 721 and Poitiers 732). But this turned out to be an episode and on the whole the Khazars resisted successfully the Arab invasions. 751 was the great battle of Talas were the Arabs and their allies clashed with the Chinese and their allies. Though the Arabs won a slight victory their expansion for the time being was stopped. Under the first Abbasid Caliphs power passed to the Persians, the Empire was reorganized on Persian and Byzantine models. After the famous Harun al-Rashid (786-809) the central power of the Caliphs decayed and the actual power passed to local landlords. These local rulers engaged Turks in their struggle for power and the Turkic element became stronger and stronger everywhere. On the other hand the groups which temporarily lost in the local struggles for power fled to the Turks.

As we have seen above, the Volga Bulgharian rulers became Muslims at the beginning of the 10th century, the Karakhanid rulers converted to Islam at the end of the 10th century (Baytash son of Bugra who died 955) and the Seljuks around 1000. The latter even conquered Baghdad in 1055. Khotan fell around 1000 and became an Islamic country. Bukhara, Samarkand and Kashghar flourished as great cities of Islam. The political and religious expansion of Islam was stopped for a short time by the Mongols.

# PART TWO
# INTRODUCTION TO THE STUDY OF OLD TURKIC

# THE SOURCES OF THE HISTORY OF THE OLD TURKIC LANGUAGES IN DIFFERENT SCRIPTS

## THE SOURCES IN RUNIC SCRIPT

The notation *Runic* is a misleading abbreviation. The name of this script was originally *Runiform*, and V. Thomsen also used this term because its outer look resembled the Germanic Runic script. But the word *Runic* comes from the Old High German word *runa* 'secret, secret script' hence 'cult script', and our script has nothing to do with the German Runic script, neither was it a secret script. It would be more appropriate to call it *Carved Script*. Nevertheless, since the name has been already accepted and used, we shall also use it with the rider, that this is the Runic Script of the Eastern Turks or the East Turkic Runic Script (ETRS, see Table I).

The first news about this script was brought to Europe by the Flemish ambassador N. Witsen (1692), the first drawings appeared in the work of Strahlenberg (1730). Many travellers collected inscriptions written in this script, mainly in Siberia in the subsequent period. All attempts to decipher the inscription failed until 1889. In that year N. M. Yadrintsev discovered the long inscriptions near the Orkhon river in Mongolia. Two expeditions were sent to bring reliable material, the Finnish expedition was led by Heikel in 1890, the expedition of the Russian Academy by Radlov in 1891. Both expeditions brought back reliable materials and were published in the same year 1892. The first attempt at deciphering by O. Donner failed. He collected and grouped the units which he thought to be words, and then tried to decipher the letters by comparing them with already known alphabets. One of the inscriptions had a text in Chinese and it was hoped that the inscription was bilingual. Later on it turned out that this was not the case. In two questions, however the Chinese text was of help. It said clearly that the deceased who was commemorated was a Turk and gave the Chinese transcription of his name.

It was Vilhelm Thomsen who first succeeded in deciphering the script and the inscriptions about which he delivered a lecture on the 13th December 1893 before the Danish Academy of Sciences. Radlov, who worked simultaneously with Thomsen and with whom Thomsen corresponded made his results public on the 19th of January 1894. Thomsen's first publication in the Bulletin of the Danish Academy of 1893 was republished with some changes in the 3rd volume of his Samlede Afhandlinger (1922) in which Thomsen added a Postscriptum written in 1920 where he tells us in detail how he deciphered the script.

Since 1893 the study of Turkic texts in Runic script has developed as a separate branch of Turkology. A good overview of studies up till 1917 can be read in Kononov (1982). The latest state of art is summarized in Kljashtornyj (1964), Tekin (1968), Kononov (1980) and Vasilev (1983a). A bibliography of the published inscriptions can be found in the DTS to which now the new data of Vasilev (1976, 1978, 1983a, b) and Kljashtornyj (1980, 1985) have to be added. In the large territory inhabited in the 8th-10th centuries by the Turks even now new inscriptions are coming to light.

The script was originally carved in wood, but most of these wooden sticks disappeared. We can distinguish the following types of texts according to the material on and with which it was written and according the contents of the texts:

1. Memorial inscriptions on steels and rocks
   a. Inscriptions commemorating a dead person
   b. Inscriptions commemorating events.
2. Short inscriptions on various objects
   a. Inscriptions on metal (dishes, coins, etc.)
   b. Inscriptions on ceramics
3. Graffiti on rocks
4. Texts written on paper or wood with a brush or calamus

The hundreds of inscriptions can be grouped in the following areal units (after Vasilev 1983a):

1. Pribaykal-Lena (cca 18)
2. Yenisei (cca 150)
3. Mongolia (cca 35)
4. Altai (cca 16)
5. East Turkestan (cca 14)
6. Northern Kirghizistan and Kazakhstan (cca 28)
7. Fergana, Alai and Northern Tocharistan (cca 18)

The discussion on the chronology of these groups seems to be finished to the extent that it seems to be clear that the earliest inscriptions are those from Mongolia, and none of them is older than the 720's. The older opinion, according to which the Yenisei inscriptions are earlier could not be substantiated. The contrary has been claimed by Kyzlasov (1960, 1965), Clauson (1962), Kormushin (1975), Vasilev (1983) and Róna-Tas (1987b). The primitive nature of these inscriptions is not due to their early age but to the fact that local lords and scribes imitated the "central" inscriptions, and thus the Yenisei inscriptions are

simultaneous with or later than those of Mongolia. The East Turkestan group seems to be the latest, from a time after the Uighurs left Mongolia, the second half of the 9th or from the beginning of the 10th century.

*The origin of the East Turkic Runic script* is a much debated question. Opinions can be divided into two groups. According to some scholars the script is of foreign most probably of Semitic origin, others think that it is autochthonous and developed from tamghas, property marks. The three most recent papers of importance are those of Clauson (1970), Livshic (1978) and Pritsak (1980) with the bibliography of earlier works.

I tried to solve the riddle in a paper read in 1986 at the 29th PIAC conference and published in 1987. My point of departure was that first we have to find the earliest set of letters, then we have to find the system of the script. Only those solutions which give an answer to the seemingly odd structure of the script can be accepted. The hypothetical solution has then to be made plausible by paleographical and historical considerations.

I have distinguished four phases of the development of the ETRS. In the first phase a Turkic group took over an alphabet of the Northern Aramaic type, which was near to Ancient Sogdian, Armazic and Pahlavi but not identical with either. This language had no *š* and *z* phonemes, thus it had to be a language of the Chuvash type, presumably Khazar. In the second phase the Turks developed new letters to meet the special needs of the Turkic language. For this they had two devices. They used pictograms and added diacritical signs to letters which already existed. In the third phase the script was forwarded to a people which spoke a common Turkic language where there also existed phonemes for *š* and *z*, thus these two letters were added and also letters for the clusters *ld*, *nd* and *nč*, perhaps the latter under the influence of the Sogdian script. In the last phase local variants emerged, among them the Yenisei inscriptions with some new letters such as the one for closed *e*. The hypothesis forwarded in Róna-Tas (1987) gives an answer to several questions hitherto unsolved. Why do the letters *a/ä* and *ï/i*, *e* denote back vocalic and front vocalic pairs, while the letters *o/u* and *ö/ü* closed and open pairs of the same synharmonic group? How did the paired letters for consonants develope if none of them can be deduced from its pair, that is $b^1$ cannot be the origin of $b^2$ or vice versa. Why do we have five characters denoting /k/ and why separate letters for *iki/ïkï*, *ökö/ükü* and *oqo/uqu*. What is the cause of the seemingly great inconsistence in the use of the letters denoting sibilants? We know also that the syllabication of the alphabet went on as *aC* and *eC* as we see in the Manichean transcription of the Runic alphabet (cf. Le Coq 1909). Though most scholars have suggested that the script goes directly back to the Ancient Sogdian script this could have been excluded because the sign for $l^1$ goes back to the Aramaic Lamed which in the earliest documents in Sogdian is the letter for *d*. We have acquired a

valuable source on the history of the Turkic language from a period from which we still have no other sources. The most important thing is, however, that after this analysis we get our feet foot on solid ground when we investigate the orthography and the graphotactics of the script as a source of the Old Turkic language.

*The orthographical rules and the graphotactics* are in certain respect very stable, in other cases we find variation and of course also mistakes. From this point of view the paper of Hovdhaugen (1974) is very instructive. He collated the identical passages of the Kül Tegin and the Bilge Kagan inscriptions. The rules are more consistent in the inscriptions of Mongolia, and a few new rules emerge in the later inscriptions and texts. In the following we shall concentrate on the usage of the larger inscriptions.

*Vowels*

The vowels are written in some positions, not in others and there exist a few positions where the fluctuation is great. The vowels are as a rule written explicitly in the following positions:

1. In the word initial position the vowels *ï, i, o, u, ü* are written with the exception of the cases of *ïq, oq/uq, ök/ük* if these sound combinations are written with the special letters Nos. 24, 25, 26 respectively. The *a/ä* is not written. In a very few cases *a-* is written if this was a long *ā* (*āč* 'hungry' *āt* 'name'), but the same word is also written without it, thus Hovdhaugen is right when he states that the spelling of long *ā* was not fixed in the orthography (1974, 61). The writing of *e-* shows some uncertainty. Words such as *et-, el, eš*, etc. are sometimes written with the vowel sign *i* and sometimes without any vowel letter. This points to the special position of the closed *e* in the phonological system. This situation had the result that in the Yenisei inscriptions a special letter was invented for the closed *e* (see Thomsen 1913).

2. In the first syllable after a consonant we find the same rules as the above. Further the *i* and *ï* are frequently not written after *y-*. The combinations *qo-, qu-, kö-, kü-* and *qï-* are in many cases written with the special letters and less frequently with $k^1$ and $k^2$ and with the vowel sign. But the same word can occur in three different writing as *qop* which is written with Nos. 26, 2, 33 $q^owp$, with 22, 33 $q^op$ and with 22, 2, 33 *qwp*.

3. The rules concerning the non first syllables have been studied by Meyer (1965-66) and checked against the parallel texts of KT and BK by Hovdhaugen (1974). They can be now refined and summed up as follows:

a. The vowels *a* and *ä* are written as a rule only if they occur in absolute word final position.

b. The labial vowels are obligatorily written plene or indicated by the letters 25 or 26 if in the preceding syllable the vowel was an illabial one.

c. The vowels *ï* and *i* are obligatorily written plene or indicated by the letter 24 if in the preceding syllable the vowel was a labial one. (But cf. *Bumn* for *Bumïn*). This rule is not valid for the vowels of most of the suffixes. We have to read *sökürtimiz* for *sökrtmz*, *yükündürtimiz* for *yükündrtmz*, because we have *boldï* and *olrtï* for *olurtï*, where in absolute final position the *-ï* is plene written.

d. The writing or indication of the labial vowels after a labial vowel of the preceding syllable is not obligatory, they may or may not be written. But if the first labial vowel is followed by a consonantal cluster the second labial vowel is, as a rule, written as in: *ortusiŋ$^{o}$ru, busgurur, urtur$^{i}$m, körgün$^{i}$n, bolčuda*, etc.

e. The writing or the indication by the consonant 24 of the vowels *ï* and *i* is not obligatory after an illabial vowel in the preceding syllable.

4. In cases where before and after a *k/q* a labial vowel occurred (type *ökün, toquz*) the first vowel was in most cases not written but the second was if the word was written with the letters 25 or 26 resp. (*t$^{o}$quz*). But if they were written with 22 or 23 resp. then the second labial vowel was in most cases not written (*tok$^{lu}$z*).

Summing up the above we can state that the vowels are not written as a rule in the following cases:

1. The vowels *a* and *ä* in absolute word initial position, in the first and further syllables with the exception of the absolute word final position where it is always written.

2. The labial vowels in non first syllables after preceding labial vowels and *ï* and *i* after preceding non labial vowels. Thus in such cases when we read 4,34,14 that is *üčn* we are entitled to read *üčün* and not *üčin*. In fact we find also the notation 4,34,4,14 that is *üčün*.

3. The vowels in absolute final position are as a rule written.

The above are, however, not strict rules, only tendencies, and a considerable number of exceptions exists.

The sign 5 for the closed *e* occurs only in a group of the Yenisei inscriptions (on the letter see Thomsen 1913, the inscriptions are enumerated in Vasilev 1983b p. 96). In the other inscriptions the closed *e* is written with the letter 3 if specially noted. Thus *bir* 'one' and *ber-* 'to give' are written in the same way.

From the graphotactics of the inscriptions the following conclusions can be drawn for the reconstruction of the vowel system of Eastern Old Turkic: The etymological length of the vowels is not marked. The existence of the closed *e* is reflected insofar as it is written in many cases with the sign for *i*, in other

cases it is implicitly written with the sign for *ā*. We can find hints on the labial or non labial quality of the non first syllables, but no answer to the question whether they were closed (*u, ü*) or open (*o, ö*).

*Consonants*

The consonants were written in all cases, and we can only suppose that the laryngeal *h* was not written in the initial position. The original or secondary long (originally reduplicated) consonants are in the most cases marked only with one consonantal letter (thus *ele-* 'to incorporate in the realm' instead of *elle-*, cf. KT E 6 *elledük*, BK E 7 *eledük*). In the case of the consonants *b, d, g, y, l, n, r, t,* and *k* the script has two series of letters (on the sibilants see later). One series can be used only with back vowels (marked with a superscript [1] in the transcription) and the other with front vowels (marked with [2]). There is seemingly one exception. In back vocalic words letters from the front vocalic series occur if there is an *i* sound after or rarely before them. This reflects two different causes. Some suffixes contained originally a front *i* (as e.g. the possessive suffix of the 3rd person) and this *i* has not yet been synharmonized, that is it was front vocalic also in back vocalic words. In such cases the preceding or following consonant may have been expressed by a letter of the front series. The other cause was that the opposition *i:ï* became neutralized and the neutral *i* stood nearer to the front *i*.

The rule governing the use of the letters for the sibilants has only gradually stabilized itself. The cause of the seemingly chaotic situation which gave ground to several hypotheses was that in the other Turkic language from which the Orkhon Turks took their script there did not exist a genuine *š*. The chancellery of the Orkhon Turks was confronted with two problems, the first was to make a difference between *s* and *š* and the second was to make a difference between their front and back variants. The most simple solution can be seen in the Ongin inscription (most probably 720, see Clauson 1957) where only No. 28 was used for the four consonants *s1, s2, š1, š2*. In the Toñuquq inscription the author(s) decided to use No. 29 for both back *s* and *š* and 28 for front *s* and *š*. In the inscriptions of Küli čur (724) Kül Tegin (732) and Bilge kagan (735) No. 29 is used only for back *s* but No. 27 is used for back and front *s* and 28 is used for front *s* and front *š*. In the Uighur inscriptions of Terh (750), Tez (752/753), Shine usu (750) and Suji (cca 840) 27 is used for both back *s* and *š* and 28 for front *s* and *š*. We find exactly the same distribution in the Turfan manuscripts written in Runic script which shows that the Uighurs took with them this script to the South after their defeat in 840 by the Kirghiz. In the Yenisei inscriptions a new sign appears, No 30 (the enumeration of the inscriptions where this letter is used see in Vasilev 1983b 133, 135). In a few cases we find this letter also in the Turfan texts but it is clear that it could not become a

stable part of the system. The letter 30 was mainly used for front vocalic *s* while back vocalic *s* and *š* are written mostly with No. 27 and front vocalic *š* with No. 28 (see Róna-Tas 1987 9). The analysis of the usage of the signs for sibilants shows that the variegated usage of the signs is due to the uncertainty of the orthography and not to some phonetical reasons of whatever origin (as was supposed, e.g. by Tenishev 1976).

Thomsen (1896) mentioned the possibility that the letter for d may have denoted a fricative, and later this became the most common opinion. We shall come back later to this question, here we shall consider only those features which are connected with Runic script. In this script there are signs which denote two consonants, namely *lt* (or *ld*), *nt* (or *nd*), and *nč* (or *nj*). Since in the same text we find in KT E 29 *k$^{1}$il$^{1}$t$^{1}$im* and in BK E 24 *k$^{1}$iltim* it has been concluded that the sign 37 had the value *lt*. In this case we have to do with the suffix *-di* and it is remarkable why the dental of this suffix became voiceless immediately after the sonorant *l*; this would be a kind of dissimilation. Johanson (1979) wrote a monograph on this question dealing with all earlier opinions and concluded that the basic opposition in Old Turkic was not unvoiced:voiced but tense:lax and stop:fricative, thus the letters *t$^{1}$* and *t$^{2}$* denoted stops in this position while the letters for *d* a fricative:

| LETTERS | ALLOPHONES | PHONEMES |
|---|---|---|
| T <t$^{1}$>, <t$^{2}$> -> | [t] | /t/ |
| | [D], [d] | /d/ |
| D <d$^{1}$>, <d$^{2}$> -> | [δ] | |

(Johanson 1979, 90)

which has to be read: the Runic letters <t$^{1}$> and <t$^{2}$> ("T") represented the phone [t], the phone [D] a voiceless lax stop and the phone [d] a voiced lax stop. While the letters d$^{1}$ and d$^{2}$ ("D") represented the fricative phone [δ], [D], [d] and [δ] were allophones of the phoneme /d/.

After *l* the *d* had a stop quality and therefore it was denoted by *t*. Johanson's reasoning is very ingenious and in some questions of detail, such as e.g. the assumption that the disappearance of the final reduced vowels was just in its last phase, (on which see later) he seems to be right. His final conclusions are, however, not fully convincing. The consonants *l, n, r* are all continuants, thus if an original fricative, as Johanson supposes the *d* was, came into contact with a continuant and then became a stop this is also a kind of dissimilation,

since all fricatives are continuants. This has been explicitly claimed by Johanson (1979 32). If the consonant after the continuant was originally a stop (as they were in the examples cited by Johanson) it can be understood why it *remained* a stop. But if it was a fricative it is not clear why it *became* a stop. Thus it seems to be more likely that the *d* was originally a stop, but later it became a fricative and the stop quality was preserved after *l, n, r* and also *d*. It is further not clear why the script has no global sign for the cluster *rt* but has such a letter for *nč*. This latter global sign is the more interesting, since we can exclude here an etymological *j* sound. In any case the appearance of the global signs 37, 38 marks very probably a Sogdian influence (supposed also by Johanson 1979 36, 37), where a stop *d* existed only after *n* and perhaps after *l, r*. It is of relevance, that the global signs 37, 38, 39 have no back and front pairs. This means that they are not from the same phase of the history of the ETRS as the second members of the paired letters. The letters 37, 38 and 39 are most probably from a later time as the second letters of the pairs. It is further remarkable that *z* has only one letter and originally, as we have seen, also *š* had one. Thus we can risk the hypothesis that the letters *z, š, lt, nt* and *nč* are from the period when the script was adapted to the needs of a common Turkic language. This was the period (and language) when and where the stop /d/ begun to change into a fricative. In this case the inner reconstruction of the history of the system of the orthography of Runic Turkic can be used as a source for the history of the Old Turkic consonants.

# THE SOURCES IN BRAHMI SCRIPT

*Historical background*

The name *Brāhmi* refers to the Indian origin of the script. The spread of Buddhism into Central Asia begun very early (see p. 47- above). During the Kushān Period (middle of the 1st century AD - 320) missionary activity increased and priests brought Sanskrit texts through Bactria to the Western oases of what was later called Turkestan. They converted the local people and developed a new Buddhist culture. The monks imported Buddhist texts written in Sanskrit from India and Afghanistan. In the Kushān Period the difference between the various scripts used in North India was not so great and therefore it is not easy to sort out the places of their origin. This changed in the time of the Gupta dynasty (320-535) and later, because the difference in type of script used in the various centres grew. Life in the Buddhist communities of Turkestan was vividly reported by Chinese travellers. Fa-hsien (399-414) crossed Central Asia and reached India from where he returned via Ceylon by sea. Chi-meng (404-424) travelled through Central Asia and visited Pataliputra, the centre of the Gupta Empire. He went back by the same route through Central Asia. The local priests employed Sanskrit script, first for copying the holy Buddhist texts in Sanskrit and later also for the rendering of their translation into the local languages. Later on this local Sanskrit script or Brāhmi was used also for other purposes. Step by step they changed Brāhmi script, adapted it to the needs of the local languages and thus the various types of Central Asiatic Brāhmi emerged. The various local scripts can be divided into two groups. In the North **Tocharian Brāhmi** was mainly used and in the South **Khotanese Brāhmi**. As we shall see the problem of Northern Brāhmi is more complex.

*Tocharian and Tocharian Brāhmi*

*Tocharian* is the name of the language or the languages which were spoken in the Northwestern part of Eastern Turkestan. The name Tocharian itself was for a long time debated. We have an early Uighur Buddhist text, the Maitreyasamiti-Nataka (see S. Tekin 1980), in which it is explicitly mentioned that the work was translated from the Tocharian into Turkic (*toχri tilintin ... türkčä äwirmiš*). On the other hand we find a people in Baktria which was called Tocharian by several Greek, Old Indian and Chinese sources and which spoke an Iranian language in contrast to the Tocharians in Turkestan who spoke a non-Iranian Indo-European language of the centum-type.

The distinction between the *centum* and *satem* types of Indo-European languages is based on the representation for the word 'hundred' in these languages. Latin had *centum* (pronounced *kentum*) and e.g. the English *hundred* developed also from a form with *k-*. Tocharian had *känte*. In the satem group of languages the palatalized initial IE *k-* became *s-* as in Old Indic *satam*, Avestan

*satəm*. Middle Persian *sad* or Russian *sto*. Khotanese had *śata* and the *satem*-form was borrowed by the Finno-Ugrian languages (see Finnish *sata*, Hungarian *száz* 'hundred').

The discussion about the question of who the "real" Tocharians were flared up after Vorobjev-Desjatovskij published in 1958 a Sanskrit-Tocharian bilingual text in which the Sanskrit name *tokharika* was rendered in the Tocharian B language as *kučanne isthake* (see Thomas 1985 15-17 with bibliography). By this it was proved that the language of Kucha, the Kuchean language was also called Tocharian. The only problem that remains is that we do not know yet which of the two groups: the speakers of the centum language in Turkestan or the speakers of the Iranian language in Baktria-Tocharistan originally bore the name, and which adopted it from the other. But we are entitled to call Tocharian the language spoken in the Northwest of Eastern Turkestan. We dispose of Tocharian texts from the 6th-8th centuries.

Tocharian has two separate dialects, which stand so far from each other that some authors speak of two Tocharian languages. Tocharian A is also called East Tocharian and was spoken in the oases of Turfan and Karashahr. Tocharian B or West Tocharian was spoken in and around Kucha, hence it is also called Kuchean. Some texts written in Tocharian B reached the territory of Tocharian A. It has also been argued that Tocharian A was a purely liturgical language in the monasteries of the East (Lane 1967 122) but this has been questioned by Thomas (1985 127). Within West Tocharian a special group can be defined, a dialect which is mainly characterized by the texts found in the caves of Ming-öi Qizil (MQ) West of Kucha. According to some scholars the language of the Ming-öi caves is a separate Tocharian language, the name Tocharian C has been also suggested. Most of the preserved Tocharian texts are Buddhist texts translated from Sanskrit, but in West Tocharian we have also original Buddhist texts, colophons, letters, poems, drafts and exercises.

The Tocharian Brāhmi script is demonstrated on Table II after Krause and Thomas 1960. To this chart we have to add some remarks which may be of importance for Turkologists.

In square brackets [ ] are those letters which were added by the Tocharians to the original Brāhmi. Most of these letters are underlined in the transcription. This means that they had to be read with *ä* instead of the "inherent" *a*, that is <u>*ka*</u> has to be read as *kä*. The notation *ä* after these letters is, however, misleading in our case, because it is used for the notation of a vowel that is unstable in Tocharian. It became under stress *a* in West Tocharian, otherwise in certain positions it appears as *i* both in West and East Tocharian and disappears in open syllables in East Tocharian and in open not accented syllables in West Tocharian. These letters occur also in Turkic texts but the "inherent" vowel is, as we shall see, not *ä*. The Tocharians used also an additional set of

letters which were used only in transcribing foreign words, mainly Sanskrit ones. Those letters which were used only for foreign words are in parentheses ().

The Tocharians played an important role in the conversion of the Turks to Buddhism. Their role also can be followed up by the Sanskrit loanwords a large group of which came to Uighur via Tocharian. After a time the Tocharians begun also to use the Brāhmi script in their missionary work. Most of our Turkic texts written in Northern Brāhmi script are wordlists, terminological glossaries, where the Sanskrit originals and their "Turkic" equivalents are given, the latter in Brāhmi script.

There does exist a very good bibliographical introduction to Tocharian studies written by W. Thomas (1985) on the literature published between 1960 and 1984. This book is also a good point of departure for earlier literature if one overlooks some personal biases. Thomas stated that many problems connected with paleography and graphotactics are yet unsolved. The works of Sander (1968, 1983) can be used to good profit. A series of published and unpublished papers of Hitch (1981, 1983, 1984, 1989, 1987 forthcoming dissertation unpublished) raised interesting problems, but his original claim that Brāhmi was first used for Turkic and only later for Tocharian and the Saka of Tumshuq cannot be accepted as he himself seems to realize it. He later suggested that the Manichean script had a key role. Realizing the difficulties involved he turned to Bactrian (1987). Very interesting and important material can be found in the paper of Couvreur (1965) where he collected writing exercises in Brāhmi. Thomas wrote in accordance with Isbaert (1983 35): "Die Kontroverse in der heutigen Forschung wurzelt natürlich darin, dass die tocharische Schrift ihrer Entstehung nach auf die indische zurückzuführen ist, ohne dass bisher gelungen ist, genau festzustellen, welche phonetische Substanz die Vokale der Brāhmi-Zeichen im Tocharischen haben" (1985 33). This situation is complicated by some incertitudes in Tocharian phonology. For the next period the monograph of Windekens (Bd. I 1976 phonology and wordlist with etymological data) will be the point of departure, which together with the detailed bibliography and critical remarks of Thomas (1985, 26-56) on divergent opinions, enables us to get a good overview. Without going into details it is clear that the problems of Tocharian phonology and those of the Turkic texts are interdependent.

*Northern Saka or Tumshuqese*

Saka dialects were spoken not only in Khotan (see below) but also in the Northwestern parts of Turkestan. One of these dialects was mentioned in the second half of the 11th century by Mahmud al-Kāshgri (see below) as *Känjäk*, and this name also occurs in Old Tibetan sources written as *Ga-'jag* (read *kanjak*) meaning 'the people of Kashghar (speaking Saka)'. NE of Kashghar a

Saka community lived around the present-day Maralbashī or Barchuq. The most important texts written in this Saka dialect were found around Tumshuq, near Maralbashi. Therefore this dialect was called Maralbashî, Barchuq or Tumshuq(ese) causing a certain confusion in the literature. Near to Tumshuq at Toquz Sarai Pelliot found a text which he handed over to Sten Konow who then had already published eight texts from Tumshuq. The text found by Pelliot turned out to be the Karmavācanā and was written in a less complicated Brāhmi showing almost no influence of Tocharian Brāhmi script and reflecting an older stage of the history of Northern Saka. The language of the Pelliot text was called by Gercenberg (1981), who misunderstood a statement by Bailey (1958 130, 147), Murtuq (Saka). In fact one text which shows similarities with later Tumshuqese was found in Murtuq, but this is not the text of Pelliot (P 140). Beside these texts three more were published by Bailey (1960) and a further three by Emmerick (1971), thus uptill now fifteen texts represent Tumshuqese. It seems to be proper to distinguish Old Tumshuqese (the text of Pelliot 140 on which see the latest publication of Emmerick 1985 with bibliography) and Late Tumshuqese. Late Tumshuqese used Tocharian Brāhmi with some additions.

Old Tumshuqese Brāhmī (Pelliot 140) has only the diacritic sign double dots on the *aksara* and the radicals *ṛ* and *ḷ*. The New Tumshuqese Brāhmi script was first analyzed by Konow (1935, 1947) and recently reanalyzed by Hitch (1981), who made later corrections to his proposals (upublished manuscript dated 9th September 1986). Unfortunately we find no material for Tumshuqese Brāhmi in the works of Sander (1968, 1983). We know about a work which she began on Saka Brāhmi but which has not yet been finished and published. Konow (1935) used for his analysis an *aksara* chart from the same region (Maralbashi TIV M 58). On one side of the partly damaged leaf there were three lines of a Brāhmi alphabet. The first two lines contained the usual Sanskrit radicals, while in the third line the new signs were written: *ka, ta, na, ma, śa, ṣa, sa, wa* and *ra*:

1...*i u ū ṛ ṝ ḷ ḹ e ai o*...

2...*ta tha da dha na : ta tha [da dha na:] pa pha ba*...

3...*ka ta na ma śa sa ṣa wa ra: (1) 2 [3] 4*

Hitch (1981 61) agreed with Konow that this represents the Tocharian B alphabet. On the opposite side of the leaf twelve previously unknown letters were found. Konow (1935, 1947) tried to interpret their value but Hitch reinterpreted them. The new interpretation of Hitch is, with some exceptions, acceptable. The "new signs" are the following (see Appendix I):

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| za | ga | =2 | ża | rla | kha | da | =4 | =7 | śa | dza | xša |

Hitch correctly stated that No. 3 was a variant of No. 2, No. 8 a variant of No. 4 and No. 9 a variant of No. 7. Hitch claimed that No. 3 was an upside down form of No. 4. The letters Nos. 2, 3 denote *g*. This sign was interpreted by Clauson (1962 95) as formed from an *r* with a subscript second *r*. This is improbable. No. 2 has the same horizontal line as diacritic which we find in *dza* and which is added in *ma, sa*. The letter No. 11 *dza* is *za* + diacritic bar or horizontal line. In an additional remark dated 9th September 1986 Hitch has identified No. 5 instead of *la* as *rra*, but it is clearly the ligature *rla* (see Sander 1968 Table 38 one but last line 6th and 7th squares). No. 6 is in fact, as Hitch corrected himself, neither *za* (so also Bailey 1958 137) nor *khu* but *kha*, and No. 10 is neither *γw* nor *su* but simply *śa*. With these we get a list of the following "new" Tumshuqese letters: <za> <ga>, <ża>, <ḏa>, <dša> and <χsa>. As we shall see that these additional letters are the same which are used for Turkic Brāhmi, only the "Turkic" letter <qa> is lacking here and and we find the Tumshuqese <χša> (in front of the Khotanese *ksa*) which does not seem to occur in Turkic Brāhmi.

More difficult is the reconstruction of the phonological system of Tumshuqese. Emmerick in his two papers on the phonology of the vowel system (1979) and the consonant system (1981) has dealt only with Old and New Khotanese, as did Gercenberg (1965, 1981). In his last paper on Tumshuqese Emmerick (1985) mainly investigated grammatical problems and referred to phonological ones only occasionally. Thus we have to rely in most cases on Bailey's remarks in the Handbuch der Orientalistik (Bailey 1958). He stated for New Tumshuqese: "The Brāhmi script has been enriched with signs to express Iranian sounds for which Khotanese used two letters. Here *ts* was used, but new signs were available for *dz* (Khot *js*), *z* (Khot *ys*), *ž*, *z*, *χš* and for *ḏ* beside *t*, and *d* for *δ*" (Bailey 1958 149). From the material one has the impression that the Late Tumshuqese letter *ḏ* denoted a stop with inherent *ä*. This can be seen e.g. in the word 'law' which is written in Early Tumshuqese (where there was no letter *ḏ*) as *data-* (P 140, K 1.7, 5.8, 10.4), but in late Tumshuqese as *daḏä-* (*dadä-*) as we have *duḏär* 'daughter' and *piḏär* 'father'. On the relationship between Tumshuqese and Tocharian see recently the remarks of von Hinüber (1988).

### *Turkic texts in Northern Brāhmi*

#### *Earlier Studies*

In 1904 H. Stönner published a paper on Central Asiatic texts in Sanskrit found at Idiqut Shahri. As an appendix he added 14 lines of a Turkic text written in Brāhmi. The published 14 lines were only a part of the 40 line text. The Stönner-text was discussed by Lewiczki (1936) and by Bailey (1938). Some

places of Turkic texts written in Brāhmī were cited by F. W. K. Müller. A. von Gabain gave a short introduction to Turkic in Brāhmī script in her *Alttürkische Grammatik* (1941, second ed. 1950, third edition 1974). In this book she gave a facsimile of a part of text I (lines 1-6) with a transcription. A few, preliminary remarks on Turkic in Brāhmī script can be read in her paper on Old Turkic "Schrifttum" (1948). The first publication of Turkic texts in Brāhmī appeared only in 1954, as volume VIII of the *Türkische Turfan-Texte* by A. von Gabain. She wrote that from the approximately 100 fragments only those published are of interest, the rest offer almost nothing which can be identified with certainty. This book contains 15 texts which are indicated by the letters A to P, three are from Sängim (A, C, I), two from Khocho (E, K), one from Yarkhoto (M) and the rest from Murtuq. Unfortunately only the facsimile of text C is added entirely as well as lines 9-16 of text F and lines 10-18 of text G. In the quoted paper of Stönner lines 25-38 of the text D have been published. Concerning the technique of the edition A. von Gabain wrote : "Auf ausdrücklichen Wunsch von Geheimrat Lüders wird jeweils in einer obersten Zeile der Text genau so abgedruckt, wie er im Mss. steht, d. h. in Aksaras abgeteilt. Das dürfte die Kritik an der vorgeschlagenen Worteinteilung erleichtern und eine klaren Eindruck vom ursprünglichen Schriftbild geben" (Gabain 1954 4). This was an excellent idea and helps in the control of those texts where we are not able to check the original. It does, however, not substitute the original. Unfortunately text D was not available for A. von Gabain when she prepared her book for publication and thus this text is published only in the linear, normal transcription. G. Clauson devoted a chapter in his book published in 1962 to texts in Brāhmī script. Though this is rarely cited, it seems to have exerted a great influence on the opinion about Brāhmī Turkic. Clauson returned to the question of the Brāhmī transcriptions in a paper written in 1967. Maue and Röhrborn published a text in 1976 (with additional remarks in 1978). L. Johanson tried to group the Brāhmī text according to the transcription of the Old Turkic dentals in 1979, to which Maue added important remarks in 1983. The representation of the gutturals in Brāhmī Turkic was dealt with by Maue in 1984. The dissertation of Maue in which he dealt with the Brāhmī Turkic texts remained unpublished but is cited by several scholars (?1982, henceforth to be cited as Maue s.a.). Maue mentions a Sogdian bilingual text (Mz 639) written in the same Brāhmī as the Uighur and the Saka texts. In his unpublished dissertation already mentioned Hitch has investigated the Turkic Brāhmī in connection with the complex problem of Central Asiatic Brāhmī (1981). Some of his results were published (1983, 1984, 1987, 1989), other remained unpublished (corrections to 1981, a paper read in 1987 Bloomington). P. Zieme, in a paper read in 1979 and published 1984, dealt with a broader question, namely the use of Brāhmī by the Uighurs to render glosses or retranscribe originally Sanskrit

words. Referring to Bazin (1974 see below) he expressed his doubts on the hitherto accepted early dating of the Brāhmi Turkic texts. From the works of Maue (1983, 1984 and unpublished) we learn that there do exist some hitherto unpublished texts which he is going to publish.

Neither the inventory of the letters nor the orthography or the graphotactics are uniform. Johanson (1979) and Maue (1984) tried to find groups of texts in which common orthographic rules can be observed. But this is not the case. One can find some predominant tendencies in certain manuscripts and we can perhaps speak of certain "orthographical isoglosses". This is of importance insofar as we can state that no uniform Brāhmi Turkic orthography existed.

Text P from Murtuq contains fragments of a calendar. Bazin(1974) investigated this calendar and came to the conclusion: "Il ne fait donc aucun doute que le fragment de calendrier uygur en caractères brahmi publié par A. von Gabain (MM 73-76) vaut pour l'anné 1277 (-1278) du calendrier sino-uygur" (Bazin 1974 449). Maue (s.a. XLIV) carefully rechecked the argumentation of Bazin and did not find it convincing in all details. He can accept a dating of the calendar between 936 and 1494 but there is surely one piece (Mz 813 unpublished) which is from the Mongolian period, i.e. after 1209 and others may be later as well (see Zieme 1984).

One can agree with Maue that the Brāhmi Turkic texts have to be dated between the 9th and the 13th centuries, I would even dare to suggest the 10th century as post quem date. He is surely also right when he assumes that the texts connected with Tocharian materials (bilinguals, etc.) are earlier, those which have only Uighur script parts are later. The use of Brāhmi could have been continued until the region lost its Buddhist character.

*The texts*

Maue (s.a.) distinguished four groups of Turkic texts written in Brāhmi:

1. Sanskrit-Uighur bilingual texts
2. Uighur texts in Brāhmi
3. Turkic texts written in Uighur script with Brāhmi glosses
4. Turkic texts written in Uighur script with inserted Brāhmi words.

According to Maue there do exist roughly 79 fragments or units, of which 57 are bilinguals, (published 18 fragments = 9 texts), 14 Uighur texts in Brāhmi (published 8 fragments = 7 texts + 1), 7 manuscripts with Brāhmi glosses and one manuscript with Brāhmi words inserted (see on this the paper of Zieme 1984). The already published texts with their place of origin and new numerations:

| *Gabain 1954 Letter* | *Place of origin* | *Mainz numbers acc. to Maue* |
|---|---|---|
| 1. A | Sängim | 680+681+683 (bilingual) |
| 2. B | Xočo | 328 (bilingual) |
| 3. C | Sängim | 642 (bilingual) |
| 4. D | Xočo | 718 (bilingual published only in linear transliteration) |
| 5. E | Xočo | 132634+636+833+834 (bilingual) |
| 6. F | Murtuq | 626 (bilingual) |
| 7. G | Murtuq | 645+646+836+837 (bilingual) |
| 8. H | Murtuq | 836 (bilingual) |
| 9. I | Sängim | 187+209 (Uighur, Maue: bilingual) |
| 10. K | Murtuq or Idiqutshahri | 632 (Uighur) |
| 11. L | Xočo | 614 (Uighur) |
| 12. M | Yarχoto | original dissapeared (Uighur) |
| 13. N | Xočo | 614 (Uighur) |
| 14. O | Murtuq | 641 (Uighur) |
| 15. P | Murtuq | 633 (Uighur) |
| 16. M-R | Sängim | 629 (bilingual) published by Maue and Röhrborn (1976, 1978) |

Two further fragments from Ming-öi Kizil have been discussed by Maue (Diss, s.a. pp CXIII, CXIV) with the original numbers MIK III 417 and MIK III 419.

*The paleographical origins*

After the pioneering works of Hoernle, F. W. Thomas, Lüders and others, Lore Sander (1968) summarized the results of earlier and her own studies on the paleography of the Central Asiatic Buddhist texts written in Sanskrit. She distinguished the earlier (3rd-4th cent.) and the later (4th-5th cent.) Gupta-type alphabets. In the 6th century a new type of Gupta script was introduced, which was independent of the simultaneous Gilgit/Bāmiyān type. The earliest Sanskrit texts in East Turkestan were written with the Gupta-type of Turkestan (column q). This *Turkestan Gupta* was replaced by *Early Turkestan Brāhmi* (cols r, s). The earliest Turkestan Mss were written in this type of script. The Early Turkestan Brāhmi developed in the North, among the Tocharians and the Northern Sakas and in the South in and around Khotan in divergent ways. In the North, *North Turkestan Brāhmi* developed fully in the 7th century into two subtypes (cols t and u). As we shall see the alphabet under

col. t is the very one which became the standard Tocharian type. Type u is somewhat later, and according to Maue (s.a.) became the basis for writing Turkic. *South Turkestan Brāhmi* (col v) served the Khotanese monks. All these occurred with Sanskrit texts. In the case of the adaptation of Brāhmi to the needs of the local languages new letters and graphotactic devices were needed and partly introduced. Some of the original letters became superfluous, other changed their value, and/or were transformed, acquired diacritical additions. All this of course did not occur on a "classical" Sanskrit basis, but departed from a kind of pronunciation which is usually called *Late Hybrid Sanskrit*. In Appendix II I reproduce the tables 29-40 from Sander's book.

*The new signs*

We are confronted with two questions, what the value of the "old" or original letters was and what the value and origin of the new signs were. Both Hitch (1981 59) and Maue (s.a. XXVIII) tried to visualize the distribution of the new signs. I give a third table:

| | Murtuq (Tumshuq) | Tocharian | Uighur | Sogdian |
|---|---|---|---|---|
| 1. *rla* | + | | | |
| 2. *kha* | + | | | |
| 3. *śa* | + | | | |
| 4. *dza* | + | | | |
| 5. *χša* | + | | | |
| 6. *za* | + | | + | + |
| 7. *gā* | + | | + | + |
| 8. *ḍā* | + | | + | + |
| 9. *ža* | + | | + | + |
| 10. *wa* | + | + | + | + |
| 11. *ḳā* | + | + | + | + |
| 12. *rā* | + | + | + | + |
| 13. *ḷā* | + | + | + | + |
| 14. *ṭā* | + | + | + | + |
| 15. *pā* | + | + | + | |
| 16. *mā* | + | + | + | |
| 17. *śā* | + | + | + | |

| | Murtuq (Tumshuq) | Tocharian | Uighur | Sogdian |
|---|---|---|---|---|
| 18. *sə̄* | + | + | + | |
| 19. *sə̲* | + | + | + | |
| 20. *n̲ə̄* | + | + | | |
| 21. *ay* | | | + | |
| 22. *ey* | | | + | |
| 23. *uyu* | | | + | |
| 24. *oy* | | | + | |
| 25. *oyo* | | | + | |
| 26. *kya* | | | + | |
| 27. *kyu* | | | + | |
| 28. *q* | | | + | |
| 29. *hk* | | | + | |

Maue writes : "In welcher Sonderschrift die in M, S und U verwendeten Sonderzeichen primär sind, lässt sich zwingend nicht nachweisen. Doch spricht der Typus der graphischen Gestaltung der Zeichen No. 1-9, wofür es keinen Anschluss an Normalzeichen der Brāhmi gibt, sowie die Verwendung im Ensemble ausschliesslich in M, dafür, dass M der gebende Part ist." (Maue s.a. XXXII-XXXIII). The opinion of Maue that the origin of the new signs 1-9 can not be connected with the "normal" Brāhmi signs is rather doubtful and is impossible to substantiate. The contrary is easy to demonstrate:

No. 1.  *rla* is an original *r* and a subscript ~ *l* that is *rl*

No. 2.  *kha* is a normal Brāhmi *kha*, its cursive type

No. 3.  *śa* is the normal Brāhmi *śa*, its cursive type

No. 4.  *dza* is the "new" sign *za* (No 6) with the vertical bar (though in its extant form is so far different, that the upper part is separate and may be concieved as  *zr*).

No. 5.  *χša* is the normal Brāhmi *ṣa* with the lower "foot" which is part of many other "new letters".

No. 10.  *wa* is the older Brāhmī *o*, while the new *o*  is *o*+*a*, i.e. long *o* (Hitch 1981). According to Maue it would denote a labiodental *v* which is rather problematical.

No. 11.  *ḳā* seems to me to be a truncated *cha* turned through 90°

No. 12.  *ṇā* is Brāhmī *ṛ* of the "type a" (Sanders col t) and the lower part is the ligature form of *ṛ*, i.e. here we have *ṛṛ*.

No. 13.  *ḷā* old Brāhmī *ḷ* (cf. Krause-Thomas 7 &)

No. 14.  *ṭā* is normal Brāhmī *dha* redefined

No. 15.  *ṗā* is normal Brāhmī *ba* redefined

No. 16.  *ṃā* is Brāhmī *ba* + vertical stroke

No. 17.  *śā* is No. 18 turned upside down

No. 18.  *ṣā* is a truncated Brāhmī *cha*

No. 19.  *sā* is Brāhmī *ṭha* and a slanting stroke

No. 20.  *ṇā* is No 14 with a cross added.

To Nos 21-27 I shall return when dealing with the Turkic Brāhmī.

No. 28.  *qa* is Brāhmī *ka* and the vertical bar.

No. 29.  *hka* is a ligature of *ha* and *ka*.

Thus we are left with four letters, Nos. 6, 7, 8 and 9, the graphical origin of which is yet unclear. These letters are common, and only common to Uighur, Sogdian and Murtuq. This has been duly recognized by Hitch and he is also right when he remarks that though it would be logically possible for different groups to have created different signs, and then later to have exchanged their respective inventions, this is far less likely than invention and unidirectional borrowing (Hitch 1981 81). Maue (s.a. XXXII) remarks that Nos. 11 and 12 are in M not used as vowelless consonantal final signs as they are in S and U. "Dadurch ist Einfluss von S und/oder U auf M und versa vice in diesem Be-

reich auszuschliessen" (s.a. XXXII). I don't see this argument as compulsive. Since No. 4 is a derivative of No. 6, i.e. the letter for [dz] is derived from a letter for [z] (either with the diacritical bar or with the subscript *r*) and with this the direction of the transfer of the writing is given, M is at the end of the line, the only open question is whether M took the writing system from U or S. The solution of this problem depends on the relationship between U and S. For the time being this question is open.

*The transcription*

General remarks

Before we go into the details of the Northern Brāhmī transcription of Turkic, we have to discuss some basic problems of the use of Northern Brāhmī script. As we have seen all consonantal letters have an "inherent" vowel, i.e. they have to be read either with *a* or with *ä* as Ca or Cā if the vowel is not marked otherwise. Vowels other than *a* and *ä* are marked with diacritical signs. In the Sanskrit alphabet there existed originally the following vowel signs: *ā, i, ī, u, ū, e, ai, o, au* and the vocalic laterals *r̄, ṛ, l̄, ḷ* were also marked by secondary signs. To this was added the vowel sign for *ä*. Thus e.g. a syllable *ka* was written by a simple letter *k* a syllable *kā* with a letter *k* and the additional sign for long *ā*. Now a syllable *kä* could have been written in two different ways. Either it was written with the new sign *ḵ* and thus it was automatically read as *kä* or they wrote a "normal" *k* and added the diacritic *ä*. We have once more to remind the reader that *ä* does not render the Turkic /ä/.

Graphotactics and transcription

In Brāhmī script the basic graphotactic rule was the rule which aimed at building up an *aksara*, i.e. a complex of consonants and a syllabic vowel, monophthong or diphthong which could be read unambigously. If one wanted to write a simple syllable containing a consonant and one of the two vowels *a* or *ä* one had only to write the consonant, the vowel being "inherent" was automatically read after the consonant. If the consonant was followed by any other vowel than *a* or *ä*, a diacritic sign was put above or below the consonantal letter. This means that the vowel sign was not written after the consonant, but "joined" to it because it was considered a modification of the otherwise existing "inherent" vowel. Whether the diacritic sign was put above or below the consonant was regulated by tradition. The diacritics for *i, e, ai, o* and *ä* were written above, the diacritics for *u, r,* and *l* were written beneath. With these rules all syllables of the type CV could be written. If the vowel was long, a diacritic sign was added in most cases above, to the right of the consonant. Thus in the case

of the type Cā or Cā only the sign of length was added e.g. *kā* . If the long vowel was other than /a/ and /ā/ the vowel diacritic and the sign of length was written. Thus e.g. the syllable *ko* became *kō* by adding the sign of vowel length: <ko> <kō>. In case of *i* the diacritic sign of *i* and the sign of length coalesced and on the surface a graphemic opposition appeared of <i> : <i>, as e.g. in the case of <ki> and <ki>. The Indian scripts, such as e.g. Devanāgari rendered Sanskrit *ī* and *ū* with signs independent of *i* and *u*.

In the case of a syllable of the type VC the diacritical signs could not have been used because they rendered only vowels after a consonant. In such a case independent vowel signs were used. The following vowels were represented by separate signs: *a, i, u, e, o, ai, au, ṛ, ḷ*. The length was marked by adding the diacritics to the vowel sign, thus e.g. <a> and <ā>. In case of *ä-* an <a> was written with the two diacritic dots above: <ā>. The practice of considering <a> as a sign with the "basic" vowel inherent [a] and modifying it with a diacritic for another vowel meant that an initial vowel *o-* could be written not only with the separate sign for <o> but also with <a> and the diacritical *o*. Thus a syllable of the type oCV could be written in two ways, either by writing the independent *o* and then the consonant or writing the independent sign for *a*, the diacritic *o* above it and then the consonant as in the following examples: <oka> or <a°ka>.

The graphotactic way of rendering initial vowels with a basic *a-* and diacritic sign gradually spread in Turkestan and the *u-* was the last vowel in the case of which the independent vowel sign other than <a-> was used. We understand the rendering of the Turkic vowels only if we can keep in mind this practice. The Tibetan transcription system was also based on this type of writing.

In her work A. von Gabain wrote: "Wenn auf die engen Vokale *ï, i* und *u* ein einzelner Konsonant und danach der gleiche Vokal (*ïdïl-, ičï, iki, ilin-, itig; uluγ, ulus, umuγ*) oder aber der Vokal *a* (*ïγač, ulati, uyat-, upasi, upasanč*) folgt, ist diese Lautverbindung *ïdï* usw. in den hier veröffentlichten Texten zuweilen in einem einzigen Aksara geschrieben worden. Es erscheint hier also zuweilen *iti lma qi, ičï nda...*". She wonders whether in such cases the writing indicates a weak pronunciation of the initial vowel. We have to exclude this possibility. As we have seen above, graphotactically the basic vowel signs were treated as basic i.e. consonantal signs. First they wrote an independent *i* sign then beneath the *i*-sign a *t* and above this unit perceived as an aksara a diacritic *i*. Thus we find e.g. *iti lma qi plo -ṛ* which we have to read *idilmaqi bolor*.

The new signs for *r* and *l* i.e. the *ḷ* and *ṛ* were used for the syllabic *r* and *l*. According to Gabain (1954 4) they denoted *lə* and *rə*. Therefore she transcribed words written as *a ṛ g* as *arəg* or *tom ḷ r g* as *tomlarəγ* being *arïγ* and *tonlarïγ*. Two dots above an aksara without the curved stroke to the left denot-

ed, according to Gabain, a reduced vowel (Gabain: Murmelvokal). This kind of notation occurs only in positions where we would expect an *ī*-sound as in such words as: *a yi ga tl g* for *ayīγ atlīγ*, and Hitch is right when he claims that this kind of writing denoted back vocalic *ī* (and not a reduced vowel).

In the case of the syllables of the type CVC the problem of the final consonant had to be solved, since all consonantal signs had to be read with a vowel after the consonant. For this purpose Sanskrit orthography used the socalled *virāma* sign. Turkestan Brāhmi and especially Tocharian Brāhmi had this virāma sign, it was put on the left upper part of the letter and bent to the left. In some of the Mss the vowelless quality was marked by an added dot above (similar to the Jezma and Sukun of the Arabic script, a small circle developed from the letter Jim). It is of special importance that in Tocharian and in Turkic Brāhmi the vowelless final consonants were overwhelmingly rendered with the "new signs" having the inherent "ā". This becames understandable if we keep in mind that this vowel disappeared in certain positions in Tocharian. The Turks, as we shall see, extended this to all new signs, also to those which had an inherent *a*. Thus e.g. a syllable of the type *kak* was written as follows:  <kak>. The fact that, if available, the new signs were used to render final consonants makes improbable the supposition of Konow that the new signs denoted a palatalized variant of the consonant.

In the Sanskrit system of writing aksaras a special case was when the vowel was followed by a nasal consonant. Originally the nasal consonants may have belonged to the following syllable, but they began to disappear through leaving behind nasalization on the preceding consonant, as e.g. in French. The nasalization of a vowel was marked by a special sign: the *anusvāra*. The anusvāra is graphically a dot above the aksara and in Latin transcription it is transcribed by an *m* and a dot beneath: *ṃ*. After the disappearence of some nasals in the pronunciation of Sanskrit, syllables ending in a nasal consonant could have been written in different ways. E.g. a Turkic word such as *sanliγ* one could write either with a *sa*, an anusvāra on it, then a *la* and an *i* on it and a *g*: *saṃ-li-g*,  or with a *sa*, a ligature *nl* and an *i* above it and a *g*: *sa-nli-g* .

Tocharian Brāhmi orthography used a special way to indicate a reduced vowel or a non syllabic part of a diphthong (see Krause-Thomas 1960 40). In such cases it was permitted to write two vowel signs on one aksara. Thus e.g. the Tocharian word *nāno* 'again' was written in the following way: two *ns* were written beneath each other and the vowel signs for *ā* and *o* were written above, or they wrote the new sign *n̲*, then beneath the normal *n* and above the diacritic *o*: *n̲no*. This rule we find in the rendering of the reduced vowel of such words as *biši γ* 'ripe' which is written as *pśi-g* or *bilig* 'knowledge' which is written as *bli-g*. This type of transcription is very interesting, because it shows that

there existed reduced vowels in the Turkic language and that the word accent was on the last syllable.

A syllable of the type $C_1C_2V$ also caused problems, because after a consonantal sign a vowel had to be pronunced. If this was not the case a graphic convention was used. If between the two consonants there was no vowel they were written in ligature, i.e. the second was written (in most cases) beaneath the first. Thus we find the rule that if two consonants are written after each other horizontally (linear writing) a vowel has to be read between them, if they are written vertically there has to be read only one vowel after the consonant written at the bottom of the group. This entailed that some frequently used letters had special ligature forms, forms which were used only if they were written beneath the preceding letter. Such a letter frequently used in Turkic transcriptions is ⟨ya⟩ which had in Tocharian Brāhmi the form:  (<  <  ). Its ligature form was earlier like  and later  (where  denotes the consonantal cluster). Other letters also developed their ligature variants.

In the case of a syllable like *-lso-* the writing practice was to write *la*, beneath *sa* and above the diacritic *o*. In the space this made the impression that since the *o*-sign was above the succession was *o-l-s*, but this was not the case, and the vowel has always to be read as the last:  .

Most of the vowel diacritics are written above the aksara, only *u*, *r* and *l* were written beneath. The *u*-diacritic had different shapes under different consonant letters ranging from  to  or  (see the tables 33, 34 of Sander). The result was that in cases such as CCu or CCCu the *u*-diacritic became more and more obscured. If the syllable had e.g. a long *u* vowel the *u*-diacritic was written below, but the marker of the length on the right upper side of the aksara, e.g. *klū*. 

We can understand the representation of Turkic sounds in Northern Brāhmi only if we take into account the pecularities of the Tocharian usage of the old and new letters. This points to the fact that the Turks used a type of Northern Brāhmi script and graphotactical rules to which they added a few more for their own purposes.

The representation of the Turkic vowels

Since Tocharian Brāhmi had the possibility to express the vowels *a*, *i*, *u*, *e*, *o*, whatever their exact phonetic value may have been, the major challenge was to find a solution for the rendering of the front vowels. The graphotactic solution was taken over from Sogdian-Uighur script. In Uighur script a *yod* is placed after a *waw* if one would like to express an *ö* or *ü*. This usage goes back to the Sogdian writing system. In Turkic Brāhmi a subscript *ya* was used for this purpose. This subscript *ya* had, however, no inherent *a*, since the aksara

under which it was placed already had another vowel indicated. Either the letter beneath which the subscript *ya* was written was itself an independent vowel sign, or if it was a consonantal sign, the consonant and the subscript *ya* together had *a* vowel, either the inherent *a*, or another vowel written with a diacritic sign.

E.g. *u* was written as 𐽆 and if *ya* was written beneath, it became graphically 𐽆𐽃 <uya> read: *ū*, or *o* was 𐽅 and if a subscript *ya* was added 𐽅𐽃 <oya> should be read as *ō*. This is the same graphotactical solution as Uighur *WY* with the additional advantage that since *waw* had two values its yodicized form also had two values, while in Brāhmi *o* became *ō* and *u* became *ū*.

If *o* or *u* was in postconsonantal position the procedure was the following: e.g. *ka* 𐽁 became *ko* 𐽁̄ with the diacritic sign of *o* ¯. If a subscript *ya* was added 𐽁𐽃̄ it had to be read as *kō* though graphically <kyo> was written. The idea behind this procedure was that the aksara originally denoting *ko* was modified by the subscript *ya* to *kō*. In the case of the syllable *ku* 𐽁𐽆 the subscript *ya* was added and thus it became 𐽁𐽆𐽃 *kū*. Here we see that the diacritic *u* had two allographs: ˒ ˓ . Technically the subscript *ya* followed the consonant and only after having written the consonantal cluster was the vowel sign added. Intended was, however to write the aksara without the subscript *ya* and then to modify it by adding a *ya*. The essence of this technique modelled on Uighur was well observed by Gabain (1964 7) but misunderstood by Maue (s.a. CXII) where he writes: "Dort [in Uighur] folgt der Palatalitätsmarker Y dem Vokal in der Brāhmi geht er voraus." In the cases with all vowels except *u* this is even not necessarily true technically. This device had no models in Tocharian or in Tumshuqese, since both had the sequence *Cy*, such as *ky, ny, py*, etc., where, as far as I could observe the subscript *ya* had no influence on the vowel of the aksara. On the other hand Turkic did not have the cluster *Cy* and thus the subscript *ya* was used only as a modifier.

Brāhmi Turkic had two options to render an E-sound. The first was to use the independent sign <e> 𐽀 or the diacritic <e> 𐽉 if it occurred in postconsonantal position. The other way of expressing an E-sound was to use the independent letter for *a* and modify it with a subscipt *ya* 𐽂𐽃. Both procedures occur in all of the published and controllable texts. The two procedures could have been contaminated. In some of the texts we find independent *e* with a subscript *ya* 𐽀𐽃 . This was transcribed by Gabain with a double dotted *ë* only for the sake of keeping the two graphical units apart. Thus 𐽀 is *e*, 𐽀𐽃 is graphically *eya* in the transciption of Gabain *ë*, and 𐽂𐽃 is *aya* in the transcription *ä*. The combination of *e* and subscript *ya* is practically not used after consonants. The hapax *kër* written as *khyer* (D 15) is most probably a loanword or a scribal error. We find it once in *tësäsär* written with *tye* (K 1) and once in *yigirmë* written with *mye* (H 8) and perhaps in one or two other cases where I could

not check the original. Therefore, until the opposite is proved I consider these *ey* writings as orthographically redundant whithout any phonetic value. We remain confronted with the question whether the writings *e* and *ä* render a phonological opposition and, if yes, which values have to be attached to the two types of writing.

The notation is not always consistenent. We find the same words with two different writings *elt-* (D:10) and *ält-* (F:9) 'to carry' or *äšit-* 'to hear' (H:10) but *ešet-* (A passim, D:19) and the passive *eštil-* (H:10). Nevertheless in the overwhelming majority of cases we find *ä* where we expect an open *ä* and *e* or *ë* is written where we expect a closed *e*. Thus the verb *är-* 'to be' occurs in the material in about 90 cases and only once is it written with *e* (*e-rsye-r* M:12), twice with *ey-* (*ërrip* E:19, O:2, the first reading is uncertain). In all the remaining places we have *ä* or *ǟ*. On the other hand the word *bēš* 'five' occurs in three texts (H, L, K) 14 times and always with *e*, *te-* 'to say' occurs more than 30 times always with *e* (once *ey*: K:1) while *tängri* 31 times only with *ä* or *ǟ*. In non first syllables *e* stands in most cases after *e* of the first syllable also in cases where we would expect otherwise *i* as in *eleg* 'king' (C:1 cf. H:2 *elig*) or the above *ešet-* 'to hear' instead of *äšit-*. Thus we can conclude that the scribes considered *ä* a counterpart of *a* and not identical with *e*. This *ä*, as has been long ago supposed, denoted the open and *e* the closed E sound. The hyperurban writing of *e* with a subscript *ya* has no significance in the reconstruction of the Old Turkic vowel system.

For rendering the vowel *ö* in a few texts (consequently only in I, once in H:4) in initial and only in initial position we find the independent letter *o* + *ya* + diacritic *o*, while in all other texts and cases we find only the independent *o* and the subscript *ya* i.e. we have in the first case  <oyo> in the second  <oya>. Gabain rendered this graphic pecularity with *ö* but she remarked: "Um diese Schreibung in der Umschrift von der sonst üblichen Schreibung *oya* zu unterscheiden, wählte ich dafür *ō*; ich glaube aber nicht, dass phonetisch etwas anderes als *oya*=*ö* der anderen Handschriften gemeint ist" (1954 60, already so in 1950 12&). She is undoubtedly right, but the transcription is misleading, so e.g. even Clauson (1972) marks this *ō* as a long vowel.

While the <oyo> writing for *ö* seems to be an exception in front of the <oya> writing, the case is different with *ü*. Here the writing <uya> is the exception if it existed at all and the normal writing is <uyu>. Since the graphical difference between initial <uya->  and <uyu->  was minimal a thorough rechecking of the originals is needed. It is, however, very likely that the picture will not esentially change. This can be inferred from the fact that in postconsonantal position we find the same if it is written in the so called liason writing (see below), and the inital *ü* is only graphically in initial position. I think that the *uyu* writing has no phonetical reasons. It may perhaps be due to

the slow change of Tocharian Brāhmī towards using as the independent initial vowel signs only *a* and *u* (and to denote all other vowels with diacritics on one of them), thus the letter *u* became a general sign for all labials and it had to be specified as front with the subscript *ya* and as closed with the subscript *u*.

The Brāhmī had one type of representation for the I-sounds, the independent <i> and the diacritic <i>. This means that it had no original letters to differentiate between the front and the back *i*-s. It would have been theoretically possible to use the *i*-signs to denote the back *i* and to denote the front *i* with one of the *i* signs and to add a subscript *ya*, but since in Tocharian (and also in Northern Saka) the *i* sound was a front vowel this procedure would have been unusual. This can be seen in such cases as *'ilintūrtāčilār* 'those who let themselves to be caught' which was written *i-lim-tyu-rtyā-či-lya-r* (A:14). In this case none of the *i*-sounds have been marked by a subscript *ya* while all other vowels have. There was only one consonant which got a subscript *ya* if it was followed by a front *i* and this was in some Mss the consonant *l* (see Maue s.a. LIV). This may have been due to the special phonetic peculiarities of the *l* sound. The almost consistent lack of the *iy(a)* notion, or as it can be otherwise expressed: the absence of the marking of the opposition of the front and the back *i* may have had phonetical, or even phonological reasons. One is entitled to suppose that the opposition of the front and the back *i* sounds was in the process of disappearing. In most phonetical positions the phonetic difference between the two *i*-sounds became minimal, or it disappeared. Thus the *i* sounds in the words *tidiγ* 'hindrance' (*ti-ti-g* G:18, *ti-dhi-g* O:1 **tīdīγ*) and *titig* 'mud' (*ti-thi-ḳ* I:7 **titig*) may have been slightly different, but were not felt as such and therefore not noted. The opposition of the front and back *i* sounds did not suddenly and at the same time disappear in all positions. There remained a restricted number of positions where the opposition had not yet vanished. In these cases the monks had more possibilities to render the back vocalic *ī*. The first was the type of writing with the double dots or with the special letters for *l̤, r̤*. As we have seen A. von Gabain suggested that some graphotactical devices would have rendered a kind of *schwa* (Murmelvokal). Such devices would be the new signs *r̤* and *l̤* which she always transcribes as *rə* and *lə* and the double dot on the non final consonants, such as e.g. *s̈* as *s* , *p̈h-* as *phə* etc. This was not accepted by Hitch (1981), who claimed that these kinds of writings denoted back vocalic *ī*, i.e we have to read *rī, lī, sī, phī* etc. If we exclude foreign words such as *čixsapad* (with *č-* three times, as *ča-* twice and with *ǰi* once) or *simnu* (*s̈* twice, *s̈* once) this kind of vowel notion occurs only after *r* and *l* or before velar stops, always in back vocalic words in most cases in the final syllable in positions where we would expect a back vocalic *ī*. Thus it is very likely that Hitch is right insofar that this kind of writing denoted a preserved back vocalic *ī* under special phonetical conditions. Against the supposition that we have to

do here with a *schwa* one also can argue with the fact that in most cases it occurs in the last syllable, where under stress the existence of a secondary *schwa* is very unlikely. But the monks also had other devices to render the back vocalic *ï*. In some texts they wrote a diacritic *e* in place of Turkic *ï* as in Text I *saweqla-* for *sawïqla-*, or used only the inherent *a*, such as e.g. *ya gla g* for *yaγlïγ* in the same text. The preservation of the back vocalic *ï* was due to the specific phonetical environment. The special kind of *l* and the influence of the gutturals could have been cumulative, and thus special transcriptions appeared. Both Maue and Hitch remarked on the special writing of the *-liγ* suffix (with *l̲*, *l̲r̲* or the former with the double dots above) which was due to this special phonetic environment.

Brāhmi was able to render long vowels by a diacritic sign. If one collects the notations of vowel length in the Brāhmi Turkic texts it becomes clear that the Turkic etymological length was not marked. A word such as *beš* 'five' occurs 14 times and always with short vowel though it had an etymological length, or *baš* 'head' is twice written with short and six times with long *ā* though it had an etymologically short vowel. On the other hand we find examples such as *at* 'name' seven times with long and once with short *a*, *ay* 'moon' twice with short and nine times with long *ā*, or *taš* 'stone' twice with short and 22 times with long *ā* (all in text L), where all three words have etymologically long vowels. It is very instructive that the final vowel of *belgü* 'sign' (3 times) is never long while the final vowel of *bilgä* 'wise' is long in all the 15 occurrences. One has the impression that the etymological length was in the process of disappearing. It has been preserved to a higher degree in open syllables and with the open vowels *a*, *ä*, to a lesser degree in closed syllables and with the closed vowels, more frequently before the laterals *-r*, *-l* and before the gutturals. Long *u* never occurs after *q-*. The distribution is the following:

| | V | CV | |
|---|---|---|---|
| ā | + | + | |
| ǟ | +* | + | *only in lieson writing |
| ē | - | - | |
| ī | +* | +** | *only in text A<br>**rarely in the texts I, P |
| ō | - | - | |
| ȫ | - | - | |

| | V | CV | |
|---|---|---|---|
| ū | - | + | |
| ū | - | +* | *only in text I |

This distribution shows that the rechecking of the originals is badly needed. But basically the Brāhmi long vowels do not reflect etymological lengths. The less so since we find them not only in cases where the vowel had no etymologycal length but also in cases where in polysyllabic words more vowels are written with a length marker. The length markers were used to render prosodic length.

The representation of the Turkic consonants

The Brāhmi alphabet was very rich in consonantal letters. The stop and affricate series had letters for voiceless unaspirated, voiceless aspirated, voiced unaspirated and voiced aspirated. E.g. in the labial series *p, ph, b, bh*. To these the new signs were added, in most cases to render the final consonants which had no inherent vowel, i.e. postconsonantal vowel. Beside the use of the original Sanskrit letters and the new ones there was a third possibility to denote a consonant. This device was used in Southern Khotanese Brāhmi, but some schools in the North also experimented with it. The original *t* became *d* in Khotanese thus the original letter for *t* denoted *d* as well. For a secondary or foreign *t* they used a duplicated *t*, (<tt>) that is two *t*s written beneath each other. It was generally, however, considered that the consonantal letters were used in a more or less random way and "...it would be rash to draw any inferences from them regarding the consonantal structure except perhaps in regard to the pronunciation of the post-palatal g" - as Clauson suggested (1962 96).

The opinion of Clauson was more or less accepted until recent years. It was only Johanson and Maue who tried to find some regularities in the seemingly capricious writing. (But see Maue s.a. LIX "Ich möchte eher mit einer spielerischen Ausnützung des graphischen Überhangs des Skt.-Alphabets rechnen" in connection with the labials). That the use of the consonantal letters was not random can be shown by the example of the labials. Though Brāhmi had the four letters *p, ph, b, bh* to which the "new sign" *pā* (in fact an "old" *ba*) was added which was used with a superposed dot and the *virāma* stroke to the left as final, the Turkic sound *p* was always written with a letter *p* (once we find *ph* in C:15 *qapïγ*). In genuine Turkic words an intervocalic *-b-* is always written with *-w-* (Maue:*v*). Exceptions such as *tvišqam* for *tabišqan* only show that the *b* was already spirant.

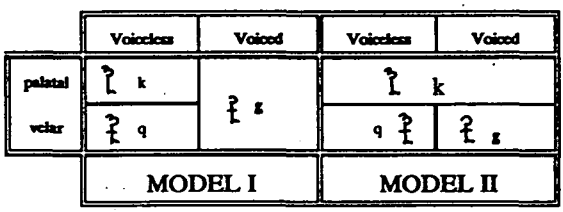
We find the general picture that Turkic consonantal phonemes were written in most cases with one consonantal letter. There exist some manuscripts which show a different usage and in some cases we find variants. The divergencies do, however, help us to reconstruct the phonetical pecularities of the Old Turkic phonemes. A good example for the situation is the rendering of the initial *b-*. As a rule it is transcribed with *p-*. In the manuscripts F, H and I we find as a rule *bh-*. The manuscript K has in most cases *ph-*. The writing with *b-* is relatively rare, but occurs in the most frequent words such as *bilig, bilge, baš, beš*. The word *burxan* is always written with *b-*, and in the two manuscripts where we find also *p-* (A, E) it occurs together with forms written with *b-*. This indicates that it was not only the orthographical usage of Brāhmi which influenced the writing, but also the originals which they copied.

Maue investigated the rendering of the gutturals in a paper published in 1984 (see also Maue s.a.). One can so far accept his results that in a certain group of manuscripts (I, K, M and unpublished Mss) the grapheme <k> was used in the same way as Keph was used in Uighur script, i.e. for fronted /k/ and /g/. On the other hand the velar *q* is in text I written with <q> but in the others with <hk>, <hq> and only in special cases with <hkh>. This points to a very deep guttural pronunciation where the [k] was pronunced at the place were the guttural /x/ or even the laringeal /h/ may have been pronounced. But I do not see any forceful argument to consider these transcriptions as evidence for the spirantisation of the velar <k>. It is very difficult to judge why we find in text D for palatal /k/ in most cases <kh> (here we do not have the aksara transcription), it can, however, not be a mere chance that the word *kir* 'dirt, filth' is written with <kh> even in such Mss (A, C, E) which otherwise write only <k->. The transcription *khgi rli g lya ri nci p* (*kirliglär inčip* E:48) undoubtedly points to the Uighur original, which was written as *kkirliglär*.

In some cases we find <h> for velar *g*. Maue correctly remarked that this type of transciption occurs in most cases before *l*, before sibilants and in intervocalic position. It is, however, disturbing and not mentioned by Maue, that if in intervocalic position, these *h*-writings occur in many cases before *i*. If the *h*-writing were a simple rendering of the spirant character of the *g*, one would expect it first between two identical vowels as in *oγol* (for *oγul*) which is in fact written twice as *ohol* in text E. From the fact that the *h*-writings occur in the overwhelming majority of cases in back vocalic words it can be concluded that the spirantisation began in back vocalic words, in special phonetical environments and the letter <h> denoted perhaps a slightly palatalized /x'/-.

The most problematical is the letter transcribed by Gabain with *Fraktur g* (and sometimes by others with gamma). It surely denoted a voiced sound. In some texts (I, K, M and unpublished Mss) it stands only for the voiced velar sound, in others also for the voiced palatal *g*.

The distribution of the letters denoting gutturals according to Maue (s.a. LX):




This does not support the assumption that the letter *g* would originally have denoted a fricative guttural. This is even not forceful in the texts where this letter was used only in back vocalic words (Maue's Model II). But we cannot wholly exclude that in some cases the *g*-writing also denoted a fricative. The letter *χ* was not used in virāma position, that is in absolut final position and thus only the letter *g* could have been used. But if the writing offered a possibility to avoid the virāma writing of *h*, it was written as in *quwraγīγ* which was written *qu wra hi g(u)* in C:5. It seems to me that *g* was the letter *gā* and therefore I transcribe it consistently as underlined *g*.

Johanson (1979) and Maue (1983, s.a.) have dealt with the graphic representation of the dental phonemes. Johanson tried to single out the Mss which used the new grapheme 𐬓 which was transcribed by Gabain as <δ> and called in the literature "delta". Maue seems to be right in stating that the writing usage of doubling *t* as *tt* and the writing *nd* do not show a direct Khotanese influence in Turkic Brāhmi, they reflect Tocharian influence which itself used this in Sanskrit words. Important is the remark of Maue that the new *t* sign, used in most cases in final position, is graphically the old *dh* sign and thus the transliteration can be confused. Though many of the results of Maue seem to be correct, the basic conclusion of Johanson, namely that the Turkic language reflected by the Brāhmi transcription did not have a spirant *d*, seems to remain valid.The representations are according to Maue:

[t] is written by <t>,<tt>, <th> (in a few cases <d>)

[d] is written by <dh>, <d> (only in *nd*), "delta", <t>.

The key to the question is Tocharian where no voiced dental stop existed. The Indo-European voiced dentals *d*, *dh* became *t* (if not affricates, see Windekens 1976 I 79-84, Thomas 1985 48-51; also *th* became *t*) and therefore the letters <th>, <d>, <dh> were not used at all, except in transcribing Sanskrit (or other foreign) words. The letter <t> denoted the Tocharian phoneme /t/ which had also voiced allophones therefore it could have been used to denote Turkic *d*. Since "delta" was used in some Mss for the dental conso-

nant which never became a spirant (or later *y*) such as e.g. the suffix *-di*, one has the impression that if "delta" denoted any special phonetic variant, it was perhaps a (more) palatalized variant. In almost half of the cases quoted by Maue (1983 Table I) "delta" occurs before or after *i*. It is further interesting that the second greatest group has an *r* before "delta" and in the third group, though "delta" does not occur after *r*, there is an *r* in the word. Whatever could have been the phonetic value of the "delta", it functions as *dā*, and its consistent transcription will be underlined *ḏ*.

There exist a few Mss where the letter *z* does not occur (N, D, P) and Turkic *z* is written with *s*, in a second group the Turkic *z* is rendered with the new letter *z*, but this letter is used only as a vowelless final (B, E, M with a very few exceptions to be controlled), and in all remaining Mss the new sign is used in all positions. This is the more interesting, since this could be how the new letters spread in the texts. /z/ is never written as <ys> which confirms that there was no Khotanese influence in this writing system.

The letters for the cerebral sounds were rarely used. An exception is *ṇ*. This is surely connected with the fact that Turkic *n* is in many cases written with *ñ*. Before *i* and *č* this is the rule in many Mss but there are many cases where there could not have occurred any kind of palatalization, as in *ñom* 'Dharma'. Before *d* the normal *n* is usual (and a few cases, such as L 23a *moñdaγ* seem to be exceptions).

Summing up the conclusions for Turkic, written in Northern Brāhmī script we can say that for the reconstruction of vocalism it offers great help. It makes it possible to differentiate between the closed and open E-sounds, it had different means to denote the back vocalic *ī*, it helps in the reconstruction of the vocalism of the non first syllables, where it could differentiate between *o* and *u* an *ö* and *ü*, resp. It reflects such changes as the beginning of the neutralisation of the opposition of the back and front *i*, it does not render (at least consistently) etymological length but denoted prosodical length etc. In the case of consonants a gradual spirantization can be observed in the case of the voiced stops. The **b* was spirantized in non initial position, there are signs of the beginning of the spirantization of the voiced velar guttural stop in back vocalic words and no signs for denoting the spirantization of the voiced dental.

*Khotanese Saka and Khotanese Brāhmī*

Khotan is a settlement in Turkestan that was conquered by the Saka tribes perhaps in the first centuries B.C. The Iranian population came into close contact with the Buddhist culture spread by monks from India in the third century A.D., but sporadic contacts can be traced with both India and China in earlier times as well. The later legends that Khotan may have been founded in the times of the son of Ashoka (3rd century B.C.) have not yet been proved.

These legends were recounted already in the biography of Hsüan Tsang, in the Turkic translation of which we read (Sa 2-8, ed. Tugusheva 1980,17):

| | |
|---|---|
| *öngrä bo Odon uluš kiši-* | Earlier this Khotan country |
| *siz quruγ ärkän Vaiśirvani-i* | was peopleless and empty. V. |
| *mqaraš munta ornaγ tutmiš* | maharaja occupied this place |
| *ärdi : qayu ödün* | in the time |
| *Ašoki-i elig qan* | when the Emperor Asboka's |
| *oγli Kunali tegin D(a)kšaš* | son Kunali tegin was ruler |
| *el balïq-ta balïq bägi ärkän...* | of the country of Taksasila... |

In later centuries Khotan became an important centre where Buddhism flourished and which was visited by monks from China and used as a refuge by monks from North India. We do not know when contacts between the Khotanese and the Turks began, but they became closer when the Uighurs became Buddhists and moved to the South after 840. The Moslem Turks from Kashghar conquered Khotan shortly after 1000, and put an end to this interesting culture. (See most recently Mayer 1990.)

The Khotanese language or Khotanese Saka belongs to the eastern group of the Iranian languages. Hercenberg (Gercenberg 1981) and others distinguish two phases in the history of Khotanese Saka, Early Khotanese and Late Khotanese. Early Khotanese served for a time as a literary language and existed together with the "vulgar" Late Khotanese. We know of several Saka dialects. Northern Saka, i.e. Tumshuqese, and Känjäk in and around Kashghar has been mentioned above (p. 65-). Some words are known from the so-called Indo-Saka, the language of the Saka tribes that conquered North India. A special dialect is the Kroraina-dialect once spoken to the east of Khotan. There exist Kroraina words in Kroraina-Prakrit texts.

Khotanese was written in Brāhmī script. In the Turkic translation of the biography of Hsüen Tsang we read about the writing of the people in Khotan: 4b:7 *iki bitigi iraq-ta [:8] änätkäk üž-ikin ayïrlayur-lar [:9] yaγuq-ta öz üž-ik-lärin [:10] tutar : öz üž-ik-läri ymä [:11] änätkäk-čäsig ök ärip* (Tugusheva 1980 16). "They have two writings. In the distance they respect the Indian letters, at home they keep their own letters, but their own letters are very similar to those Indian ones". According to Sander (1968 column v in her work) the so called Southern Turkestan Brāhmi was used in Khotan. The script was not homogeneous, there existed a rather official and a more cursive type. The phonetical evaluation of the script is hampered by several facts. Most of the texts are Buddhist texts and here the writing of the Sanskrit original and the actual pronunciation of the word cannot always be separated. The early Khotanese orthography was in use for a long time after the Late Khotanese pronunciation had become standard.

As to the Khotanese language the following peculiarities may be of importance for the evaluation of the Turkic material: The stops /b/, /g/ and

/d/ became fricatives but the letters, <b>, <g> and <d> were used to render them. However in words of foreign origin and in some allophonic positions stops occurred and this caused inconsistencies. The situation was even more obscured when in Late Khotanese the fricatives in intervocalic position disappeared *via* a guttural fricative but they were often still written. Since earlier voiceless stops, such as /t/ became voiced, the letter <t> was used to denote /d/, to render an unvoiced /t/ a double letter <tt> was introduced. On the other hand <gg> denoted the stop /g/. For the transcription of Turkic it is important that /z/ was written with the combination <ys>. The cerebral <d> denoted a retroflex /d/, as the latest investigations of Emmerick show (personal communication). It occurs for Turkic /l/ after or before /i/. With the exception of a few old texts the older voiced sibilants became devoiced and therefore they wrote the voiceless sibilants with double <ss> and <śś> and the simple letters <s> and <ś> were used for foreign voiced sibilants. In Late Khotanese these changed and simple sibilant letters were also used to render unvoiced sibilants.

The letter <h> denoted a laryngeal spirant. If a colon like diacritical sign was added, it served to render the Turkic *g*, both velar and palatal. In most cases it also got a hook, which is transcribed with a half circle below the letter, *ḫ*:. (On the details see later). The anusvāra is either a dot or a hook above the aksara, the latter is very similar to a short *i*.

The system of writing the vowels was essentially the same as in Tocharian or Tumshuqese. The independent vowel signs were reduced to *a* and *u* (with perhaps rare exceptions for *ai*). Here we also find a vowel which is transcribed by *ä*. This *ä* was a reduced vowel, and its graphic rendering cannot always be clearly distinguished from *i*. Seemingly in Khotanese only open syllables existed. If they had to transcribe a final consonant, in most of the cases they added this *ä*, which may already have in fact disappeared in Late Khotanese in unstressed word finals. If *ä* was not added, an *a* was automatically read. In Late Khotanese the long <ā> had the phonetic value of *o*; thus Turkic /o/ was written with long <ā>.

The reconstruction of the Khotanese phonological system is a field where only a few scholars are working. Two texts have evoked a broader interest. The first is the so-called Staël-Holstein Scroll (see below) and the other a Chinese text written in Khotanese Brāhmi. This latter text, earlier studied by F. W. Thomas, Bailey and Csongor is now going to be edited by Emmerick and Pulleyblank. A Tibetan text in Khotanese Brāhmi twice published by Bailey has been reanalyzed by me (presented in Sopron 1987, yet unpublished).

A good introduction to Khotanese textology was written by Bailey in the fourth volume of his Khotanese Texts (Bailey 1979 reprint). A more recent

survey can be found in Bailey (1982). For a phonology and the orthographical usage the two papers of Emmerick (1979: vocalism, 1981: consonantism) and the chapter written by Hercenberg (Gercenberg 1981) in the "Soviet Handbook of Iranian linguistics" can be recommended (see also Gercenberg 1965). An excellent bibliographical survey has been published by Emmerick (1979a) with later additions (1983).

In the two tables No. V, VI we give the Khotanese Brāhmī rendering of the vowels and the consonants.

*Turkic in Khotanese Brāhmī*

The Khotanese documents which contain Turkic words can be divided into three groups (see Hovdhaugen 1971):

1. Official documents with Turkic proper names (personal names, geographical names, ethnonyms, titles etc.).

2. A list of Turkic persons, titles and peoples which can be found in the Staël-Holstein Scroll dated to the 9th March 925.

3. The Turkic word list registered under P(elliot) 2892 (new number). This word list contains technical terminology on archery, horse-breeding and parts of the body. Following or beneath the Turkic words in some cases there occur Khotanese translations. Unfortunately the words are so specialized that many of the Khotanese words occur only in this word list.

For a detailed enumeration of the material and bibliography see Hovdhaugen (1971), Clauson (1973) and Hamilton (1977).

Most of the texts were first published by Bailey. The Word list first appeared in 1943/47, then it was re-edited in *Khotanese Texts vol. III* (1945, second ed. 1969, reprint 1980). The facsimile was published in 1973 with some amendments to the earlier publication. The entire material was collected and evaluated from the Turkological point of view by Hovdhaugen (1971). This has remained the best publication, though in some details it is now outdated and it is difficult to use for somebody who is not acquainted with the Khotanese background. In 1973 Clauson published his paper on the Word list, in which he suggested a series of corrections to the readings of Hovdhaugen. Unfortunately Clauson's paper is full of printing errors and it is not always clear how Clauson read the word because he refers only to the main entry of his Dictionary (1972), where the words of the List have not been included. A new edition of the Word list is being prepared by me (presented in June 1989, Oslo) in collaboration with R. E. Emmerick.

The Staël-Holstein Scroll has a larger bibliography. It was first published by F.W. Thomas and Sten Konow in 1929. The most important papers later devoted to the Scroll are by Clauson (1931), Konow (1948), Bailey (1951), Pulleyblank (1954) and Hamilton (1958 with further bibliography on p. 116). The Khotanese text itself was published for the last time by Bailey (1979 in Khotanese Texts IV, reprint).

*The representation of the Turkic vowels*

In the first syllable Turkic /a/ is written with independent or inherent <a>. The very few exceptions, all written with <ā>, are problematical, thus 45. *yaŋaq*, (the numeration is new and is identical with that found in the new edition, see here the Appendix V) until now read as *yīŋaq*, but it has to be the same word as Kāshgari *ya aq*, 56. *tamγaq* until now read *tīmγaq*, 54. *art*, 60. *qari*, 62. *aya*. In first syllables Turkic *e/ä* is always written with <e> the only exception is 84. *kärgök* in non first syllables it is always rendered by <a>. Turkic /o/ is written with long <ā>, the two exceptions are 33. *tomo*, where the first *o* is written with <o>, the only case of the use of the diacritic <o>, and 70. *boγun* written with <au>. Turkic /ö/ is in most cases written with <au>, those written with long <ā> are 21. *kömüldürüγ*, the very problematical 32. *tōli*, further 65. *köküz*, 91. *ö öč*. Turkic /u/ is written with long <ū>, the two exceptions are 10. *tutasi* which is obscure, and 35. *tulun*, /ü/ is written with long <ū> the only real exception is 2. *tüpi*, which is written with short *u*, because 53. *ün* is very problematical. The front vocalic /i/ is without exception written with long <i>, while the back vocalic /ï/ is written either with short <i> as in 6. *yïγïrγoq*, 26. *qïrïγ*, 66. *sa ïrsaγ* or with *ai* as in 14. *yïγ* or 95. *bīqïn*.

| Turkic vowel | Main representation | | Occasionally | |
|---|---|---|---|---|
| | first syllable | further | first syllable | further |
| a | a | a | ā | |
| e/ä | e | a | ā | ā |
| o | ā | ā | o, au | |
| ö | au | au | ā | au |
| u | ū | ū | ū | u |
| ü | u | ū | ū | u |
| i | i | i | | i, ai |
| ï | i | i, ai | ai | |

Some regularities can be observed. The opposition of long and short vowels is not marked. This is clear from the fact that front /i/ and back /ï/ are graphically distinguished as short and long <i>. There is no distinction between the open and closed *e*-sounds, <e> is always written in the first and <a> is always written in the non-first syllables. In non-first syllables we find open and closed labials as well, which corroborates the reality of the picture we get from the Uighur texts written in Northern Brāhmi.

*The representation of the Turkic consonants*

The consonant /b/ occurs only in word initial position and is always written <b>. Turkic /p/ which occurs only in non initial positions is always written <p>. Turkic /t/ is consistently written <tt>. Turkic /k/ whether in front or in back vocalic words is always written with <k>. A few exceptions turned out to be wrong identifications. So 17. *iliγ* is not *ilik* 'marrow', but *ilig* 'attachment', 20. *emziγ* is not *emzik* 'nipple', but *emzüg* 'tip of a saddle-tree on front or back' (Kāshgari's entry was read hitherto as *ümzük*). In a few cases the <ḫ:> writing of a final *-k/q* seems to render a secondary development in the underlying Turkic dialect, as in 21. *kömüldrüγ*, 66. *siŋarsaγ*, 86. *baγirsoγ*, cf. 46. *qasiγ*, where Kāshgari has also *-γ* but in most languages and dialects we find a *-q*. The /g/ in back vocalic words is always written <ḫ:> even in phonetic positions where one may expect a stop allophone as in 5. *qapγaq*. In front vocalic words the /g/ is written in most cases <ḫ:> with the following exceptions of where we find <g>: 11. *yügün*, 13. *tizgin*, 88. *yürgāk*(?). This raises the problem of whether the <ḫ:> writings in fact reflect a fricative as supposed by Hovdhaugen (op.cit. 175)? The hook beneath the <ḫ:> is missing in 17., 20., 21., 25., 44., 48., 64., 66.and 97. for reasons that are not quite clear. It is written four times beneath an initial <a> on which diacritical signs were placed (twice <e> 20., 59. and twice <au> =*ö* 78., 87.), and once it is used in the dubious 33. *tomo* beneath *t* and *m*. One has the impression that these hooks did not have a special phonetic value, but denoted only that the complex beneath which it was written had some peculiarities not expressed by the writing. This may be its function beneath the initial <a>. In Late Khotanese these hooks were written under letters denoting syllables which had lost intervocalic consonants. It is difficult to suppose that the <ḫ:> writings denoted only the voiced or sonorant character of the /g/, since the letter <g> was used as well. Therefore we have to suppose that the <ḫ:> writings denoted a fricative, and, if that, a voiced one. This is seemingly in contradiction to the material reflected by Kāshgari, where the /g/ in front vocalic words is denoted by Kāf and Kāshgari writes that [g] is denoted by the "thin" (*rikka*) Kāf, i.e. with a Kāf which

has to be pronounced voiced (cf. DK I 54). In some words (3., 49.) we find a secondary *ŋ* in place of *γ* (on the "unstable *ŋ*" see Hamilton 1977).

The Turkic /t/ is written without exception as <tt> and /d/ as <d>. Since in Khotanese <t> denoted the stop [d], and the old /d/ became a fricative denoted by the letter <d>, we have to suppose that the <d > denoted a fricative dental. Two words seem to be problematical, 71. *yoda* 'thigh' and 81. *bidi* 'face', 71. has in all sources and dialects *-t-*, 81. is problematical only if it would be *bet* 'face' which is unlikely.

Interesting are the words with palatalized *ñ*, 72. *baqañoq*, 85. *sarqañaq* 'third stomach', because exactly the same words have in Kāshgari's Divān a curious notation which hints at the same feature. The fact that the /l/ is rendered by <d> before (9., 17., 24., 32., 34.) and after *i* (17., 24., 51., 61., 90.) and only there, points to the fact that the /l/ had a palatalized allophone.

| TURKIC CONSONANTS | KHOTANESE BRAHMI LETTERS |
|---|---|
| /p/ | <p> |
| /b/ | <b> |
| /t/ | <tt> |
| /d/ [ð] | <d> |
| /k/ | <k> |
| /g/ [γ] | <ḫ: ??; g> |
| /m/ | <m> |
| /n/ | <n> |
| /ŋ/ | <ṅ> |
| /ñ/ | <ñ> |
| /s/ | <s> |
| /z/ | <ys, s> |
| /š/ | <ś> |
| /č/ | <c> |
| /y/ | <y> |
| /r/ | <r; rr> |
| /l/ | <l> |
| [l'] | <d> |

# THE SOURCES IN TIBETAN SCRIPT

*Historical background*

The Tibetan Empire was founded after a long period of regional consolidation in the 7th century by the kings of the Yarlung valley. They united several petty local chieftains and Tibet entered world history. They occupied great parts of Turkestan, West China and the Himalayas. They took part in the great battle of the Talas valley on the side of the Arabs against the Chinese, and if only for a short time, they occupied the Chinese capital in 763. They concluded peace treaties with the Chinese from which the most famous is that of 821/822, because the Tibetan and Chinese texts were carved in rock and have remained to our days. The Tibetans had contacts with the Turks since the the 7th century. We read in the Kül tegin inscription that the Khans of the first Turkic Khanate lead their troops "almost until Tibet" and that at the mourning ceremonies of Kül tegin the special envoy of the Tibetan king the *Bölön* (Tibetan *blon* 'Minister') was present. Another Old Turkic inscription (at the river Abakan, Altyn Köl II) mentions that the deceased "because of his virtues has been sent to the Khan of the Tibetans". The close contacts with the Turks is reflected in several early Tibetan sources, among them the Tibetan Annals and the Old Tibetan Chronicle. The Turks (*Dru-gu*), the Qarluqs (*Gar-logs*) and the Türgesh (*Dur-gyis*) are frequently mentioned in these sources. After the defeat of the Uighurs by the Kirghiz (Tib. *Gir-kis*), the Uighurs of Kocho and Kansu came in much closer contact with the Tibetans. Along with the military and diplomatic contacts Budhism also linked the two peoples. In 842 there was an Anti-Buddhist uprising in Tibet the result of which was the temporary collapse of the Tibetan Empire. But in the border areas the Tibetan Buddhist colonies lived further. This was also the case in Tunhuang, which was reoccupied by the Chinese from the Tibetans in 848, but where the Tibetan community lived and worked also later and had close contacs with the Uighurs.

*Old Tibetan literature as a source of the Turkic linguistic history*

Tibetan literature is known from the middle of the 7th century. Important are the great inscriptions in Central Tibet from which the earliest is the Zhol-Inscription written in 764 and which have been preserved until our days. Though the Annals and the Royal Chronicle had Chinese models, they were not only in their content but also much in an independent form. In the military governments of East Turkestan official administrative documents, letters and other texts were written in Tibetan, and since contacts with the Turks were very close many Turkic names, titles, ethnical names etc. can be found in the docu-

ments. With the exception of the inscritpions and some excavated documents the Tibetan texts found are mostly copies or excerpts, abstracts etc., but even they are no later than the 11th century.

Buddhist literature also began in the 7th century. Most of the texts are translations from Sanskrit or Chinese, some from Khotanese. The great activity of the translators made it necessary to issue a Royal Edict in 815 which fixed the "new rules" for orthography and translation. Also a terminological dictionary the Mahāvyutpatti was compiled to help the translators. The literary language before the Edict is called the "Old language" (*skad rñiṅ*) while the language which already followed the new norms was called the "New language" (*skad gsar*). Some colophones of these Buddhist texts also contain Turkic personal, geographical or ethnic names. A few Old Turkic texts have been translated from Turkic into Tibetan, or have been thoroughly excerpted as the famous Uighur intelligence report (Pelliot tib. 1283).

However important the various Turkic names, glosses may be, their value is only complementary to the Turkic texts written in Tibetan. We know of a few fragments from Turfan, some unedited, and a complete text from Tun-huang, a cathecism (Pelliot tib. 1292). The three types of material, the glosses in Tibetan original texts, the Turkic elements in texts translated from Turkic and Turkic texts written in Turkic differ insofar as the latter try to follow a more strict system of transcription.

*The Tibetan alphabet and Tibetan orthography*

According to the mediaeval tradition of the Tibetans they had no writing before the King Srong-btsan sgam-po.This founder of the Tibetan Empire, who died in 649, sent his Minister Thonmi Sambhota to India, and after his return the Minister took as a model certain Indian scripts and created the Tibetan alphabet. The latest researches have revealed that this tradition is only reliable insofar as Tibetan script has its origin in the Central Asian branch of the Brāhmi script. We also know that the formation of the Tibetan script and orthographical system went through several stages.

The Tibetan language typologically belongs to the so called isolating monosyllabic languages. The basic unit of the language is the monosyllabic word, which can acquire prefixes, infixes and suffixes. All di- or polysyllabic words are, or were originally composed of monosyllables. The adaptation of the Indic system of writing had to match this structure. The writing is "syllabic" in that all consonantal letters have the inherent *a* if not otherwise indicated.

In Table VII we give the Tibetan alphabet in the usual Tibetan sequence which follows the Indian prototype. In the Latin transcription those letters are marked by a + sign which are new and do not originate directly from the Indic prototype, which is important for their early phonetical value. <ga> is a turned

<kha>, <ka> is a modified <ga>, <da> is a modified <ta>. <da> and <ta> are in manuscripts not always clearly distinguishable. The letter <ba> denoted originally a [v] and it stood in some early alphabets in place of <wa>, but later it has been reinterpreted. The letter <wa> was originally an *a-churi* and beneath a <ba>, this became the letter for the spirant, and <ba> was placed back in the series of stops. The picture was later complicated by the fact that the Tibetan stop /b/ became in certain positions itself [v] and thus the letter <b> denoted both the stop and the spirant. In postconsonantal position [v] is written with the subscript <ba> which had a triangular form and was called *wa-zur*. The letters <tsa>, <tsha> and <dza> have been formed from <ca>, <cha>, and <ja> respectively with the small diacritic stroke. The letter <zha> was formed from <sha>, <za> from <ja>. As the latest researches have shown (Róna-Tas 1985) the *a-churi* or the "small a" goes back to the Khotanese Brāhmi syllabic cluster <gā>.

Though the Tibetan writing developed through several stages in the time which is important for the Turkic material it was almost fully developed. Later the Tibetans added several new letters to the alphabet to render Sanskrit sounds which had no equivalent in the original alphabet. There were also some graphotactic devices which were used only for transcribing Sanskrit, but with a few exception none of them was necessary for writing Turkic. In the only case, where it would have been important, the denotion of vowel length with subscribed *a-churi* was not used.

The letters given in Table VII are the so-called *dbu-can* letters or letters "with head". They got their name because most of the letters have an upper horizontal stroke. Already in Old Tibetan times a so called "headless script" (*dbu-med*) was in use, a more informal way of writing, which later reached a stage approaching a type of shorthand. It is in order to mention that the so called *Phags-pa* script of the Yüan Emperor Kubilai was formed after a type of the Tibetan script used for seals.

Tibetan graphotactics used a superscript dot called *tseg* to mark the monosyllabic unit. In the script there are the following possible places:

| | EXAMPLE | READ |
|---|---|---|
| V | | |
| $C_2$ | s | |
| $C_1C_3C_5C_6$ | bgbs | *bsgrubs* |
| $C_4$ | r | |
| (V) | u | |

From the given places only the place $C_3$ has obligatorily to be filled, because if the vowel is the inherent *a* no vocalic letter is needed. In the Tibetan graphematics $C_3$ is called Radical, $C_1$ Praescript, $C_2$ Superscript, $C_4$ Subscript, V Vocal, $C_5$ Final and $C_6$ Postfinal. From the point of view of phonological analyzis $C_1$ and $C_2$ are praeradicals, $C_3$ is the radical, $C_4$ postradical and $C_5$ final, $C_6$ postfinal. Morphologically the praeradicals are praefixes, the postradical is an infix and the postfinal, and in some cases the final is a suffix. One syllable can have only one vocalic element which is either a monophthong or a diphthong. The vocalic sign <u> is written beneath the radical or subscript all other vowels <o>, <e>, and <i> were written above. In Old Tibetan orthography an "inverted" *i*-sign was also used, on which see later.

In the position of $C_3$ all consonants can occur, but in the other places there exist certain well defined constraints.

*The most important tendencies in the history of Tibetan*

The most important tendency is the simplification of the consonantal clusters in the initial and final position. In the so-called non-archaic dialects all initial clusters have been simplified to one consonant. The postradical /r/ and the radical developed into a cerebral sound (*gr-, dr-, br-* became *ṭ* or *ḍ* respectively), the postradical /y/ palatalized the radical (*ky-, gy-, phy-, by-* became *c-, j-* or *kʻ*). The original opposition voiced:voicless changed into a register opposition of low:high (pitch). The postfinal *–d* disappeared early, the *–s* later, most finals also disappeared while they changed the preceding vowel and generated a new suprasegmental feature the contour which may be even or falling if the syllable was in a stressed position.

These radical changes began in the Old Tibetan period. On the other hand there exist at present so called archaic dialects which have preserved many traits from the structure reflected by the script. Besides the very different dialectal background the evaluation of the Tibetan material is hampered also by the fact that there existed a so called learned monastery reading style which was neither identical with the spoken language nor with the letter-by-letter reading. From the 13th century on we know that even between the monastery reading styles there existed differences, but it is very likely that these differences were to in some extent also present earlier.

*Turkic texts in Tibetan Script*

Though Turkic texts written with Tibetan letters had been mentioned earlier by P.Pelliot (1921) and others, it was G. Clauson who first dealt with some of them. In his book (Clauson 1962) he mentioned four unpublished texts the photographs of which he got from F.W. Thomas "...of what he beleived to

be the whole collection..". In the first text there are parts of lines 142 to 161 of the Sekiz Yükmek Sutra the text of which in Uighur script was published in *Türkische Turfantexte VI* and this text now bears the Sigla Mainz 329 and was quoted by Clauson as *SY*. The second is a text with the invocation of Boddhisatvas similar to that found in the Altan Yaruq (see Müller 1908), its present Sigla is Mainz 712 and it was quoted by Clauson as *NK*. The third is a short list of proper names, some Turkic some Tibetan (Clauson: *PN*), the fourth is a small fragment of a Buddhist text, Mainz 194 (Clauson: *Fr*). Clauson (1962 97-100) gave a list of Turkic words and names to be found in the four fragments. Some mistakes crept into the transcription and the readings. In 1942 when A. von Gabain was in Budapest they read together with Ligeti Turkic texts in Tibetan transcription. Ligeti got the transcription of two texts from Gabain. In a letter of 11 Januray 1961 Gabain quoted from five texts some Turkic words in Tibetan transcription. The texts had the following Sigla Mainz 194, Mainz 196, Mainz 329, Mainz 619 and Mainz T II Y 35. Ligeti quoted some of these words with reference to A. von Gabain in his paper (Ligeti 1961 210-211), unfortunately in some cases Gabain's transcription was either inaccurate or it misled Ligeti for other reasons. In their paper published in 1984 Maue and Röhrborn mentioned 7 fragments which are in the Staatsbibliothek Preussischer Kulturbesitz (West Berlin) and add "Sie sind von Clauson: *Turkish and Mongolian studies*, London 1962, S.96-100, in Form einer Wortliste ausgewertet worden" it is unclear to them why Hamilton (1981 15) thought that there existed one or two Mss in London. By courtesy of the Staatsbibliothek Preussischer Kulturbesitz (West Berlin) I have photocopies of the following fragments:

1. Mainz 127, TII Y.35 (Second Turfan expedition Yarkhoto)
2. Mainz 194a, 194b Tu.1o9, Clauson Fr
3. Mainz 196, Tu 113
4. Mainz 329, Tu 110, Clauson SY
5. Mainz 619, TII. Y.59, Tu 117 ? Clauson PN
6. Mainz 637
7. Mainz 712, Clauson NK

Mainz 619 in fact consists of two fragments, and in the second only the name *gut-lug tog-r[i]l* is present. Among the proper names quoted from the third fragment by Clauson there existed a *gut-lug tog-ril*, and other names as well. It may be that when Thomas's photocopies were made this fragment was larger.

The most important text is a Buddhist catechism written in the Uighur language but in Tibetan script. The text was brought by Pelliot from Tunhuang,

got the number Pelliot tib. 1292 and was first quoted by Pelliot in 1921, and on some occasions by Ligeti. The facsimile was published in [Macdonald] Spanien - Imaeda (II 1979 Table 608). It was dealt with by me in 1983/84 in my lectures given in Vienna (see Róna-Tas 1985, 355-359) and independently published by Maue and Röhrborn in 1985 and 1986 resp. In 1985 appeared the publication of Moriyasu. The text consists of 44 lines, it is written in a good readable *dbu-med* script. The edition of Maue and Röhrborn (M/R) solved most of the problems, the transcription needs very few corrections which are of minor importance.

*The transcription and representation of Turkic vowels*

The vowels in initial position

There exists a very good, comprehensive study on Tibetan script and its use in transcribing Turkic witten by G. Uray in the vol. III of Hajdú - Kristó - Róna-Tas (1980 95-112). Here we shall sum up the most relevant problems concerning the transcription of Turkic with special respect to the text P.1292.

As we have seen Tibetan script could render five vowels: *a, e, i, o,* and *u*. If these vowels occurred in postconsonantal position *a* was inherent, that is not marked, and the other four were marked by diacritic signs, <e>, <i>, <o>, were written above and <u> beneath the radical. The first problem arose with the fact that in Old Tibetan there did not exist any syllable with a vocalic initial. This was the reason why Tibetan originally did not take over from Brāhmi any of the independent vocalic signs. Later on when because of the transcription of Sanskrit the rendering of a vocalic initial was needed they used the so called "great *a*" or *a-chen*. The *a-chen* was the independent <a> of Brāhmi, but from the fact that in the Tibetan alphabet it stands as the last letter (in contrast to the Sanskrit alphabet where it is the first letter) we may infer that it had been taken over later and not together with the other letters. The *a-chen* had been used as a consonantal sign insofar that if the vowel was not *a*, the vocalic signs were put above or beneath. This graphotactic usage can be observed also in Khotanese Brāhmi. The phonetical value of the *a-chen* was a voiceless glottal or laryngeal stop. This we know from the Old Tibetan transcriptions of Chinese where *a-chen* transcribed the so called *ying*-initial, the voiceless glottal stop, in opposition to the *yū*-initial which was transcribed by the *a-churi* and which was a voiced glottal stop. In the transcription of Turkic the *a-chen* was used to render the smooth vocalic ingress of the vocalic initial.

The initial *a-* was transcribed by the simple *a-chen*, the initial *i-, o-* and *u-* with the *a-chen* and the vocalic signs. Somewhat more complicated was the situation with the *e*-sounds. As we have seen Tibetan had a special *e*-sign, but originally the Turkic had two *e*-sounds, an open *ä* and a closed *e*. The open *ä*

was considered in Brāhmī as a front variant to the *a*-sound and thus written with the subscribed *ya*. This usage was taken over also in the system of the Tibetan transcription of Turkic. Accordingly under the *a-chen* they wrote a subscribed <ya>, which is called in Tibetan *ya-btags* (*yata*), and thus it became an *ä*. The solution is the same as in Brāhmī:

| BRAHMI | TIBETAN |
|---|---|
| *A + ya* /ä/ | *a-chen + y* /ä/ |

There are a few cases where the initial *ä* was written with an *a-chen* but without the subscribed *ya*. If we collect these words we see that all words had a Kāf in the original Uighur script. The use of the Kāf was a sure indication of frontness in the Uighur script unlike in Brāhmī. If Turkic *äzüg* was written as *"a-(g)zug*, or *ägsüg* as *"ag-sug* it is clear that the Uighur model was followed. This was an orthographical, formal process. In the case of *ämgänürlär* the pronounciaton was [*ämŋänürlär*] and therefore it was correctly written in Tibetan as *"am-ṅan-nur-lar*, i.e. without a letter <g>, but in the first syllable the *a-chen* did not get a subscribed *ya* because in the Uighur original the word was written with a Keph as *ämgänürlär* and the rule "where there is a Keph the subscribed *ya* is not obligatory for expressing the frontness" operated. Tibetan also had the sign for <e> and one would have expected that this would be written for the closed /e/, while *ä* or *a* for the open /ä/. It may have been the tendency but it was not a rule. The scribe was surely perplexed about the opposition, and we find the same word written in two different ways as in the case of *ämgän-*, once as *"am-'ṅan* and once as *"e-mṅan*.

A special problem is connected with the Tibetan *i*-sign. When the Tibetans took over the Brāhmī script there existed a sign for short *i* and another for the long *i*. In Tibetan there was no long and short opposition of the *i* vowel and for some reason or other they accepted both, and the graphical opposition had no phonetical marker function in the script. The Brāhmī sign for long *i* became the standard Tibetan *i*-sign, but the other, the sign of short *i* called "the inverted *i*-sign" was ocasionally used as well. The idea of R.A.Miller (1966) that the two *i*-signs were used to denote two *i*-sounds differentiated by vowel harmony, could not be proved. In the earliest texts at our disposal the distribution of the two *i*-signs is only graphical, since they resemble as mirror forms, in certain manuscripts it is avoided for two identical signs to occur side by side,(in some types of manuscipts this would be very similar to one o-sign). The inverted *i*-sign was used only in Old Tibetan documents and later disappeared. It exists in our text, but has no phonetical marker function, it is not used to distinguish

front and back *i*, as can be seen from such example as *iki* written as *ĭ-ki*, the first *i* written with an "inverted", the second with a normal *i*-sign:

The rendering of the Turkic *ö* and *ü* was achieved by writing an *a-chen* then a subscribed <ya> and an <o> or <u> respectively.

In Appendix VI we compare the independent vocalic letters of Brāhmi, Manichean, Sogdian, Uighur, Arabic and Tibetan. In the upper part we see the independent letters, in the lower part the combinations with which they try to transcribe the Turkic vowels in initial position.

Vowels in postconsonantal position

Since the Tibetan *a-chen* is graphotactically considered as a consonantal sign the rendering of the vowels in postconsonantal position does not esentially differ. The vowels /a/, /i/, /o/ and /u/ are noted as "inherent" or with vowel signs. In the case of the front vowels there existed three types of solution. The front character was either marked by the subscribed <ya> or in the original a Kaf secured the front pronunciation, or the front character was not marked at all. In polysyllabic words it was sufficient to mark the first syllable, thus *dyag-gzin-dur-da-ci = tägzindürdäči*, where the "Kaf" in the original would have been in itself sufficent, or *tyab-pra-mas = täprämäz*.

It caused a problem that in back vocalic words front vocalic *i* occurred, as e.g. in the possessive suffix of the third person after consonants. On the other hand the phonetic realization of the back vocalic *ï* was different according to its environment. Especially in the vicinity of the guttural consonants or the sonorants (*r, l, n*) it still preserved its original back character. This allophone needed a special transcription. The Tibetans used the following solutions:

1. The use of the *e*-sign: *"a-reg = arïγ, "og-lan-ne = oγlanï, "oy'i-ga-le = oyγalï, "a-ge = aγï* etc.
2. Writing as *a* or *ā*: *"a-rag = arïγ, yab-pa'r-gāg-ke = yapïrγaqï, "ad-glāg = atlïγ* etc.
3. Writing as *a* and adding a subscribed *ya*: *gā'-tyag-lan-mag = qatïγlan-maq, "ad-glyag = atlïγ, yar-lyag-gĭn = yarlïγïn* etc.
4. Writing with *i* (in most cases in words which have *q*): *gir-kin-nin = qïr-kïnïn, za-kin-cin = saqïnčïn* etc. or *zi-mas = sïmaz*.

We know that in the present dialects of Turkestan the *i*-sound in mono- or polysyllabic words became neutralized, and this neutral *i* is phonetically nearer to the front *i*. Such words could get front and/or back vocalic suffixes. The beginning of these changes we find in our text: *din-lgig = tinliγiγ, din-lāg, dyin-lag = tinlïγ*. Where the alternative writing *di-* or *dyi-* shows that the first *i* was perceived as front but the second as back because of the following guttural.

Table III shows the transcription of the vowels.

The functions of the *a-chuṅ*

The *a-chuṅ* has as its model the letter <ga> of Khotanese Brāhmī script, and its function can also be derived from it (see Róna-Tas 1985 259). To understand the role played by the *a-chuṅ* in the transcription of Turkic we have to sum up what we know about its functions in the graphotactics of the Old Tibetan. It had the following functions:

1. If it was used as a praeradical and written in the praeradical position it denoted a nasal consonant homorgan with the following one, or it was used to express a prenasalisation of the consonant. In any case it secured the voiced pronunciation of the following consonant whatever its original voicing has been.
2. If it stood in the radical position it denoted a voiced glottal or laringeal fricative consonant.
3. If it stood after the radical there were three possibilities:

   a. if it was written after the radical it had an orthographical function. In cases where the word had a final (and maybe a postfinal) closed by the syllable marker dot (*tseg*) it may have happened that the reading was not unambiguous. Thus in the case where, e.g. a <db> was written, this could have been read both as *dab* and *dba*, because *-b* was a final permitted by the orthographical rules and thus it was not clear whether the inherent *a* should or should not be read after the first consonant. In such cases if the reading should be *bda* an *a-chuṅ* was written after the <b> to ensure the reading: *dba'*. In analogy to this there are some cases where the *a-chuṅ* was written even if there was no ambiguity, as e.g. in the case of *mtha* written as though *-th* was not permitted as syllable final.

   b. If the consonant was followed by a diphthong the second element needed a "supporter", and this was the *a-chuṅ*. Thus in such cases as *mde'u* or *mi'i* the *'u* or the second *'i* respectively was written with an *a-chuṅ*, and the vowel sign written in addition.

   c. In the case of the transcription of Sanskrit words the long vowel was indicated by writing the *a-chuṅ* under the radical. This was the rule in classical Tibetan, and it was also used in Old Tibetan. In Old Tibetan, however, at least in the transcription of a text dated from the 10th century (Hackin 1924) it could have been observed that the *a-chuṅ* did not denote the etymological length of the Sanskrit word but one has the imporession that it denoted a length due to phraseological stress (cf. Róna-Tas 1985 349).

In what now concerns the role of the *a-chuṅ* in the transcription of Turkic we see the following:

1. If the *a-churi* occurs in a place which was considered by the Tibetan graphotactics as praeradical it served to ensure the voiced feature of the following consonant. This is the more interesting, since Turkic /k-/ and /t-/ are written in most cases with <g> and <d> respectively, but in none of the cases is the initial consonant be it <g>, <k> or <kh> or <d> <t> or <th> preceded by an *a-churi*. On the other hand in the case of such words which are loanwords and had a voiced initial they are always written with an *a-churi* like *dhyana* 5x, *dendar* 1x, *dharma* 2x, *didimliγ* 1x without counterexample. Quite different is the case with Turkic /b-/. This was been never written with <p> or <ph>, exlusively with <b->, and in the overwhelming majority of cases with *a-churi* + *b*. Where we find an initial <p-> we have a foreign word which was pronounced with a voicless [p] as in *paramid* or *purhan*. The voiced character is indicated in a very few cases with a prescript *d-* as in *dbyis-ka-le* = *bičγalï*.

The voiced consonants which occur in non initial position are rarely written with a prescribed *a-churi* (*"a-'bag* in line 20 is obscure) and this also has its origin in Tibetan graphotactics. As we have seen the prescribed *a-churi* represented a nasal element. In cases where Tibetan had a disyllabic word (originally a composed word) the praeradical nasal element of the second word has been preserved even in most modern dialects. Thus a word like Written Tibetan *dge-'dun* is pronounced as [*gendun*]. Therefore we find the *a-churi* in non initial position in such cases as in *am-'ńag* where the *a-churi* is the marker of the nasal consonant, as can be seen from the parallel forms where the same word is written as *"e-mńag*, and *"am-'ńag*. Both transcribe Turkic *ämgäk* [ämŋäk]. This nasal occurs also in another type of the cases: the Turkic word *andaγ* is written three times with <-n-'d->, three times with <-n-d-> and three times with <-n-t->, *ičindä* is written as *"i-cin-'da'*. On the other hand *ikinti* is always written with <-n-t-> as is Ksāntipāla. Such words as *üntürüp, basintur, tägzintür, käntü* are written with <-n-d->. It cannot be doubted that the transcriptions *"an-'dag, "i-cin-'dä, den-'da'-rag* contain a cluster [nd], and if there had not been a drastic change since Runic script was in use, the special Runic sign had to denote [nd] and not [nt]. To this we have to add that the cluster *-lt-* is in most cases written with <-l-lt> as in *altun, altinč, qaltï*. If the cluster *nč* is not in the intervocalic position, then Tibetan could not transcribe it. If it was rendered somehow, it was always written as <-n-c->.

In three Turkic words *a-churi* denotes an initial *h-*: these are *härdi* (*'ardi*), *härsär* (*'ar-sar* twice, the same word also with *a-chen*), *hol* (*'ol* 11 times, four times with *a-chen*) and the enclitic particle *hök* (*temin hök* = *dye/de-min 'og*).

*The transcription and representation of the Turkic consonants*

Consonants in initial position

As we have seen the voiced character of the initial consonants was ensured by writing an *a-chuṅ*. We know that the originally voiced stops and affricates lost their voiced character in many Tibetan dialects and the following word got a suprasegmental low pitch while the words with an original unvoiced initial got a high pitch. This is a feature characteristic to a large region in East Asia and characteristic to all tonal languages there, among them also Chinese. Further, for several reasons Tibetan words did not begun with simple (not prefixed) nonsapirated unvoiced consonants as *t-, k-, p-, č-*. For these reasons the transcription of Turkic /t-/, /k-/ and /č-/ by <d->, <g-> and <j-> does not reflect any Turkic feature, its causes have to be found in peculiarities on the Tibetan side. The uncertainty of the Tibetan scribe can also be seen in such variations as *tyeb, dyeb = tep, tyag, dag = täg* but never *b- = <p->*.

Though as a rule the Turkic initials /t-/ and /k-/ are transcribed by Tibetan <d-> and <g-> there is a group of words in which this rule does not work and the initials are transcribed by a letter denoting an unvoiced sound. Those words pertain to this category in which the second consonant is a tense unvoiced consonant (stop or sibilant) in Turkic, such as *tutar* (*thud-tar*), *tutsar* (*thud-sa*), *tutmasar* (*thud-ma-sar*), *tükel* (*thu-kāl*), *tükellig* (*thu-kal-lig*), *Tïši* (*ti-ši*), *qutïŋa* (*kod-tiṅ-ṅa*), *kaš* (*kās*), *kiši* 'human being' (*khi-ši*), *kiši* 'woman' (*khri-ši*), *Kišantipale* (*khi-šaṅ-ti-pa-le*). There exist a very few exceptions to this rule, such as *toquz* written as *do-kos* or *kätäyin* written as *gyed-ta-yin*. The affricate *-č-* does not seem to pertain to the category as we find for *küčintä gud-cin-da'*, or for *qačïγ gā-cag*. It is very likely that the phonetical pecularities of those initials which were followed by a consonant of the tense category were different and this was perhaps due to either the assimilatory effect of the following consonant or this was the original situation which has been preserved only in cases where there existed a second tense consonant. A slight aspiration can perhaps be supposed in opposition to the strong aspiration of Tibetan *kh, th, ph* etc. If this were true we would make a great step forward in the reconstruction of the Old Turkic phonological and phonetical system. We could suppose that in word initial position the following phonetical units were present:

| TURKIC PHONETICAL UNITS | | TIBETAN TRANSCRIPTION |
|---|---|---|
| I. VOICED LAX | [d] in foreign words<br>[g] in foreign words<br>[b] | <'d><br><'g> no example<br><'b> |
| II. UNVOICED LAX | [D]<br>[G] | <d><br><g> |
| III. UNVOICED TENSE | [t]<br>[k]<br>[p] | <t, th><br><k, kh><br><p> |

It remains unclear whether in category III there was also aspiration and, if so, what its phonological relevance was. As regards the tripartite structure the situation is very similar to that of Tibetan and Chinese, nevertheless the units do not match each other. In Tibetan we have on the graphical level <t>, <th>, <d>, but the Turkic system is rendered by <t, th>, <d>, <'d>. In Indo-European and in Mongolian we know that between two tense aspirated consonants there may have occurred dissimilation. In any case the distributive rules between categories II and III may have been the following:

/t/ [d] all cases with the exception of where [t]
[t] if the second consonant is + tense and either + obstruent or + sibilant.

The same may have been true for /k/.

The transcription of the sibilants reflects an early Tibetan development. Both Tibetan /z/ and /ž/ became voiceless at least in the dialects which may have played here a role. Therefore it was possible to write the Turkic /s-/ with <z-> as in *sab* written as *zab*, *sögüt* written as *zu-gūd* etc., in all cases with one exception (*sub*). We have, however, to put the question why they did not use the letter <s>. In the later Tibetan dialects the difference was in pitch, words with initial /z/ had a low, those with /s/ had a high register, with the low register a breathy voice coocurred.

The transcription of the very few nasal consonants in initial position did not cause special problems. The curious Tibetan transcription of *miŋ* 'thousand' as *dmyiṅ* is a Tibetan orthographical feature. In the Old Tibetan orthography a special category was written with <my-> which later became in Central Tibet *m-*, in the East Tibetan dialects *ñ*. The prescribed <d--> was needed to ensure the high register category, therefore e.g. Classical Tibetan *mig* was written in Old Tibetan as *dmyig*. The transcription *dm-* points to a kind of unvoiced *m-*.

Consonants in non initial position

The transcription of the stops in intervocalic position caused no problems as can be seen from the following examples: *tägilä = dya-gi-lya', tükel = thu-kāl, qadaγ = gā-dag, qatïγlanmaq = gā-tyag-lan-mag*. If certain consonantal clusters were present in the set permitted by Tibetan graphotactical rules the Turkic cluster was transcribed accordingly as *artoqï = "ar-rtog-ke, ötrö = "oy-trō, "oyd-tro, "od-tro*, where the Turkic *-tr-* was rendered by the Tibetan *tr-*.

In most Tibetan dialects the final consonants disappeared. The first final which disappeared was *-s*. Since we find in our transcription everywhere final *-s* the text must have been written at a time and in a dialect where the final *-s* have not yet disappeared: *az = "as, altmiš = "al-(ti)-mis, beš = 'bes, üzüksüz = "u-zug-sus*. It would be a possible argument against this hypothesis that the scribe copied the Uighur text and therefore wrote a Tibetan <s> on place of an Uighur <s> even when the Tibetan *-s* was not pronounced. But such an argument cannot be accepted, because in some cases another writing would have been possible as well, e.g. *asqančulamasar* is written *"a-skan-cu-la-ma-sa*, while *kästi* is written as *gyās-ti*. Since there exists a Tibetan pattern *sti* this could have been written as **gya-sti*.

The assumption that the disappearence of the final consonants had not yet begun is important if we try to get the answer to the question of why we have curious reduplications. Why do we find *kōŋül* written as *kor-ñol, eligin* 'of the hand' as *"el-lig-gin, elig* 'king' as *"el-lig* (in this case etymologically correct) or *oqar* as *"og-kar*? In the last example the final of the first syllable and the initial of the second is not the same. But this is due to the Tibetan rules, according to which only <g, d, b, l, s, r, m, n, ṅ, '> can occur in final position. Therefore we have for *ötrü "od-tro*, for *täprämäz dyab-pra-mas*, for *qutiŋa kod-tiṅ-ṅa* or for *yapūrγaqï yab-pa'r-gāg-ke*. It is impossible to suppose in all these and the similar cases that this transcription reflects a reduplicated or long Turkic consonant. The most reasonable answer to this question is that the Tibetan scribe had to write closed syllables. This may have happened because the vowel of the closed syllable was shorter, and the prosodic shortness was rendered with this graphotactical device, since in fact all Tibetan syllables which had no consonantal finals have a longer vowel than those which have. What did

the Tibetan scribe do when he had to indicate a long consonant, as in the case of *bütti*? One of the possible solutions was that he indicated that the vowel is long, and this happened, the word is written as *būd-ti* with the *a-chuṅ* under the <b>. This means that the transcription technique could distinguish between short and long vowels and between short and long consonants. As we have seen above the opposition of the long and the short vowels was not identical with the etymological shortness and length, it depended on prosodic reasons.

There are no direct traces of the fricative pronunciation of the consonants /d/, /g/ or /b/. The word *'ye-vin* in line 11 is obscure. The <h> in *yumur-hā-da* (for *yumurtγada*) is due to the combinatory influence of the preceding [r]. Two cases are of interest. Turkic *egid* is written as *"ey-gil* instead of the expected and possible **"ey-gid*, and we find for *yutuz* 'wife' once *yul-tu-zi-ña* and once *yul-tus-sin* (the *-d* in the transcription as *yuld-* is a misprint, Moriyasu has here the correct reading), where we would have expected *yud-tu[z]*. These cases are perhaps the first signs of the spirantization of the Tibetan final which surely preceded its later disappearance. It may be that the transcription of *kūčsüz* as *gud-sus* also belongs here. The word *adirmaz* is written as *"a-drir-mas*, this is the only possible sign for a fricative character of the Turkic /d/.

The Turkic word *tört* is consequently written as *dyor* i.e. without the final *-t*. This important insofar as it shows that the disappearance of the Tibetan postfinal *-d* or *da-drag* was complete (not so in Ms Mainz 194).

The transcription was able to distinguish between the consonant /y/ and the semivocalic *i* element of the diphthongues. For *qoyin go-yin* is written, but in the case of *äyrig* [*äirig*] the transcription is *"ey'i-rig* where the first, subscribed <ya> only marks the front character of the [e] and the *a-chuṅ* and <i> denote the semivowel [i]. The accusative of *bay* [bai], *bayïγ* is written as *'ba'i-yag*, that of *čiγay* [ciγai] as *ji-gā'i-yag*, where we find not only the rendering of the diphthong, but also the graphotactic device to produce a closed syllable. The writing <a'i-y-> is the same as the writings <-d-t-> or <-g-k>.

In the Tibetan writing system we find a subscribed *-r* which is called *ra-btags*. Originally it was used to denote a postconsonantal *r*, later the clusters *Cr* developed in most cases into a cerebral diphthong or even monophthong. In words such as *ötrö* (*'yo/'yod/'od*)*-tro* this subscribed *r* denotes an original postconsonantal [r]. But in several cases this subscribed *r* denoted a special quality of the preceding consonant as in: *arhan* = *"ar-hrin*, *čixsabut* = *chig-šra-bud*, *kišide* (or *-da*? locative of *kiši* 'woman') = *khri-ši-da*. In a few other texts we find for the back vocalic [q] the writing <gr> as *quruγ* = *gru-rug*, *alqo* = *"al-gro* etc., see below.

Some problems were involved in writing the sibilants in non initial position. There were no problems with the sibilants in intervocalic position. They were written in most cases with the letter originally denoting voiced

sibilants, but we also have examples for the use of unvoiced: *öltäsitä = "oyl-ta-ži-da, äsirkämäz = "ay-zir-ka-mas*, and also *kiši = khi-ši, tesär = de-sar*. The final *-s* caused also no problem as in *üstün = "us-stun, kästi = gyās (or phyās)-ti*, only in postconsonantal position as in *tärs*, in line 14 there is *tyer-zi* but the syllable *-zi* seems to be crossed out; in line 40 only *dyar* is written. This is understandable, because the original Tibetan orthography knew the postfinal *--s* but it was not permitted after *-r*. It was not permitted to write a final *-š* or *-z* and therefore this Turkic final was rendered by the final *-s*. We find even for *išläyürlär "is-slya-yur-lar*. The choice was perhaps also influenced by the Uighur script where the two final sibilants were distinguished only by diacritical dots. The fact that cases such as the rendering of *beš* 'five' as *bes, 'bes, 'bēs* we have not to do with any Turkic phonetical pecularity can be seen in the case of *bešinč* which is written as *'bes-sin, bēs-šin*, or *täŋäš+i+siz* with an *-š* originally in final position is transcribed as *dyaṅ-ṅa-ši-sis*. The suffix *-miš* is written as *-mis*, both in the deverbal suffix as in *ämgänmištä = "em-ṅan-mis-ta*, and the ordinal suffix as in *altmiš = "al-ti-mis*, in both case a Turkic *-š* is transcribed by the Tibetan *-s*.

The voiced Turkic /z/ caused some problems to the Tibetan scribes. The Old Tibetan /z/ became early unvoiced therefore its voiced character had to be ensured by writing a prescribed consonant which was, however, not pronounced. This was the case for *uzati = 'u-gza'-ti* or *uzun = 'u-gzun*. But seemingly the unvoiced pronunciation of <z> was not wholly accomplished in certain non initial positions and we find also *sezig = ze-zig, tözi = dyo-zyi*. The final *-z* was always written with *-s* but this is not a Turkic pecularity, the privative suffix is always *-sis*, such words as *tuz* 'salt' and *tüz* 'equal' are both transcribed as *dus* and this has nothing to do with the also otherwise unlikely hypothesis of Shcherbak (1970) that Early Turkic final *-s* became *-z* after long vowels.

A great problem was caused by Turkic *č* in non initial position. Tibetan could not write final *-č* therefore we find for *küčsüz gud-sus* or for *üč 'uys*, but for *üčünč* it was possible to write *"uys-cun*, i.e. to express the first *-č-* while the second had to be omitted. That the latter was a problem of Tibetan transcription and not of Turkic pronunciation can be seen in the case of *törtünč(ü)* which was witten as *dyor-tun* and *dyor-tun-chu* resp., and all Turkic *-nč* finals were written with a simple *-n* as *basinč = bas-sin, bešinč = bēs-šin, saqinč = za-kin*. We may suppose that in some clusters, as Turkic *-čt-*, *-čk-*, *-čg-*, the [č] lost its stop component. In the case of *bičti = byis-ti, bičγali = dbyis-ka-la* it has to be kept in mind that in Tibetan syllable-final *-č* cannot be written. In case of *yinčgä* the transcription was *yiṅ-skya*, where the *-ṅ* is not a mistake (cf. Kazakh *jiŋiške*, the *ŋ* was also present in Üzbek and some Kipchak languages) and the transcription may have rendered something like [yiŋškä] or [yiŋskä].

In three foreign words [z] occur. The word for "Rsi" *arži* was written as *"a-rje, "ar-rje,* perhaps under the analogical influence of the Tibetan word *rje* 'lord', otherwise *ažun* 'existence' is *"a-žon-(ne)* and the voiced pronunciation of [z] is secured by a praescript which was not pronunced in the transcription of *Tušit* [tožit] written as *do-gžid.*

There are a few cases where the Turkic final *-r* was not transcribed as in *ärür = "a-ru*, and mostly in the suffix *-sar/ser*: *barmasar = bar-ma-sa, čašurmasar = ja-šur-ma-sa* etc., but written in *tüšmäsär, turγumasar.* This is a Turkic feature. The disappearance of the final *-r* in Tibetan was surely a later change.

One Old Tibetan feature is consistently missing, the graphic complex <myi>, which we would expect, e.g. in *išlämiš = "is-sla-m[i]s*, and in *barmiš* or *altmiš.*

*On the date of Pelliot tibetain 1292*

Maue and Röhrborn dated this text in the timespan between the 8th and 10th centuries mainly relying on Taube 1980 and on the analogies which they found with the Tibetan text of the Staël-Holstein Roll dated 925.

It is unlikely that we have to do here with a copy, the corrections made in the text show that we have to do with an original, autograph version.

On the Tibetan side we can observe the following:

1. The oral praeradicals were already in the process of disappearing, *d--*, *g--* and *b--* had only orthographic functions. The praeradicals *s--*, *r--* and *l--* were still present as we see from *"a-skan-cu-la-ma-sa* (*asqančumalasar*), *"og-rla-ma-sa* (*oqrīlmasar*) and *"a-lko* (*alqo*). On the disappearance of the oral praeradicals through fricatives see (Róna-Tas 1966).
2. The two nasal praeradicals *a-churi* and *m--* became homorgan nasals before stops.
3. The postfinal *--d* or *da-drag* had already disappeared.
4. The final consonants, including *-s*, were all pronounced, perhaps *-d* had begun its spirantization.
5. There are no traces of the Old Tibetan orthographical *myi* where later in classical orthography we find *mi.*

With the exception of the last feature all these point to the North-Eastern Archaic Tibetan dialects and nothing speaks directly against the 10th century, but it cannot be earlier, and a later dating in the early 11th century cannot be excluded. It is unlikely that it belongs to those texts which were added later to the Cave Library after its opening, but we have no criteria to exlude this. If it belongs to the original stock of the Cave Library, the latest possible years are the first years of the 11th century (1000-1005).

*Some remarks on the Turkic language of Pelliot tibetain 1292*

There is only a slight tendency to distinguish closed /e/ and open /ä/, the back /ï/ was on its way to being neutralized, but was still preserved in the neighbourhood of gutturals and sonorants. Etymological vowel length is not reflected, but the opposition of prosodically short and long vowels could have been expressed. In the case of some words we can identify the quality of the vowel for the first time as in the case of *öz* 'hatred' (Clauson *üz*).

A strong tendency can be observed in the assimilation according to the open:closed opposition:

First category: *ö+ü > ü+ü*

*'uy-lyur-ma-sa = ülürmäsä(r)*, cf. *ölürmäsär*
*du-ru-sin = türüsin*, cf. *törüsin*
*dur-lug, dur-lyug = türlüg*, cf. *törlüg*, also *dyor-lug*
*zu-gūd = sügüt*, cf. *sögüt*

Second category: *ö+ü > ö+ö*

*'oy-tro = ötrö*, cf. *ötrü*
*do-kos = toqoz*, cf. *toquz*
*koṅ-ṅol = köŋöl*, cf. *köŋül*

I do not see here any possibility to explain the difference betwen the two categories from differences in the original vocalism or in the phonetic environment. The less so, since e.g. in Northern Brāhmi we have *törö* and *ötrö*. It cannot be excluded that we have here to do with movable stress and the assimilation occurred according the stressed syllable.

There operated a tendency that *u* changed to *o* if *a* and *ï* followed:

*gol-lkag-kiṅ = qolqaqin*, cf. *qulqaqin*
*kod-tiṅ-ṅa = qotiŋa*, cf. *qutiŋa*
*do-gžid = Tožit*, cf. *Tušit*

or preceded:

*'a-gō = aγo*, cf. *aγu*
*'a-lko = alqo*, cf. *alqu*
*'a-žon = ažon*, cf. *ažun*
*da-mo = tamo*, cf. *tamu*
*ya-rog = ya-roq*, cf. *yaruq*
*dur-ka-ro = turqaro*, cf. *turqaru*

The demonstrative pronoun is always *bo*.

In certain suffixes we already find the fourfold vocalism:

| ī | i |
|---|---|
| *atlīγ* (*"ad-glāg/glyag*)<br>*tinlīγ* (*din-lāg/dyin-lāg*)<br>*yarlīγ* (*yar-lyag-*)<br>*didimlīγ* (*'di-dyim-glāg*) | *erlig* (*"er-rlig*)<br>*tükällig* (*thu-kal-lig*)<br>*täŋäšisiz* (*dyaṅ-ṅa-ši-sis*)<br>*bešinč* (*'bes/bēs-sin*) |
| **u** | **ū** |
| no example | *törlüg* (*dyor/dur-lug/lyug*)<br>*köŋüllüg* (*go[ṅ]-ṅol-lug*)<br>*üzüksüz* (*"u-zug-sus*)<br>*törtünčü* (*dyor-tun-chu*) |

but in other suffixes this had not yet occurred, as in *tolti* (*dol-tī*), *körti* (*gor-di*), *oqmiš* (*"og-mis*), *tözi* (*dyo-zyi*), *türüsin* (*du-ru-sin*), *küčindä* (*gud-cin-da'*) etc.

The rudimentary presence of the initial *h-* is surely an archaic trait.

As we have seen there are no traces of the fricative quality of the earlier stops. The *-r* in the suffix *-sar/ser* begun to disappear.

We have no indication to see in Pelliot tibetain 1292 traits of the early Manichean texts as was claimed by Maue and Röhrborn (1985 77, on the "early" date of the Manichean texts see later). Their observation is correct that in the text we find instead of the ablative case the locative governed by several verbs.

*Some special traits of other Turkic texts written in Tibetan*

One of the special graphic renderings used in the texts Mainz 329 and Mainz 712 was the transcription of the /k/ in back vocalic words and non-final position by Tibetan <gr-> such as : *qayu* = *gra-yu, quruγ* = *gru-yug, turqaru* = *tur-gra-yu, alqo* = *"al-gro* etc., perhaps also for back vocalic /g/: *aγirlayurlar* = *"a-gri-[la-y]ur, yaγidin* = *[ya]-gri-din*, and from a chronological point of view it is of interest that in the tetx Mainz 194 we find *tört* = *tyord*, that is the postfinal *d* in Tibetan was pronunced, unlike in the other texts.

*The Turkic material in the translated texts: Pelliot tibetain 1283*

The pecularities of the text (translated and excerpted) were discussed in detail by Ligeti (1971). I would mention here only some important traits. The initial *h-* is consistently transcribed with <h> as in: *hadaqlīγ* = *ha-dag-leg, hala* = *ha-la, hirkin* = *hir-kin*. To this category pertains the transcription of the Uighur ethnonym as *ho-yo-hor, ho-yo-'or* but here we find also *a-chuṅ 'u-yi-kor*. The vocalic initials are written with *a-chen* as in *"Er-myis* (until now read *Yer-mis*), *"og-rag* = *oγraγ, "ud* = *ud*. Turkic /k-/ and /t-/ are written with <g>

and <d> respectively: *ga-ra = qara, Gu-log = Külüg, gar-logs = Qarluq, dad-byi = Tatbi, du-rgyus = Türgeš*, /b-/ is written with <b>: *ba-ker-ba-lig = Baqïrba-lïq, be-ca-nag = Bečeneg*, and if it is written with *a-chuṅ* the nasalization was present as in *'bug-chor* = Chin. *Mo-cho*. The back vocalic [ï] is in most cases written with <e> as in *Khe-rged = Qïrgïz, Ba-smel = Basmïl, ha-dag-leg = hadaglïγ*. The labial variants already appear as in *yun-log = yunlüγ*.

On the Tibetan side we do not find the posfinal *--d* (*yun-log* could have been written as **yund-log*), but *myi* is consistently used, cf. *"Er-myis hir-kin = Ermiš hirkin*. The *a-chuṅ* denoted in intervocalic position a voiced fricative as in *ho-yo-'or = Hoyoγor* i.e. Uighur or *ku-chu-'ur = Küčüγür*. The praeradicals were in the process of disappearing as in *ba-yar-bgo = Bayïrqo* but *he-bdal = Heftal, bas-mel, ba-mel = Basmïl, khe-rged = Kïrgïz*. Whether *hi-kil-rkor* and *"i-byil-kor* were the same or not, *-rkor* and *kor* had to be, and *-r-* was surely not pronunced in *-rkor*. All finals seem to have been pronounced, even *-s* as in *ges-dum = Keštim*, the variant *ba-sme* for *ba-smel* may be a mistake, but the final *-d* may have been on its way to becoming a fricative, cf. *khe-rged = Qïrγïz*.

The Tibetan language of the text of Pelliot tibetain 1283 shows a somewhat older stage than the text of the Pelliot tibetain 1292.

*Turkic names and titles in Old Tibetan texts*

In the rich Old Tibetan literature we can find many Turkic geographical and personal names and titles. A Turkic ruler Kha-gan Ton Ya-bgo is mentioned in the Royal Annals from the years 694 and 700. In 708 we read about the burial of the Ga-tun that is the Qatun. There exists a fragment of a legendary Chronicle of the Turks where we read that the kings of the Great Turks ruled 72 years. After they had flourished 72 years the Eastern Turks (*ñi-ma sar logs-kyi Dru-gu*) and the Western Turks (*ñi-ma nub-pa logs-kyi Dru-gu*) fought with each other. We learn from the Old Tibetan Annals that in 739 the Tibetan King visited the country of the Beg, and from the Chinese sources we know that he paid this visit to the ruler of the Turk Shahi dynasty in Gandhara on the occasion of the enthronement of the young King From Gesar. The daughter of this king became the wife of the Khotanese ruler and had the name in Tibetan transcripiton *'U-roṅ-ga*. The name *Uroŋγa, Uroŋγo* occurs in many Turkic sources (hitherto read as *Oronγa*). The word is written with *a-chuṅ* and based on this we can reconstruct a form **Huronγa*. The word has the meaning "flag, standard" and in fact had an inital *h-* which can be found in the Old Russian sources. This word dealt with by many scholars (see e.g. Ligeti 1949) is neither Avar nor Mongolian as formerly was claimed by Menges (1979 157-166), it is Turkic. An exhaustive collection of the scattered Turkic material in the Old Turkic sources and its linguistic and historical evaluation is a future task. Good preliminary work has been done by F. W. Thomas (1931, reedited in 1951).

# TABLE I

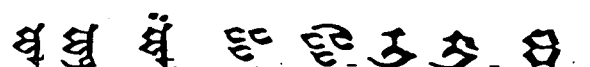
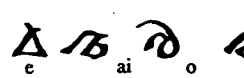
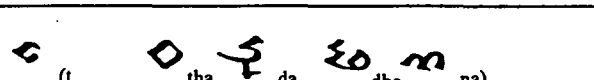
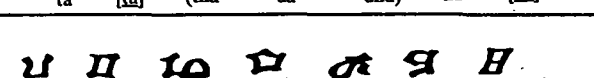
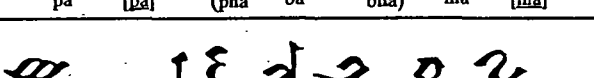
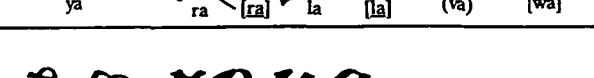
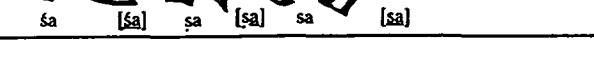
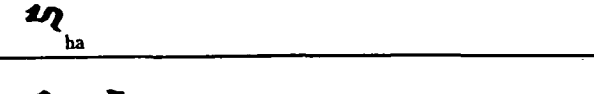
*The inventory of the graphemes of the East Turkic Runic Script*

| VOWELS | | |
|---|---|---|
| 1 | 𐰀 , 𐰁 | a, ä |
| 2 | 𐰂 | o, u |
| 3 | 𐰃 | ï, i, e |
| 4 | 𐰆 | ö, ü |
| 5 | 𐰇 | e (only Yenisei) |
| **CONSONANTS** | | |
| *A. "Paired"* | | |
| 6 | 𐰉 , 𐰊 | $b^1$ |
| 7 | 𐰋 , 𐰌 | $b^2$ |
| 8 | 𐰍 | $d^1$ |
| 9 | 𐰏 | $d^2$ |
| 10 | 𐰐 , 𐰑 | $g^1$ |
| 11 | 𐰒 , 𐰓 | $g^2$ |
| 12 | 𐰔 , 𐰕 | $y^1$ |
| 13 | 𐰖 , 𐰗 | $y^2$ |
| 14 | 𐰘 , 𐰙 | $l^1$ |
| 15 | 𐰚 | $l^2$ |
| 16 | 𐰛 | $n^1$ |
| 17 | 𐰜 , 𐰝 | $n^2$ |
| 18 | 𐰞 , 𐰟 | $r^1$ |
| 19 | 𐰠 | $r^2$ |
| 20 | 𐰡 , 𐰢 , 𐰣 | $t^1$ |
| 21 | 𐰤 | $t^2$ |

| *B. Letters denoting k* | | |
|---|---|---|
| 22 | 𐰴 | $k^1$ |
| 23 | 𐰅 | $k^2$ |
| 24 | 𐰼 | $k^3$ ($\check{k}^1$) |
| 25 | 𐰽 , 𐰾 | $k^4$ ($\breve{k}^a$) |
| 26 | 𐰿 , 𐱀 | $k^5$ ($\breve{k}^u$) |
| *C. Sibilants* | | |
| 27 | 𐰦 , 𐰧 , 𐰨 | $ś^1$, $ś^2$, $s^1$ |
| 28 | 𐰩 | $ź^1$, $ź^2$, ($s^1$, $s^2$) |
| 29 | 𐰪 | $s^1$ |
| 30 | 𐰫 | $ś^1$, $ś^2$ (only Yenisei) |
| *D. "Singles"* | | |
| 31 | 𐰃 , 𐰃 | m |
| 32 | 𐰭 , 𐰮 | ŋ |
| 33 | 𐰯 | p |
| 34 | 𐰱 | č ( 𐰧 $\check{c}^1$) |
| 35 | 𐰲 , 𐰲 | z |
| 36 | 𐰳 | ń |
| *E. Clusters* | | |
| 37 | 𐰵 | ld |
| 38 | 𐰶 , 𐰷 , 𐰸 , 𐰹 | nd |
| 39 | 𐰺 , 𐰻 | nč (nǰ) |

# TABLE II

*The Tocharian Brāhmī alphabet according to Krause and Thomas 1960*

| VOWELS | |
|---|---|
| Simple vowels |   |
| Diphthongs |   |
| **CONSONANTS** | |
| Velars |   |
| Palatals |   |
| Cerebrals |   |
| Dentals |   |
| Labials |   |
| Semi-vowels and liquids |   |
| Sibilants |   |
| Laringal |   |
| Affricates |   |

# TABLE III

*The graphic representation of the Turkic vowels in Tibetan*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| a- |  | "a | | | | | | |
| -Ca- |  | C | | | | | | |
| ä- |  | "ay |  | "a + 'k' | | | | |
| -Cä- |  | Cy |  | C + 'k' | | | | |
| e- |  | "e |  | "ey | | | | |
| -Ce- |  | Ce |  | Cye |  | C + 'k' | | |
| i- |  | "i |  | "ĭ |  | "iy | | |
| -Ci- |  | Ci |  | Cĭ |  | Cyi |  | Ci + 'k' |
| ĭ- |  | "i |  | "ĭ | | | | |
| -Cĭ- |  | Ci |  | Cĭ | | | | |
| o- |  | "o | | | | | | |
| Co |  | Co | | | | | | |
| ö- |  | "oy- |  | "o + 'k' | | | | |
| Cö- |  | Coy |  | Co + 'k' | | | | |
| u- |  | "u | | | | | | |
| Cu- |  | Cu | | | | | | |
| ü- |  | "uy |  | "u + 'k' | | | | |
| Cü- |  | Cuy |  | "u + 'k' | | | | |

| | | | |
|---|---|---|---|
|  | consonant with inherent /a/ |  | ĭ |
|  | "a-čhen |  | o |
|  | ya-btags |  | u |
|  | e | 'k' | Keph in the word |
|  | i | | |

# TABLE IV

*The graphic representation of Turkic consonants in Northern Brāhmī*

| | The original letters | The new Tocharian letters | The new letters used for Turkic | Vowelless final forms |
|---|---|---|---|---|
| k | <ka> <kha> | <kā> | <qa> | <-k> |
| g | <ga> <gha> | | <γa> | |
| t | <ta> <tha> <tta> | <ṭā> | | <-ṭ> |
| d | <da> <dha> | | <ź a> | |
| p | <pa> <pha> | <pā> | | <-p> |
| b | <ba> <bha> | | | |
| c | <ca> <cha> <ja> <jha> | <tsa> <ṭṣā> | (?)<dza> | |
| n | <na> | <nā> | | |
| ŋ | <ṅa> | | | |
| ñ | <ña> | | | |
| m | <ma> | <mā> | | <-m> |
| s | <ṣa> | <ṣā> | | <-ṣ> |
| z | | | <za> | |
| ś | <śa> | <ṣā> | (?) <źa> | <-ś> |
| ṣ | <ṣa> | <ṣā> | | <-ṣ> |
| y | <ya> | | | |
| r | <ra> | <ṛā> | | <-ṛ> |
| l | <la> | <ḷā> | | <-ḷ> |
| h | <ha> | | | |
| v | <va> | <wa> | | |

| | |
|---|---|
| Anusvāra | |
| The Cerebrals | <ṭa> <ṭha> <ḍa> <ḍha> <ṇa> |

# TABLE V

*The graphic representations of the vowels in Khotanese Brāhmī*

| | | | |
|---|---|---|---|
| a- |  | Ca- |  |
| ā- |  | Cā- |  |
| i- |  | Ci- |  |
| ī- |  | Cī- |  |
| u- |  | Cu- |  |
| ū- |  | Cū- |  |
| e- |  | Ce- |  |
| ai- |  | Cai- |  |
| au- |  | Cau- |  |
| o- |  | Co- |  |
| Anusvāra |  V, V̇, Ċ<br>V̲, C̲ |  |   (also -ṃ) |

Symbol of an Aksāra: 

# TABLE VI

*The graphic representations of the consonants in Khotanese Brāhmī*

| LETTERS | PHONETIC VALUES |
|---|---|
| p | p |
| ph | ɸ |
| b | b, β |
| v | w |
| m | m |
| k | k |
| kh | $k^c$, χ |
| g | g, ɣ |
| gg | g |
| h | h |
| h: | G, χ (?) |
| ḫ: | g, ɣ (?) |
| ṅ | ŋ |
| t | d |
| tt | t |
| th | θ |
| d | ð, d |
| n | n |

| LETTERS | PHONETIC VALUES |
|---|---|
| ṭ | $t^j$ |
| ḍ | $d^j$, $l^j$ |
| ṭh, ṭhṭh | $θ^j$ |
| ṇ | $n^j$ |
| tc | ts |
| js | dz |
| c, ky | k', tš |
| j, gy | g', dž |
| ts | $ts^j$ |
| ch | $tš^j$ |
| s | s |
| ys | z |
| ṣ | ẓ |
| ṣṣ | š |
| ś | $z^j$, ž, ś |
| śś | $s^j$ |
| r | r |
| l | l |

For the graphic forms see Appendix II, alphabet *v*.

# TABLE VII

*The Tibetan alphabet*

| CONSONANTS | | | |
|---|---|---|---|
| +ka | kha | +ga | ṅa |
| ca | cha | ja | ña |
| ṭa | ṭha | +ḍa | ṇa |
| pa | pha | ba | ma |
| +tsa | +tsha | +dza | +wa |
| +zha | +za | 'a | ya |
| ra | la | śha | sa |
| ha | "a | | |

| VOWELS | | | | |
|---|---|---|---|---|
| a | i | e | o | |
| "inverted" i | | | | |
| SUBSCRIBED CONSONANTS | | | | |
| ya-btags | ra-btags | | w'a-zur | |

# REFERENCES

*5D Pentaglott dictionary* see *Xan-i araxa*

ADAMS, D. Q. (1988). *Tocharian historical phonology and morphology*. American Oriental Series 71. New Haven. [See the review of RINGE, D. A. (1990). Language 66, 400-408.]

ALMÁSSY Gy. (1901). *Kara-kirgiz nyelvészeti jegyzetek*. KSz 2, 108-122.

ARAT, R. R. (1957, 1979r.). *Kutadgu bilig I. Metin* [repr. ed. in: *Türk Dil Kurumu Yayinlari* vol. 458]. Ankara.

- (1979). *Kutadgu Bilig III. Indeks*. Istanbul.

BAILEY, H. W. (1937-39). *Hvatanica III*. BSOAS 9, 521-543.

- (1938). *Indo-Turcica*. BSOAS 9, 289-302.
- (1939). *Vajra-prajña-paramitā*. ZDMG 92 (N.F.17), 579-593.
- (1940-42). *Ttagutta*. BSOAS 10, 599-605.
- (1943-46). *Iranica*. BSOAS 9, 1-5.
- (1943-47). *A Turkish-Khotanese vocabulary*. BSOAS 11, 290-296.
- (1945, 1969, 1980r). *Khotanese Texts I-III*. London.
- (1945-1967). *Khotanese texts I-VI*. Cambridge. (I-III ed. 2 in one vol. 1969, repr. 1980, IV repr. 1979, V repr. 1980).
- (1949-50). *A Khotanese text conceming the Turks in Khantsou*. Asia Major N. S. 1, 28-52.
- (1950). *The Tumshuq Karmavācanā*. BSOAS 13, 649-670, 809-810.
- (1951-52). *The Staël-Holstein miscellany*. Asia Major. N. S. 2, 1-45.
- (1958). *Languages of the Saka*: Handbuch der Orientalistik IV. Band Iranistik, I. Abschnitt Linguistik. Leiden-Köln, 131-154.
- (1960). *Saka Documents I*. London.
- (1973a). *Dictionary of Khotan Saka*. Cambridge.
- (1973b). *Taklamaklan miscellany*. BSOAS 36, 224-227 + VIII plates.
- (1979). *Dictionary of Khotan Saka*. Cambridge.
- (1982). *The culture of the Sakas in ancient Iranian Khotan*. New York.
- (1985). *Indo-Scythian studies (Khotanese texts vol VII)*. Cambridge.

BANG, W.; GABAIN, A. von and RACHMETI, G. R. (1934). *Das buddhistische Sutra Sākiz Yükmek*. Sitzungsberichte der Preussischen Akademie der Wissenschaften 1934, 93-192.

BASKAKOV, N. A. (1952). *Klassifikacija tjurkskih jazykov v svjazi s istoričeskoj periodizaciej ih razvitija i formirovanija*. Trudy Instituta jazykoznanija AN SSSR 1, 7-57.

- (1962, 1969). *Vvedenie v izučenie tjurkskih jazykov*. Moskva.

BAZIN, L. (1974). *Les calendriers turcs anciens et medievaux*. [Paris.]

BENZING, J. (1953). *Einführung in das Studium der altaischen Philologie und der Turkologie*. Wiesbaden.

BLOOMFIELD, L. (1933). *Language*. New York.

BROCKELMANN, C. (1928. 1964r). *Mitteltürkischer Wortschatz*. Budapest.

CERETELI, G. V. (1948-49). *Armazskoe pis'mo i problema prishoždenija gruzinskogo alfavita I. II.* Épigrafia Vostoka 2, 90-101; 3, 59-71.

CINCIUS, V. I. (1975. 1977). *Sravnitel'nyj slovar' tunguso-man'čžurskih jazykov*. Materialy k etimologičeskomu slovarju. I-II. Leningrad.

CLAUSON, G. (1931). *The geographical names in the Staël-Holstein Scroll.* JRAS 297-309.

- (1957). *The Ongin inscription.* JRAS 177-192.

- (1962). *Turkish and Mongolian studies.* London.

- (1967). *Three notes on Early Turkish.* Türk Dili Araştirmalari Yilliği Belleten 1966 [sic. published in 1967], 1-37.

- (1970). *The origin of the Turkish "Runic" alphabet.* Acta Orientalia Havniae 32, 51-76.

- (1973). *The Turkish-Khotanese Vocabulary reedited*: Zeki Velidi Togan'in armağani. Islam Tetikleri Enstitüsü Dergisi. 5, 37-45.

COUVREUR, W. (1965). *Koetsjische schrifttabellen in Slanting Gupta.* Orientalia Gandensia 2, 111-143.

CSONGOR, B. (1960). *Some Chinese texts in Tibetan script from Tun-huang.* AOH 10, 97-140.

- (1972). *A Chinese Buddhist text in Brāhmi script.* Unicorn (Chi-lin) 10. 36-77.

DANKOFF, R. and KELLY, J. (1982-1985). *Mahmud al Kāšγari. Compendium of the Turkic Dialects.* Harvard Univ. Printing Office. Sources of the Oriental languages and literatures 7.

*DK* see DANKOFF, R. and KELLY, J. (1982-1985).

DOERFER, G. (1963-1975). *Türkische und mongolische Elemente im Neupersischen I-IV.* Wiesbaden.

- (1971a). *Bemerkungen zur Methodik der türkischen Lautlehre.* OLZ 66, 326-344.

- (1971b). *Khalaj materials.* Indiana University Publications. Uralic and Altaic Series 115. Bloomington.

- (1975-76). *Proto-Trukic: Reconstruction problems.* Türk Dili Araştirmalari Yilliği Belleten 1-59.

DONNER, K. (1924). *Zu den ältesten Berührungen zwischen Samojeden und Türken.* JSFOu 40, 3-42.

- (1896). *Sur l'origin de l'alphabet turc du Nord de l'Asie.* JSFOu 15:1, 21-46.

DUL'ZON, A. P. (1961). *Slovarnye materialy XVIII v. po ketskim narečijam.* Učenye Zapiski. Tomskij Gosudarstvennyj Pedagogičeskij Institut. 19:2, 152-189.

EDGERTON, F. (1953). *Buddhist hybrid Sanskrit grammar and dictionary.* New Haven - London.

*EDict* see Clauson (1972).

EMMERICK, R. E. (1971). *Saka Documents V.* London.

- (1979a). *The vowel phonemes of Khotanese*: Amsterdam Studies in the theory and history of linguistic science IV. Current issues in linguistic theory 11. 239-250.

- (1979b). *A guide to the literature of Khotan*: Studia philologica Buddhica. Occasional paper series III. The Reyukai Library. Tokyo.

- (1981). *The consonant phonemes of Khotanese*: Monumentum Georg Morgenstierne I. Leiden, 185-209.

- (1984). *Research on Khotanese: A survey (1979-1982)*: Middle Iranian Studies. Ed. W. Skamowski and A. van Tangerloo. Leuven, 127-145.

- (1985). *The Tumshuqese Karmavācanā text.* Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse. Akademie der Wissenschaften und der Literatur. Mainz. Jhrg. 1985. Nr.2. Stuttgart.

- (1987a). *Buddhism in Central Asia*: The Encyclopedia of Religion. vol 2. Ed. M. Eliade. New York. 400-404.
- (1987b). *The transition from Old to Late Khotanese*: Transition periods in Iranian history. Actes du symposium de Fribourg-en-Brisgau (22-24 Mai 1985). Studia Iranica 5. 33-42.
- (1989). *Khotanese and Thumshuqese*: Compendium Linguarum Iranicarum. Ed. R. Schmitt. Wiesbaden, 204-229.

EMMERICK, R. E. and SKJAERVO, P.O. (1982, 1985). *Studies in the vocabulary of Khotanese I. II.* Wien.

EMRE, A. C. (1938). *Sur l'origine de l'alphabet vieux-turc*. Istanbul.

FUJIEDA, A. (1966. 1969). *The Tunhuang manuscripts: a general description*. Zinbun. Memoires of the Research Institute for Humanistic Studies. Kyoto University 9, 10. 1-32, 17-39.

GABAIN, A. von (1948). *Alttürkisches Schrifttum*. Sitzungsberichte der Preussischen Akademie der Wissenschaften. Berlin. Philosophisch-historische Klasse. 1948.
- (1954a). *Türkische Turfan-Texte VIII. Texte in Brāhmi-Schrift*. Abhandlungen der Deutschen Akademie der Wissenschaften in Berlin. Jhrg. 1952. No. 7. Berlin.
- (1954b). *Buddhistische Türkenmission*: Asiatica. Weller-Festschrift. Leipzig. 161-173.
- (1955). *Alttürkische Datierungsformen*. UAJb 27. 191-203.
- (1964). *Die Schreiber der alt-türkischen Brāhmi-Texte*. Studia Orientalia 28:5, 3-11.
- (1967). *Die Drucke der Turfan-Sammlung*. Sitzungsberichte der Deutschen Akademie der Wissenschaften. Klasse für Sprache. Literatur und Kunst. Jhrg. 1967. 1, 1-40.
- (1973a). *Das Leben im uigurischen Königreich von Qočo (850-1250)*. Textband. Veröffentlichungen der Societas Uralo-Altaica. Bd. 6. Wiesbaden.
- (1973b). *Das uigurische Königreich von Kotscho 850-1250*. Sitzungsberichte der Deutschen Akademie der Wissenschaften. Klasse für Sprachen, Literatur und Kunst. Jhg 1961. Nr. 5. Berlin.
- (1974). *Alttürkische Grammatik*. 3. Aufl. Wiesbaden.

GAMKRELIDZE, T. and IVANOV, Vjač. Vs. (1984). *Indoevropejskij jazyk i indoevropejcy. Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokultury I-II*. Tbilisi.

GREENBERG, J. H. (1957). *Essays in Linguistics*. Chicago.

GERCENBERG, L. G. (1965). *Hotanosakskij jazyk*. Moskva.
- (1981). *Hotanosakskij jazyk*: Osnovy iranskogo jazykoznanija. Sredneiranskie jazyki. otv. red. V. S. Rastorgueva. Moskva. 233-313.

*H* see HOVDHAUGEN, E. (1971).

HACKIN, J. (1924). *Formulaire sanscrit-tibétain du 10e siècle*. Paris.

HAJDÚ P. and DOMOKOS P. (1978). *Uráli nyelvrokonaink*. Budapest.

HAJDÚ P.; KRISTÓ Gy. and RÓNA-TAS A. (eds.) (1980). *Bevezetés a magyar őstörténet kutatásának forrásaiba III*. Budapest.

HAMILTON, J. R. (1958). *Autour du manuscrit Staël-Holstein*. TP 46, 115-159.
- (1977). *Nasales instables en turc khotanais du Xe siècle*. BSOAS 40, 508-521.
- (1979). *Les regnes khotanais entre 851 et 1001*: Contributions aux études sur Touen-houang. Genève-Paris. 49-54.

(1981). *Les manuscrits turcs anciens de la grotte murée de Touen-Hoang*: Manuscrits et inscriptions de Haute Asie du Ve au XIe siècle. Colloque international. organisé par l'ERA 438 (Paris. 2-4 octobre 1979). Actes. Paris. 15-17.

HAUSSIG, H. W. (1988). *Ueberlegungen zum Namen der Türken*: Turcica and Orientalia [Jarring Festschrift] Swedish research Institute in Istanbul. Transactions 1. Stockholm. 45-50.

HAZAI, G. and ZIEME, P. (1970). *Zu einige Fragen der Bearbeitung türkischer Sprachdenkmäler*. Acta Orientalia Havniae 32, 125-140.

HENNING, W. B. (1949-50). *The name of the "Tokharian" language*. Asia Major N. S. 1, 158-162.

HINÜBER, O. v. (1988). *EMMERICK, Ronald E.: The Tumshuqese Karmavācanā Text*. Mainz [1985]. OLZ 83, 619-620.

HITCH, D. (1981). *Central Asian Brāhmī Paleography: The relationships among the Tocharian. Khotanese and Old Turkic Gupta scripts*. A thesis submitted to the Faculty of Graduate Studies in partial fulfilment of the requirements for the degree of Master of Arts. Calgary. Alberta (unpublished).

- (1983). *A Brāhmi-Manichean hybrid script*. Orientalia Lovanensia Periodica 14, 293-312.
- (1984). *Kharosti influences on the Saka Brāhmī scripts*: Middle Iranian studies. Proceedings of the International Symposium organized by the Katholjeke Universiteit Leuven. Leuven. 187-202.
- (1987a). *Bactrian Brāhmī*. Paper read at the 30th PIAC meeting. Bloomington, 1987 June (unpublished).
- (1987b). *Tumshuqese and Turkic Brāhmī ru*. Journal of Turkish Studies 11, 249-252.
- (1989a). *Brāhmī*: Encyclopaedia Iranica IV. fasc. 4. London & New York. 432-433.
- (1989b). *Old Turkic 'b' and 'p'*: Religious and lay symbolism in the Altaic World and other papers. Proceedings of the 27th meeting of the PIAC. Ed. K. Sagaster. Wiesbaden. 130-146.

HOVDHAUGEN, E. (1971). *Turkish words in Khotanese texts. A linguistic analysis*. Norsk Tidskrift for Sprogvidenskap 24. 163-209.

- (1974). *The relationship between the two Orkhon inscriptions*. Acta Orientalia Havniae 36, 55-82.

*Inscriptions* (1892). *Inscriptions de l'Orkhon. Recueillies par l'expedition finnoise 1890*. Helsinkfors.

ILLIČ-SVYTIČ, V. M. (1971). *Opyt sravnenija nostratic"eskih jazykov. (Semitohamijskij, kartvelskij, indoevropejskij, ural'skij, dravidijskij, altajskij)*. Vvedenie. Sravnitel'nyj slovar I. [II. 1976. III. 1984]. Moskva.

ISBAERT, L. (1983). *Der Beitrag der indischen Lehnwörter zu Problemen der tocharischen Phonologie*: Sprachen des Buddhismus in Zentralasien. Ed. K. Röhrborn. and W. Veenker. Wiesbaden. 35-43.

IVANOV, V. V. (1988). *K probleme toharo-altajskih leksic"eskih svjazej*. Voprosy Jazykoznanyja 1988:4, 99-102.

JAKOBSON, R.; HÜTTL-WORTH, G. and BEEBE, J. (1957). *Paleosiberian peoples and languages*. New Haven.

JANHUNEN, J. (1977). *Samoyed-Altaic Contacts. Present state of research*. MSFOu 158, Helsinki, 123-129.

JARRING, G. (1941). *Zurufe an Tiere bei den Osttürken*: Die Welt des Islams. Sonderband. Festschrift Friedrich Giese. Berlin. 46-52.

- (1964). *An Eastern Turki-English Dictionary*. Lunds Univeristets Årskrift N. F. Avd. 1. Bd. 56. Nr. 4. Lund.

JOHANSEN, U. (1987). *Zur Geschichte des Schamanismus*: Synkretismus in den Religionen Zentralasiens. Hrsg. W. Heissig. H.-J. Klimkeit. Studies in Oriental Religions 13. Wiesbaden. 8-22.

JOHANSON, L. (1979). *Alttürkisch als "dissimilierende Sprache"*. Wiesbaden.

*K (Kāshgarī)* see DANKOFF, R. and KELLY, J. (1982-1985).

*Kh = Khotanese*

KLJAŠTORNYJ, S. G. (1964). *Drevnetjurkskie runičeskie pamjatniki kak istočnik po istorii Srednej Azii*. Moskva.

- (1980). *Terhinskaja nadpis'*. Sovetskaja Tjurkologija 1980:3, 82-95.
- (1982). *The Terkhin inscription*. AOH 26, 335-366.
- (1985). *The Tes inscription of the Uighur Bögü Qaghan*. AOH 39, 137-156.

KLJAŠTORNYJ, S. G. and LIVŠIC, V. A. (1972). *The Sogdian inscription of Bugut revised*. AOH 26, 69-102.

KONONOV, A. N. (1980). *Grammatika jazyka tjurkskih runičeskih pamjatnikov VII-IX vv*. Leningrad.

KONOW, St. (1948). *The Khotanese text of the Staël-Holstein Scroll*. Acta Orientalia Havniae 20, 133-160.

- (1935). *Ein neuer Saka-Dialekt*. Sitzungsberichte der preussischen Akademie der Wissenschaften. Philosophisch-Historische Klasse. 20. Berlin.
- (1941-42). *Note sur une nouvelle forme aberrante du khotanais*. JA 233, 83-102.
- (1947). *The oldest dialect of Khotanese Saka*. Norsk Tidskrift for Sprogvidenskap 14, 156-190.

KORMUŠIN, I. V. (1975). *K osnovnym ponjatijam tjurkskoj runičeskoj paleografii*. Sovetskaja Tjurkologija 1975:2, 25-47.

KŐHALMI, K. U.-. (1953). *Über die pfeifenden Pfeile der innerasiatischen Reitemomaden*. AOH 3, 45-71.

- (1956). *Der Pfeil bei den innerasiatischen Reitemomeden und ihren Nachbarn*. AOH 6, 109-161.
- (1962). *Der Abschnitt der Waffenbehälter und des Waffengürtels in den polyglotten Wörterbüchern der Ch'ing-Epoche*. AOH 15, 195-206.
- (1972). *A steppék nomádja lóháton. fegyverben*. Kőrösi Csoma Kiskönyvtár 12. Budapest.

KRAUSE, F. and THOMAS, W. (1960). *Tocharisches Elementarbuch. Bd. I. Grammatik*. Heidelberg.

KYZLASOV, L. R. (1960). *Novaja datirovka jenisejskoj pis'mennosti*. Sovietskaja Arheologiia 1960:3, 93-120.

- (1965). *O datirovke pamjatnikov jenisejskoj pis'mennosti*. Sovetskaja Arheologija 1965:3, 38-49.

LANE, G. S. (1967). *On the interrelationship of the Tocharian dialects*: Studies in historical linguistics in honor of George Sherman Lane. University of North Carolina Studies in the Germanic Languages and Literatures No. 58. 105-134.

LAUFER, B. (1914). *Bird divination among the Tibetans*. TP 15, 1-110.

LE COQ, A. A. (1909). *Köktürkisches aus Turfan*. Sitzungsberichte der Preussischen Akademie der Wissenschaften 1909, 1047-1061.

LEWICZKI, M. (1936). *O tekście sanskrycko-tureckim w pismie brāhmī wydanym przez Stönnera*. Rocznik Orientalisticzny 12. 194-208.

LI XUE WEI and STUART, K. (1989). *Population and culture of the Mongols, Tu, Baoan, Dong-xiang and Yugu in Gansu* Mongolian Studies 12, 71-93.

LIGETI L. (1949). *L'étymologie mongole du slave chorǫgy 'drapeau'*. Études Slaves et Roumaines 2, 45-56.

- (1950). *Mots de civilisation de Haute Asie en transcription chinoise*. AOH 1, 141-188.

- (1953). *Az uráli és altaji nyelvek viszonyának kérdése*. Az MTA I. Osztályának Közleményei 4, 333-364.

- (1961). *Trois notes sur l'écriture 'phags-pa*. AOH 13, 201-237.

- (1963). *Uráli török jövevényszavaink kérdéséhez*. MNy 59, 381-393.

- (1966). *Un vocabulaire sino-ouigur des Ming. Le Kao-tch'ang houan yi-chou du Bureau des Traducteurs*. AOH 19, 117-199.

- (1971). *A propos du "rapport sur les rois demeurant dans le nord"*: Études tibétaines dédiées a la mémoire de Marcelle Lalou. Paris. 166-189.

LIVŠIC, V. A. (1990). *O proishoždenii drevnetjurkskoj runičeskoj pis'mennosti*. Sovetskaja Tjurkologija 1978:4, 84-98.

*LKh = Late Khotanese*

MACKENZIE, D. N. (1986). *A concise Pahlavi dictionary*. Second ed. London.

MALOV, S. E. (1957). *Jazyk želtyh ujgurov*. Alma-Ata.

MAUE, D. (1983). *Zu den Dentalen im Brāhmi-Uigurischen*: Sprachen des Buddhismus in Zentralasien. Ed. K. Röhrborn. and W. Veenker. Wiesbaden. 53-64.

- (1984). *Vorläufige Bemerkungen zu den Gutturalgraphemen in der alttürkischen Brāhmi*. UAJb N. F. 4, 90-96.

- (s. a.). *Dissertation* (unpublished). s.l.

MAUE, D. and RÖHRBORN, K. (1976). *Ein zweisprachiges Fragment aus Turfan*. CAJ 20, 208-221. See also CAJ 22 [1978], 134-135.

- (1984-85). *Ein "buddhistischer Katechismus" in alttürkischer Sprache und tibetischer Schrift*. ZDMG 134, 286-313; 135, 68-91.

MAYER, A. L. (1990). *Die Gründunkslegende Khotans*: Buddhistische Erzählliteratur und Hagiographie in türkischer Überlieferung. ed. J. P. Laut and K. Röhrborn (Veröffentlichungen der Societas Uralo-Altaica Bd. 27.). Wiesbaden, 37-65.

MENGES, K. H. (1951). *The Oriental elements in the vocabulary of the oldest Russian epos. The Igor' Tale "Slovo o P"lku Igoreve"*. Supplement to Word. New York.

- (1968). *The Turkic languages and peoples*. Wiesbaden.

- (1979). *Vostočnye élementy v "Slove o polku Igoreve"*. Leningrad. [Updated transl. of Menges 1951].

MEYER, I. R. (1965-66). *Bemerkungen über Vokal- und Schriftsystem des Runentürkischen*. Acta Orientalia Havniae 29:1-2, 183-202.

MILLER, R. A. (1966). *Early evidence for vowel harmony in Tibetan*. Language 42, 252-277.
- (1968). *Review of Róna-Tas (1966)*. Language 44, 252-277.
MORIYASU, T. (1980). *La nouvelle interpretation des mots Hor et Ho-yo-hor dans le manuscrit Pelliot tibetain 1283*. AOH 34, 171-184.
- (1985). *Étude sur un cathecisme bouddhique ouigour en écriture tibetaine (P.t. 1292)*. Memoirs of the Faculty of Letters. Osaka University 24, 1-85 [in Japanese].
MÜLLER, F. W. K. (1908). *Uigurica I*. Abhandlungen der Preussischen Akademie der Wissenschaften 1908 Nr. 2, 3-60.
NÉMETH, J. (1926). *Der Volksname türk*. KCsA 2, 275-281.
- (1942-1947). *Probleme der türkischen Urzeit*: Analecta Orientalia memoriae Alexandri Csoma de Kőrös dicata. Bibliotheca orientalis Hungarica V. Budapest. 57-102.
PEDERSEN, H. (1903). *Türkische Lautgesetze*. ZDMG 57, 535-561.
PELLIOT, P. (1959). *Notes on Marco Polo I-II*. Paris.
POLIVANOV, E. O. (1925). *Ideografic'eskij motiv v formacii orhonskogo alfavita*. Bjuleten Sredneaziatskogo gosudarstvennogo Instituta 1925:9, 177-179.
POPPE, N. (1965). *Introduction to Altaic linguistics*. Wiesbaden.
PRITSAK, O. (1980). *Turkology and the comparative study of the Altaic languages. The system of the Old Runic script*. Journal of Turkish Studies 4, 83-100.
PULLEYBLANK, E. G. (1954). *The date of the Staël-Holstein Roll*. Asia Major N. S. 4, 90-97.
- (1961-1962). *The consonantal system of Old Chinese*. Asia Major N. S. 9, 58-142; 206-265.
RADLOFF, W. (1892-1899). *Atlas der Altertümer der Mongolei*. Petersburg.
RAMSTEDT, G. J. (1946-47). *The relation of the Altaic languages to other language groups*. FUF 53:1, 15-26.
- (1957) *Einführung in die altaische Sprachwissenschaft I. Lautlehre*. MSFOu 104:1. Helsinki.
RÁSONYI, L. (1957). *Review on Benzing (1953)*. AOH 7, 309-320.
RÄSÄNEN, M. (1949). *Materialien zur Lautgeschichte der türkischen Sprachen*. Studia Orientalia 15.
- (1953). *Contributions au classement des langues turques*. Rocznik Orientalisticzny 17, 92-104.
REDHOUSE, J. W. (1890). *A Turkish and English Lexicon*. Constantinople.
ROUX, J.-P. (1956). *Tängri. Essai sur le ciel-dieu de peuple altaïques*. RHR 149, 49-82; 197-250; 150, 27-54; 173-212.
- (1958). *Notes additionelles à Tängri*. RHR 154, 48-50.
- (1962). *La religion des turcs de l'Orkhon des VII^e et VIII^e siècles*. RHR 161. 1-24, 199-231.
RÓNA-TAS, A. (1956). *Tally-stick and divination-dice in the iconography of Lha-mo*. AOH 6, 163-179.
- (1966). *Tibeto-Mongolica. The Tibetan loanwords of Monguor and the development of the Archaic Tibetan dialects*. Budapest-The Hague.
- (1968). *A brief note on the chronology of the Tun-huang collections*. AOH 21, 313-316.
- (1970). *Some problems of Ancient Turkic*. Acta Orientalia Havniae 32, 209-229.
- (1973). *Did the Proto-Altaic people know the stirrup?* Studia Mongolica 1973:13, 169-171.

- (1974a). *Tocharische Elemente in den altaischen Sprachen?*: Sprache, Geschichte und Kultur der altaischen Völker. Protokollband der XII. Tagung der Permanent International Altaistic Conference 1969 in Berlin. Schriften zur Geschichte und Kultur des alten Orients 5. Berlin. 499-504.
- (1974b). *Obščee nasledie ili zaimstvovanija? (K probleme rodtsva altajskih jazykov)*. Voprosy Jazykoznanija 1974:2, 31-45.
- (1976). *On the meaning of "Altaic"*: Tractata Altaica. Denis Sinor sexagenario optime de rebus altaicis merito dedicata. red. W. Heissig etc. Wiesbaden. 549-556.
- (1978). *A nyelvrokonság. Kalandozások a történeti nyelvtudományban*. Budapest.
- (1980). *On the earliest Samoyed-Turkish contacts*: Congressus Quintus Internationalis Fenno-Ugristarum III. Turku. 377-385.
- (1981). *The character of the Hungarian--Bulgaro-Turkic relations*: Studia Turkologica 5. ed. Gy. Káldy-Nagy. 119-128.
- (1982a). *The periodization and sources of Chuvash linguistic history*: Chuvash studies. Ed. A. Róna-Tas. Bibliotheca Orientalis Hungarica 28. Asiatische Forschungen 79. Budapest-Wiesbaden. 113-169.
- (1982b). *On the history of the Turkic and Finno-Ugrian affricates*. AOH 36, 429-447.
- (1983a). *De hypothesi Uralo-Altaica*. MSFOu 185, 235-251.
- (1983b). *Linguistic notes on an Amdowa text*: Contributions on Tibetan Language, history and culture. Ed. E. Steinkellner. H. Tauscher. Wien. 243-280.
- (1983c). *"Rokonságszerű ősi kapcsolat". Németh Gyula nézetei az uráli és a török nyelvek viszonyáról*: Urálisztikai Tanulmányok. Ed. Berecki G. and Domokos P. Budapest, 331-346.
- (1984a). *Some remarks on the vowel system of Spoken Tibetan*: Tibetan and Buddhist Studies. Ed. L. Ligeti. Budapest. 215-231.
- (1984b). *The daughter of Dru-gu Ge-sar* (Paper read at the Csoma de Kőrös Symposium, Visegrád 1984. Yet unpublished).
- (1985). *Wiener Vorlesungen zur Sprach- und Kulturgechichte Tibets*. Wiener Studien zur Tibetologie und Buddhismuskunde 13. Wien.
- (1986). *Language and history. Contributions to comparative Altaistics*. Studia Uralo-Altaica 25. Szeged.
- (1987a). *Materialien zur alten Religion der Türken*: Synkretismus in den Religionen Zentralasiens. Ed. W. Heissig and H.-J. Klimkeit. Studies in Oriental Religions 13. Wiesbaden 33-45.
- (1987b). *On the development and origin of the East Turkic "Runic" script*. AOH 41, 7-14.
- (1988a). *Ethnogenese und Staatsgründung. Die türkische Komponente in der Ethnogenese des Ungarntums*: Studien zur Ethnogenese. Bd. 2. Rheinisch-Westfälische Akademie der Wissenschaften. Abhandlungen 78. Düsseldorf. 107-142.
- (1988b). *Turkic influence on the Uralic languages*: The Uralic languages. Description. History and Foreign Influences - Handbuch der Orientalistik. Achte Abteilung - Handbook of Uralic Studies. Ed. D. Sinor. Leiden. 742-780.
- (1988c). *Tibetan linguistics. History. State-of-the-Art. Tasks*: Tibetan studies. Proceedings of the 4th Seminar of the International Association for Tibetan Studies

Schloss Hohenkammer - Munich 1985. Ed. H. Uebach and I. L. Panglung. Studia Tibetica II. München. 399-403.

- (1990). *Altajskij i indoevropejskij (Zametki na poljah knigi T. V. Gamkrelidze i Vjač. Vs. Ivanova).* Voprosy Jazykoznanyja 1990:1, 26-37.

RÖHRBORN, K. (1977-1981). *Uigurisches Wörterbuch. Sprachmaterial der vorislamischen Türkischen Texte aus Zentralasien.* Lfg 1. (1977). 2 (1979). 3 (1981)

- (1981-1982). *Zu einem Dialekt-Differenzierenden Lautübergang im Alttürkischen.* Materialia Turcica 7-8, 295-305.

SAMOJLOVIČ, A. N. (1922). *Nekotorye dopolnenija k klassifikacii tureckih jazykov.* Petrograd.

SANDER, L. (1968). *Paläographisches zu den Sanskrithandschriften der Berliner Turfansammlung.* Verzeichnis der Orientalischen Handschriften in Deutschland. Supplementband 8. Wiesbaden.

- (1983). *Einige neue Aspekte zur Entwicklung der Brāhmī und Bamiyan (ca 2.-7. Jh. n. Chr.):* Sprachen des Buddhismus in Zentralasien. Hrsg. K. Röhrborn. and W. Veenker. Veröffentlichungen der Societas Uralo-Altaica 16. Wiesbaden. 113-124.

- (1984). *Zu dem Projekt "Paläographie khotan-sakischer Handschriften":* Middle Iranian studies. Proceedings of the International Symposium organized by the Katholjeke Universiteit Leuven. Leuven. 159-186.

SCHLEICHER, A. (1871). *Compendium der vergleichenden Grammatik der indogermanischen Sprachen.* 3. Aufl. Weimar.

SCHMIDT, J. (1872). *Die Verwandschaftsverhältnisse der indogermanischen Sprachen.* Weimar.

*SDD* see *Türkiye'de*

SERTKAYA, O. F. (1990). *Do the Köktürk characters have an alphabetical order?* [Ms. Paper read at the 33rd meeting of the PIAC. Budapest. 1990 June.] Istanbul.

SEVORTJAN, É. V. (1962). *Affiksy glagoloobrazovanija v azerbajdžanskom jazyke.* Moskva.

- (1966). *Affiksy imennogo slovoobrazovanija v azerbajdžanskom jazyke.* Moskva.

- (1971). *O soderžanii termina "obšeturjurkskij".* Sovetskaja Tjurkologija 1971:2, 3-12.

SCHMIDT, K. T. (1988). *Ein Beitrag des Tocharischen zur Entzifferung des Tumšuqsakischen.* Altorientalischen Forschungen 15:2, 306-314.

ŠČERBAK, A. M. (1970). *Sravnitel'naja fonetika tjurkskih jazykov.* Leningrad.

SIMS-WILLIAMS, N. (1981). *The Sogdian sound-system and the origins of the Uighur script:* Manuscrits et inscriptions de Haute Asie du V$^{e}$ au XI$^{e}$ siècle. Colloque international. organisé par l'ERA 438 (Paris. 2-4 octobre 1979). Actes. Paris. 347-360.

SINOR, D. (1988). *The problem of the Ural-Altaic relationship:* The Uralic languages. Ed. D. Sinor. Leiden. 706-741.

SPANIEN, A. and IMAEDA, Y. (1979). *Choix de documents tibétains conservés à la bibliothéque Nationale complété par quelques manuscrits de l'India Office et du British Museum II.* Paris.

STÖNNER, H. (1904). *Zentralasiatische Sanskrittexte in Brāhmischrift aus Idiqutsahri. Chinesisch-Turkistan. I.* Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften 44, 1282-1290. (Anhang: Uigurische Fragmente in Brāhmischrift, 1288-1290.)

- (1905). *Über die Kultur- und sprachgeschichtliche Bedeutung der Brāhmitexte in der Turfan-Handschriften.* Zeitschrift für Ethnologie 1905, 415-420.

STRAHLENBERG, Ph. J. (1730, 1975r). *Das Nord und Ostliche Theil von Europa und Asia.* Stockholm.

SÜMER, F. (1983). *Türkler'de atcılık ve bincilik*. [Istanbul].

TAKATA, T. (1988). *Chinese words in Khotanese texts - A phonological analysis*: Kangoshi no shomondai. [Kyoto]. 71-128.

TAUBE, M. (1980). *Die Tibetica der Turfansammlung. Schriften zur Geschichte und Kultur des Alten Orients*. Berliner Turfantexte 10. Berlin.

TEKIN, Š. (1980). *Maitresimit nom bitig. Die uigurische Uebersetzung eines Werkes der buddhistischen Vaibhasika-Schule. I-II*. Berlin.

- (1985). *Some thoughts on the etymology of the Turkish üzeŋgi 'stirrup'*. Journal of Turkish Studies 9. [Cleaves-Festschrift]. 237-241.

TEKIN, T. (1968). *A grammar of Orkhon Turkic*. Uralic and Altaic series. 69. The Hague.

TENIŠEV, É. R. (1971). *Pereboj s/š v tjurskih runiceskih pamjatnikah*: Struktura i istorija tjurskih jazykov. Moskva, 289-295.

- (1976a). *O naddialektnom haraktere jazyka tjurkskih runiceskih pamjatnikov*: Turcologica. K semidesjatiletiju akademika A. N. Kononova [Kononov-Festschrift]. Leningrad. 164-172.

- (1976b). *Stroj saryg-ujgurskogo jazyka*. Moskva.

THOMAS, F. W. (1931) see (1951).

- (1937). *A Buddhist Chinese text in Brāhmī script*. ZDMG 91 (N. F. 16), 1-48.

- (1939). *Notes in consideration of Professor Bailey's critical observations. Postscript*. ZDMG 92 (N. F.17), 594-610.

- (1951). *Tibetan literary texts and documents concerning Chinese Turkestan. Part II. Documents*. London.

- (1954). *Brāhmī script in Central-Asian Sanskrit manuscripts*: Asiatica. Festschrift Friedrich Weller. Leipzig. 667-700.

- (1985). *Die Erforschung des Tocharischen (1960-1984)*. Stuttgart.

THOMAS, F. W. and CLAUSON, G. L. M. (1927). *A second Chinese Buddhist text in Tibetan characters*. JRAS 1927, 281-306.

THOMAS, F. W. and KONOW, St. (1929). *Two Mediaeval documents from Tun-huang*. Oslo Etnografisk Museums Skrifter Bd. III. Hft 3. Oslo.

THOMSEN, V. (1893). *Déchiffrement des inscriptions de l'Orkhon et de l'Iénisséi. Notice préliminaire*. Bulletin de l'Académie Royal des Sciences et des Lettres de Danemark. 285-299.

- (1894-96/1922) *L'alphabet runiforme turc*. MSFOu 5. Reprinted with additions: Samlede Afhandlinger III. København. 27-82.

- (1896) *Inscriptions de l'Orkhon déchiffrées*. MSFOu 5. Helsinki.

- (1913). *Une lettre méconnue des inscriptions de l'Iénissei*. JSFOu 30, 1-9.

- (1922). *Samlede Afhandlinger III*. København.

*TMEN* see DOERFER (1963-1975).

TRUBETZKOY, N. S. (1933). *Gedanken über das Indogermanenproblem*. Acta Linguistica Havniae 1, 81-89.

TUGUŠEVA, L. Ju. (1980). *Fragmenty ujgurskoj versii biografii Sjuan'-Czana*. Moskva.

*Türkiye'de* (1963-). *Türkiye'de halk ağzïdan derleme sözlüğü*. Ankara.

URAY, G. (1988). *Eine Einführung in die frühen tibetischen Quellen der Geschichte und Kultur Tibets bis zum Anfang des 11. Jh's*. [Manuscript]. Wien.

VASIL'EV, D. D. (1976). *Pamjatniki tjurkskoj runiceskoj pis'mennosti aziatskogo areala*. Sovetskaja Tjurkologija 1976:1, 71-81.

- (1978). *Pamjatniki tjurkskoj runič'eskoj pis'mennosti aziatskogo areala II.* Sovetskaja Tjurkologija. 1978:5, 92-95.
- (1983a). *Grafič'eskij fond pamjatnikov tjurkskoj runič'eskoj pis'mennosti aziatskogo areala.* Moskva.
- (1983b). *Korpus tjurkskih runič'eskih pamjatnikov bassejna Eniseja.* Leningrad.

VDOVIN, I. S. and TEREŠČENKO, N. M. (1959). *Očerki istorii izučenija paleoaziatskih i samodijskih jazykov.* Leningrad.

VOROBEV-DESJATOVSKIJ, V. S. (1958). *Pamjatniki central'noaziatskoj pis'mennosti.* Učennye zapiski Instituta Vostokovedenija 16, 280-308.

WINDEKENS, A. J. von. (1976). *Le tokharien confronté avec les autres langues indo-européennes I. La phonetique et la vocabulaire.* Louvain.

WITSEN, N. (1692. 1705. 1785) *Noord and Oost Tartarye.* Amsterdam.

*Xan-i araxa.* (1957). *Xan-i araxa sunja xacin-i xergen qamcixa manju gisun-i buleku bitxe I-III.* Peking.

YUDAHIN, K. K. (1965). *Kirgizsko-russkij slovar'.* Moskva.

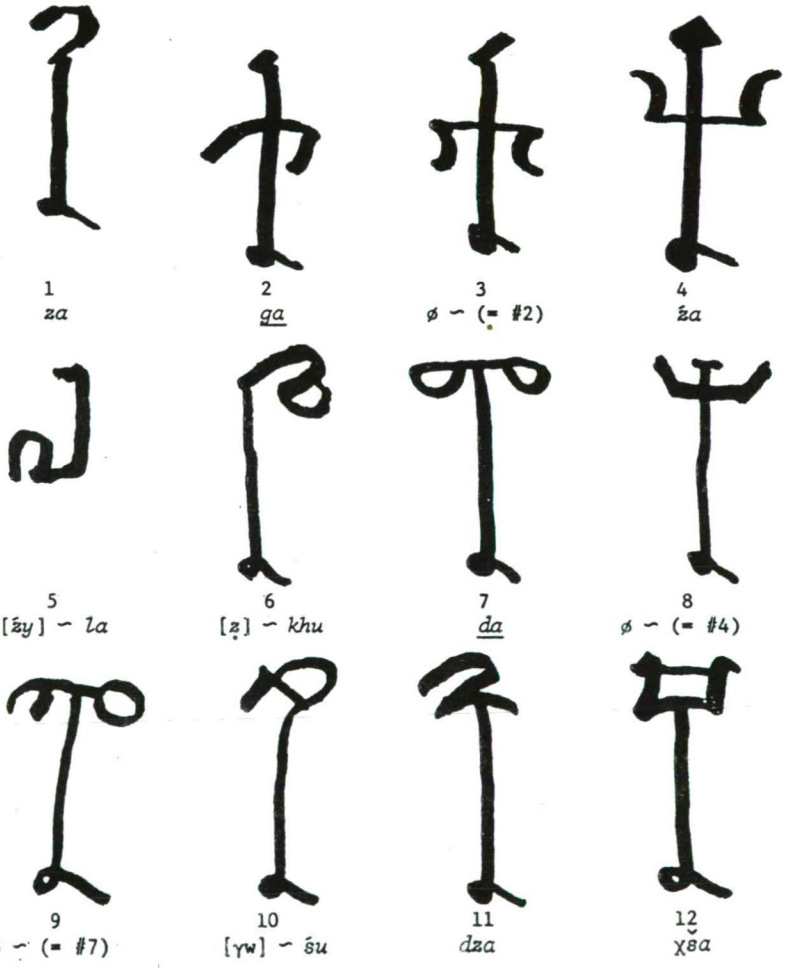
ZIEME, P. (1984). *Zur Verwendung der Brāhmi-Schrift bei den Uiguren.* Altorientalische Forschungen 11, 331-346.

# APPENDICES

# APPENDIX I

*The first attempt of Hitch to identify the new Tumshuquese letters after Hitch 1981*

The List of Twelve Symbols



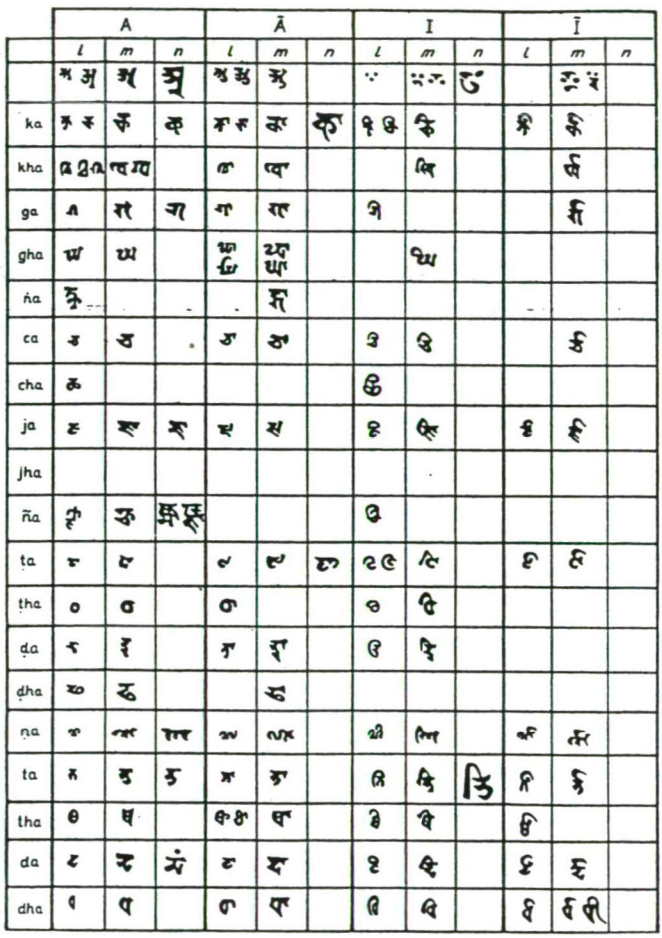

LEGEND: ø = Konow thought these were absent from the eight Maralbashi documents

() = Repeated signs

# APPENDIX II

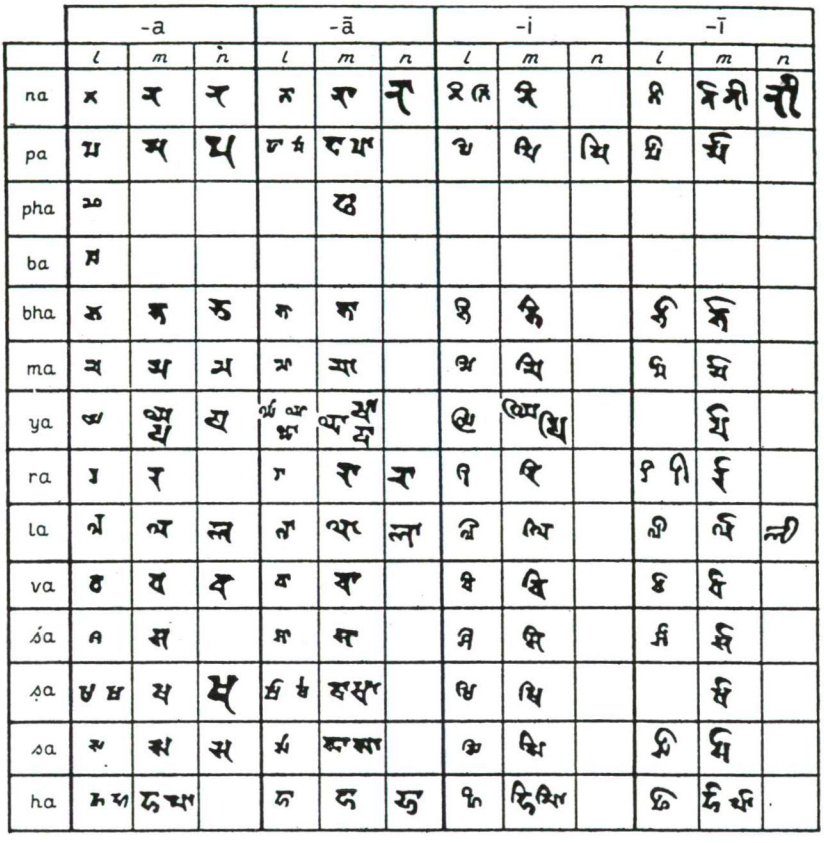
*The Indian alphabets from Sander 1968*

Indische Alphabete TAFEL 21



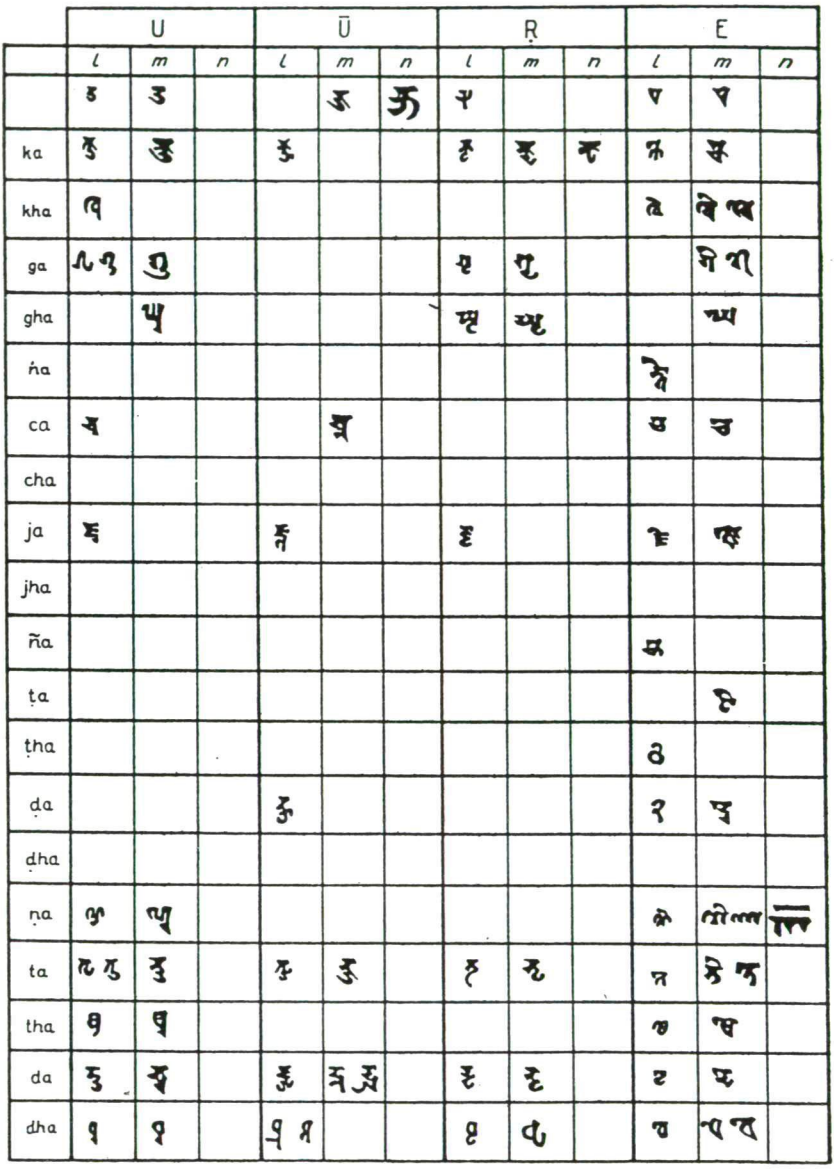

1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] I); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

TAFEL 22 Indische Alphabete




1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] I); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

Indische Alphabete TAFEL 23




1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] I); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

TAFEL 24 Indische Alphabete

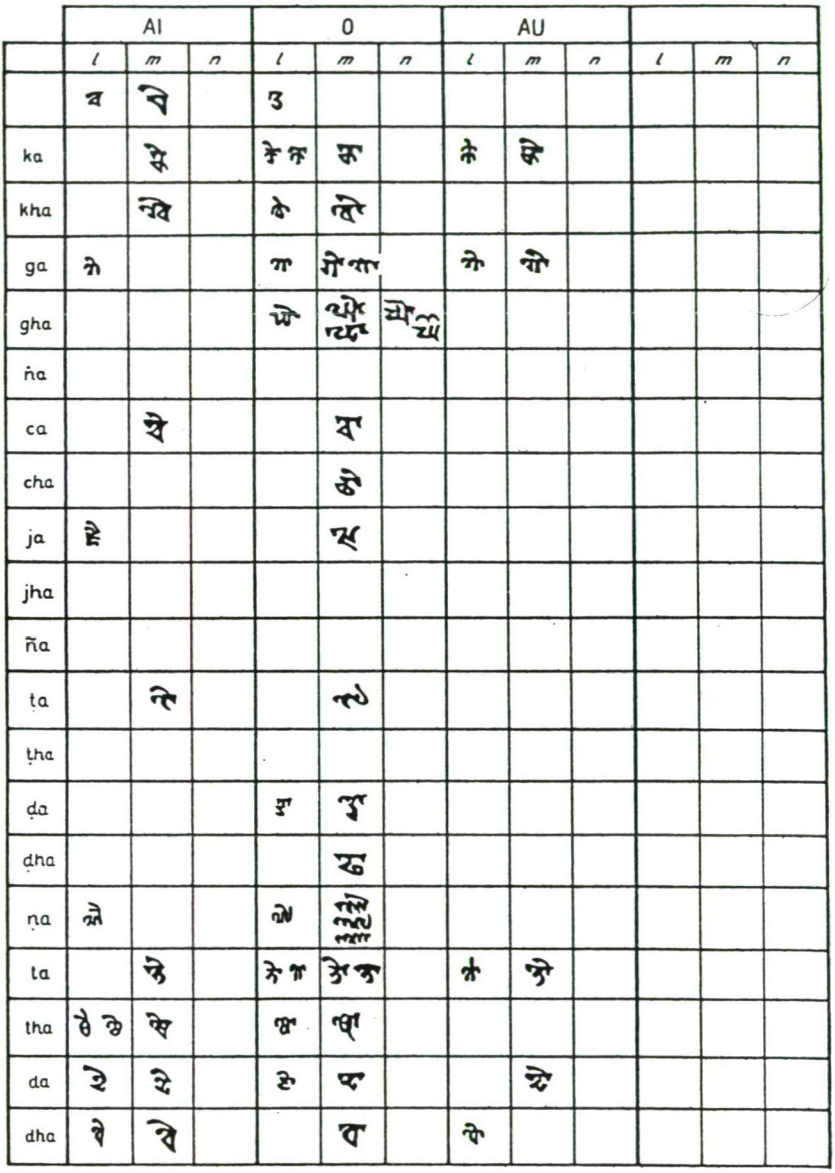
| | -u | | | -ū | | | -ṛ | | | -e | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *l* | *m* | *n* | *l* | *m* | *n* | *l* | *m* | *n* | *l* | *m* | *n* |
| *na* | | | | | | | | | | | | |
| *pa* | | | | | | | | | | | | |
| *pha* | | | | | | | | | | | | |
| *ba* | | | | | | | | | | | | |
| *bha* | | | | | | | | | | | | |
| *ma* | | | | | | | | | | | | |
| *ya* | | | | | | | | | | | | |
| *ra* | | | | | | | | | | | | |
| *la* | | | | | | | | | | | | |
| *va* | | | | | | | | | | | | |
| *śa* | | | | | | | | | | | | |
| *ṣa* | | | | | | | | | | | | |
| *sa* | | | | | | | | | | | | |
| *ha* | | | | | | | | | | | | |

Ligaturen

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *r-a* | *trā* | | *d-a* | *drā* | | *ndha* | *ndhā* | | *bdha* | *rdi* | |
| | *ttra* | *ttra* | *trya* | | | *dhmā* | *ndra* | *ndra* | | *rtha* | *rtha rthā* | *rtha* |
| | *ddhe* | *ddhi* | *ddhā* | | *dhya* | *dhya* | *pta* | *pta* | | *rya* | *rya* | |
| | *dya* | *dya* | | *nta* | *nta* | *nta* | *pra* | *pra* | *pra* | *rtra* | *rra* | |

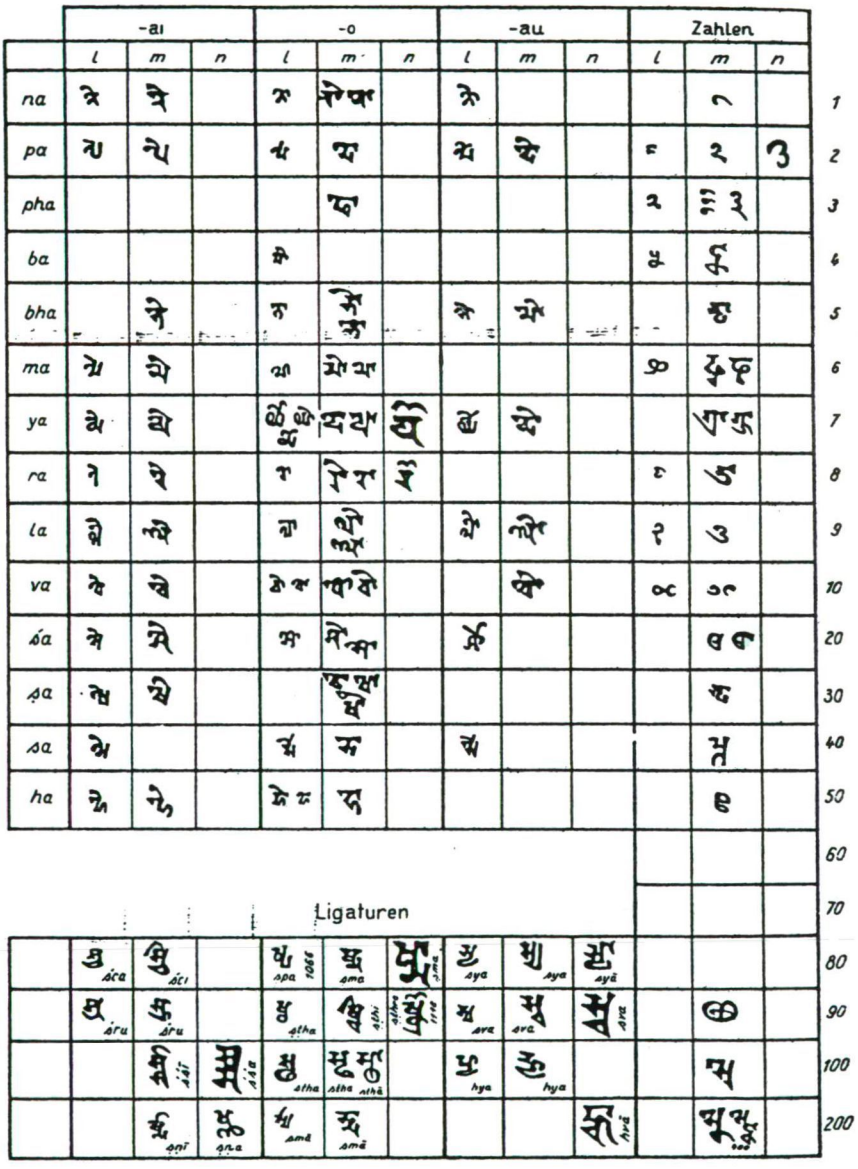
1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] I); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

Indische Alphabete TAFEL 25



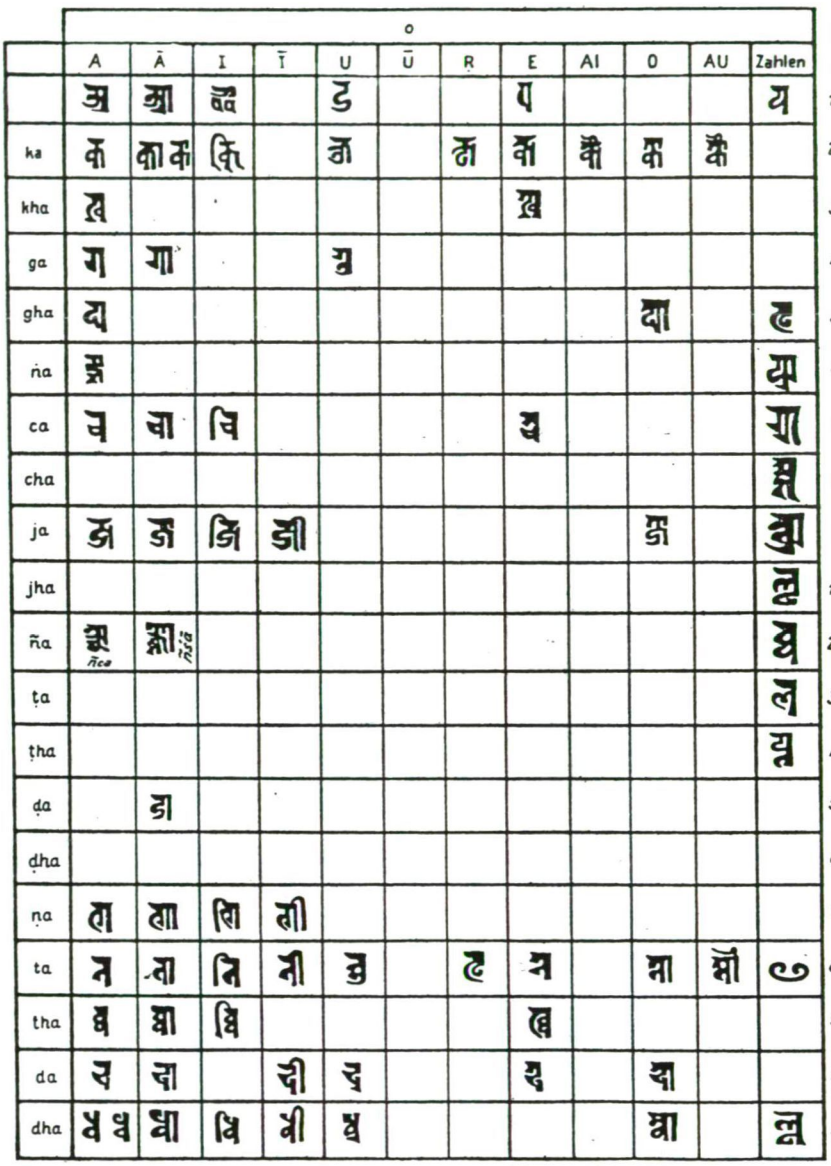

1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] I); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

TAFEL 26 Indische Alphabete




1. Spätes Gupta-Alphabet l (Schrifttypus II); ca. 6. Jh. n. Chr.
2. Gilgit/Bamiyan-Typ II, Alphabet m (Schrifttypus S [= Sonderschrift] 1); ca. vom 6. Jh. n. Chr. an.
3. Śāradā-Schrift, Alphabet n (Schrifttypus S II); ca. 13. Jh. n. Chr.

Indische Alphabete TAFEL 27



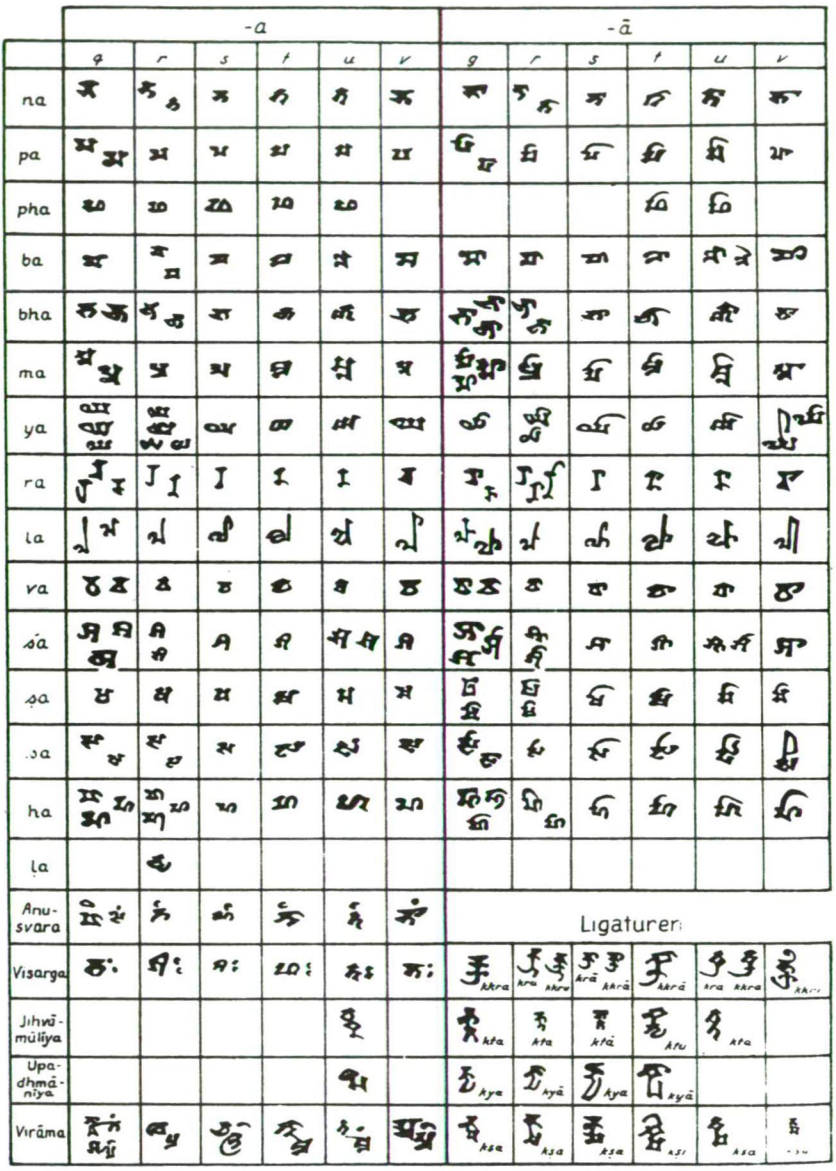

Pāla-Schrift, Alphabet o (Schrifttypus S [= Sonderschrift] III); ca. 12. Jh. n. Chr.

Turkistanische Alphabete TAFEL 29



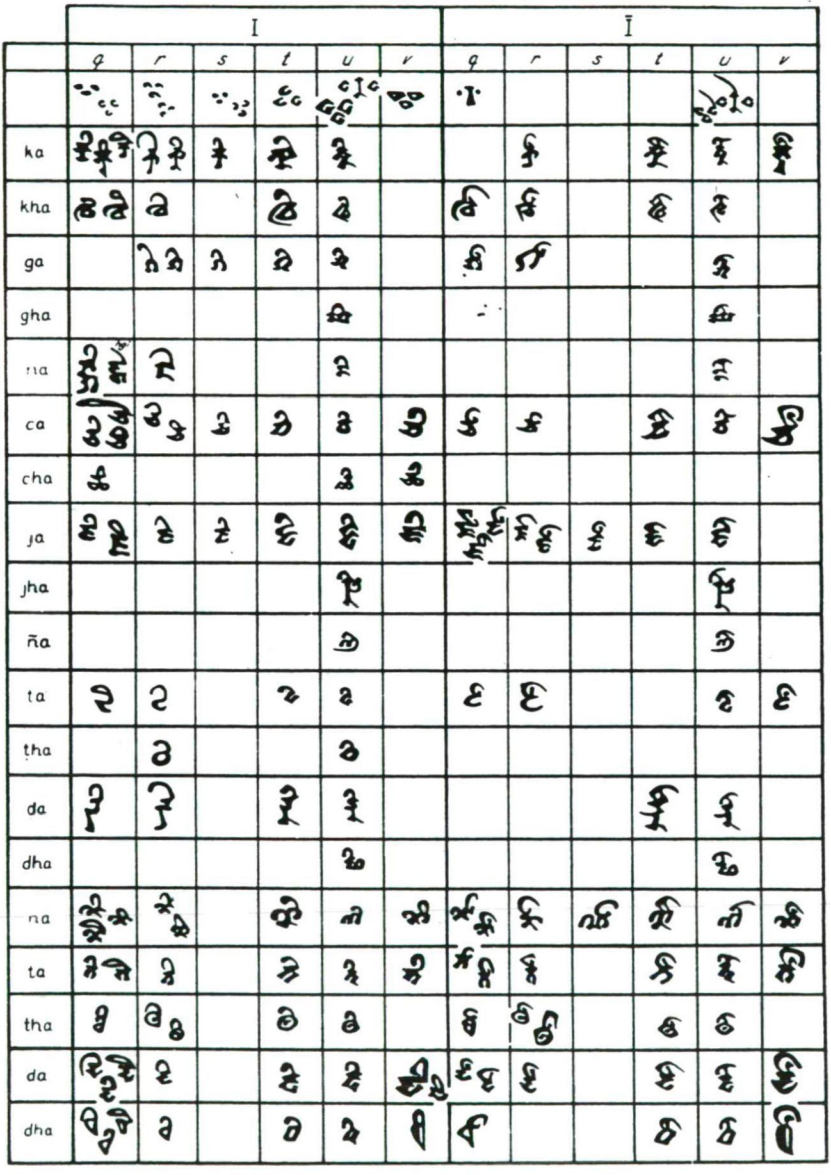

Turkistanischer Gupta Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmi (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmi, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmi, Typ b (Schrifttypus VI), Alphabet u Südturkistanische Brāhmi (Schrifttypus VII), Alphabet v

TAFEL 30 Turkistanische Alphabete



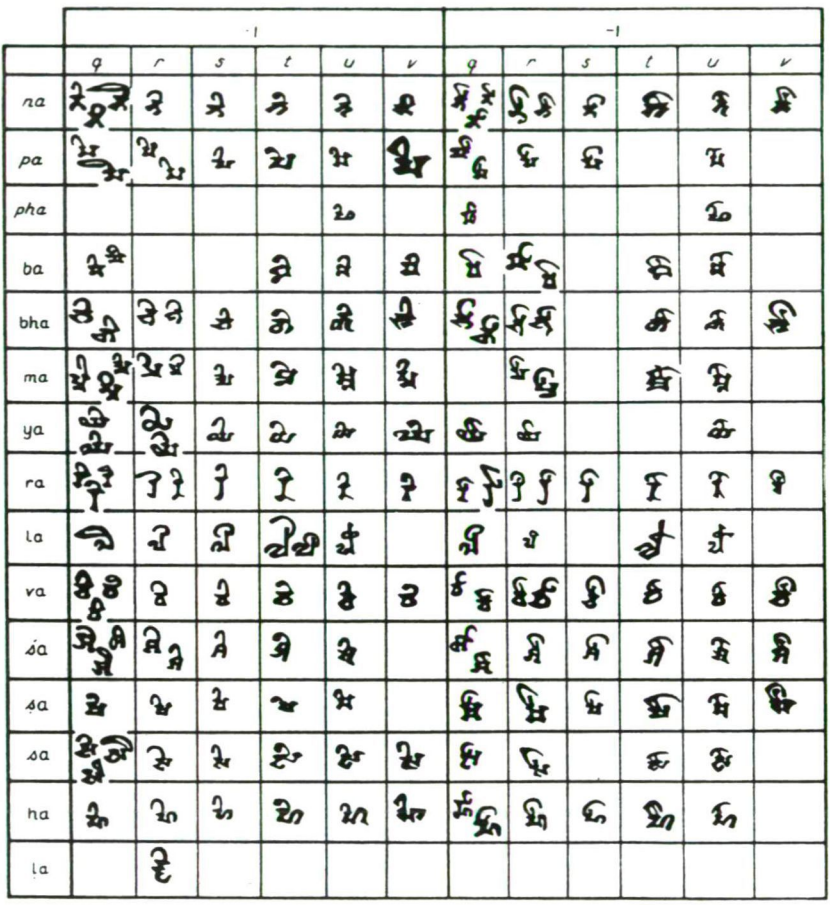

Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v

Turkistanische Alphabete TAFEL 31

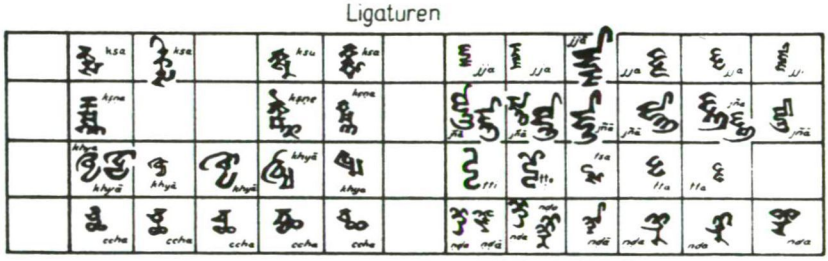



**Turkistanischer Gupta-Typ (Schrifttypus III). - Alphabet q. Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s. Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u. Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.**

TAFEL 32 Turkistanische Alphabete

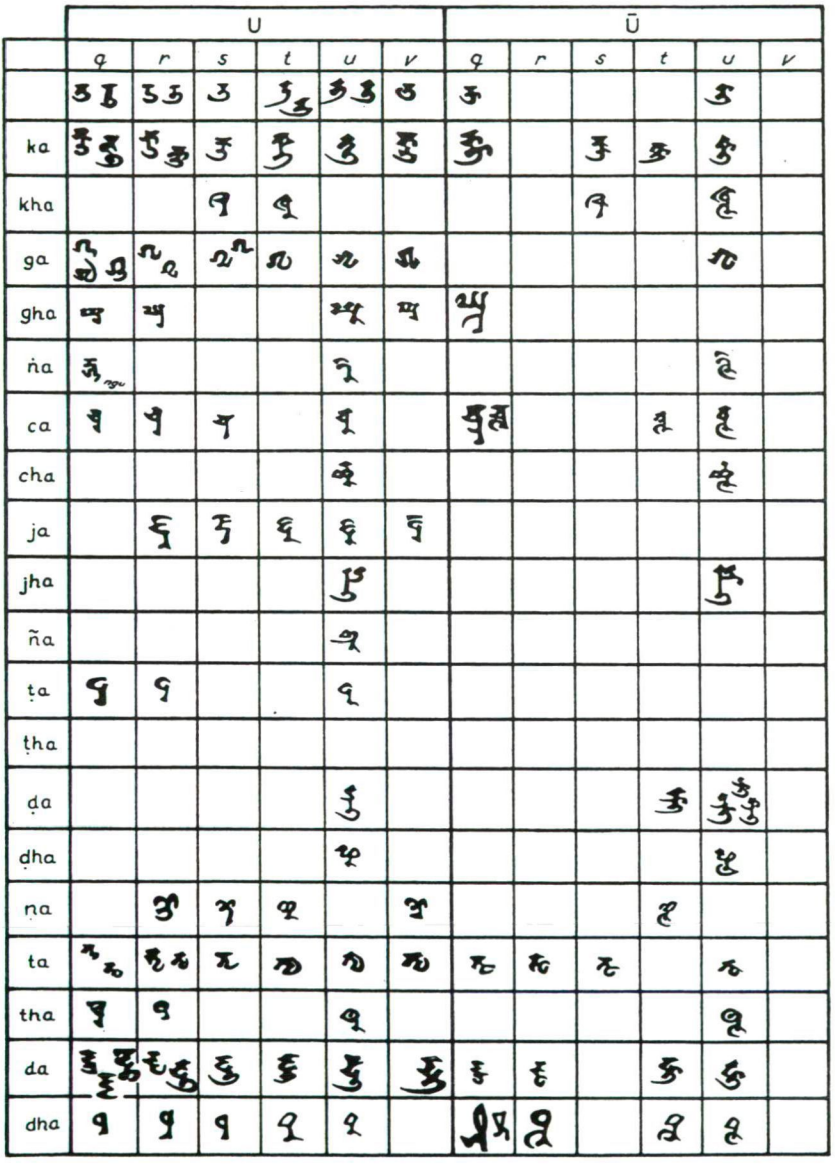



Ligaturen



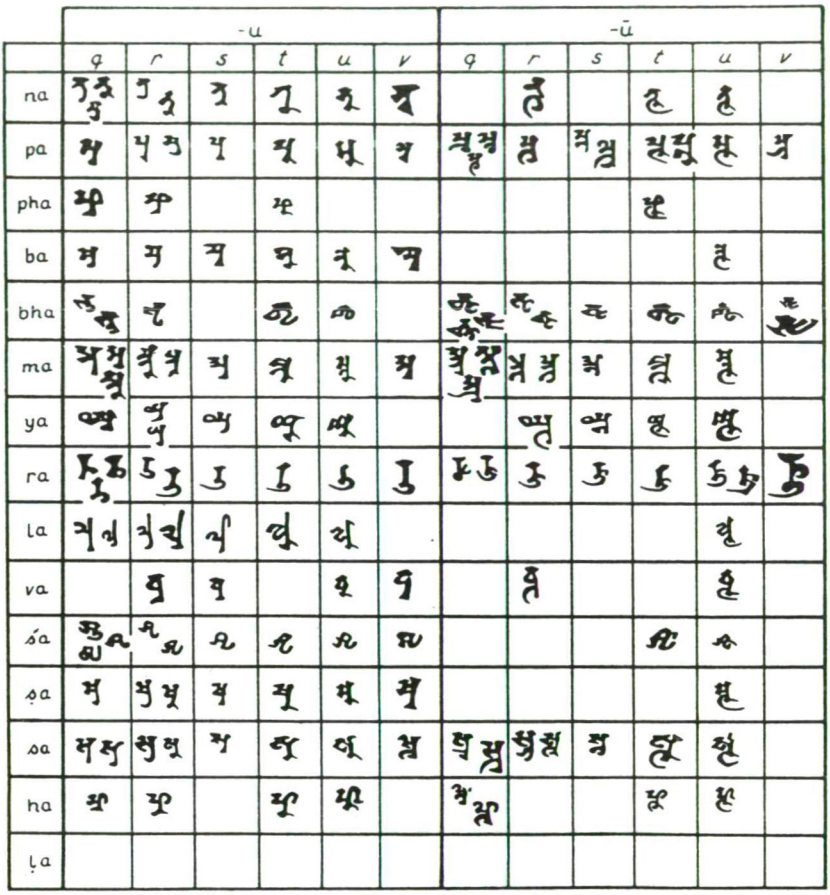
Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmi (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmi, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brâhmi, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmi (Schrifttypus VII), Alphabet v

Turkistanische Alphabete TAFEL 33




**Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI) Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.**

TAFEL 34 Turkistanische Alphabete

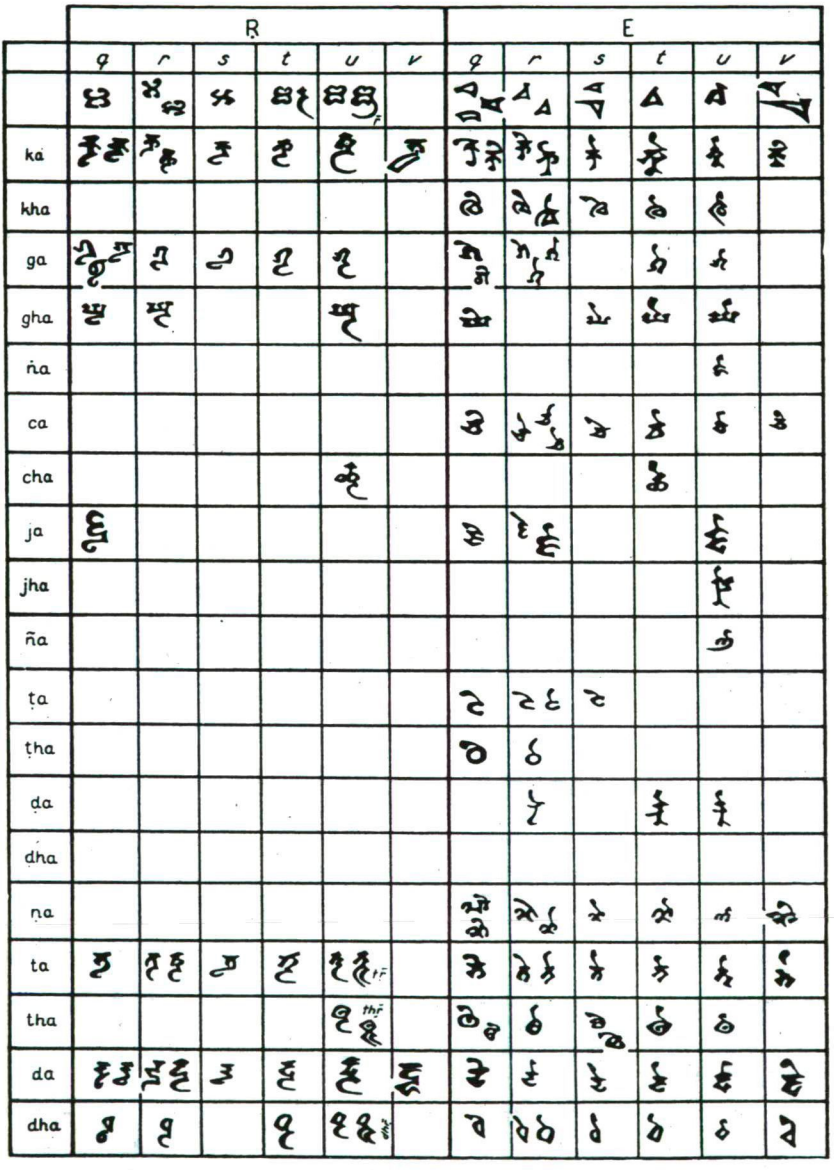



Ligaturen



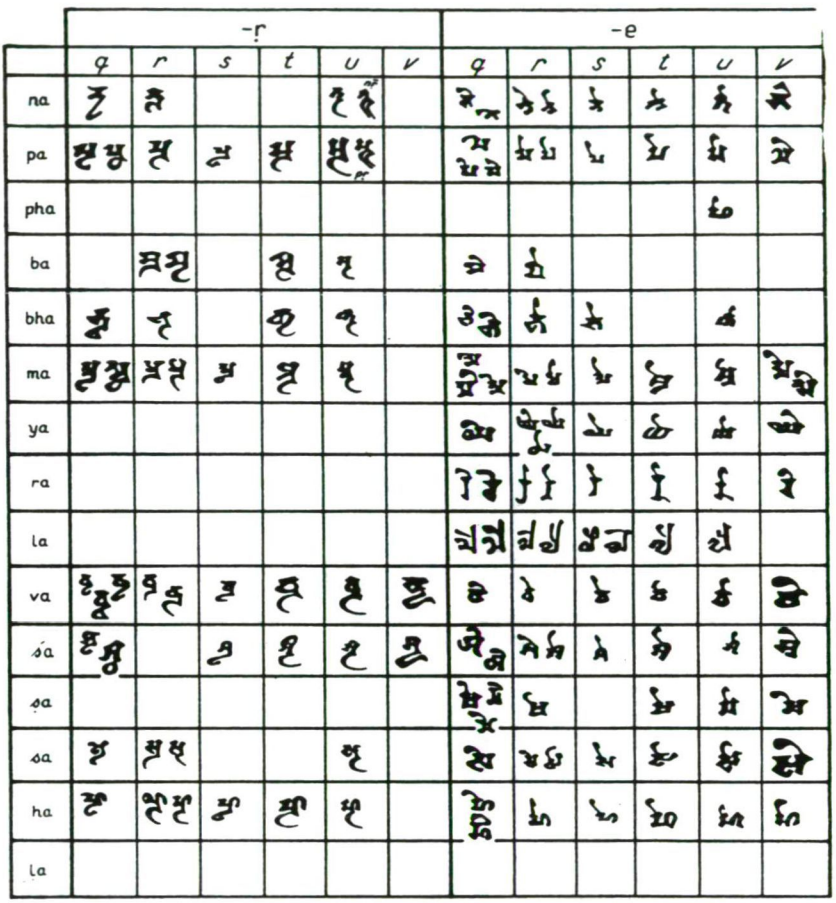
Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmi (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmi, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmi, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmi (Schrifttypus VII), Alphabet v.

## Turkistanische Alphabete TAFEL 35



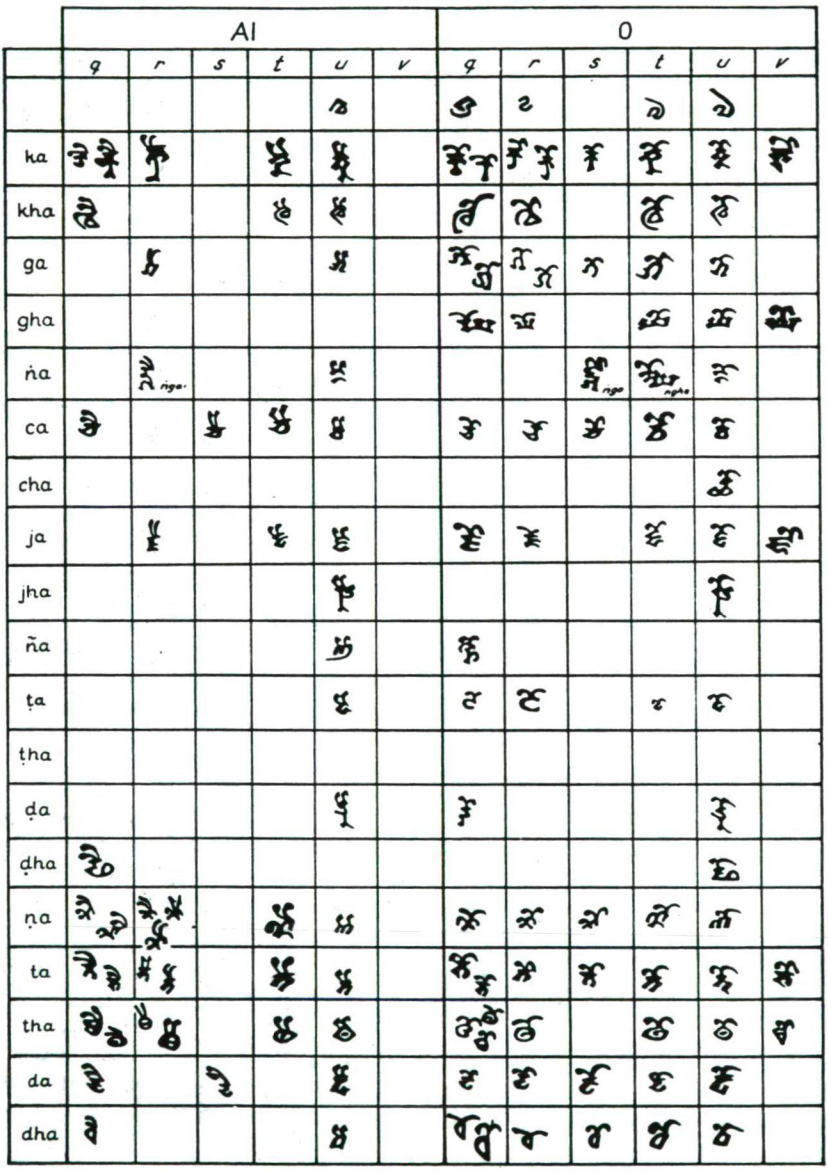

**Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.**

TAFEL 36 Turkistanische Alphabete




Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.

Turkistanische Alphabete TAFEL 37




**Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.**

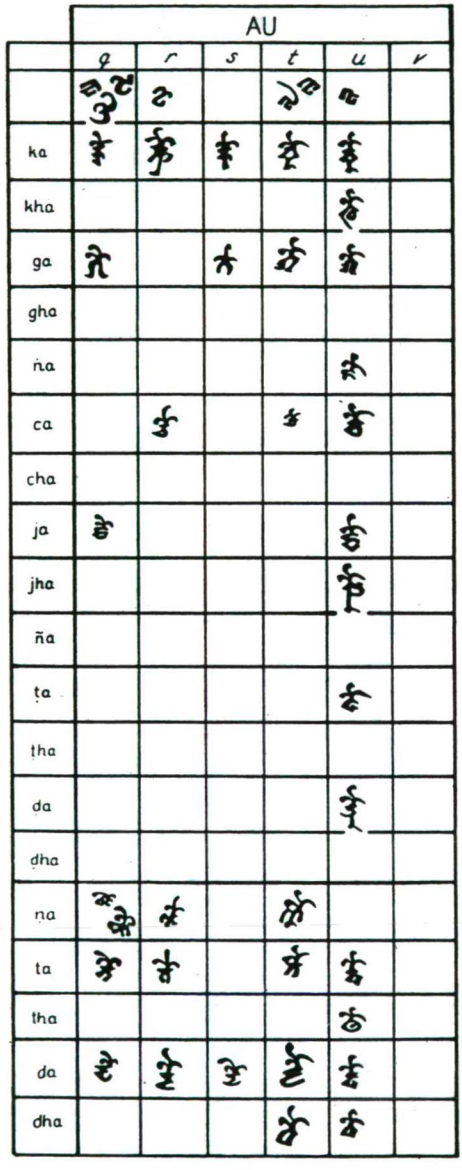
TAFEL 38 **Turkistanische Alphabete**




Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.

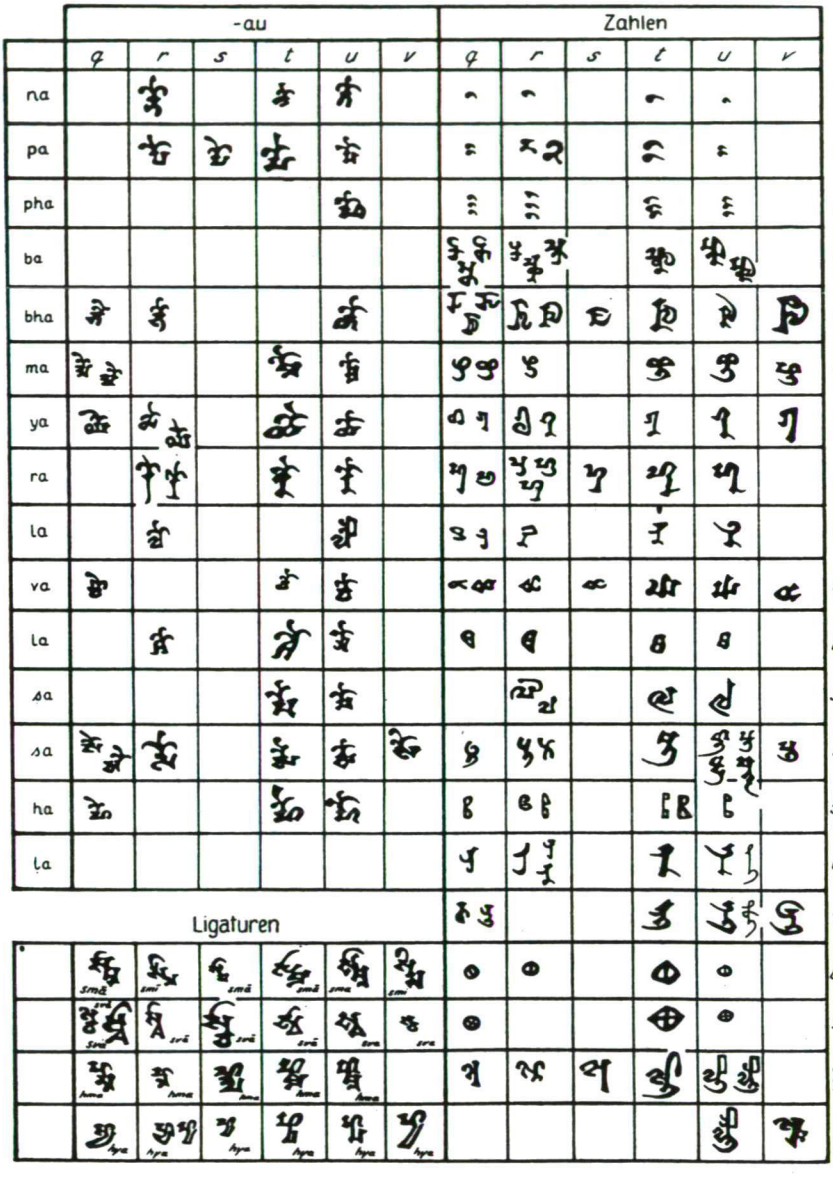
## Turkistanische Alphabete

TAFEL 39




Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmi (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmi, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmi, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmi (Schrifttypus VII), Alphabet v.

TAFEL 40 Turkistanische Alphabete




Turkistanischer Gupta-Typ (Schrifttypus III), Alphabet q; Frühe turkistanische Brāhmī (Schrifttypus IV), Alphabet r und s; Nordturkistanische Brāhmī, Typ a (Schrifttypus V), Alphabet t; Nordturkistanische Brāhmī, Typ b (Schrifttypus VI), Alphabet u; Südturkistanische Brāhmī (Schrifttypus VII), Alphabet v.

TAFEL 41

| Fremdzeichen | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Tocharisch* [1] |  ṭa |  ṭa(ṭha) |  ṇa |  ṗa |  ṃa |  ṛa |  ḷa |  ś́a |  ṣa |  ṣa |  wa |
| *Uigurisch* [2] |  -k | | |  -p |  -m |  -r |  -l |  -ś |  -ṣ |  -ṣ |  wa |
| *Uigurisch* |  ga |  ga |  ba |  dža (ǰ) |  za |  ža |  a |  u |  e |  ö |  ô |
| *Sakisch* [3] |  rə | | | | | | | | | | | |

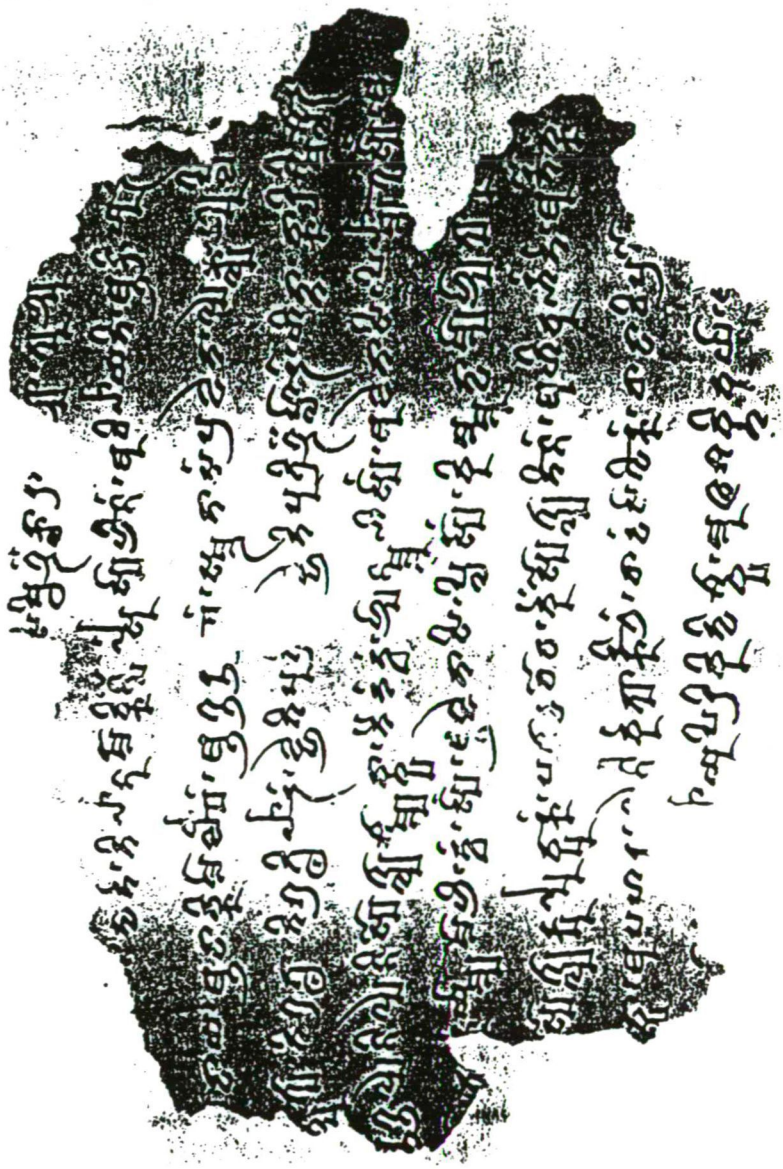
1 Die Fremdzeichen wurden J. FILLIOZAT, Fragments de Textes Koutchéens, 1948, entnommen. Vgl. auch E. SIEG/W. SIEGLING, Tocharische Sprachreste, Sprache B, Heft 2, 1953.

2 Die Fremdzeichen wurden A. v. GABAIN, Alttürkische Grammatik, 1950, entnommen. Aus der Gegenüberstellung mit den tocharischen Fremdzeichen wird deutlich, daß die Uiguren die Brāhmī von den tocharisch sprechenden Bewohnern der Turfan-Oase übernommen haben. Vgl. hierzu auch A. v. GABAIN, Die Schreiber der alttürkischen Brāhmī-Texte, Studia Orientalia Fennica, Vol. XXVIII, Heft 5, Helsinki 1964.

3 Die sakischen Fragmente weisen gegenüber den Sanskrithandschriften nur das Fremdzeichen ə auf. Vgl. M. LEUMANN, Sakische Handschriften, 1934.

# APPENDIX III

*Text G, lines 10-18 from Gabain 1954 (Tafel II)*

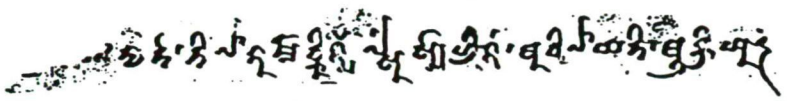


10



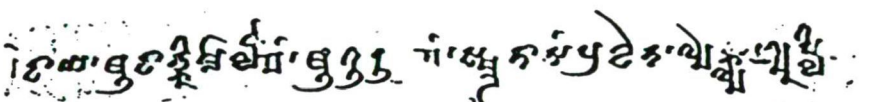
si ṣṭhi lr glā rā [ṁ]yā tyuṃ ñyu[ṁ]
nā tümñü[ŋ]

11



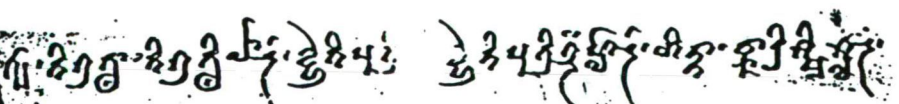
[ā] t tu · t te · ti l lā ṁ rmāqi p polsuṃ s syā ñiṅ · a dh dhi l lā y ya · ti t tiläṇur ā āgri g gṛiyur y yur

12



[ā] d dā y · a ā d dā q qi m mā l li p · a ā z zu r ru p · s ṣ m ṃ t tā ṡ p p j jā t t o oya g gl lya ñ ñyu p pli
ādā ya · ā dā qi mā lip · ā zu ru p smṛ tā saṃ pra jā ta oya glya ñyu pli
ādāqim ālip āzurup öglāñü pli [...]

13

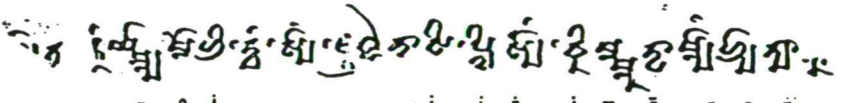


[u] z zā i y yu · n g th n · ā n · i n th i l lā r · i j ñ t i p t pu r · ā j ñ i t p u t i l lā r · i bh n nā · n · q qā i r ṛ i ś m ś lā r ṛ ·
[u]yuzyā · nigranthā · nigranthilāṛ jñātiputri[ka] jñāti putri ḷṛ hlār · bhinnā · qārriśmiś lāṛ
āzā · nigranthilār jñātiputriḷṛhlār · qārrišmišlāṛ ·

14 

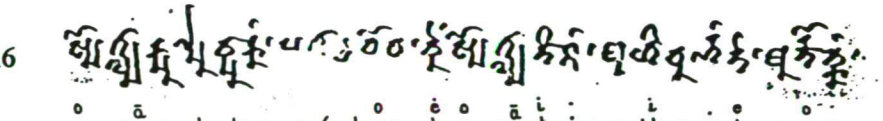
ā z o ṅ o ṅ s z ś k · ñ t ṭ · ñ ṅ ṇ s · a j n ā s · p l z s · a
ā u y y ly my ky v mū my y y

ā zu oya ṅi oya ṅi syo zlyā śmyā kkyā · ñaiṭaṃṭvaṃ · ñyāṅmū ṇi syāṃ · a jā nā si · pi lmya z syāṃ a
āzu ōṇi ōṇi sözlāšmāškkā ñāṇ mūṇi sän pilmäz sän

15 

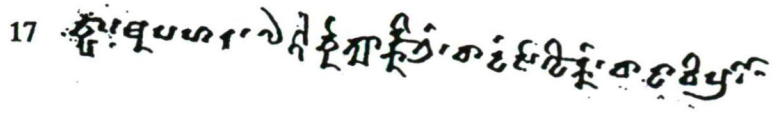
ā · t g y m ṅ · ṭ · s · a j n ā s · p s · k m n ṁ ñ t · g
my v y ā ly u d my ny ty

ā ymyā mo ñi tvaṃ · syaṃ · ā jā nā si · pli rsyaṃ · ki mmu ndā myām ñyā tyā · g
ymā moñi sän plirsän kim mundā mān ñātōg

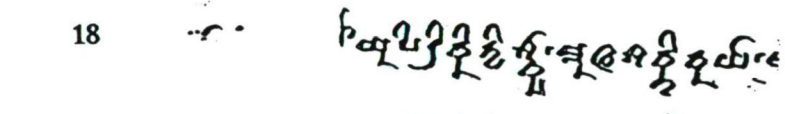
16 

s z g l k g · p ś d v c · k s z t ṅ · a bh c r t · a r q ·
y ly yu yu yu [c] y ly ṇ t t

syo zlyā gyu lyu kyu g · pas[c]ā[d q]voca · kem sya zlyā ti ṅ · a bhi cu rnaṃ te · a rto q
sözlāgülükü kem sözlātiŋ artoq

17 

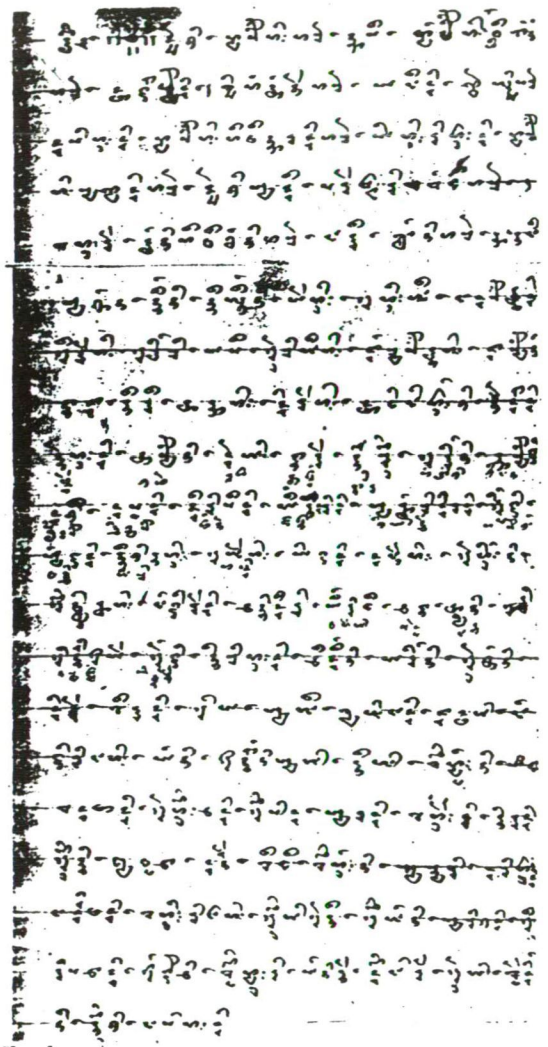
kyā·a pa ha ra o[y]a ṅi ke ṯyā rgi ḻ · vā daṃ sā wi g · vā dā a pra[mo]
kā ōṇi ketārgil sawig

18 

[l]āyu pi rri ki ndi śkyā· mu kha śa kti ka yā
... lāyu pirr ikindiśkā

# APPENDIX IV

*The Turkic Khotanese Wordlist from Bailey 1974*

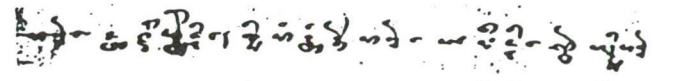


[Photo Bibliothèque Nationale, Paris

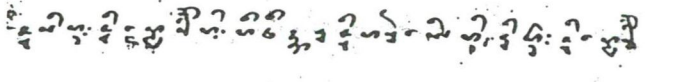
P 2892, 166 84

165 

kye śä hū lai hä : hame · ttu pī · hu lai hä : hī vī bana
keš tüpi

166 

hame kū rṅā lū kä byi ha¸ dū¸ nai hame ya šī kä ṇve sdyä hame
qurŋoluq yašiq

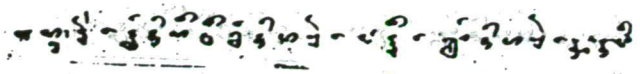
167 

ka pä ha : kä hu lai hä hī vī ttu ra kä ha mä · yi hi : ra hā : kä · hu lai
qapχaq yiɣirɣoq

168 

hä · mū nū kä ha me · kye śä yū kī · pa rai ṣkha rä · baṃ da hame
keš yüki

169



b ḫ:r · ḏ n h v th n h m · s ḏ b n h m · t t s

ba ha:rai · duᵢnä hīvī tham nä hame · sadī brrāᵢnä hame ttuttasī
bayarī sali tutasi

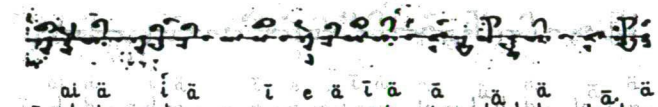
170



y ǵ n · t n · t ys n · y ḫ: a ḫ: ys · s k l ḏ k

yūguᵢna · ttīᵢnä · ttiysgīnä yaiḫä: aḫä·ysī · sakalä drrūkä
yügün tin tizgin yīy ayzi saqaldruq

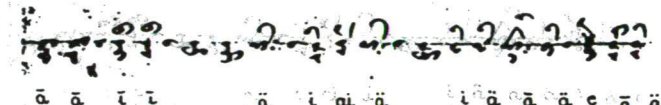
171



ī ḏ h · ḏ m · y p · ḍ m ys h · k m l ḏ h · k l n

ī ḏai hä · adrriᵢmä · yapī · emäysīhä · kämülädrrühä · kälünä
iliy adrim yapi emziy kömüldrüy qolun

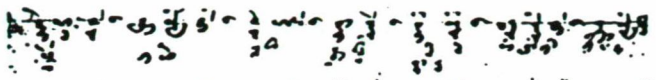
172



t k · t ḏ · ū t h: k r h · k d s ḫ: n t r k

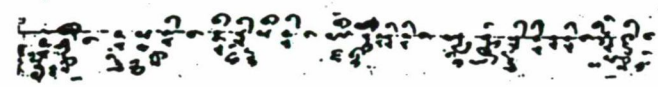
ttākä · ttīdī · uttu hä: ki rai hä · kūdisäḫā:nä tterkäkä
toqo tili utuy qiriy qudisyon terkök

173



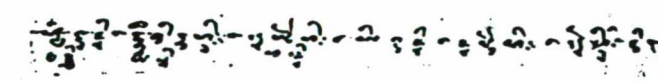
tte ḫa : kä · ülü₁nä · keysä · ttäḍai · ṯṯoṃau · aḍi₁nä ttulü₁nä
skai gahe rahä chaskaṃ ttära haṃdrräṃgä gïsana
teɣäk ulun kez töli tomo alin tulun

174



kaśï · kapakä · kïräpïkä · yïtti karäkä · yürümä kärakä ïṅä
brraukalai teiṃmülä häme jastä siyätceṃ halaśa
qašï qapaq kirpik yiti qaraq yürüm qaraq ïŋ

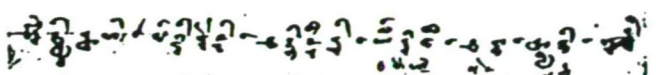
175



vtä ïä sä
yɛ m h k

yü₂ṅakä · ttïsättaḫä : aysaihä : yäṅakä kasai₂hä : ehi₁nä
vatcä ysïmä haksä
yuŋaq tištaɣ azïɣ yaŋaq qašïɣ eɣin

176

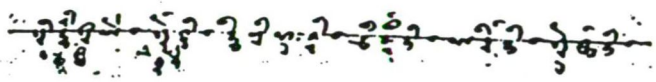


au ä ä äi a ä i ä ī ä ī ä ä
a v r h: ṡ ṅ r k · c ṯ k r · y ḍ d c ṯ n
ü u t t ü

ai ä e e
v h ys s k s b
y

au vüäruhä : saɻnäraikä cattäkirä yīdadī cattä üɣünä
vahai ysä sakye sabaɻ
ovruɣ saŋrïq čatqïr yildi cato ün

177

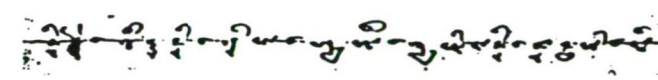


ä ä äi e ä ä ä ī ṯ ä ṯ ä e ä
ä r a y · a ṅ ṯ m ḫ: k c k n y r n aɻ ś n
a ṯ ṯ u

n m ś a ph ṯ n
ü th ṯ

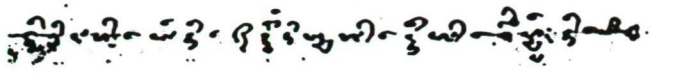
ä rttä äyai · ẹṅä ttämäḫa: kä cïkïɻnä yarïɻnä ẹśuɻnä
na mü sthaṃ phattanai
art oyi eŋ tamɣaq čikin yarin ešün

178



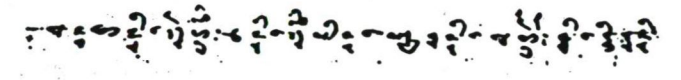
ä äi ī ī ä ä ä ä ä
k r b ḍ k a y · y ü ys · b ü h: s k · k k ü ys s
ä ä

kärai bïḍakä äya · yüysï · bühä: sakä · käküysä sä-
qarï biläq aya · yüzi · büɣsäk köküz si-

179 

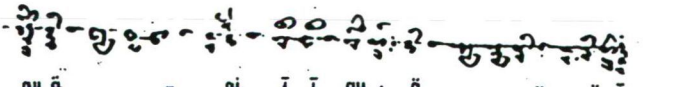
ä ä ä . ä ā ā ā ī ā au . ā ā
ṅ r s h: y n · a ṛ n y y t ys b ḥ: n y d
ṭ u s ṭ u

ṅäräsahä: yaᵢnä ärtläᵢnä yuysä tlī ysä bauhuᵢ: nä yädа
ŋirsaɣ yan orton yüz tiz boɣun yoda

180 

ā e au ā au ä ai ä i ä
b k ñ k · a ḥ: c k · a p k · y ṛ k · b ḥ: r · tt ḍ k
u ū

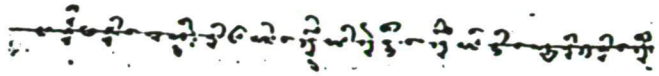
baqa ñakä · e ḥau: cakä · aupäka · yürakä · baḥaiᵢ: rä ttiḍakä
baqañaq eɣöčäk öpkä yüräk baɣïr tilaq

181 

au ä ai ī ī au . ä ā
ọ ṭ · s v c ā · k ṛ · b d · b ḥ: n · y ṃ r ä k r ä ḥ ā: ä
ṭ ū v n u u k

ọuttä · süɣ cä · karnai · bïdï · bauhaᵢ: nä · yu mu rä karäḥä: kä
öt suvčo qarnï bïti böɣän yumur kärɣök

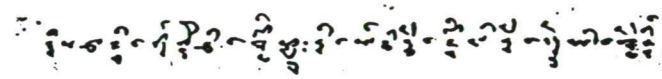
182



s r ñ k ä · b h : r s h : · ą ys a t · a ys n · y ą r g k a

sarka₁ñakä · bahä : räsähä : ąuysä etti au₁ysa₁nä · yürägakäï
sarqañaq bayïrsoy öz eti özän yürgäk ï-

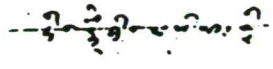
183



ḍ p c k · a ṅ c · b h : r · y n t · k s r · ą y b k

ḍäpacäkä · ä₁ṅacä · bau₁hü : rä · ya₁nätlai kausärai ęyä bai ki
lpačak öŋöč böɣür yän eti kösri ey bïqïn

184



n t ś · s p h : k

nä tlau₁śä · sapäha : kä
-n töš sapɣaq

*The quiver and the bowcase*

(165) 1. ***keš*** **<kyeśā>** 'quiver'. K *keš*.
2. ***tūpi*** **<ttupi>** 'its bottom'. K *tüp*.
(166) 3. ***qur oluq*** **<kūrnālūkā>** 'bowcase'. K *quruγluq* also *keš q*. 'quiver and bowcase'
4. ***yasiq*** **<yasikā>** 'bow case'. K *yasïq* 'Türk dialect. The Oguz and the Qifčaq do not know this word, they call it *qurman*'.
(167) 5. ***qapγaq*** **<kapāḫa:kā>** 'cover of the quiver'. K *qapγaq*.
6. ***yiγirγoq*** **<yiḫi:rāḫā:kā>** 'something which draws together the quiver'.Cf. K *yiγril-* 'to draw together (man his arms from cold, garment shrunk from being washed)'.
(168) 7. ***keš yūki*** **<kyeśā yūki>** 'the binding or fastening of the quiver'. Cf. Kirg. *džük* 'blankets or pillows used for piling against the door', K *ük-* 'to pile up', *ükäk* 'box'.

*The bow*

(169) 8. ***baγari*** **<baḫa:rai>** 'its handgrip'. K *yā baγri* 'the middle of the bow'.
9. ***sali*** **<sadi>** 'its flat side (of the bow's handgrip)'. K *sāl* 'raft'.
10. ***tutasi*** **<ttuttasi>** 'its handle'. K *tut-* 'to catch'.

*The harness*

(170) 11. ***yūgün*** **<yūgu,na>** 'bridle'. K *yügün*.
12. ***tin*** **<tti,nā>** 'halter'. K *tin*.
13. ***tizgin*** **<ttisgi,nā>** 'reins'. K *tin tizgin* 'halter and bridle'.
14. ***yiγ*** **<yaiḫā:>** 'horse bit'. K *yügün yigi* 'bit of the bridle' (DK read yik), to **yig-*'press', *yigi* 'dense'.
15. ***aγzi*** **<aḫā:ysi>** 'its opening hole(mouth)'. K *aγzi* to *aγiz*.
16. ***saqaldruq*** **<sakalādrrūkā>** 'the throatstrap of a headstall'. K *saqalduruq* 'name for the thread woven from silk which is attached to caps in order to fasten the cap underneath the chin so it does not fall off' cf. further 5D 1137.
(171) 17. ***iliγ*** **<idaihā:>** 'attachment (on the harness)'. Uigh *ilig tutuγ* 'attachment and ties', Kirg *ilik* 'zacepka'.

*The saddle*

18. ***adrim*** <adrri,mā> 'saddle pad'. Perhaps to K *adir-* 'to separate (? horseback and saddle)'.
19. ***yapi*** <yapi> 'horsblanket'. K *yapi* 'saddlecloth, Čigil dialect.'
20. ***emziγ*** <ęmāysihā:> 'the extremity of a saddle-bow'. Cf. K *emzüg* 'tip of a saddle-tree on front or back' (DK I 134 add 'first U altered from original A(?)'.
21. ***kömüldrüγ*** <kā,mūlādrrūhā:> 'breast strap'. K *kömüldürük* 'the breast-girth on the saddle'.
22. ***qolun*** <kalunā> '(saddle) girth'. K *qolan* 'saddle girth'.

(172) 23. ***toqo*** <ttāka> 'belt buckle (with a tongue)'. K *toqu* 'belt buckle' and CC toγa 'Schnalle; Ring am Zaumzeug'.

24. ***tili*** <ttidi> 'its (the buckle's) tongue'. K *til*.
25. ***utuγ*** <ūttuhā:> 'a broad strap on the left side of the saddle'. Cf. K *utγun* 'a broad strap on the saddle from the left side; the ring of the girth is attached to it and fastened with a tongue(!)'.
26. ***qiriγ*** <kiraihā:> 'selvage (of the saddle)'. K *qirγaq* 'the selvages of a garment' to *qiriγ* 'edge, border'.
27. ***qudisγon*** <kūdisāḫā,:nā> 'crupper strap'. K *quduzγun* (so DK I 13, until now misread), Kirg *kuzuskan* (Almássy 1901 438). Osm. *kuskun* < **quyuzqun* < **quduzqun*.
28. ***terkök*** <tterkakā> 'saddle strap'. Cf. K *terkü* [*terkö*?] (DK read *tergü*) 'saddle strap'.

*The arrow*

(173) 29. ***teγāk*** <tteḫa:kā> 'incison on arrow'. Teleut *täk* 'der Kerb am Ende des Pfeils an den man ihn gegen die Sehne drückt'.

30. ***ulun*** <ūlū,nā> 'arrow shaft'. K *ulun*.
31. ***kez*** <keysā> 'arrow notch'. K *käz*.
32. ***töli*** <ttādai> 'its progeny?'. K *töl*, (or read *toli* 'a part of the arrow?').
33. ***tomo*** <ţţomąu> 'the round top of the (arrow) head(?)'. Cf. Kirg *tomo* 'komel' roga; the root of the horn' to **tom* 'round thing', K *tomur-* 'to cut rounded', Tara *tom* 'walzenrund, rund', Chag *tomalaq* 'round', Khaz *tomar* 'stump of a tree, block, thick log' etc.

*The horse*

34. *alin* <adi,nā> 'forehead'. K *alin*.
35. *tulun* <ttulū,nā> 'temples'. K *tulun* 'temple (of the head)'.
(174) 36. *qaši* <kaši> 'its brim, eyebrow'. K *qaš*.
37. *qapaq* <kapakā> 'eyelid'. K *qapaq*.
38. *kirpik* <kirāpikā> 'eyelash'. K *kirpik*.
39. *yit qaraq* <yi,ttā karākā> 'pupil of the eye'. K *üt* 'but the Oguz say *üd*'.
40. *yürüm qaraq* <yüru,mā kārakā> 'the white of the eye '. K *ürün qaraq*.
41. *iŋ* <i,ṅā> 'slime'. K *yin*.
(175) 42. *yuŋaq* <yū,ṅakā> 'small hair'. K *yun*.
43. *tiš tay* <ttišāttahā:> 'molar tooth, crown of the tooth'. K *tiš, tay*.
44. *azīy* <aysaihā:> 'canine tooth'. K *azïy*.
45. *yaŋaq* <yāṅakā> 'cheekbone'. K *yaŋaq*.
46. *qasīy* <kaṣaihā:> 'inside of the cheeks'. K *qasiy* 'id., jaws'.
47. *eyin* <eḫi,:nā> 'shoulder'. K *ägin*.
(176) 48. *ovuruy* <auyūārūhā:> 'vertebra'. K *owruy* 'first vertebra of the neck, the more correct form ... is *oyruy*' (without specifying the dialect).
49. *saŋrīq* <sa,ṅāraikā> 'rump'. K *sayri* 'hide', Osm. 'rump'.
50. *čatqūr* <cattākirā> '?'
51. *yildi* <yi,dādi> 'descended?'
52. *čato* <cattā> 'ladder?'. Uig. *šatu*.
53. *ūn* <ūvūnā> 'rise!?'
(177) 54. *art oyi* <ārttā āyai> 'the hollow of the nape'. K *art* 'nape', *oy* 'ravine'.
55. *eŋ* <ęṅā> 'cheek'. K *äŋ*.
56. *tamyaq* <ttāmāha:kā> 'throat'. K *tamyaq* 'so the Turks, the Oguz and Kipchak say *tamaq* 'throat, larynx", Turki *tamyaq* 'palate'.
57. *čikin* <cikī,nā> 'upper shoulder'. Osm *čigin* (Redhouse).
58. *yarin* <yari,nā> 'shoulder blade'. K *yarin*.
59. *ešūn* <esu,nā> 'the point of the shoulder'. K *ösün* 'shoulderblade'.
(178) 60. *qarī* <kārai> 'foreleg, upper arm'. K *qarï*.
61. *bilāk* <bidakā> 'wrist, forearm'. K *biläk*.
62. *aya* <āya> 'palm of the hand (?)'. K *aya*.
63. *yūzi* <yūysi> 'its (sur)face'. K *yüz*.
64. *būysāk* <būhā:sakā> 'upper chest'. K *bügsäk* (DK:*böksek*).

65. *kökūz* <kākuysā> 'breast'. K *köküz* (DK *kögüz*).
(179) 66. *sa arsaγ* <sa,ṅārāsahā:> 'the hindquarters of a horse where a second rider sits'. K *sinarsuq*.
67. *yan* <ya,nā> 'socket or tips of the hipbone'. K *yan* (only this meaning, not 'side').
68. *orton yūz* <ārttā,nā yūysā> 'the middle (sur)face'. K *ortu, yüz*.
69. *tiz* <ttiysā> 'knee'. K *tiz*.
70. *boγun* <bauḫū,:nā> 'knuckle'. K *boγun, boγum*, (acc. to DK a later hand corrected on K 201 the *kesre* to *damma*, i.e *boγin* to *boγun* (more likely *vice versa*). [Read *bögün* ?]).
71. *yoda* <yāda> 'thigh'. Yellow Uighur, New Uighur *yota*, Turki *yote, yöte*.
(180) 72. *baqañoq* <bakañākā> 'the frog in the horse's hoof'. K *BAQ'Y(N)AQ* 'the space between the two sides of a cloven hoof or one of the two sides of a cloven hoof', *BAQA'Y(N)UQ* 'the frog of a horse's hoof' *Y* has two dots beneath and one above. Cf. No. 85.
73. *eγöčäk* <eḫau,:cakā> '?' ? to K *egin* 'shoulder', cf. 47.
74. *öpkā* <aupāka> 'lung'. K *öpkä, öwkä*.
75. *yürāk* <yūrakā> 'heart'. K *yürük*.
76. *baγūr* <baḫai:rā> 'liver'. K *baγur*.
77. *tilaq* <ttidakā> 'clitoris'. K *tilaq*.
(181) 78. *öt* <ąuttā> 'gall bladder'. K *öt*.
79. *suv(a)čo(?)* <sūyacā> '?'
80. *qarnī* <karnai> 'its belly'. K *qarin*.
81. *bidi* <bidi> 'its face(?)'. Uigh. *bet*, or cf. Turki, Khot *biǰek, biǰik* 'nipple, female breast' ?
82. *bögän* <bauḫa,:nā> 'caecum'. K *bögün* (so DK, while, Cl read *büken*).
83. *yumur* <yūmūrā> 'stomach'. K *yumur* 'caecum of an animal'.
84. *kärγök* <karāḫā:kā> 'paunch (?)'. K *kärgük* 'something like the third stomach (sheep)'. [Read *qarγoq*?].
(182) 85. *sarqañaq* <sarka,ñakā> 'third stomach'. K *SAR.QIY(N)Q* 'third stomach of a ruminant, the form with *nun* is a variant, as in Arabic *mizab, minzab, misar, minsar*', cf. 72. *baqañoq* above.
86. *baγirsoγ* <bahā:rāsahā:> 'entrails'. K *baγirsuq*.
87. *öz eti* <ąuysā etti> 'internal flesh(?)'. K *üz* is 'grease', but *öz* 'the heart and what is inside the belly'.
88. *özän* <auysa,nā> 'internal part'. Uigh *özen*.
89. *yürgäk* <yūrāgakā> '?'
(183) 90. *ülpačoq/ilpäčök* <idāpacakā> '?'

91. *ō ōc* <a,ṅācā> 'larynx'. Uighur *ö üč* 'id.', Kirg. *ö göč* 'gullet'.
92. *bōγūr* <ƀauɣu:rā> 'kidney'. K *bögür.*
93. *yān eti* <ya,nāttai> 'flesh'. K *et yen* 'body flesh'.
94. *kösri* <kausārai> 'side of the chest'. K *kösri* (DK *küsri*) 'ribcage'.
95. *ey* <ɐyā> 'rib?'. Cf. K *eyegü* 'rib (of animal)'?
96. *bīqin* <ƀaiki,nā> 'hip'. K *biqin.*

(184) 97. *tōš* <ttąu,ṣā> 'sternum, breast bone'. K *tös.*

98. *sapγaq* <sapāha:kā> 'waist'. K *sapγaq.*
99. *yūk* see 7.
100. *qaraq* see 39, 40.
101. *taγ* see 43.
102. *oy* see 54.
103. *eti* see 86, 91.

# APPENDIX V

*The fragment of the Turkic Sekiz Yükmek*

The Uighur text of the *Sekiz Yükmek* was published by W.Bang in collaboration with A.von Gabain and G.R.Rachmati.This edition is based on several fragments, among others a Scroll from the Stein Collection in London (B. M. 8212: 104), two texts from Kyoto, many fragments kept in Berlin, and two pieces from Petersburg, published by Radlov. The publication of Bang appeared in the Türkische Turfantexte VI (SPAW 1934 93-192, repr. 1971).

The fragment in Tibetan script is now kept in West Berlin under the code number Mainz 329, its word stock has been published by Clauson (1962 97-100). The ten lines of the fragment are on the right side of a page torn in the middle. The text corresponds to the lines 142-162 of the one edited by Bang. There are some discrepancies and perhaps also misreadings in Bang's version.

Mainz 329 [1]. [..] **bo-di-si-byid-lar** [...]
*bodisiwidlar*
SY ed Bang [142]. bodistwlar

[2]. **'iyn-ca 'a-[ya]-yur 'a-gri-[rla -y]ur ta-pyi-nur 'u-du-nu[r]**
*inča ayayur aγirlayur tapinur udu-nur*
[144]. inca ayayur aγrilayurlar tapinur udu-nur

[3]. **-si-byid kyim gra-yu tyo-zun-lar 'og-li | [...]tyo-zun**
*(bodi)siwid kim qayu tözünlär oγli tözün[...]*
[146]. bodistw qayu tözünlär oγli tözünlär

[4]. **'eyr-tya-ṅu the-riṅ tyoz yil-diz no-mug 'ug-sar 'byi[...]**
*ärtiŋü teriŋ töz yiltiz nomug uqsar bi(lig)*
[148]. ärtingü tering töz yiltiz nomuγ [149]uqsar ol

[5]. **kyim 'ol tin-lig tur-gra-ru | 'byil-kya 'byi-lig-li[g][...]**
*kim ol tinliγ turqaru bilgä biliglig*
[150]. tisär ol tinliγ turqaru [151] bilgä bilig közün

[6]. **t[i]r 'ol yo-grug 'byi-lir | 'byil-kya 'byi-li[g][...]**
*(ti)tir ol yoquγ bilir bilgä bilig*

[152]. titir ol yoquγ [153] biligli bilgä bilig

[7]. "al-gro tyo-rog yog gru-rug kyor-sar "ol tin-lig[...]
*alqo törög yoq quruγ körsär ol tinliγ*
[154]. alqo törlügüg yoq quruγ körsär[155] ol tinliγ

[8]. kyorg "u-la'-di | "al-di tyor-lu[g] [ya]-gri-din grud-ru
*körk ulati alti törlüg yaγidin qutru*
[158]. körk ulati alti törlüg yaγilarta üni qurtulur

[9]. gru-rug [...]ir | "oyn kyog yi-ma "[...] "og "e[...]
*quruγ [tit]ir öngög yimä [ol] oq ä[rür]*
[158]. quruγ titir öng körk[159]yimä ol oq ärür

[10]. sa-[.]i[.] [...]inc byi-lig yi-ma "al-[...]"og[...]byil
*sa[qinč] [qil]inč bilig yimä al[qo] oq[miš] bi[l-]*
[161]. saqinč qilinč[162] bilig yükmäkig alqo insä bilmiš oqmiš



Készítette a JATEPress

# studia uralo-altaica

Attila József University, Szeged

No. 1. Róna-Tas A. – Fodor S.: *Epigraphica Bulgarica*. 1973.

No. 2. *Die erste sölkupische Fibel aus dem Jahre 1879*. Eingeleitet von P. Hajdú. 1973.

No. 3. Novickij, Gr.: *Kratkoe opisanie o narode ostjackom (1715)*. 1973.

No. 4. Paasonen, H.: *Tschuwaschisches Wörterverzeichnis*. Eingeleitet von A. Róna-Tas. 1974.

No. 5. A. Molnár, F.: *On the History of Word-Final Vowels in the Permian Languages*. 1974.

No. 6. Hajdú, P.: Samojedologische Schriften. 1975.

No. 7. *N. Witsens Berichte über die uralischen Völker*. Aus dem Niederländischen ins Deutsche übersetzt von T. Mikola. 1975.

No. 8. Ph. Johann von Strahlenberg: *Das Nord und Ostliche Theil von Europa und Asia*. (Stockholm, 1730.) With an introduction by I. R. Krueger. 1975.

No. 9. Kiss, J.: *Studien zur Wortbildung und Etymologie der finnisch-ugrischen Sprachen*. 1976.

No. 10. U. Kőhalmi, K.: *Chrestomatia Sibirica. Auswahl aus der Volksdichtung der sibirischen Urvölker*. 1977.

No. 11. Mikola, T.: *Materialen zur wotjakischen Etymologie*. 1977.

No. 12. Popova, Ja. N.: *Nenecko-russkij slovar' lesnoe narečie)*. 1978.

No. 13. Tardy, L.: *Beyond the Ottoman Empire*. 1978.

No. 14. Clauson, G.: *An Etymological Dictionary of Pre-Thirteenth-Century Turkish, Index Vol. I*. With a preface by A. Róna-Tas. 1981.

No. 15. Clauson, G.: *An Etymological Dictionary of Pre-Thirteenth-Century Turkish, Index Vol. II*. With a preface by A. Róna-Tas. 1982.

No. 16. Honti, L.: *Nordostjakisches Wörterverzeichnis*. 1982.

No. 17. *Studies in Chuvash Etymology I*. Edited by A. Róna-Tas. 1982.

No. 18. Domokos, P.: *Handbuch der uralischen Literaturen*. 1982.

No. 19. M. Korchmáros, V.: *Definiteness as Semantic Content and its Realization in Grammatical Form*. 1982.

No. 20. Tardy, L.: *Sklavenhandel in der Tarterei*. 1983.

No. 21. Berta, Á.: *Die russischen Lehnwörter der Mundart der getauften Tataren*. 1983.

No. 22. Helimsky, E.: *The Language of the First Selkup Books*. 1983.

No. 23. Pusztay, J.: *Die Pur-Mundart des Waldjurakischen. Grammatikalischer Abriss aufgrund der Materialien von T. V. Lehtisalo*. 1984.

No. 24. Radics, K.: *Typology and Historical Linguistics. Affixed Person-Marking Paradigms*. 1985.

No. 25. Róna-Tas, A.: *Language and History. Contributions to Comparative Altaistics*. 1986.

No. 26. Keresztes, L.: *Geschichte des mordwinischen Konsonantismus II. Etymologisches Belegmaterial*. 1986.

No. 27. Keresztes, L.: *Geschichte des mordwinischen Konsonantismus I*. 1987.

No. 28. *Papers on Derivation in Uralic. Szegeder und Turkuer Beiträge zur uralischen Derivation*. 1987.

No. 29. Schulze, B.: *Der Wortparallelismus als ein Stilmittel der (nord-)ostjakischen Volksdichtung*. 1988.

No. 30. *Tatarische etymologische Studien II*. 1988.

No. 31. Berta, Á.: *Lautgeschichte der tatarischen Dialekte*. 1989.

No. 32. Zimonyi, I.: *The Origins of the Volga Bulghars*. 1990.

No. 33. Róna-Tas, A.: *An Introduction to Turkology*. 1991.

ISSN 0133-4239